

Nama : Aura Liyanti Fani

NPM : 2413031089

ANALISIS BUTIR SOAL

A. Analisis Aspek Pengetahuan

1. Karakteristik Instrumen

Jenis Soal	: Pilihan Ganda
Jumlah Butir	: 20 Soal
Jumlah Responden	: 30 Siswa
Skor Ideal	: 100

2. Uji Reliabilitas

Berdasarkan hasil jawaban siswa, diperoleh Gambaran statistic sebagai berikut :

- a) Mean = 85
- b) Median = 87
- c) Modus = 90
- d) Nilai Tertinggi = 98
- e) Nilai Terendah = 38

Sebagian besar siswa memperoleh nilai tinggi, sehingga menunjukkan bahwa instrumen memiliki konsistensi yang cukup baik. Namun demikian, kondisi ini juga menunjukkan adanya fenomena *ceiling effect*, yaitu soal yang terlalu mudah sehingga sebagian besar siswa mendapatkan skor tinggi. Akibatnya, variasi nilai menjadi kecil dan instrumen kurang mampu membedakan kemampuan siswa secara optimal, sehingga reliabilitas instrumen dapat dikatakan cukup tetapi belum maksimal.

3. Uji Validitas

a) Validitas Isi

Instrumen yang digunakan dalam penelitian ini mencakup seluruh indikator pembelajaran yang berkaitan dengan topik badan usaha, seperti definisi, jenis-jenis, dan karakteristik masing-masing badan usaha. Pertanyaan-pertanyaan disusun berdasarkan kisi-kisi sistematis yang selaras dengan tujuan pembelajaran yang telah ditetapkan. Setiap pertanyaan dirancang untuk mencerminkan kompetensi inti yang diharapkan dikuasai oleh siswa, dengan demikian secara teoretis instrumen ini memenuhi persyaratan **validitas isi**. Oleh karena itu, dapat disimpulkan bahwa instrumen tersebut **tepat dan sesuai** untuk mengukur kemampuan kognitif siswa terkait materi pelajaran tersebut.

b) Validitas Empiris

Berdasarkan hasil jawaban peserta didik, sebagian besar peserta didik menjawab pertanyaan dengan benar, yang menunjukkan bahwa instrumen tersebut secara umum berhasil mengukur kompetensi yang ditargetkan. Namun, butir soal tertentu khususnya nomor **17 dan 20** menunjukkan tingkat kesalahan yang relatif tinggi dan tidak dijawab secara konsisten oleh peserta didik. Hal ini mengindikasikan bahwa butir-butir soal tersebut **kurang memiliki validitas empiris**, karena gagal mengukur kemampuan siswa secara akurat. Kondisi ini mungkin disebabkan oleh rumusan kalimat yang kurang jelas, tingkat kesulitan yang terlalu tinggi, atau pilihan jawaban pengecoh (distractor)

yang tidak efektif; oleh karena itu revisi diperlukan untuk mengoptimalkan kualitas instrumen tersebut.

4. Tingkat Kesukaran

Tingkat kesukaran soal menunjukkan seberapa sulit suatu soal bagi siswa.

Rumus : $P = B/N$

Keterangan :

P : indeks Tingkat kesukaran

B : jumlah siswa yang menjawab benar

N : jumlah seluruh siswa

Berdasarkan hasil perhitungan, rata-rata siswa mampu menjawab benar sekitar 17 soal dari 20 soal, sehingga diperoleh

$P = 17/20 = 0,85$ (85%)

Tingkat kesukaran suatu soal menunjukkan seberapa mudah atau sulit suatu butir soal bagi siswa untuk menjawabnya dengan benar. Berdasarkan perhitungan yang dilakukan, nilai indeks kesukaran sebesar 0,85 menunjukkan bahwa sebagian besar siswa menjawab soal tersebut dengan benar. Menurut kriteria yang digunakan, nilai ini termasuk dalam kategori mudah, tepatnya berada pada batas atas skala klasifikasi tingkat kesulitan.

Jika dianalisis lebih mendalam sebagian besar soal khususnya nomor 1, 2, 5, 6, 7, 9, 10, 14, dan 15 termasuk dalam kategori mudah. Hampir seluruh siswa menjawab soal-soal ini dengan benar, hal ini menunjukkan bahwa materi pelajaran telah dikuasai sepenuhnya atau bahwa soal-soal tersebut tidak memberikan tingkat tantangan yang memadai. Sebaliknya, soal nomor 3, 4, 8, 11, 12, 13, 16, 18, dan 19 termasuk dalam kategori sedang, yang menunjukkan kemampuan efektif dalam membedakan berbagai tingkat kompetensi siswa. Terakhir, soal nomor 17 dan 20 termasuk dalam kategori sulit, karena hanya sedikit siswa yang dapat menjawabnya dengan benar.

Dominasi soal berkategori mudah menunjukkan bahwa instrumen evaluasi ini mungkin tidak memberikan tantangan kognitif yang memadai bagi siswa, khususnya dalam menguji keterampilan berpikir tingkat tinggi (HOTS). Kondisi ini juga mengakibatkan terbatasnya variasi skor, karena sebagian besar siswa memperoleh hasil yang tinggi. Oleh karena itu, diperlukan penyesuaian untuk menyeimbangkan distribusi soal khususnya dengan menambah jumlah soal berkategori sedang dan sulit serta merancang butir soal yang menuntut keterampilan analisis, evaluasi, dan pemecahan masalah sehingga instrumen tersebut dapat mengukur kompetensi siswa secara lebih komprehensif.

5. Daya Pembeda

Daya pembeda digunakan untuk mengetahui kemampuan soal dalam membedakan siswa yang berkemampuan tinggi dan rendah.

Rumus : $D = PA - PB$

Keterangan :

D = indeks daya pembeda

PA = proporsi jawaban benar kelompok atas

PB = proporsi jawaban benar kelompok bawah

Kelompok atas dan kelompok bawah masing-masing diambil sebesar 27% dari jumlah siswa, yaitu sekitar 8 siswa. Berdasarkan hasil perhitungan diperoleh :

$D = (19/20) - (16/20) = 0,95 - 0,80 = 0,15$

Daya pembeda menunjukkan kemampuan suatu butir soal dalam membedakan antara siswa yang memiliki kemampuan tinggi (kelompok atas) dan siswa yang memiliki kemampuan rendah (kelompok bawah). Berdasarkan hasil perhitungan, diperoleh nilai daya pembeda sebesar 0,15, yang termasuk dalam kategori rendah. Hal ini berarti bahwa perbedaan kemampuan antara kelompok atas dan kelompok bawah dalam menjawab soal tidak terlalu signifikan.

Kelompok atas rata-rata mampu menjawab benar 19 dari 20 soal, sedangkan kelompok bawah juga mampu menjawab benar sekitar 16 soal. Selisih yang kecil ini menunjukkan bahwa soal belum mampu secara efektif mengidentifikasi perbedaan tingkat penguasaan materi antar siswa. Kondisi ini biasanya disebabkan oleh tingkat kesukaran soal yang terlalu mudah, sehingga baik siswa berkemampuan tinggi maupun rendah sama-sama dapat menjawab dengan benar.

Akibatnya, instrumen kurang optimal dalam fungsi evaluatifnya, khususnya dalam mengelompokkan siswa berdasarkan kemampuan. Oleh karena itu, perlu dilakukan perbaikan dengan menyusun soal yang memiliki variasi tingkat kesulitan yang lebih seimbang, serta menambahkan soal yang menuntut kemampuan analisis dan pemecahan masalah agar daya pembeda menjadi lebih baik dan instrumen mampu mengukur kemampuan siswa secara lebih akurat.

6. Distraktor Efektif

Distraktor merupakan pilihan jawaban yang salah tetapi dirancang untuk mengecoh siswa. Berdasarkan hasil analisis, sebagian besar distraktor sudah berfungsi cukup baik karena terdapat variasi jawaban siswa. Namun, pada soal yang tergolong mudah, distraktor cenderung tidak berfungsi secara optimal karena sebagian besar siswa langsung memilih jawaban yang benar. Selain itu, pada **soal nomor 17 dan 20 ditemukan bahwa distraktor kurang efektif** karena tidak cukup logis atau terlalu mudah dikenali sebagai jawaban yang salah. Oleh karena itu, distraktor perlu diperbaiki dengan menyusun pilihan jawaban yang lebih logis dan berasal dari kesalahan umum yang sering dilakukan siswa.

B. Analisis Aspek Keterampilan

1. Data Hasil Penilaian

Berdasarkan hasil penilaian keterampilan siswa, diperoleh distribusi nilai sebagai berikut :

No	Kategori	Rentang Nilai	Jumlah Siswa	Persentase
1	Sangat Baik (A)	86-100	18 Siswa	60%
2	Baik (B)	71-85	12 Siswa	40%
3	Cukup (C)	56-70	0 Siswa	0%
4	Kurang (D)	<55	0 Siswa	0%
Total			30 Siswa	100%

Berdasarkan tabel di atas, dapat diketahui bahwa sebagian besar siswa berada pada kategori **sangat baik**, yaitu sebanyak 18 siswa atau 60%, sedangkan sisanya sebanyak 12 siswa atau 40% berada pada kategori **baik**. Tidak terdapat siswa yang berada pada kategori cukup maupun kurang. Hal ini menunjukkan bahwa secara umum kemampuan keterampilan siswa tergolong tinggi dan telah mencapai kompetensi yang diharapkan. Rentang nilai yang diperoleh siswa berkisar antara

75 hingga 100, yang mengindikasikan bahwa hampir seluruh siswa mampu menyelesaikan tugas dengan baik.

2. Uji Reliabilitas

Reliabilitas pada aspek keterampilan menunjukkan konsistensi hasil penilaian. Berdasarkan data yang diperoleh, sebagian besar siswa mendapatkan nilai tinggi dengan distribusi yang relatif merata, di mana mayoritas siswa berada pada kategori sangat baik dan baik. Kondisi ini menunjukkan bahwa instrumen memiliki konsistensi yang cukup baik. Namun demikian, karena tidak terdapat variasi nilai yang signifikan ke arah rendah, maka instrumen cenderung kurang sensitif dalam membedakan kemampuan siswa secara detail. Hal ini menunjukkan bahwa reliabilitas instrumen tergolong cukup, tetapi masih perlu ditingkatkan agar mampu memberikan hasil yang lebih variatif dan akurat.

3. Uji Validitas

Instrumen penilaian aspek keterampilan memiliki validitas yang baik, baik dari segi isi maupun konstruk. Dari segi isi, instrumen telah mencakup indikator keterampilan yang relevan dengan pembelajaran ekonomi, seperti kemampuan menganalisis masalah, mengidentifikasi solusi, dan menarik kesimpulan. Selain itu, instrumen juga selaras dengan tuntutan Kurikulum Merdeka yang menekankan pada keterampilan berpikir kritis dan pemecahan masalah. Dari segi konstruk, instrumen telah mampu mengukur keterampilan yang seharusnya diukur, sehingga dapat disimpulkan bahwa instrumen memiliki validitas yang baik.

4. Tingkat Kesukaran

Rumus : $P = \text{Rata-rata skor} / \text{Skor Maksimal}$

Dari data diperoleh rata-rata nilai keterampilan sebesar 88 dari 100, maka :

$$P = 88 / 100 = 0,88$$

Berdasarkan hasil perhitungan, nilai tingkat kesukaran sebesar 0,88 menunjukkan bahwa tugas yang diberikan tergolong mudah bagi siswa. Hal ini terlihat dari sebagian besar siswa mampu memperoleh nilai tinggi, bahkan tidak ada yang berada pada kategori rendah. Kondisi ini menunjukkan bahwa instrumen kurang memberikan tantangan yang memadai, sehingga belum mampu mengidentifikasi perbedaan kemampuan keterampilan siswa secara optimal. Oleh karena itu, diperlukan peningkatan kompleksitas tugas, misalnya dengan memberikan kasus yang lebih variatif, menuntut analisis lebih mendalam, serta melibatkan situasi nyata yang lebih kompleks.

5. Daya Pembeda

Rumus : $D = PA - PB$

Maka diperoleh :

$$D = (95/100) - (80/100) = 0,95 - 0,80 = 0,15$$

Nilai daya pembeda sebesar 0,15 menunjukkan bahwa kemampuan instrumen dalam membedakan siswa berkemampuan tinggi dan rendah masih tergolong rendah. Hal ini terlihat dari selisih nilai antara kelompok atas dan kelompok bawah yang tidak terlalu besar. Kondisi ini mengindikasikan bahwa hampir semua siswa mampu menyelesaikan tugas dengan baik, sehingga instrumen kurang efektif dalam mengidentifikasi variasi kemampuan keterampilan siswa. Oleh karena itu, perlu dilakukan perbaikan dengan menyusun tugas yang lebih menantang dan bervariasi agar daya pembeda meningkat.

C. Analisis Aspek Sikap

1. Data Hasil Penelitian

No	Kategori	Rentang Nilai	Jumlah Siswa	Presentase
1	Sangat Baik	86-100	10 Siswa	33,3%
2	Baik	71-85	20 Siswa	66,7%
3	Cukup	56-70	0 Siswa	0%
4	Kurang	<55	0 Siswa	0%
TOTAL			30 Siswa	100%

Berdasarkan tabel di atas, dapat diketahui bahwa sebagian besar siswa berada pada kategori **baik**, yaitu sebanyak 20 siswa atau 66,7%, sedangkan 10 siswa atau 33,3% berada pada kategori **sangat baik**. Tidak terdapat siswa yang berada pada kategori cukup maupun kurang. Hal ini menunjukkan bahwa secara umum sikap siswa dalam pembelajaran ekonomi sudah tergolong positif dan sesuai dengan nilai-nilai yang diharapkan dalam proses pembelajaran.

Namun demikian, distribusi nilai yang terkonsentrasi pada kategori tinggi menunjukkan bahwa instrumen penilaian sikap belum mampu memberikan variasi penilaian yang cukup tajam. Hal ini dapat disebabkan oleh indikator yang masih bersifat umum atau teknik penilaian yang belum mampu menangkap perbedaan sikap siswa secara mendalam. Oleh karena itu, diperlukan pengembangan instrumen yang lebih rinci agar mampu mengidentifikasi variasi sikap siswa secara lebih akurat.

2. Uji Reliabilitas

Reliabilitas instrumen sikap berkaitan dengan konsistensi hasil penilaian. Berdasarkan hasil yang diperoleh, sebagian besar siswa menunjukkan nilai yang relatif tinggi dan tidak jauh berbeda satu sama lain. Hal ini menunjukkan bahwa instrumen memiliki tingkat konsistensi yang cukup baik. Namun demikian, karena tidak terdapat variasi nilai yang signifikan pada kategori rendah, instrumen cenderung kurang sensitif dalam membedakan tingkat sikap siswa. Selain itu, penilaian sikap yang dilakukan melalui observasi berpotensi mengandung subjektivitas penilai, sehingga dapat mempengaruhi hasil penilaian. Oleh karena itu, diperlukan penggunaan teknik penilaian yang lebih beragam, seperti penilaian diri dan penilaian antar teman, untuk meningkatkan objektivitas dan reliabilitas instrumen.

3. Uji Validitas

Instrumen penilaian sikap dalam penelitian ini memiliki validitas yang baik, karena telah mencakup berbagai indikator sikap yang relevan dengan tujuan pembelajaran. Indikator yang digunakan meliputi aspek tanggung jawab, kejujuran, kerjasama, dan kedisiplinan, yang merupakan bagian penting dalam pembentukan karakter siswa. Selain itu, instrumen juga telah disesuaikan dengan dimensi Profil Pelajar Pancasila, seperti gotong royong, mandiri, dan berakhlak mulia. Dengan demikian, secara konseptual instrumen telah mampu mengukur sikap yang seharusnya dikembangkan dalam pembelajaran ekonomi.

4. Tingkat Kesukaran

Rumus : $P = \text{Rata-rata} / \text{Skor maksimal}$

Data menunjukkan rata-rata nilai sikap siswa Adalah 82 dari 100, maka :

$$P = 82 / 100 = 0,82$$

Nilai tingkat kesukaran sebesar 0,82 menunjukkan bahwa instrumen penilaian sikap tergolong mudah bagi siswa. Hal ini terlihat dari dominasi nilai pada kategori baik dan sangat baik. Kondisi ini mengindikasikan bahwa sebagian besar siswa mampu memenuhi indikator sikap yang dinilai tanpa mengalami kesulitan yang berarti. Namun demikian, tingkat kesukaran yang terlalu rendah menyebabkan instrumen kurang mampu memberikan tantangan dalam menilai variasi sikap siswa secara lebih mendalam. Selain itu, kondisi ini juga berdampak pada rendahnya variasi skor, sehingga sulit untuk mengidentifikasi perbedaan tingkat sikap antar siswa. Oleh karena itu, diperlukan pengembangan indikator yang lebih spesifik dan mendalam agar instrumen mampu mengukur sikap secara lebih komprehensif.

5. Daya Pembeda

Rumus : $D = PA - PB$

Maka Diperoleh :

$$D = (90/100) - (75/100) = 0,90 - 0,75 = 0,15$$

Nilai daya pembeda sebesar 0,15 menunjukkan bahwa instrumen penilaian sikap memiliki kemampuan yang rendah dalam membedakan siswa dengan tingkat sikap yang berbeda. Hal ini terlihat dari perbedaan skor antara kelompok atas dan kelompok bawah yang relatif kecil. Kondisi ini mengindikasikan bahwa hampir seluruh siswa memperoleh penilaian yang baik, sehingga instrumen kurang mampu mengidentifikasi siswa yang benar-benar memiliki sikap unggul dibandingkan dengan siswa lainnya. Rendahnya daya pembeda ini dapat disebabkan oleh indikator penilaian yang terlalu umum atau kurang spesifik, serta teknik penilaian yang belum mampu menangkap variasi perilaku siswa secara detail. Oleh karena itu, perlu dilakukan perbaikan dengan menyusun indikator yang lebih terukur dan menggunakan teknik penilaian yang lebih variatif agar daya pembeda meningkat.

D. Kesimpulan

Berdasarkan hasil analisis terhadap aspek pengetahuan (kognitif), dapat disimpulkan bahwa instrumen yang digunakan telah memiliki validitas isi yang baik karena sesuai dengan indikator pembelajaran. Secara empiris, sebagian besar soal mampu mengukur kemampuan siswa, meskipun terdapat beberapa butir yang kurang optimal. Tingkat kesukaran soal didominasi kategori mudah dengan indeks sebesar 0,85, yang menunjukkan bahwa siswa relatif mampu menjawab soal dengan baik. Namun, daya pembeda yang rendah (0,15) menunjukkan bahwa instrumen kurang mampu membedakan kemampuan siswa secara optimal, sehingga perlu perbaikan pada tingkat kesulitan soal dan kualitas distractor.

Pada aspek keterampilan (psikomotor), hasil penilaian menunjukkan bahwa sebagian besar siswa berada pada kategori sangat baik dan baik, sehingga dapat dikatakan bahwa siswa telah mampu menunjukkan keterampilan yang diharapkan. Instrumen memiliki validitas yang baik, namun tingkat kesukaran yang cenderung mudah dan daya pembeda yang rendah menunjukkan bahwa instrumen belum mampu mengidentifikasi variasi kemampuan siswa secara mendalam. Oleh karena itu, diperlukan peningkatan kompleksitas tugas serta penyempurnaan rubrik penilaian agar hasil yang diperoleh lebih akurat.

Pada aspek sikap (afektif), seluruh siswa berada pada kategori baik dan sangat baik, yang menunjukkan bahwa sikap siswa selama pembelajaran sudah berkembang secara positif. Instrumen penilaian telah memiliki validitas yang baik, namun distribusi

nilai yang homogen menunjukkan bahwa instrumen kurang sensitif dalam membedakan tingkat sikap siswa. Secara keseluruhan, instrumen penilaian sudah cukup baik, tetapi masih perlu penyempurnaan pada variasi tingkat kesukaran, daya pembeda, serta pengembangan instrumen yang lebih kompleks agar mampu mengukur kemampuan siswa secara lebih komprehensif dan optimal.