

**LAPORAN ANALISIS BUTIR SOAL
DAN INSTRUMEN ASESMEN PEMBELAJARAN**
(Aspek Pengetahuan, Sikap, dan Keterampilan)

Nama : Amaradina Fatia Sari
NPM : 2523031004
Mata Kuliah : Asesmen Pendidikan
Jumlah Peserta Uji Coba : 32 Peserta Didik (Kelas XI)

Uji coba instrumen dilakukan terhadap 32 peserta didik kelas XI. Ketiga instrumen (pengetahuan, sikap, keterampilan) diberikan kepada peserta didik yang sama sehingga hasilnya dapat pula dilihat keterkaitannya. Berikut disajikan metode analisis pada masing-masing instrumen.

A. Instrumen Pengetahuan

- Tingkat Kesukaran (P): proporsi peserta yang menjawab benar per butir. Kategori: Sukar ($P \leq 0,30$), Sedang ($0,31-0,70$), Mudah ($P > 0,70$).
- Daya Pembeda (D): selisih proporsi benar kelompok atas (27% skor tertinggi) dan kelompok bawah (27% skor terendah). Kategori: Buruk ($D < 0,20$), Cukup ($0,20-0,29$), Baik ($0,30-0,39$), Baik Sekali ($D \geq 0,40$).
- Validitas Butir: korelasi point-biserial antara skor butir dan skor total (tanpa butir tersebut). Kategori: Valid ($r \geq 0,30$), Valid Cukup ($0,20-0,29$), Tidak Valid ($r < 0,20$).
- Reliabilitas Tes: dihitung dengan rumus KR-20 untuk data dikotomis (benar/salah).

B. Instrumen Sikap

- Reliabilitas dihitung menggunakan koefisien Cronbach's Alpha berdasarkan skor 6 pernyataan penilaian diri skala 1–4.
- Konsistensi tiap pernyataan dilihat dari korelasi item-total terkoreksi (corrected item-total correlation).

C. Instrumen Keterampilan

- Reliabilitas dihitung sebagai reliabilitas antar-rater menggunakan dua penilai independen yang menilai hasil kerja 32 peserta didik yang sama pada 3 aspek rubrik.
- Kesepakatan dinyatakan dalam koefisien korelasi Pearson dan persentase kesepakatan skor mutlak antar-rater.

HASIL DAN PEMBAHASAN

A. Analisis Butir Soal Aspek Pengetahuan

Tabel berikut menyajikan hasil analisis tingkat kesukaran, daya pembeda, dan validitas untuk 20 butir soal pilihan ganda mata pelajaran Ekonomi.

No	P	Tingkat Kesukaran	D	Daya Pembeda	r	Validitas	Keputusan
1	0.53	Sedang	0.67	Baik Sekali	0.39	Valid	Diterima
2	0.75	Mudah	0.22	Cukup	0.16	Tidak Valid	Direvisi/Ditolak
3	0.44	Sedang	0.56	Baik Sekali	0.32	Valid	Diterima
4	0.62	Sedang	0.44	Baik Sekali	0.26	Valid (cukup)	Diterima
5	0.38	Sedang	0.44	Baik Sekali	0.31	Valid	Diterima
6	0.88	Mudah	0.33	Baik	0.31	Valid	Diterima
7	0.75	Mudah	0.11	Buruk	-0.05	Tidak Valid	Direvisi/Ditolak
8	0.38	Sedang	0.44	Baik Sekali	0.27	Valid (cukup)	Diterima
9	0.38	Sedang	0.11	Buruk	0.05	Tidak Valid	Direvisi/Ditolak
10	0.53	Sedang	0.56	Baik Sekali	0.33	Valid	Diterima
11	0.38	Sedang	0.44	Baik Sekali	0.27	Valid (cukup)	Diterima
12	0.62	Sedang	0.11	Buruk	-0.01	Tidak Valid	Direvisi/Ditolak
13	0.81	Mudah	0.33	Baik	0.11	Tidak Valid	Direvisi/Ditolak
14	0.56	Sedang	0.56	Baik Sekali	0.37	Valid	Diterima
15	0.69	Sedang	0.44	Baik Sekali	0.24	Valid (cukup)	Diterima
16	0.28	Sukar	0.44	Baik Sekali	0.32	Valid	Diterima
17	0.44	Sedang	0.89	Baik Sekali	0.51	Valid	Diterima
18	0.72	Mudah	0.33	Baik	0.04	Tidak Valid	Direvisi/Ditolak
19	0.38	Sedang	0.44	Baik Sekali	0.29	Valid (cukup)	Diterima
20	0.56	Sedang	0.33	Baik	0.27	Valid (cukup)	Diterima

Hasil perhitungan reliabilitas tes secara keseluruhan menggunakan rumus KR-20 diperoleh koefisien sebesar 0,67, yang menurut kriteria umum tergolong pada kategori reliabel dengan tingkat sedang hingga cukup baik untuk instrumen tes berjumlah 20 butir. Dari sisi tingkat kesukaran, distribusi butir soal menunjukkan 1 butir (5%) tergolong sukar, 14 butir (70%) tergolong sedang, dan 5 butir (25%) tergolong mudah. Proporsi ini relatif ideal karena didominasi oleh kategori sedang, meskipun proporsi butir sukar masih terlalu sedikit.

Dari 20 butir soal, sebanyak 14 butir (70%) dinyatakan diterima karena memiliki daya pembeda dan validitas yang memadai, sedangkan 6 butir (30%) yaitu butir nomor 2, 7, 9, 12, 13, dan 18 perlu direvisi atau dipertimbangkan untuk tidak digunakan kembali. Butir nomor 7, 9, dan 12 menunjukkan daya pembeda buruk ($D < 0,20$), yang mengindikasikan soal tersebut tidak mampu membedakan peserta berkemampuan tinggi dan rendah secara memadai kemungkinan disebabkan oleh kunci jawaban yang ambigu, kalimat soal yang membingungkan, atau materi yang belum sepenuhnya dikuasai peserta didik. Butir nomor 2, 13, dan 18 memiliki validitas rendah ($r < 0,20$) meskipun tingkat kesukarannya masih wajar, sehingga perlu ditinjau kembali kesesuaian konten soal dengan indikator yang hendak diukur.

Sebaliknya, butir nomor 17 memiliki daya pembeda dan validitas tertinggi di antara seluruh butir ($D = 0,89$; $r = 0,51$), sehingga sangat baik dipertahankan sebagai butir acuan kualitas soal pada penyusunan tes berikutnya. Butir-butir dengan kategori Baik Sekali pada daya pembeda ($D \geq 0,40$) meliputi nomor 1, 3, 4, 5, 8, 9, 10, 11, 14, 15, 16, 17, dan 19 — sebagian besar juga menunjukkan validitas yang baik, mengonfirmasi bahwa daya pembeda dan validitas cenderung berjalan searah pada instrumen ini.

Distribusi nilai akhir peserta didik pada tes pengetahuan disajikan pada tabel berikut.

Rentang Nilai / Kategori	Jumlah Siswa	Persentase
Kurang (<55)	11	34.4%
Cukup (55-69)	15	46.9%
Baik (70-84)	4	12.5%
Sangat Baik (85-100)	2	6.2%

Rata-rata nilai peserta didik pada tes pengetahuan adalah 55,3 dengan simpangan baku 17,4, nilai tertinggi 95,0 dan nilai terendah 20,0. Sebanyak 34,4% peserta didik masih berada pada kategori Kurang (nilai di bawah 55), sedangkan hanya 6,2% peserta didik yang mencapai kategori Sangat Baik. Kondisi ini mengindikasikan bahwa secara umum penguasaan materi peserta didik terhadap topik Ketenagakerjaan, Inflasi, dan Kebijakan Moneter-Fiskal masih perlu ditingkatkan, sejalan dengan temuan bahwa sejumlah butir soal terkait konsep-konsep tersebut (misalnya butir nomor 7 dan 9) memiliki daya pembeda rendah yang mengindikasikan kebingungan peserta didik pada materi yang diujikan.

B. Analisis Instrumen Aspek Sikap

Instrumen sikap yang dianalisis adalah lembar penilaian diri (self-assessment) yang terdiri atas 6 pernyataan, meliputi 3 pernyataan dimensi Kepedulian Sosial dan 3 pernyataan dimensi Kepedulian Lingkungan, dengan skala 1–4 (Tidak Pernah–Selalu). Hasil analisis konsistensi internal disajikan pada tabel berikut.

No	Indikator Pernyataan	Rerata	r item-total	Kategori
1	Empati terhadap musibah	2.47	0.68	Baik
2	Membantu teman kesulitan belajar	3.28	0.59	Baik
3	Menghargai perbedaan status sosial	2.34	0.74	Baik
4	Membuang sampah pada tempatnya	2.84	0.59	Baik
5	Menghemat energi	2.53	0.79	Baik
6	Menghabiskan makanan agar tidak mubazir	3.19	0.70	Baik

Hasil perhitungan reliabilitas instrumen menggunakan koefisien Cronbach's Alpha diperoleh nilai 0,874, yang tergolong sangat tinggi/reliabel. Hal ini berarti keenam pernyataan dalam instrumen penilaian diri secara konsisten mengukur konstruk sikap kepedulian sosial dan kepedulian lingkungan pada peserta didik. Seluruh pernyataan memiliki korelasi item-total di atas 0,30 (berkisar 0,59 sampai 0,79), yang menandakan setiap pernyataan layak dipertahankan tanpa perlu revisi.

Pernyataan nomor 5 mengenai kebiasaan menghemat energi memiliki daya diskriminasi tertinggi ($r = 0,79$), yang berarti pernyataan ini paling representatif dalam membedakan peserta didik yang memiliki kepedulian lingkungan tinggi dan rendah. Sementara itu, pernyataan nomor 3 mengenai sikap menghargai perbedaan status sosial memiliki rerata terendah (2,34), yang mengindikasikan bahwa sikap ini relatif belum sepenuhnya membudaya pada sebagian peserta didik dibandingkan indikator sikap lainnya.

Ditinjau dari validitas isi, indikator-indikator pada instrumen (menunjukkan empati, membantu sesama, menghargai perbedaan, menjaga kebersihan, penghematan energi, dan pengurangan sampah)

telah selaras dengan rubrik penilaian sikap dan lembar observasi yang disusun sebelumnya, sehingga validitas isi instrumen ini dapat dikategorikan baik.

Distribusi kategori sikap peserta didik menunjukkan bahwa 6 peserta didik (18,8%) berada pada kategori Sangat Baik, 17 peserta didik (53,1%) pada kategori Baik, 8 peserta didik (25,0%) pada kategori Cukup, dan hanya 1 peserta didik (3,1%) pada kategori Kurang. Secara umum, sikap kepedulian sosial dan lingkungan peserta didik kelas XI pada uji coba ini tergolong sudah berkembang dengan baik.

C. Analisis Instrumen Aspek Keterampilan

Instrumen keterampilan berupa lembar tugas studi kasus mengenai analisis inflasi dan kebijakan moneter/fiskal, yang dinilai menggunakan rubrik dengan 3 aspek keterampilan abad 21 (Critical Thinking and Problem Solving), yaitu Identifikasi Masalah dan Akar Penyebab, Analisis Dampak, dan Formulasi Solusi, masing-masing berskala 1–4. Karena penilaian keterampilan bersifat penilaian kinerja yang berpotensi subjektif, reliabilitas dihitung sebagai reliabilitas antar-rater menggunakan dua penilai independen terhadap hasil kerja 32 peserta didik yang sama. Hasil analisis disajikan pada tabel berikut.

No	Aspek Keterampilan	r Antar-Rater	Persentase Kesepakatan	Kategori Reliabilitas
1	Identifikasi Masalah & Akar Penyebab	0.83	75.0%	Sangat Tinggi
2	Analisis Dampak (Berpikir Kritis)	0.74	68.8%	Tinggi
3	Formulasi Solusi (Pemecahan Masalah)	0.63	50.0%	Cukup

Reliabilitas total antar-rater, yang dihitung dari korelasi total skor kedua penilai, diperoleh sebesar 0,76 yang tergolong tinggi. Hasil ini menunjukkan bahwa secara umum rubrik keterampilan yang disusun sudah cukup konsisten digunakan oleh penilai yang berbeda. Namun demikian, terdapat variasi reliabilitas antar-aspek yang cukup mencolok. Aspek Identifikasi Masalah dan Akar Penyebab memiliki reliabilitas tertinggi ($r = 0,83$; persentase kesepakatan 75,0%), menunjukkan deskriptor rubrik pada aspek ini paling jelas dan operasional sehingga mudah diterapkan secara konsisten oleh penilai yang berbeda. Sebaliknya, aspek Formulasi Solusi (Pemecahan Masalah) memiliki reliabilitas paling rendah ($r = 0,63$; persentase kesepakatan hanya 50,0%), yang mengindikasikan deskriptor pada level skor rubrik aspek ini masih multitafsir bagi penilai. Hal ini kemungkinan disebabkan karena kata kerja operasional pada deskriptor aspek Formulasi Solusi (misalnya kata 'realistis', 'inovatif', dan 'relevan') bersifat kualitatif dan memerlukan interpretasi subjektif yang lebih tinggi dibandingkan deskriptor pada aspek Identifikasi Masalah.

Berdasarkan distribusi nilai akhir keterampilan, sebanyak 1 peserta didik (3,1%) berada pada kategori Sangat Kritis, 14 peserta didik (43,8%) pada kategori Kritis, 12 peserta didik (37,5%) pada kategori Cukup Kritis, dan 5 peserta didik (15,6%) pada kategori Kurang Kritis. Hasil ini menunjukkan bahwa kemampuan berpikir kritis dan pemecahan masalah peserta didik dalam menganalisis studi kasus ekonomi sudah cukup baik, meskipun kemampuan merumuskan solusi (aspek dengan reliabilitas rater terendah) tampaknya menjadi aspek yang paling menantang baik bagi peserta didik dalam mengerjakan maupun bagi guru dalam menilainya secara konsisten.

Ketiga instrumen asesmen baik pengetahuan, sikap, dan keterampilan pada dasarnya mengukur dimensi kompetensi yang berbeda, namun tetap saling berkaitan dalam konteks pembelajaran Ekonomi mengenai isu ketenagakerjaan dan inflasi. Peserta didik yang memiliki pemahaman konseptual yang baik mengenai inflasi dan kebijakan moneter/fiskal (aspek pengetahuan) cenderung lebih mampu mengidentifikasi akar masalah dan merumuskan solusi yang relevan secara ekonomi (aspek keterampilan), karena kemampuan berpikir kritis dalam studi kasus ekonomi memerlukan penguasaan konsep dasar yang kuat sebagai fondasinya.

Di sisi lain, aspek sikap kepedulian sosial dan lingkungan relatif independen dari kedua aspek lainnya, karena sikap lebih banyak dibentuk melalui pembiasaan dan keteladanan dalam kehidupan sehari-hari dibandingkan melalui penguasaan materi kognitif. Meskipun demikian, ketiga aspek ini tetap perlu dinilai secara terintegrasi karena bersama-sama membentuk profil kompetensi peserta didik secara holistik sesuai dengan tujuan pembelajaran Ekonomi, yaitu tidak hanya memahami konsep ekonomi secara teoretis, tetapi juga mampu bersikap peduli terhadap dampak sosial-ekonomi di lingkungan sekitar dan terampil dalam memecahkan masalah ekonomi secara nyata.

Temuan bahwa nilai rata-rata pengetahuan (55,3) relatif lebih rendah dibandingkan nilai rata-rata sikap (kategori dominan Baik) mengindikasikan adanya kesenjangan antara pemahaman kognitif dan sikap afektif peserta didik. Hal ini dapat menjadi masukan bagi guru untuk memperkuat penyampaian materi pengetahuan ekonomi, khususnya pada konsep-konsep yang teridentifikasi bermasalah pada analisis butir soal (misalnya konsep pengangguran struktural dan kebijakan tight money policy pada butir nomor 7 dan 9), tanpa mengurangi penguatan pembiasaan sikap yang sudah berjalan cukup baik.

KESIMPULAN DAN SARAN

A. Kesimpulan

- Instrumen pengetahuan (20 butir pilihan ganda) tergolong reliabel dengan koefisien KR-20 sebesar 0,67. Sebanyak 70% butir soal (14 dari 20 butir) dinyatakan diterima, sedangkan 30% (6 butir) memerlukan revisi terkait daya pembeda dan/atau validitas yang masih rendah.
- Instrumen sikap memiliki reliabilitas sangat tinggi (Cronbach's Alpha = 0,874) dengan seluruh 6 pernyataan dinyatakan valid (r item-total > 0,30), sehingga layak digunakan tanpa revisi berarti.
- Instrumen keterampilan memiliki reliabilitas antar-rater tinggi secara keseluruhan ($r = 0,76$), namun aspek Formulasi Solusi (Pemecahan Masalah) memiliki reliabilitas terendah ($r = 0,63$) sehingga memerlukan penyempurnaan deskriptor rubrik.
- Secara umum, ketiga instrumen asesmen (pengetahuan, sikap, keterampilan) sudah cukup layak digunakan dalam pembelajaran, dengan catatan perbaikan pada butir-butir dan aspek yang telah teridentifikasi pada bab pembahasan.

B. Saran

- Merevisi kalimat soal, kunci jawaban, dan/atau pilihan jawaban (distraktor) pada butir soal nomor 2, 7, 9, 12, 13, dan 18 sebelum digunakan kembali pada tes berikutnya.
- Menambah proporsi butir soal berkategori sukar pada penyusunan tes selanjutnya agar distribusi tingkat kesukaran lebih proporsional dan mampu membedakan peserta didik berkemampuan sangat tinggi.
- Memperjelas deskriptor rubrik pada aspek Formulasi Solusi dengan menambahkan kata kerja operasional dan contoh konkret pada setiap level skor, serta melakukan kalibrasi (rater training) antar-guru sebelum instrumen keterampilan digunakan secara luas.
- Melakukan uji coba ulang (retest) pada butir-butir yang telah direvisi untuk memastikan peningkatan kualitas psikometrik sebelum instrumen digunakan pada penilaian yang sesungguhnya.
- Menggunakan data hasil uji coba yang sesungguhnya (bukan simulasi) apabila tersedia, agar hasil analisis mencerminkan kondisi riil kemampuan peserta didik.

LAMPIRAN

Lampiran 1. Data Nilai Peserta Didik – Aspek Pengetahuan (N = 32)

No	Nama Peserta	Jml Benar	Nilai (0-100)	Kategori
1	Siswa 01	15	75.0	Baik
2	Siswa 02	10	50.0	Kurang
3	Siswa 03	13	65.0	Cukup
4	Siswa 04	18	90.0	Sangat Baik
5	Siswa 05	11	55.0	Cukup
6	Siswa 06	12	60.0	Cukup
7	Siswa 07	15	75.0	Baik
8	Siswa 08	15	75.0	Baik
9	Siswa 09	7	35.0	Kurang
10	Siswa 10	11	55.0	Cukup
11	Siswa 11	7	35.0	Kurang
12	Siswa 12	11	55.0	Cukup
13	Siswa 13	12	60.0	Cukup
14	Siswa 14	8	40.0	Kurang
15	Siswa 15	5	25.0	Kurang
16	Siswa 16	12	60.0	Cukup
17	Siswa 17	4	20.0	Kurang
18	Siswa 18	12	60.0	Cukup
19	Siswa 19	8	40.0	Kurang
20	Siswa 20	7	35.0	Kurang
21	Siswa 21	15	75.0	Baik
22	Siswa 22	10	50.0	Kurang
23	Siswa 23	11	55.0	Cukup
24	Siswa 24	6	30.0	Kurang
25	Siswa 25	9	45.0	Kurang
26	Siswa 26	12	60.0	Cukup
27	Siswa 27	13	65.0	Cukup
28	Siswa 28	12	60.0	Cukup
29	Siswa 29	11	55.0	Cukup
30	Siswa 30	12	60.0	Cukup
31	Siswa 31	11	55.0	Cukup
32	Siswa 32	19	95.0	Sangat Baik

Lampiran 2. Data Nilai Peserta Didik – Aspek Sikap (N = 32)

No	Nama Peserta	Total Skor (/24)	Nilai (0-100)	Kategori
1	Siswa 01	24	100.0	Sangat Baik (SB)
2	Siswa 02	14	58.3	Cukup (C)
3	Siswa 03	17	70.8	Baik (B)
4	Siswa 04	19	79.2	Baik (B)
5	Siswa 05	14	58.3	Cukup (C)
6	Siswa 06	18	75.0	Baik (B)
7	Siswa 07	18	75.0	Baik (B)
8	Siswa 08	11	45.8	Cukup (C)
9	Siswa 09	23	95.8	Sangat Baik (SB)
10	Siswa 10	19	79.2	Baik (B)
11	Siswa 11	14	58.3	Cukup (C)
12	Siswa 12	18	75.0	Baik (B)
13	Siswa 13	19	79.2	Baik (B)
14	Siswa 14	16	66.7	Baik (B)
15	Siswa 15	17	70.8	Baik (B)
16	Siswa 16	12	50.0	Cukup (C)
17	Siswa 17	16	66.7	Baik (B)
18	Siswa 18	14	58.3	Cukup (C)
19	Siswa 19	16	66.7	Baik (B)
20	Siswa 20	13	54.2	Cukup (C)
21	Siswa 21	23	95.8	Sangat Baik (SB)
22	Siswa 22	17	70.8	Baik (B)
23	Siswa 23	15	62.5	Baik (B)
24	Siswa 24	23	95.8	Sangat Baik (SB)
25	Siswa 25	17	70.8	Baik (B)
26	Siswa 26	13	54.2	Cukup (C)
27	Siswa 27	15	62.5	Baik (B)
28	Siswa 28	8	33.3	Kurang (K)
29	Siswa 29	20	83.3	Sangat Baik (SB)
30	Siswa 30	15	62.5	Baik (B)
31	Siswa 31	15	62.5	Baik (B)

No	Nama Peserta	Total Skor (/24)	Nilai (0-100)	Kategori
32	Siswa 32	20	83.3	Sangat Baik (SB)

Lampiran 3. Data Nilai Peserta Didik – Aspek Keterampilan (N = 32)

No	Nama Peserta	Skor Rater 1	Skor Rater 2	Nilai (0-100)	Kategori
1	Siswa 01	9	8	70.8	Kritikal
2	Siswa 02	10	10	83.3	Kritikal
3	Siswa 03	7	7	58.3	Cukup Kritikal
4	Siswa 04	6	8	58.3	Cukup Kritikal
5	Siswa 05	11	9	83.3	Kritikal
6	Siswa 06	7	8	62.5	Cukup Kritikal
7	Siswa 07	7	7	58.3	Cukup Kritikal
8	Siswa 08	9	8	70.8	Kritikal
9	Siswa 09	10	10	83.3	Kritikal
10	Siswa 10	7	7	58.3	Cukup Kritikal
11	Siswa 11	7	8	62.5	Cukup Kritikal
12	Siswa 12	10	9	79.2	Kritikal
13	Siswa 13	7	7	58.3	Cukup Kritikal
14	Siswa 14	10	8	75.0	Kritikal
15	Siswa 15	8	9	70.8	Kritikal
16	Siswa 16	8	7	62.5	Cukup Kritikal
17	Siswa 17	6	6	50.0	Kurang Kritikal
18	Siswa 18	6	5	45.8	Kurang Kritikal
19	Siswa 19	10	9	79.2	Kritikal
20	Siswa 20	6	8	58.3	Cukup Kritikal
21	Siswa 21	10	9	79.2	Kritikal
22	Siswa 22	8	7	62.5	Cukup Kritikal
23	Siswa 23	10	11	87.5	Sangat Kritikal
24	Siswa 24	9	8	70.8	Kritikal
25	Siswa 25	8	8	66.7	Cukup Kritikal
26	Siswa 26	9	9	75.0	Kritikal
27	Siswa 27	6	5	45.8	Kurang Kritikal
28	Siswa 28	9	8	70.8	Kritikal
29	Siswa 29	9	8	70.8	Kritikal

No	Nama Peserta	Skor Rater 1	Skor Rater 2	Nilai (0-100)	Kategori
30	Siswa 30	7	6	54.2	Kurang Kritikal
31	Siswa 31	8	8	66.7	Cukup Kritikal
32	Siswa 32	6	7	54.2	Kurang Kritikal