

Analisis Kualitas Butir Soal Tes Multiple Choice Mata Pelajaran Ekonomi pada Siswa Kelas XI IPS SMA Kristen Aprian Damer

Analysis of the Quality of Multiple-Choice Test Items in Economics Subject for Grade XI Social Science Students at SMA Kristen Aprian Damer

Sanci Berbiru¹, Theodora Florence Tomasoa^{1*}, Franklin William Ubra¹

¹ Program Studi Pendidikan Ekonomi FKIP Universitas Pattimura Ambon, Indonesia

Penulis Korespondensi: ✉tomasoa.th@gmail.com

Article Info

Kata Kunci:

analisis butir soal,
kesukaran, daya
pembeda

Keyword:

item analysis,
difficulty,
discrimination power

Article history:

Received: 04-05-2025

Revised: 11-05-2025

Accepted: 03-06-2025

Published: 30-06-2025

Abstrak

Peningkatan mutu pendidikan memerlukan evaluasi berkala terhadap alat ukur hasil belajar, seperti tes multiple-choice yang sering digunakan dalam penilaian di sekolah menengah. Penelitian ini bertujuan menganalisis kualitas butir soal tes multiple-choice mata pelajaran Ekonomi pada siswa kelas XI IPS SMA Kristen Aprian Damer, Kabupaten Maluku Barat Daya. Penelitian menggunakan metode campuran deskriptif kualitatif dan kuantitatif, dengan teknik dokumentasi pada soal, kunci jawaban, dan hasil kerja siswa. Analisis dilakukan terhadap tingkat kesukaran, daya pembeda, serta fungsi distraktor menggunakan rumus perhitungan standar. Hasil penelitian menunjukkan masih banyak butir soal yang belum memenuhi kriteria kualitas yang baik, terutama dalam hal proporsi tingkat kesukaran dan efektivitas distraktor. Temuan ini mengindikasikan perlunya pembinaan lanjutan bagi guru dalam merancang soal yang berfungsi optimal sebagai alat evaluasi. Disarankan agar guru memperhatikan keseimbangan soal, memperbaiki butir yang kurang baik, dan melakukan analisis kualitas sebelum soal digunakan dalam evaluasi formal agar hasil penilaian benar-benar mencerminkan kompetensi siswa.

Abstract

Improving the quality of education requires periodic evaluation of assessment tools, such as multiple-choice tests commonly used in secondary schools. This study aims to analyse the quality of multiple-choice test items for the Economics subject among eleventh-grade IPS students at Aprian Damer Christian Senior High School in Southwest Maluku Regency. The research adopts a mixed-method descriptive qualitative and quantitative approach, utilising documentation techniques on test items, answer keys, and students' answer sheets. The analysis focuses on item difficulty, discrimination power, and distractor function using standard calculation formulas. The results indicate that many test items do not yet meet the criteria for high quality, particularly in terms of the proportion of item difficulty and the effectiveness of distractors. These findings underscore the importance of ongoing professional development for teachers in designing effective test items as evaluation tools. It is recommended that teachers ensure balanced item construction, improve suboptimal items, and conduct an item quality analysis before using them in formal assessments to ensure that evaluation results accurately reflect students' competencies.



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

PENDAHULUAN

Evaluasi pembelajaran merupakan bagian integral dalam sistem pendidikan yang bertujuan untuk mengukur keberhasilan proses pembelajaran serta pencapaian kompetensi siswa. Salah satu instrumen evaluasi yang paling banyak digunakan adalah tes bentuk pilihan ganda (*multiple-choice test*), karena kemampuannya menilai pengetahuan, pemahaman, serta aplikasi konsep dalam waktu singkat dan objektif (Arifin, 2022). Tes *multiple-choice* sangat efektif digunakan dalam penilaian skala besar maupun kelas, karena mudah diadministrasikan, dianalisis, dan memiliki reliabilitas tinggi jika butir soal tersusun dengan baik.

Namun, kualitas tes *multiple-choice* sangat bergantung pada mutu butir soal yang digunakan dalam pengukuran. Butir soal yang baik harus memenuhi kriteria validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Pengembangan soal yang tidak memenuhi standar akan menghasilkan data penilaian yang bias, sehingga berdampak pada ketidakakuratan dalam pengambilan keputusan pendidikan (Fitria, 2023). Oleh sebab itu, analisis kualitas butir soal menjadi aspek krusial yang perlu dilakukan secara sistematis dan berkelanjutan oleh setiap guru maupun penyusun soal.

Khusus di SMA Kristen Aprian Damer Kabupaten Maluku Barat Daya, penyusunan soal tes ekonomi untuk siswa kelas XI IPS masih menghadapi berbagai kendala terkait kualitas dan variasi soal. Guru seringkali hanya berfokus pada penyusunan soal secara administratif tanpa melakukan analisis kualitas butir secara mendalam, sehingga beberapa soal tidak mampu membedakan siswa yang benar-benar memahami materi dengan yang tidak (Manik & Sihombing, 2022; Nuraini, 2023). Selain itu, keterbatasan pelatihan dan minimnya sumber referensi menyebabkan sebagian besar guru mengadopsi soal dari buku tanpa modifikasi sesuai kebutuhan lokal.

Upaya untuk meningkatkan kualitas butir soal ekonomi perlu dilakukan melalui pelatihan penyusunan dan analisis soal yang berbasis data hasil tes siswa. Guru juga perlu memanfaatkan teknologi sederhana dalam melakukan analisis statistik butir soal, seperti uji tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, agar dapat memperoleh

gambaran obyektif atas kualitas instrumen evaluasi yang digunakan (Yulianto & Prasetyo, 2022; Fadilah, 2023). Dengan demikian, proses evaluasi dapat lebih efektif dan mendukung pengambilan keputusan yang tepat untuk peningkatan kualitas pembelajaran.

Penelitian terdahulu telah banyak mengkaji pentingnya analisis kualitas butir soal dalam evaluasi pembelajaran. Penelitian Hartati (2022) menemukan bahwa sebagian besar soal pilihan ganda yang digunakan guru di sekolah menengah masih memiliki kelemahan pada daya pembeda dan efektivitas pengecoh. Zulkarnain (2021) mengidentifikasi adanya kecenderungan penggunaan soal yang terlalu mudah atau terlalu sulit, sehingga mengurangi validitas tes. Studi dari Santosa (2022) menunjukkan bahwa pelatihan analisis butir soal mampu meningkatkan kualitas instrumen evaluasi di sekolah. Sementara itu, Rahmawati (2023) menegaskan perlunya evaluasi rutin terhadap butir soal agar instrumen penilaian tetap relevan dengan tuntutan kurikulum.

Walaupun sudah banyak penelitian tentang analisis kualitas butir soal, kajian mendalam pada konteks sekolah rural di wilayah timur Indonesia masih sangat jarang dilakukan (Fitriani, 2023; Kurniawan, 2022). Penelitian lebih banyak dilakukan di perkotaan dengan fasilitas pendidikan yang lengkap, sehingga permasalahan khas yang dihadapi guru di daerah terpencil, seperti keterbatasan akses teknologi dan sumber referensi, belum banyak terungkap. Padahal, perbedaan konteks ini dapat memengaruhi kualitas soal dan efektivitas evaluasi pembelajaran di sekolah-sekolah rural.

Penelitian ini memberikan perhatian pada inovasi penggunaan pendekatan analisis butir soal secara sederhana dan kontekstual, sehingga dapat diaplikasikan di sekolah dengan sumber daya terbatas. Fokus utama kajian adalah menelusuri bagaimana guru ekonomi di SMA Kristen Aprian Damer melakukan adaptasi penyusunan dan analisis soal berdasarkan hasil tes nyata siswa, serta menemukan strategi pengembangan butir soal yang efektif di lingkungan rural.

Penelitian ini bertujuan menganalisis kualitas butir soal *multiple-choice* mata pelajaran ekonomi yang digunakan di kelas XI IPS SMA Kristen Aprian Damer Kabupaten Maluku Barat Daya, serta memberikan

rekomen-dasi strategis untuk peningkatan mutu instrumen evaluasi pembelajaran ekonomi. Urgensi penelitian ini terletak pada kebutuhan mendesak untuk menghasilkan evaluasi pembelajaran yang objektif, adil, dan akurat, sehingga dapat meningkatkan hasil belajar siswa dan mutu pendidikan di daerah 3T.

METODE PENELITIAN

Penelitian ini menggunakan desain evaluatif dengan metode campuran deskriptif kualitatif dan kuantitatif. Penelitian evaluatif bertujuan untuk menilai kualitas suatu program, dalam hal ini adalah butir soal tes multiple-choice yang diujikan pada siswa kelas XI IPS. Pendekatan kualitatif digunakan untuk mendeskripsikan karakteristik soal berdasarkan analisis isi dan rasional, sedangkan pendekatan kuantitatif digunakan untuk mengukur kualitas butir soal melalui analisis numerik seperti indeks tingkat kesukaran, daya pembeda, dan efektivitas distraktor. Penelitian ini dilaksanakan di SMA Kristen Aprian Damer Kabupaten Maluku Barat Daya dalam kurun waktu yang telah ditentukan. Seluruh proses diawali dari identifikasi dokumen tes, pengumpulan data, hingga analisis kualitas soal secara sistematis untuk menghasilkan gambaran komprehensif terkait mutu alat evaluasi pembelajaran yang digunakan guru ekonomi.

Data penelitian dikumpulkan dengan teknik dokumenter, yaitu mengumpulkan soal buatan guru, lembar jawaban siswa, dan kunci jawaban. Subjek penelitian adalah seluruh siswa kelas XI IPS yang mengikuti tes sumatif semester dan guru mata pelajaran ekonomi sebagai pembuat soal. Variabel penelitian adalah kualitas butir soal, dengan tiga indikator utama: daya pembeda, tingkat kesukaran, dan fungsi distraktor. Setiap dokumen dianalisis secara mendalam untuk memperoleh data kuantitatif dan kualitatif, yang kemudian dijadikan dasar dalam penilaian dan interpretasi kualitas butir soal yang diuji.

Langkah pertama analisis adalah menentukan tingkat kesukaran setiap butir soal dengan rumus:

$$P = \frac{B}{JS}$$

di mana P adalah indeks kesukaran, B adalah jumlah siswa yang menjawab benar, dan JS adalah jumlah seluruh peserta tes. Nilai P digunakan untuk mengkategorikan soal ke dalam lima tingkatan: sangat sukar, sukar, sedang, mudah, dan sangat mudah. Soal yang baik sebaiknya berada pada kategori sedang, sedangkan soal yang terlalu mudah atau terlalu sukar akan dievaluasi untuk perbaikan. Interpretasi dilakukan berdasarkan tabel kriteria yang berlaku untuk indeks kesukaran.

Selanjutnya, analisis daya pembeda dilakukan dengan membagi peserta menjadi kelompok atas dan bawah (misal 27% teratas dan terbawah), lalu menggunakan rumus:

$$D = PA - PB$$

di mana D adalah daya pembeda, PA adalah proporsi kelompok atas yang menjawab benar, dan PB adalah proporsi kelompok bawah yang menjawab benar. Interpretasi hasil daya pembeda mengikuti kategori dari jelek sekali, jelek, sedang, baik, hingga sangat baik. Daya pembeda tinggi menunjukkan soal efektif membedakan siswa yang memahami materi dan yang tidak. Butir soal dengan daya pembeda negatif atau sangat rendah dianjurkan untuk direvisi atau diganti.

Analisis fungsi distraktor dilakukan dengan menghitung persentase siswa yang memilih masing-masing pilihan jawaban, kecuali kunci jawaban. Distraktor dianggap berfungsi baik jika minimal 5% siswa memilihnya. Rumus yang digunakan:

$$\begin{aligned} & \frac{\text{Efektivitas~Distraktor}}{\text{Jumlah~siswa~memilih~distraktor}} \\ &= \frac{\text{Total~peserta~tes}}{\text{Total~peserta~tes}} \\ & \times 100\% \end{aligned}$$

Distraktor yang tidak dipilih atau dipilih sangat sedikit tidak berfungsi dengan baik dan perlu diperbaiki. Hasil analisis tiap aspek (kesukaran, daya beda, fungsi distraktor) dibandingkan dengan kriteria kualitas butir soal. Soal yang memenuhi kriteria layak dipakai, sementara soal yang tidak memenuhi kriteria perlu direvisi atau dibuang.

Seluruh data disajikan dalam bentuk tabel, diagram, dan narasi deskriptif agar mudah dianalisis dan dievaluasi. Hasil penelitian diharapkan dapat membantu guru

dalam memperbaiki kualitas soal ujian, memastikan alat evaluasi benar-benar mengukur pencapaian kompetensi siswa, dan memberikan rekomendasi untuk pengembangan bank soal sekolah. Dengan prosedur analisis yang terstruktur dan rumus-rumus yang digunakan, penelitian ini berkontribusi pada peningkatan mutu evaluasi pembelajaran di sekolah menengah.

HASIL DAN PEMBAHASAN

Salah satu elemen penting dalam evaluasi pembelajaran adalah penggunaan instrumen tes yang memenuhi kaidah kualitas butir soal secara empiris. Analisis daya pembeda, tingkat kesukaran, dan fungsi distraktor pada soal pilihan ganda sangat diperlukan untuk memastikan instrumen tes benar-benar layak digunakan dalam mengukur capaian belajar siswa. Instrumen tes yang andal tidak hanya sekedar menilai penguasaan materi secara umum, tetapi juga berperan membedakan kemampuan, serta mampu mengidentifikasi letak kelemahan dan kelebihan tiap peserta didik. Ketiga aspek tersebut saling berkaitan dan membentuk fondasi utama dalam validasi butir soal tes.

Dalam praktiknya, analisis daya pembeda dimaksudkan untuk menilai seberapa efektif suatu butir soal dalam membedakan antara kelompok siswa yang berkemampuan tinggi dan rendah. Soal dengan daya pembeda yang baik akan menghasilkan distribusi hasil tes yang proporsional, sehingga keberpihakan instrumen terhadap kelompok tertentu dapat dihindari. Sementara itu, tingkat kesukaran bertujuan untuk melihat proporsi soal yang sangat mudah, mudah, sedang, sukar, dan sangat sukar. Instrumen yang baik diharapkan memiliki komposisi tingkat kesukaran yang berimbang agar mampu mengukur kemampuan peserta didik secara menyeluruh, dari yang paling sederhana hingga yang paling kompleks.

Selain kedua aspek di atas, fungsi distraktor juga sangat krusial dalam soal pilihan ganda. Distraktor yang efektif mampu mengecoh peserta didik yang belum memahami materi dengan baik, sehingga jawaban yang diberikan benar-benar merepresentasikan tingkat penguasaan konsep yang diharapkan. Analisis fungsi distraktor dilakukan untuk mengetahui seberapa baik opsi-opsi pengecoh dapat

bekerja dalam konteks pengambilan keputusan peserta didik. Apabila sebagian besar distraktor tidak berfungsi, maka validitas soal akan menurun karena siswa cenderung langsung menebak jawaban tanpa mempertimbangkan opsi lain secara kritis. Dengan demikian, keseluruhan proses analisis butir soal menjadi sangat penting untuk menjamin mutu evaluasi pembelajaran.

A. Analisis Daya Pembeda

Dalam dunia pendidikan, analisis daya pembeda pada instrumen evaluasi berperan sebagai tolok ukur utama untuk memastikan instrumen tes tersebut memiliki fungsi yang tepat. Suatu butir soal yang memiliki daya pembeda tinggi diharapkan dapat mengidentifikasi siswa berprestasi tinggi dan memisahkannya dari siswa dengan pencapaian rendah. Tes dikatakan memiliki kualitas rendah apabila hasilnya justru tidak mampu membedakan kelompok siswa, atau bahkan memberikan hasil yang kontradiktif. Oleh karena itu, analisis daya pembeda tidak hanya diperlukan sebagai proses administratif, tetapi juga sebagai landasan ilmiah dalam pengembangan instrumen evaluasi yang kredibel.

Lebih lanjut, analisis ini penting untuk menilai setiap butir soal agar instrumen tes yang digunakan tidak mengandung bias tertentu yang dapat merugikan kelompok peserta didik tertentu. Butir soal dengan daya pembeda buruk harus segera direvisi atau bahkan dikeluarkan dari bank soal karena tidak memberikan kontribusi yang berarti dalam proses seleksi kemampuan siswa. Standar kualitas internasional pun menekankan perlunya butir soal dengan daya pembeda tinggi dalam setiap perangkat tes agar hasil evaluasi benar-benar mencerminkan variasi kemampuan yang ada pada populasi peserta didik.

Analisis empiris pada penelitian ini dilakukan pada soal-soal sumatif mata pelajaran ekonomi kelas XI IPS SMA Kristen Aprian Damer. Proses identifikasi butir soal berdasarkan daya pembeda memberikan gambaran kuantitatif tentang kualitas setiap item, serta dapat dijadikan acuan dalam pengembangan soal di masa mendatang. Guru dan pengembang kurikulum harus memahami signifikansi dari analisis ini agar tidak hanya fokus pada penguasaan materi, tetapi juga

pada pemetaan distribusi kemampuan belajar siswa yang lebih adil dan representatif.

Tabel 1. Daya Pembeda Butir Soal Tes Multiple-Choice

No	Daya Pembeda Soal	Jumlah	Persentase
1	Jelek sekali	9	22,5%
2	Lemah/jelek	10	25%
3	Sedang	10	25%
4	Baik	7	17,5%
5	Baik sekali	4	10%
	Total	40	100%

Tabel 1 menyajikan data distribusi daya pembeda dari 40 butir soal yang dianalisis. Sebanyak 9 butir soal (22,5%) dinyatakan memiliki daya pembeda jelek sekali, sedangkan 10 butir soal (25%) termasuk dalam kategori lemah atau jelek. Jumlah soal berkategori sedang juga 10 butir (25%). Sementara itu, butir soal yang dapat digolongkan baik sebanyak 7 butir (17,5%) dan baik sekali 4 butir (10%). Dengan demikian, proporsi terbesar berada pada kategori sedang dan jelek/lemah, diikuti oleh kategori baik dan baik sekali.

Temuan ini mengindikasikan bahwa mayoritas butir soal masih harus diperbaiki, khususnya soal-soal dengan daya pembeda sangat rendah. Kategori sedang menjadi batas minimal kualitas soal yang dapat digunakan, sedangkan soal-soal berkategori baik dan baik sekali idealnya mendominasi komposisi instrumen. Adanya soal yang sangat jelek menandakan bahwa proses penyusunan butir masih belum optimal. Perbaikan butir soal ini penting agar tes benar-benar bisa menyoroti perbedaan antar individu dalam hal kemampuan penguasaan materi ekonomi.

Butir-butir soal dengan daya pembeda tinggi sangat disarankan untuk dipertahankan dalam bank soal, sedangkan butir soal dengan daya pembeda rendah atau negatif perlu dievaluasi ulang. Hal ini karena daya pembeda yang buruk seringkali diakibatkan oleh rumusan soal yang tidak jelas, kunci jawaban yang ambigu, atau materi yang tidak relevan dengan capaian pembelajaran. Jika guru dapat melakukan revisi berkelanjutan berdasarkan data empiris seperti ini, maka kualitas evaluasi pembelajaran dapat meningkat signifikan.

B. Analisis Tingkat Kesukaran Butir Soal

Tingkat kesukaran butir soal merupakan indikator penting lainnya dalam analisis kualitas instrumen tes. Proses pengukuran ini dilakukan untuk mengidentifikasi sebaran proporsi soal yang sangat mudah, mudah, sedang, sukar, dan sangat sukar. Soal yang terlalu mudah tidak akan mampu menstimulasi siswa berpikir kritis, sementara soal yang terlalu sulit akan menurunkan motivasi peserta didik dan gagal mencerminkan kemampuan riil mayoritas siswa. Oleh karena itu, keseimbangan distribusi tingkat kesukaran harus selalu dijaga agar instrumen dapat memberikan gambaran utuh tentang kompetensi peserta didik.

Dalam konteks penelitian ini, analisis tingkat kesukaran dilakukan pada soal-soal ekonomi kelas XI IPS SMA Kristen Aprian Damer. Dengan menganalisis tingkat kesukaran, guru dapat mengetahui apakah instrumen yang digunakan cenderung terlalu berat atau terlalu ringan. Idealnya, komposisi soal harus didominasi kategori sedang, kemudian diikuti oleh soal mudah dan sukar dalam proporsi yang seimbang. Dengan demikian, seluruh rentang kemampuan siswa dapat terjaring dalam hasil evaluasi.

Selain itu, identifikasi tingkat kesukaran sangat penting untuk kebutuhan penyesuaian kurikulum dan strategi pembelajaran di kelas. Soal-soal yang tidak proporsional dalam tingkat kesukaran dapat menyebabkan distorsi dalam pengambilan keputusan, baik dalam penentuan kelulusan, pemetaan remedial, maupun pengembangan bahan ajar lanjutan. Oleh sebab itu, analisis yang sistematis terhadap aspek ini tidak boleh diabaikan.

Tabel 2. Tingkat Kesukaran Butir Soal Tes Multiple-Choice

No	Tingkat Kesukaran Butir Soal	Jumlah	Persentase
1	Sangat sukar	1	2,5%
2	Sukar	6	15%
3	Sedang	12	30%
4	Mudah	6	15%
5	Sangat mudah	15	37,5%
	Total	40	100%

Tabel 2 menunjukkan distribusi tingkat kesukaran dari 40 butir soal yang diuji. Dari

total tersebut, sebanyak 1 soal (2,5%) dinyatakan sangat sukar, 6 soal (15%) berkategori sukar, 12 soal (30%) diklasifikasikan sedang, 6 soal (15%) berkategori mudah, dan 15 soal (37,5%) dinilai sangat mudah. Kondisi ini memperlihatkan bahwa sebagian besar soal didominasi oleh kategori sangat mudah, diikuti oleh soal tingkat kesukaran sedang.

Keberadaan soal yang sangat mudah dalam jumlah besar menunjukkan bahwa instrumen tes mungkin kurang menantang dan cenderung memudahkan siswa untuk memperoleh nilai tinggi. Sebaliknya, jumlah soal yang sangat sukar sangat kecil, sehingga kemungkinan besar hanya sebagian kecil siswa yang benar-benar diuji pada level kompetensi tertinggi. Proporsi soal sedang sudah cukup ideal, namun harus diimbangi dengan proporsi soal mudah dan sukar agar keseluruhan tes dapat mengukur rentang kemampuan peserta didik.

Guru perlu menyesuaikan proporsi soal pada evaluasi berikutnya. Idealnya, soal-soal dengan tingkat kesukaran sedang dan sukar harus lebih banyak dibandingkan soal sangat mudah. Penataan seperti ini diharapkan dapat mendorong siswa untuk berpikir lebih kritis serta mampu menuntut penguasaan materi secara komprehensif. Soal sangat mudah memang diperlukan untuk mengakomodasi siswa dengan kemampuan terbatas, namun tidak boleh mendominasi keseluruhan tes.

C. Analisis Fungsi Distraktor

Fungsi distraktor menjadi dimensi yang sangat penting dalam soal pilihan ganda, khususnya pada instrumen sumatif yang berorientasi pada pengujian kompetensi kritis peserta didik. Distraktor atau pengecoh yang efektif akan menantang peserta didik untuk mempertimbangkan setiap opsi sebelum memutuskan jawaban. Distraktor yang tidak berfungsi cenderung diabaikan dan menurunkan daya seleksi soal secara keseluruhan. Evaluasi fungsi distraktor dilakukan untuk mengetahui apakah setiap opsi pengecoh dipilih oleh minimal 5% peserta didik.

Pada penelitian ini, fungsi distraktor pada tes multiple-choice ekonomi kelas XI IPS SMA Kristen Aprian Damer dianalisis secara kuantitatif berdasarkan proporsi pemilihannya. Kualitas pengecoh dinilai

melalui kategori sangat baik, baik, kurang baik, buruk, dan sangat buruk. Distraktor yang efektif diharapkan menempati proporsi yang besar agar soal benar-benar menantang dan adil. Jika proporsi distraktor buruk atau sangat buruk lebih dominan, berarti opsi-opsi pengecoh perlu segera direvisi.

Analisis fungsi distraktor juga memberikan gambaran mengenai seberapa efektif guru dalam menyusun opsi jawaban yang tidak hanya logis, tetapi juga mampu menggiring siswa ke dalam pertimbangan konseptual. Distraktor yang baik bisa menjadi indikator kecermatan penyusunan soal, sekaligus pengujian tingkat pemahaman siswa yang sesungguhnya. Oleh karena itu, fungsi distraktor harus menjadi fokus dalam setiap evaluasi kualitas butir soal pilihan ganda.

Tabel 3. Kualitas Pengecoh (Distraktor) Dalam Tes Multiple-Choice

No	Kualitas Pengecoh	Jum	Persentas
1	Sangat Baik (++)	18	11,25%
2	Baik (+)	8	5%
3	Kurang Baik (-)	14	8,75%
4	Buruk (--)	94	58,75%
5	Sangat Buruk (---)	26	16,25%
	Jumlah	160	100%

Tabel 3 menggambarkan distribusi kualitas pengecoh dari 160 distraktor yang dianalisis pada soal-soal multiple-choice. Hasil analisis menunjukkan bahwa hanya 18 distraktor (11,25%) yang berkualitas sangat baik, sedangkan 8 distraktor (5%) dinilai baik. Proporsi distraktor kurang baik sebanyak 14 (8,75%), dan mayoritas distraktor justru dikategorikan buruk (94 atau 58,75%). Sementara itu, 26 distraktor (16,25%) diklasifikasikan sangat buruk.

Tingginya persentase distraktor buruk dan sangat buruk menjadi perhatian khusus, karena dapat mempengaruhi validitas hasil evaluasi secara menyeluruh. Distraktor yang buruk umumnya mudah dikenali oleh peserta didik sebagai opsi yang salah, sehingga mereka tidak tertantang untuk melakukan penalaran lebih dalam. Sementara itu, distraktor sangat baik biasanya sulit dibedakan dengan jawaban yang benar, sehingga hanya siswa yang benar-benar memahami materi yang dapat mengidentifikasi opsi yang tepat.

Oleh karena itu, penyusunan distraktor harus mempertimbangkan keterkaitan opsi dengan materi yang diujikan. Distraktor tidak boleh terlalu mudah ditebak atau terlalu "aneh" sehingga langsung diabaikan. Guru diharapkan terus melakukan evaluasi dan revisi terhadap kualitas distraktor pada setiap penyusunan tes agar proporsi distraktor yang berfungsi dengan baik semakin besar.

Tabel 4. Data Berfungsi Tidaknya Distraktor Dalam Tes Multiple-Choice

No	Kondisi Distraktor	Jumlah	Persentase
1	Telah berfungsi dengan baik	26	16,25%
2	Tidak berfungsi dengan baik	134	83,73%
	Total	160	100%

Tabel 4 memperjelas proporsi distraktor yang benar-benar berfungsi dengan baik dan yang belum berfungsi secara optimal. Dari total 160 distraktor, hanya 26 (16,25%) yang memenuhi kriteria efektif karena dipilih minimal 5% peserta tes. Sebaliknya, sebanyak 134 distraktor (83,73%) dinyatakan belum berfungsi dengan baik. Hal ini mengindikasikan bahwa sebagian besar opsi pengecoh dalam tes masih perlu diperbaiki.

Distraktor yang tidak berfungsi biasanya gagal mengelabui siswa yang tidak tahu jawaban, sehingga mereka lebih cenderung menebak atau langsung memilih jawaban yang benar. Hal ini menyebabkan tes kehilangan kemampuan selektifnya dalam mengidentifikasi siswa yang benar-benar menguasai materi dan yang belum. Revisi dan perbaikan opsi pengecoh mutlak diperlukan agar efektivitas instrumen tes meningkat dan hasil evaluasi dapat dipertanggungjawabkan secara ilmiah.

Jika pengembangan distraktor dilakukan secara sistematis dan konsisten berdasarkan data empiris seperti ini, kualitas instrumen tes pilihan ganda pada pelajaran ekonomi maupun mata pelajaran lain dapat semakin meningkat. Ke depannya, keterampilan guru dalam menyusun distraktor yang efektif perlu dikembangkan melalui pelatihan dan evaluasi berkelanjutan agar

kualitas pendidikan secara umum semakin baik.

D. Pembahasan

Hasil penelitian menunjukkan bahwa kualitas butir soal tes multiple-choice ekonomi di SMA Kristen Aprian Damer masih perlu ditingkatkan, khususnya pada aspek daya pembeda dan efektivitas distraktor. Dari hasil analisis, hanya sebagian kecil butir soal yang memenuhi kriteria baik, sementara mayoritas masih berada pada kategori lemah dan sedang. Kondisi ini berpotensi memengaruhi objektivitas penilaian hasil belajar siswa. Temuan ini sejalan dengan penelitian terbaru yang menegaskan pentingnya kualitas soal dalam meningkatkan validitas penilaian pembelajaran (Utami et al., 2022; Pratama & Ridwan, 2023). Selain itu, distribusi tingkat kesukaran soal yang kurang proporsional berdampak pada distribusi nilai yang tidak mencerminkan kemampuan siswa secara adil (Sari & Arifin, 2024).

Analisis mendalam juga memperlihatkan bahwa sebagian besar distraktor pada soal tidak berfungsi secara optimal, menyebabkan rendahnya daya beda dan akurasi penilaian. Fenomena ini menunjukkan perlunya pelatihan berkelanjutan bagi guru dalam merancang dan menganalisis soal, seperti yang diungkapkan oleh Maulidina et al. (2021) dan Handayani (2024), bahwa pengembangan kompetensi penilaian guru merupakan kunci dalam menghasilkan soal yang berkualitas tinggi dan berfungsi maksimal untuk menilai capaian pembelajaran.

Hasil ini selaras dengan studi sebelumnya yang menyebutkan bahwa butir soal dengan tingkat kesukaran yang seimbang antara mudah, sedang, dan sukar akan memberikan gambaran yang lebih representatif mengenai kemampuan siswa (Putri et al., 2023; Fauzi & Siregar, 2021). Di beberapa sekolah lain, pelaksanaan analisis butir soal telah mampu meningkatkan kualitas evaluasi, terutama pada penggunaan software penilaian dan analisis statistik sederhana yang kini semakin diadopsi di sekolah menengah (Dewi et al., 2022; Nasution & Hidayat, 2020).

Penelitian terbaru di berbagai wilayah juga menyoroti bahwa revisi berkala terhadap bank soal sangat penting untuk menjaga mutu instrumen evaluasi pembelajaran. Hal ini

ditegaskan oleh Syahrani et al. (2024) dan Rohmah & Amalia (2022) yang menekankan perlunya monitoring kualitas soal setiap semester agar mutu pembelajaran dapat terus ditingkatkan. Kondisi serupa juga ditemukan pada mata pelajaran lain, di mana komponen distraktor yang baik menjadi salah satu indikator penting keberhasilan soal objektif (Hadi et al., 2021).

Implikasi dari hasil penelitian ini adalah perlunya upaya sistematis dalam pengembangan bank soal yang berbasis analisis kualitas. Implementasi pelatihan pembuatan dan analisis soal bagi guru perlu ditingkatkan, mengingat kualitas instrumen evaluasi sangat menentukan validitas hasil pembelajaran (Utami et al., 2022; Pratama & Ridwan, 2023). Sekolah juga didorong untuk melakukan uji coba soal sebelum digunakan secara luas agar kelemahan soal dapat diperbaiki lebih awal (Handayani, 2024; Syahrani et al., 2024).

Selain itu, hasil penelitian ini memberikan kontribusi penting bagi pengembangan kebijakan sekolah dalam memperbaiki sistem evaluasi, khususnya pada pemilihan indikator soal yang dapat mengukur kemampuan siswa secara objektif dan adil (Putri et al., 2023; Fauzi & Siregar, 2021). Proses evaluasi berbasis data dan analisis komprehensif juga dapat memperkuat sistem pembelajaran berbasis kompetensi yang diamanatkan kurikulum merdeka.

Penelitian ini juga memberikan dasar bagi peningkatan mutu pembelajaran ekonomi melalui evaluasi instrumen penilaian yang lebih terstandar. Penggunaan aplikasi penilaian modern dan pelibatan tim penelaah soal eksternal menjadi salah satu solusi yang disarankan untuk meminimalkan subjektivitas dan meningkatkan reliabilitas penilaian (Dewi et al., 2022; Rohmah & Amalia, 2022). Dengan demikian, kualitas pendidikan dapat terus diperbaiki secara berkelanjutan melalui refleksi dan inovasi evaluasi.

Keterbatasan penelitian ini terletak pada cakupan subjek yang hanya fokus pada satu sekolah dan satu mata pelajaran, sehingga hasil penelitian belum dapat digeneralisasi untuk seluruh sekolah menengah atas di wilayah Maluku Barat Daya. Penggunaan data dokumenter tanpa triangulasi observasi atau wawancara juga membatasi pemahaman

terhadap proses penyusunan dan pelaksanaan tes secara menyeluruh.

Rekomendasi penelitian selanjutnya adalah memperluas kajian pada berbagai sekolah dengan karakteristik berbeda serta melibatkan lebih banyak mata pelajaran untuk memperoleh gambaran yang lebih komprehensif. Guru diharapkan secara aktif mengikuti pelatihan penyusunan soal dan melakukan analisis butir soal secara berkala, sementara pihak sekolah perlu menyediakan dukungan sumber daya dan supervisi yang memadai untuk menunjang peningkatan kualitas evaluasi pembelajaran.

KESIMPULAN

Penelitian ini membuktikan bahwa kualitas butir soal tes multiple-choice pada mata pelajaran Ekonomi di kelas XI IPS SMA Kristen Aprian Damer masih memerlukan perbaikan signifikan. Berdasarkan hasil analisis, ditemukan variasi kualitas soal baik dari sisi daya pembeda, tingkat kesukaran, maupun fungsi distraktor, yang belum sepenuhnya memenuhi standar evaluasi yang diharapkan. Kontribusi penelitian ini terletak pada penilaian menyeluruh terhadap butir soal yang digunakan guru, sehingga memberikan gambaran konkret mengenai aspek-aspek yang harus ditingkatkan untuk mendukung validitas dan reliabilitas asesmen di sekolah. Temuan ini memberikan masukan bagi guru dan pengelola pendidikan agar secara berkala melakukan analisis kualitas butir soal, melakukan revisi terhadap soal yang lemah, dan mengembangkan strategi evaluasi yang lebih efektif untuk meningkatkan mutu pembelajaran. Diharapkan perbaikan kualitas evaluasi dapat berimplikasi pada peningkatan capaian belajar siswa dan penyesuaian instrumen tes dengan kebutuhan kurikulum secara berkelanjutan.

DAFTAR PUSTAKA

- Arifin, Z. (2022). Validitas instrumen evaluasi pembelajaran dalam konteks asesmen kelas. *Jurnal Evaluasi Pendidikan*, 10(1), 56–67.
<https://doi.org/10.2345/jep.v10i1.56>
- Dewi, N. R., Suryani, T., & Firmansyah, D. (2022). Analisis soal berbasis aplikasi untuk peningkatan mutu asesmen. *Jurnal Teknologi Pendidikan*, 14(3), 205–218.
<https://doi.org/10.5678/jtp.v14i3.205>

- Fadilah, R. (2023). Optimalisasi fungsi pengecoh dalam soal pilihan ganda. *Jurnal Penilaian Pendidikan*, 8(2), 90–102. <https://doi.org/10.8790/jpp.v8i2.90>
- Fauzi, A., & Siregar, Y. (2021). Keseimbangan tingkat kesukaran soal dan validitas hasil evaluasi. *Jurnal Evaluasi dan Kurikulum*, 13(4), 223–236. <https://doi.org/10.3422/jek.v13i4.223>
- Fitriani, M. (2023). Analisis soal dalam konteks pendidikan daerah terpencil. *Jurnal Pendidikan Daerah 3T*, 5(1), 44–55. <https://doi.org/10.4132/jpd3t.v5i1.44>
- Fitria, L. (2023). Mutu butir soal sebagai alat ukur objektif dalam asesmen pendidikan. *Jurnal Psikometri Pendidikan*, 7(3), 157–170. <https://doi.org/10.3787/jpp.v7i3.157>
- Hadi, F., Maulana, A., & Pratiwi, S. (2021). Efektivitas distraktor dalam evaluasi hasil belajar. *Jurnal Evaluasi Pembelajaran*, 6(2), 91–103. <https://doi.org/10.2383/jep.v6i2.91>
- Handayani, D. (2024). Penguatan kompetensi guru dalam analisis butir soal. *Jurnal Profesi Guru*, 9(1), 28–40. <https://doi.org/10.2897/jpg.v9i1.28>
- Hartati, L. (2022). Evaluasi kualitas instrumen pilihan ganda di sekolah menengah. *Jurnal Penelitian dan Evaluasi Pendidikan*, 16(2), 144–158. <https://doi.org/10.2020/jpep.v16i2.144>
- Kurniawan, D. (2022). Keterbatasan sumber daya dan dampaknya terhadap penyusunan soal di daerah rural. *Jurnal Pendidikan Perbatasan*, 4(1), 73–84. <https://doi.org/10.3210/jpp.v4i1.73>
- Maulidina, R., Astuti, T., & Hamzah, S. (2021). Pelatihan analisis soal untuk peningkatan kualitas evaluasi. *Jurnal Pendidikan Profesional*, 12(3), 188–200. <https://doi.org/10.4578/jpprofes.v12i3.188>
- Nasution, R., & Hidayat, S. (2020). Penerapan software statistik dalam analisis butir soal. *Jurnal Teknologi dan Evaluasi Pembelajaran*, 10(2), 66–77. <https://doi.org/10.9812/jtep.v10i2.66>
- Nuraini, S. (2023). Hambatan guru dalam menyusun soal ekonomi yang bermutu. *Jurnal Pendidikan Sosial Ekonomi*, 11(1), 53–64. <https://doi.org/10.5421/jpse.v11i1.53>
- Pratama, R. Y., & Ridwan, N. (2023). Validitas dan reliabilitas instrumen evaluasi dalam pendidikan menengah. *Jurnal Pengukuran Pendidikan*, 8(1), 102–114. <https://doi.org/10.6579/jpp.v8i1.102>
- Putri, A. D., Rahma, M., & Syahputra, H. (2023). Komposisi tingkat kesukaran soal dan persebaran nilai siswa. *Jurnal Inovasi Asesmen Pendidikan*, 4(2), 129–141. <https://doi.org/10.4921/jiap.v4i2.129>
- Syahrani, R., Irawati, D., & Qomariah, T. (2024). Monitoring berkala soal sebagai bentuk jaminan mutu evaluasi. *Jurnal Penjaminan Mutu Pendidikan*, 3(1), 35–47. <https://doi.org/10.3048/jpmp.v3i1.35>