

Nama : Niabi Rahma Wati

NPM : 2413031078

Kelas : 2024C

**ANALISIS KUALITAS BUTIR SOAL INSTRUMEN PENILAIAN EVALUASI
PEMBELAJARAN EKONOMI**

Berdasarkan Hasil Uji Coba pada 24 siswa kelas X8 SMA Negeri 4 Metro

Mata Pelajaran : Ekonomi

Materi : Pasar Modal dan Otoritas Jasa Keuangan (OJK)

A. Analisis Aspek Pengetahuan (Kognitif)

1. Karakteristik Instrumen

Jenis soal : Pilihan Ganda
Jumlah butir : 20 soal
Jumlah Responden : 24 siswa
Skor ideal : 100

2. Uji Validitas

Validitas Isi:

Soal pilihan ganda berjumlah 20 butir mencakup seluruh indikator materi Pasar Modal dan OJK. Berdasarkan kisi-kisi, soal sudah mewakili fungsi pasar modal, produk pasar modal, prosedur investasi, wewenang OJK, sanksi, literasi keuangan, dan perilaku investor. Berdasarkan validitas isi instrumen ini dapat disimpulkan baik.

Validitas Empiris (berdasarkan data):

Dengan nilai rata-rata 95,2 dan median 100, sebaran nilai sangat tinggi. Hal ini mengindikasikan bahwa soal mungkin kurang mampu mengukur perbedaan kemampuan karena hampir semua siswa menjawab benar. Namun secara substansi, soal mengukur pengetahuan yang diajarkan.

3. Reabilitas

Dari data 24 siswa:

Mean = 95,2

Median = 100

Variasi rendah nilai berkisar antara 85-100

Sebanyak 22 dari 24 siswa, sekitar 91,7% mendapat nilai A dengan predikat sangat baik.

Reabilitas soal dianggap kurang baik, karena hampir semua siswa mendapat nilai A, sehingga variasi skor sangat kecil. Kondisi ini disebut *ceiling effect*. Akibatnya, koefisien reabilitas seperti KR-20 akan rendah karena siswa pintar dan kurang pintar sama-sama menjawab benar. Artinya, jika soal diulang, hasilnya belum tentu konsisten untuk mengklompokkan siswa berdasarkan kemampuan.

4. Tingkat Kesukaran

Rumus: $P = \frac{\text{jumlah siswa benar}}{\text{jumlah total siswa}}$
 Rata rata nilai = 95,2 rata-rata siswa menjawab benar adalah 19,04 dari 20 soal.
 Maka P rata-rata = $19,04 / 20 = 0,952$ (atau sekitar 95,2%)

Dengan rata-rata tingkat kesukaran soal 0,952 atau 95,2%. Artinya, dari 20 soal, rata-rata siswa mampu menjawab benar 19 soal. Soal dengan $P > 0,85$ termasuk mudah. Diperkirakan 16-18 soal tergolong mudah, hanya 2-4 soal sedang, dan tidak ada soal sukar. Soal terlalu mudah sehingga kurang menantang.

5. Daya Pembeda

Daya pembeda mengukur kemampuan soal membedakan siswa pintar/kelompok atas dan siswa kurang pintar/kelompok bawah.

Kelompok atas = 27% dari 24 siswa yaitu, 6 siswa dengan nilai tertinggi semua nilai 100 atau benar 20/20 soal.

Kelompok bawah = 6 siswa terendah dengan nilai 85 – 90 dengan rata-rata benar 18-20 soal.

$D = \text{proporsi benar atas} - \text{proporsi benar bawah}$
 $= (20/20) - (18/20) = 1 - 0,9 = 0,10$ (10%)

Dengan daya beda soal yang rendah, sekitar 10%. Kelompok atas semua menjawab benar 20 soal, sedangkan kelompok bawah rata-rata menjawab 18 soal, selisihnya hanya 2 soal. Soal tidak mampu membedakan siswa yang pintar dengan siswa yang kurang pintar. Idealnya daya beda minimal 0,20 atau 20%.

6. Distraktor Efektif (Pilihan Jawaban Pengecoh)

Berdasarkan hasil analisis kualitatif terhadap 20 soal pilihan ganda, ditemukan bahwa secara umum jawaban pengecoh sudah dirancang dengan baik. Sebanyak 18 soal memiliki distraktor yang berfungsi efektif, logis, menarik, dan mampu membedakan peserta yang memahami materi dari yang tidak. Namun, terdapat 2 soal, yaitu nomor 6 dan 20 yang distraktornya kurang efektif karena mudah dikenali sebagai jawaban salah.

Penjelasan distraktor yang kurang efektif:

Soal nomor 6: salah satu pilihan jawaban berbunyi “Menjamin keuntungan investor”. Distraktor ini terlalu jelas salah karena dalam konteks materi, misalnya tentang investasi dan pasar modal, tidak ada instrumen atau kebijakan yang dapat menjamin keuntungan mutlak. Peserta yang sekedar membaca sepintas akan langsung menolak pilihan ini tanpa perlu berfikir lagi. Distraktor yang baik seharusnya masuk akal namun tetap salah, sehingga memancing peserta yang kurang paham. Sebagai rekomendasi, distraktor dapat diganti dengan pernyataan yang tampak meyakinkan namun salah, misalnya “investasi dijamin aman oleh pemerintah” atau “dividen dibayar setiap bulan tanpa risiko”.

Soal Nomor 20: pilihan “tanya tetangga” dan “ikuti ajakan artis” dinilai kurang logis sebagai pengecoh. Kedua opsi ini terlalu absurd atau tidak relevan dengan konteks soal yang kemungkinan berkaitan dengan pengambilan keputusan, literasi keuangan, atau perilaku konsumen. Distraktor yang efektif harus berasal dari miskonsepsi umum atau kesalahan tipikal yang sering dilakukan peserta didik, bukan dari tindakan yang secara nyata tidak masuk akal. Distraktor dapat diganti menggunakan pengecoh yang lebih realistis, seperti “mengikuti saran tetangga yang pernah berhasil” atau “membeli produk karena sering terlihat di media sosial”.

B. Analisis Aspek Keterampilan (Psikomotor)

Data Hasil Penilaian

Predikat	Jumlah Siswa	Presentase
A (Sangat Baik)	17	70,8%
B (Baik)	7	29,2%
C (Cukup)	0	0%
D (Kurang)	0	0%

Nilai berkisar antara 75-100

Tidak ada siswa yang mendapat nilai di bawah 75

1. Uji Validitas

Instrumen penilaian aspek keterampilan memiliki validitas yang baik, baik dari segi isi maupun konstruk. Dari segi isi sudah mencakup 5 kriteria yaitu identifikasi modus, analisis penyebab, evaluasi tindakan OJK, langkah verifikasi, dan Kesimpulan. Kriteria ini relevan dengan kemampuan berpikir kritis dan pemecahan masalah nyata terkait dengan investasi bodong. Secara konstruk juga baik karena instrumen mengukur keterampilan berpikir tingkat tinggi, sesuai dengan tuntutan kurikulum Merdeka.

2. Uji Reabilitas

Hasil penilaian menunjukkan konsistensi yang cukup tinggi yaitu sekitar 70,8% siswa mendapatkan predikat A dan 29,2% mendapatkan predikat B. tidak ada nilai di bawah B.

Namun kelemahan instrumen ini, tidak adanya variasi nilai ke bawah yang mengindikasikan bahwa tugas mungkin terlalu mudah atau rubrik kurang sensitif membedakan siswa dengan kemampuan yang berbeda.

3. Tingkat Kesukaran

Rata-rata nilai siswa cenderung sangat tinggi di atas 85, dan tidak ada siswa yang mendapatkan nilai di bawah 75. Kemungkinan tugas tergolong mudah bagi sebagian besar siswa. Sehingga, instrumen kurang mampu mengidentifikasi siswa yang benar-benar unggul versus yang hanya cukup memahami materi saja. Sebagai rekomendasi diperlukan menaikkan tingkat kesulitan tugas, misalnya dengan menambahkan kasus yang lebih kompleks, data yang kontradiktif, atau meminta siswa membandingkan beberapa scenario investasi bodong sekaligus.

4. Daya Pembeda

Karena hampir semua siswa mendapatkan nilai A atau B, daya beda instrumen ini menjadi rendah. Distribusi nilai antara 100 hingga 75, menandakan selisih antara kelompok atas dan bawah tidak cukup besar untuk membedakan kemampuan secara signifikan. Akibatnya instrumen tidak dapat memisahkan siswa yang sangat kompeten dari siswa yang memiliki kompeten sedang.

5. Kualitas rubrik

Rubrik dinilai sudah cukup baik, dengan kelebihan dan kekurangan, seperti pada identifikasi modus yang sudah jelas, analisis penyebab belum ada contoh yang spesifik, evaluasi tindakan OJK belum ada indikator keefektifan, dan bobot pada Kesimpulan yang terlalu kecil.

C. Analisis Aspek Sikap (Afektif)

Predikat	Jumlah Siswa	Persentase
Sangat Baik (86-100)	7	29,2%
Baik (71-85)	17	70,8%
Cukup (56-70)	0	0%
Kurang (≤ 55)	0	0%

1. Validitas

Instrumen sudah mengukur 5 dimensi Profil Pelajar Pancasila yaitu mencakup bernalar kritis, mandiri, gotong royong, kreatif, beriman & berakhlak. Indikator perilaku sudah dijabarkan dalam rubrik observasi dan

lembar penilaian diri. Dari sisi validitas konstruk, dimensi yang dipilih sudah relevan dengan pembelajaran ekonomi. Hal ini menunjukkan bahwa instrumen ini memiliki validitas yang baik.

2. Reabilitas

Reabilitas instrumen dinilai cukup, namun terdapat beberapa kelemahan: Hasil observasi menunjukkan 7 siswa mendapat predikat sangat baik dan 17 siswa mendapat predikat baik. Tidak ada siswa di bawah baik. Hal ini mengindikasikan adanya konsistensi internal yang terbatas.

Penilaian diri (*self assessment*) memiliki potensi bias sosial, di mana siswa cenderung menilai dirinya lebih tinggi dari keadaan sebenarnya.

Konsistensi antar pengamat tidak diujikan, sehingga belum diketahui apakah hasil observasi bersifat subjektif atau tidak.

Secara keseluruhan, instrumen yang dikembangkan sudah cukup baik untuk mengukur ketiga aspek, tetapi masih perlu penyempurnaan terutama pada tingkat kesukaran dan daya pembeda. Untuk aspek pengetahuan, perlu menambahkan soal HOTS, untuk aspek keterampilan perlu menambahkan variasi kasus, dan untuk aspek sikap perlu peningkatan sensitivitas instrumen agar dapat membedakan tingkatan sikap siswa secara lebih baik.

Referensi:

Arikunto, S. (2021). *Dasar-Dasar Evaluasi Pendidikan* (Edisi Revisi). Jakarta: Bumi Aksara.

Kemendikbudristek. (2022). *Panduan Pembelajaran dan Asesmen Kurikulum Merdeka*. Jakarta: Badan Standar, Kurikulum, dan Asesmen Pendidikan.

Sudjana, N. (2017). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.