

ANALISIS KUALITAS SOAL PILIHAN GANDA MELALUI SIMULASI TUGAS PROYEK PADA MAHASISWA PENDIDIKAN EKONOMI MENGGUNAKAN SPSS

Epelima Sinaga¹, Mizael Demak Sitohang², Emia Perangin Angin³, Agnes Dea Lita
Br Sitepu⁴

^{1,2,3,4}Pendidikan Ekonomi FE Universitas Negeri Medan

¹epelimasinaga@gmail.com, ²mizaelsitohang11@gmail.com,

³emiamei29@gmail.com, ⁴agnesdealita20@gmail.com,

ABSTRACT

This study is motivated by the importance of the quality of evaluation instruments in accurately measuring students' abilities, particularly multiple-choice items developed by students through project-based assignments. However, the instruments produced often have not undergone a comprehensive item analysis, so their validity and reliability cannot be assured. This study aims to analyze the quality of multiple-choice test items constructed by Economics Education students through a project-based simulation using the SPSS application. The method employed is a quantitative approach with a descriptive research design. The subjects of this study consisted of 35 students, while the research object comprised 20 test items on the topic of market failure. Data were collected through an online instrument trial and analyzed using tests of validity, reliability, difficulty level, discrimination index, and distractor effectiveness. The results showed that 75% of the items were valid and 25% were invalid. The reliability coefficient of 0.802 indicates a high category. The level of difficulty was dominated by easy items (80%) and moderate items (20%), with no difficult items identified. The discrimination index was mostly categorized as very good (65%), although some items were still in the fair and poor categories. The effectiveness of distractors varied, with some distractors not functioning optimally. Therefore, in general, the quality of the test items can be considered good; however, revisions are still needed to improve the balance of difficulty levels and the effectiveness of distractors.

Keywords: Item Analysis, Instrument Quality, Multiple-Choice Questions

ABSTRAK

Penelitian ini dilatarbelakangi oleh pentingnya kualitas instrumen evaluasi dalam mengukur kemampuan peserta didik secara akurat, khususnya pada soal pilihan ganda yang disusun oleh mahasiswa melalui tugas proyek. Namun, instrumen yang dihasilkan seringkali belum melalui analisis kualitas butir soal secara komprehensif sehingga validitas dan keandalannya belum dapat dipastikan. Penelitian ini bertujuan untuk menganalisis kualitas butir soal pilihan ganda yang disusun oleh mahasiswa Pendidikan Ekonomi melalui simulasi tugas proyek dengan menggunakan aplikasi SPSS. Metode yang digunakan adalah pendekatan

kuantitatif dengan jenis penelitian deskriptif. Subjek penelitian terdiri dari 35 mahasiswa, dengan objek penelitian berupa 20 butir soal pada materi kegagalan pasar. Data diperoleh melalui uji coba instrumen secara daring dan dianalisis menggunakan uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas pengecoh. Hasil penelitian menunjukkan bahwa 75% butir soal dinyatakan valid dan 25% tidak valid. Nilai reliabilitas sebesar 0,802 menunjukkan kategori tinggi. Tingkat kesukaran didominasi soal mudah (80%) dan sedang (20%), tanpa adanya soal sukar. Daya pembeda sebagian besar berada pada kategori sangat baik (65%), namun masih terdapat soal dengan kategori cukup dan tidak baik. Efektivitas pengecoh menunjukkan variasi, dengan sebagian distraktor belum berfungsi optimal. Dengan demikian, secara umum kualitas soal tergolong baik, namun masih diperlukan revisi untuk meningkatkan keseimbangan tingkat kesukaran dan efektivitas pengecoh.

Kata Kunci: Analisis Butir Soal, Kualitas Instrumen, Soal Pilihan Ganda

A. Pendahuluan (12 pt dan Bold)

Pendidikan merupakan salah satu pilar utama dalam meningkatkan kualitas sumber daya manusia yang mampu beradaptasi dengan perkembangan ilmu pengetahuan dan teknologi. Dalam proses pendidikan, keberhasilan pembelajaran tidak hanya ditentukan oleh penyampaian materi, tetapi juga oleh efektivitas evaluasi yang dilakukan. Evaluasi pembelajaran berfungsi untuk mengukur tingkat pencapaian kompetensi peserta didik sekaligus menjadi dasar dalam menentukan langkah perbaikan pembelajaran. Sejalan dengan itu, Arikunto (2018) menyatakan bahwa evaluasi merupakan kegiatan untuk mengumpulkan informasi guna mengetahui keberhasilan suatu

program pendidikan. Dengan demikian, evaluasi yang baik harus didukung oleh instrumen yang mampu mengukur kemampuan peserta didik secara tepat dan akurat.

Salah satu bentuk evaluasi yang paling umum digunakan dalam pembelajaran adalah tes, khususnya tes pilihan ganda. Jenis tes ini memiliki keunggulan dalam hal objektivitas penilaian, efisiensi waktu, serta kemudahan dalam proses pengolahan hasil. Menurut Sudjana (2017), tes pilihan ganda mampu mengukur berbagai aspek kemampuan kognitif peserta didik apabila disusun dengan baik dan sesuai dengan indikator pembelajaran. Namun demikian, keunggulan tersebut hanya dapat dicapai apabila butir soal yang

digunakan memiliki kualitas yang baik. Sebaliknya, apabila butir soal disusun tanpa memperhatikan prinsip-prinsip penyusunan instrumen yang benar, maka hasil evaluasi berpotensi menjadi bias dan tidak mencerminkan kemampuan peserta didik secara sebenarnya .

Dalam upaya memastikan kualitas instrumen tes, diperlukan suatu proses analisis yang dikenal sebagai analisis butir soal. Analisis ini merupakan kegiatan sistematis untuk mengkaji karakteristik setiap item soal berdasarkan berbagai indikator, seperti validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Menurut Elizanoviyanti & Saepurokhman (2025), analisis butir soal merupakan langkah penting dalam pengembangan instrumen evaluasi karena dapat memberikan informasi mengenai kualitas setiap item yang digunakan dalam tes. Dengan analisis ini, pendidik dapat mengidentifikasi butir soal yang layak digunakan, perlu direvisi, atau bahkan harus dibuang. Dengan demikian, analisis butir soal menjadi langkah penting dalam meningkatkan kualitas evaluasi pembelajaran secara keseluruhan.

Berbagai penelitian sebelumnya menunjukkan bahwa kualitas instrumen evaluasi yang digunakan dalam pembelajaran masih belum optimal. Sebagian besar butir soal yang dianalisis dalam penelitian-penelitian tersebut masih memerlukan perbaikan, baik dari segi validitas maupun daya pembeda, sehingga belum mampu mengukur kompetensi peserta didik secara maksimal.

Penelitian yang dilakukan oleh Auliya et al. (2025), menunjukkan bahwa meskipun soal telah disusun berdasarkan kisi-kisi, masih ditemukan butir soal yang tidak valid serta memiliki daya pembeda yang rendah, yang pada akhirnya dapat memengaruhi keakuratan hasil evaluasi. Temuan serupa juga diungkapkan oleh Basri et al. (2021), yang menyatakan bahwa masih terdapat sejumlah soal yang belum mampu membedakan kemampuan peserta didik secara efektif. Hal ini menunjukkan bahwa penyusunan instrumen tes tidak hanya membutuhkan pemahaman teoritis, tetapi juga keterampilan praktis dalam merancang soal yang berkualitas.

Dalam konteks pendidikan tinggi, khususnya pada program studi Pendidikan Ekonomi, kemampuan

menyusun instrumen evaluasi yang baik merupakan kompetensi penting yang harus dimiliki oleh mahasiswa sebagai calon pendidik. Mahasiswa tidak hanya dituntut memahami konsep evaluasi pembelajaran, tetapi juga harus mampu mengimplementasikannya dalam praktik nyata. Salah satu bentuk implementasi tersebut adalah melalui tugas proyek pada mata kuliah Evaluasi Hasil Belajar, di mana mahasiswa diberi kesempatan untuk menyusun instrumen tes secara mandiri berdasarkan kisi-kisi yang telah ditentukan.

Namun demikian, instrumen yang dihasilkan dalam kegiatan tugas proyek tersebut pada umumnya belum melalui proses analisis kualitas butir soal secara komprehensif. Instrumen tersebut lebih sering digunakan sebagai alat penilaian akademik semata, tanpa dilakukan pengujian lebih lanjut terhadap kualitas setiap butir soal. Kondisi ini menimbulkan pertanyaan penting mengenai sejauh mana kualitas instrumen yang disusun oleh mahasiswa tersebut telah memenuhi standar evaluasi pembelajaran yang baik. Tanpa adanya analisis yang sistematis, kualitas butir soal tidak dapat

diketahui secara pasti, sehingga berpotensi menghasilkan evaluasi yang kurang akurat.

Jika ditinjau dari kajian penelitian terdahulu, sebagian besar penelitian tentang analisis butir soal masih berfokus pada instrumen evaluasi yang digunakan oleh guru di tingkat sekolah. Penelitian-penelitian tersebut umumnya bertujuan untuk mengevaluasi kualitas soal dalam konteks pembelajaran formal di sekolah, serta memberikan rekomendasi perbaikan terhadap instrumen yang digunakan. Sementara itu, penelitian yang mengkaji instrumen evaluasi yang disusun oleh mahasiswa dalam konteks pembelajaran di perguruan tinggi, khususnya melalui pendekatan tugas proyek, masih relatif terbatas.

Dari aspek metodologis, sebagian besar penelitian sebelumnya cenderung menggunakan perangkat lunak tertentu seperti Anates dalam melakukan analisis butir soal. Pemanfaatan aplikasi SPSS sebagai alat analisis dalam mengevaluasi kualitas butir soal, khususnya pada konteks simulasi tugas proyek mahasiswa, masih belum banyak dieksplorasi secara mendalam.

Padahal, SPSS memiliki kemampuan analisis statistik yang lebih luas dan fleksibel, sehingga berpotensi memberikan hasil analisis yang lebih komprehensif.

Berdasarkan uraian tersebut, dapat dilihat bahwa terdapat kesenjangan penelitian yang terletak pada masih terbatasnya kajian mengenai kualitas instrumen evaluasi yang disusun oleh mahasiswa serta belum optimalnya pemanfaatan SPSS sebagai alat analisis dalam mengevaluasi kualitas butir soal. Oleh karena itu, penelitian ini dilakukan untuk menganalisis kualitas butir soal pilihan ganda yang disusun oleh mahasiswa Pendidikan Ekonomi melalui simulasi tugas proyek ditinjau dari aspek tingkat kesukaran, daya pembeda, validitas, reliabilitas, dan efektivitas pengecoh menggunakan aplikasi SPSS. Penelitian ini juga bertujuan untuk memberikan rekomendasi perbaikan terhadap butir soal yang belum memenuhi kriteria kualitas yang baik.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis

penelitian deskriptif. Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada pengukuran dan analisis data numerik untuk mengetahui kualitas butir soal secara objektif melalui teknik statistik. Menurut Sugiyono (2017), penelitian deskriptif kuantitatif bertujuan untuk mendeskripsikan fenomena secara sistematis berdasarkan data yang diperoleh di lapangan.

Subjek dalam penelitian ini adalah 35 mahasiswa Program Studi Pendidikan Ekonomi Universitas Negeri Medan yang terlibat sebagai responden dalam uji coba instrumen tes. Adapun objek penelitian ini adalah butir soal pilihan ganda yang disusun oleh mahasiswa melalui tugas proyek pada mata kuliah Evaluasi Hasil Belajar, yang terdiri dari 20 butir soal pada materi kegagalan pasar.

Data dalam penelitian ini diperoleh dari hasil uji coba instrumen yang dilakukan secara daring melalui *Google Form*. Instrumen tes disusun terlebih dahulu oleh mahasiswa berdasarkan kisi-kisi yang mengacu pada indikator pencapaian kompetensi, kemudian diujikan kepada responden untuk memperoleh data berupa jawaban peserta tes. Teknik pengumpulan data

menggunakan metode dokumentasi, yaitu dengan mengumpulkan data berupa naskah soal, kunci jawaban, serta hasil respon mahasiswa dari *Google Form*. Data tersebut selanjutnya dianalisis untuk mengetahui kualitas setiap butir soal. Teknik analisis data dilakukan menggunakan bantuan aplikasi SPSS. Menurut Ghozali (2016), SPSS merupakan perangkat lunak statistik yang mampu membantu peneliti dalam melakukan analisis data secara cepat dan tepat, khususnya dalam penelitian kuantitatif. Analisis yang dilakukan meliputi uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil analisis tersebut kemudian diinterpretasikan untuk menentukan kualitas butir soal, apakah termasuk kategori baik, perlu direvisi, atau tidak layak digunakan.

C. Hasil Penelitian dan Pembahasan

Analisis butir soal merupakan tahapan penting dalam proses evaluasi penggunaan instrumen tes karena hasil analisis tersebut dapat dijadikan dasar dalam melakukan perbaikan dan penyempurnaan soal agar kualitasnya semakin baik (Ashayidatina and Alber 2023).

Analisis ini berperan dalam menilai mutu instrumen penilaian secara menyeluruh dengan menyediakan informasi empiris terkait tingkat ketepatan, konsistensi, serta efektivitas soal dalam mengukur kompetensi peserta didik. Berdasarkan hasil analisis butir soal yang dilakukan menggunakan IBM SPSS Versi 25 diperoleh gambaran kualitas soal baik secara kualitatif maupun kuantitatif yang meliputi aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas distraktor atau pengecoh.

1. Uji Validitas

Validitas menunjukkan tingkat ketepatan suatu tes dalam mengukur kemampuan atau aspek yang seharusnya diukur. Data atau informasi dapat dinyatakan valid apabila sesuai dengan kondisi nyata atau fakta yang sebenarnya (Saputri et al. 2023). Suatu tes dikatakan memiliki validitas yang baik apabila instrumen tersebut mampu mengukur tujuan yang telah ditetapkan secara tepat. Butir soal dinyatakan valid apabila nilai Sig. (2-tailed) < 0,05 (Rosadi, Basuki, and Hadi 2022).

Tabel 1 Uji Validitas

No Soal	Sig. (2 Tailed)	Kesimpulan
1.	0,110	Tidak Valid
2.	0,275	Tidak Valid
3.	0,000	Valid
4.	0,000	Valid
5.	0,000	Valid
6.	0,003	Valid
7.	0,000	Valid
8.	0,000	Valid
9.	0,275	Tidak Valid
10.	0,012	Valid
11.	0,087	Tidak Valid
12.	0,000	Valid
13.	0,023	Valid
14.	0,000	Valid
15.	0,000	Valid
16.	0,016	Valid
17.	0,043	Valid
18.	0,137	Tidak Valid
19.	0,002	Valid
20.	0,002	Valid

Berdasarkan hasil analisis validitas butir soal diperoleh 15 soal atau 75 % dinyatakan Valid dan 5 Soal atau 25 % dinyatakan tidak valid sesuai dengan Syarat. Butir soal yang dinyatakan valid terdiri dari 3,4,5,6,7,8,10,12,13,14,15,16,17,19,20. Sedangkan untuk butir soal yang tidak valid terdiri dari 1,2,9,11,18. Temuan ini mengindikasikan soal mampu mengukur kompetensi yang seharusnya diukur sesuai dengan tujuan pembelajaran. Hasil penelitian ini sejalan dengan penelitian Auliya et al. (2025) dan Nafs et al. (2023), yang menegaskan bahwa soal dominan valid sehingga memiliki kualitasnya yang baik dari segi validitasnya. Hal ini diperkuat oleh Arifin (2017), yang

menyatakan suatu instrumen dikatakan valid apabila mampu mengukur apa yang diinginkan. Khasanah et al. (2023) juga menyatakan bahwa butir soal yang memiliki nilai validitas tinggi secara jelas menunjukkan kemampuan dalam mengukur keterampilan siswa dengan tepat dan akurat. Soal yang valid (75%) berarti butir soal tersebut sudah dapat menjalankan fungsinya yaitu mengukur apa yang dapat diukur. Sedangkan soal yang tidak valid (25%) dapat disebabkan dari beberapa faktor. Terdapat tiga faktor yang mempengaruhi validitas hasil tes yaitu faktor instrumen yang digunakan untuk tes, faktor administrasi dan penskoran, serta faktor dari jawaban siswa. Ramadhan et al. (2023), menyatakan bahwa penggunaan butir soal yang tidak valid berpotensi menghasilkan data penilaian yang tidak akurat dan bias, sehingga instrumen penilaian tersebut tidak mampu mengukur kompetensi yang seharusnya diukur.

Dari penjelasan diatas, diketahui bahwa soal Pilihan Ganda pada materi kegagalan pasar untuk mahasiswa Program Studi Pendidikan Ekonomi Universitas Negeri Medan

memiliki kualitas baik dari segi validitas.

2. Uji Reliabilitas

Reliabilitas tes digunakan untuk mengukur tingkat konsistensi suatu instrumen penilaian sehingga hasil yang diperoleh menunjukkan kestabilan (Khasanah et al. 2023). Setelah butir soal dinyatakan valid, tahap berikutnya adalah melakukan uji reliabilitas. Reliabilitas yang dihitung pada penelitian ini merupakan reliabilitas komposit. Pengujian reliabilitas terhadap butir soal pilihan ganda yang telah valid dilakukan dengan bantuan perangkat lunak IBM Statistics 25 for Windows, kemudian hasilnya dikonversikan menggunakan rumus alpha strata untuk memperoleh nilai reliabilitas komposit dari kedua jenis soal tersebut. Model yang digunakan adalah Alpha Cronbach dengan kriteria keputusan bahwa nilai koefisien alpha $> 0,70$ menunjukkan reliabilitas berada pada kategori baik (Fiska et al. 2021).

Tabel 2 Uji Reliabilitas

Reliability Statistics	
Cronbach's Alpha	N of Items
0,802	16

Berdasarkan hasil analisis Reabilitas butir soal nilai Cronbach's

Alpha sebesar 0,802, Hal tersebut mengindikasikan bahwa nilai diperoleh memenuhi syarat yaitu Cronbach's Alpha $> 0,70$. Oleh karena itu dapat disimpulkan bahwa butir soal tersebut terbukti reliabel. Hasil ini juga menandakan bahwa instrumen memiliki tingkat reliabilitas yang tinggi. Hal tersebut sejalan dengan pendapat Soesana et al. (2023), yang menyatakan bahwa koefisien reliabilitas sebesar 0,80 atau lebih menunjukkan bahwa tes memiliki reliabilitas tinggi. Temuan ini juga didukung oleh penelitian Basri et al. (2021), yang menjelaskan bahwa instrumen dengan nilai reliabilitas di atas 0,70 termasuk dalam kategori reliabilitas tinggi.

Perhitungan reliabilitas tersebut menunjukkan bahwa instrumen penilaian dapat dipercaya untuk mengukur kemampuan peserta didik. Artinya, skor yang dihasilkan relatif konsisten dan stabil meskipun pengujian dilakukan kembali dalam kondisi yang serupa.

3. Tingkat Kesukaran

Tingkat kesukaran butir soal merupakan indikator yang digunakan untuk mengetahui kualitas suatu soal apakah termasuk kategori mudah,

sedang, atau sukar. Suatu soal dikategorikan mudah apabila peserta didik mampu menjawabnya dengan benar sedangkan soal dikatakan sukar apabila sebagian besar peserta didik tidak mampu memberikan jawaban yang benar. Terdapat Kriteria pengambilan Keputusan dalam menentukan Tingkat kesukaran soal (Zahiroh 2020).

Kriteria Pengambilan Keputusan:

- Jika $> 0,70$ maka soal dinyatakan Mudah.
- Jika $0,30 - 0,70$ maka soal dinyatakan Sedang.
- Jika $< 0,30$ maka soal dinyatakan Sukar.

Tabel 3 Tingkat Kesukaran

No Soal	Mean	Kesimpulan
1.	0,91	Mudah
2.	0,97	Mudah
3.	0,97	Mudah
4.	0,97	Mudah
5.	0,89	Mudah
6.	0,74	Mudah
7.	0,69	Sedang
8.	0,91	Mudah
9.	0,97	Mudah
10.	0,66	Sedang
11.	0,54	Sedang
12.	0,86	Mudah
13.	0,74	Mudah
14.	0,86	Mudah
15.	0,71	Mudah
16.	0,91	Mudah
17.	0,66	Sedang
18.	0,89	Mudah
19.	0,89	Mudah
20.	0,77	Mudah

Berdasarkan hasil analisis butir diklasifikasikan sesuai kategori tingkat

kesukaran, dari total 20 butir soal ditemukan bahwa terdapat 80% butir soal dengan kategori mudah dan 20 % butir soal dengan kategori sedang. Dari hasil tersebut diperoleh bahwa tidak terdapat keseimbangan antar soal karena tidak terdapat soal yang sukar. Suatu instrumen penilaian dapat dikatakan baik apabila terdapat keseimbangan antara soal sehingga mampu mengukur kemampuan peserta didik secara optimal (Saputri et al. 2023). Ambarwati & Ismiyati (2021), juga menyatakan bahwa soal yang berkualitas adalah soal yang tidak terlalu mudah maupun terlalu sulit. Soal yang terlalu mudah tidak mendorong peserta didik untuk meningkatkan kemampuan pemecahan masalah sedangkan soal yang terlalu sulit dapat menurunkan motivasi belajar karena dianggap sulit untuk dicapai.

Temuan ini menunjukkan bahwa tingkat kesukaran soal masih didominasi oleh kategori mudah sehingga perlu dilakukan revisi untuk meningkatkan variasi tingkat kesukaran. Dominasi soal mudah sebesar 80% mengindikasikan adanya kelemahan dalam perancangan instrumen penilaian. Tingkat kesukaran yang kurang

bervariasi tidak hanya menyebabkan instrumen gagal membedakan kemampuan peserta didik tetapi juga berpotensi menurunkan motivasi belajar. Instrumen evaluasi yang ideal seharusnya mampu merepresentasikan keragaman tingkat kognitif serta memberikan tantangan yang seimbang bagi peserta didik dengan kemampuan yang berbeda-beda.

4. Daya Pembeda

Daya pembeda adalah pengukuran sejauh mana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu. Untuk menentukan baik tidaknya butir soal dilihat dari indeks daya beda terhadap beberapa acuan yang dapat digunakan. Jika dibuat suatu rentangan, interpretasi indeks daya pembeda suatu butir tes adalah sebagai berikut:

Tabel 4 Kriteria Daya Pembeda

Index Daya Beda	Kriteria Butir
$0,40 < D < 1,0$	Sangat Baik
$0,3 < D < 0,4$	Baik
$0,2 < D < 0,3$	Cukup
$0,0 < D < 0,2$	Tidak Baik
$D < 0,0$	Sangat Tidak Baik

Sumber: Iskandar & Rizal (2018)

Tabel 5 Daya Pembeda

No Soal	Pearson correlation	Kesimpulan
1.	0,275	Cukup
2.	0,190	Tidak Baik
3.	0,568	Sangat Baik
4.	0,568	Sangat Baik
5.	0,737	Sangat Baik
6.	0,487	Sangat Baik
7.	0,575	Sangat Baik
8.	0,628	Sangat Baik
9.	0,190	Tidak Baik
10.	0,420	Sangat Baik
11.	0,293	Cukup
12.	0,709	Sangat Baik
13.	0,384	Baik
14.	0,581	Sangat Baik
15.	0,581	Sangat Baik
16.	0,403	Sangat Baik
17.	0,345	Baik
18.	0,256	Cukup
19.	0,511	Sangat Baik
20.	0,495	Sangat Baik

Berdasarkan hasil analisis butir soal dan diklasifikasikan sesuai kategori tingkat kesukaran, dari total 20 butir soal Pilihan Ganda ditemukan bahwa terdapat 65% butir soal dengan kategori sangat baik, 10% butir soal dengan kategori baik, 15% butir soal dengan kategori Cukup, dan 10% butir soal dengan kategori tidak baik. Menurut Waminton (2015), daya pembeda soal merupakan kemampuan suatu butir soal dalam membedakan peserta didik yang telah menguasai materi dengan peserta didik yang belum menguasainya. Pernyataan tersebut sejalan dengan pendapat Sudijono (2015), yang menegaskan bahwa daya pembeda pada setiap butir soal sangat penting

untuk diperhatikan karena kemampuan peserta didik bersifat beragam sehingga penyusun soal perlu memastikan bahwa setiap item mampu mengidentifikasi perbedaan kemampuan tersebut.

Hasil analisis menunjukkan bahwa sebagian butir soal masih belum optimal dalam membedakan tingkat kemampuan siswa. Butir soal yang berada pada kategori cukup dan tidak baik cenderung kurang dalam menangkap perbedaan kemampuan peserta didik sehingga dapat memengaruhi ketepatan skor dalam merepresentasikan kompetensi sebenarnya. Oleh karena itu, butir soal yang termasuk kategori cukup dan tidak baik perlu dilakukan revisi agar memiliki sensitivitas yang lebih baik terhadap variasi kemampuan mahasiswa.

5. Distraktor/Efektivitas Pengecoh

Analisis distraktor (pengecoh) merupakan proses penilaian terhadap jawaban yang dirancang sehingga berpotensi membuat peserta didik keliru dalam menentukan jawaban pada soal tes (Firdaus and Sri Rahayu 2020). Distraktor memiliki fungsi untuk mengalihkan perhatian peserta didik yang belum memahami materi dengan

baik sehingga mereka tidak langsung memilih jawaban yang benar (Radja, Bano, and Ina 2023). Suatu pengecoh dikatakan efektif apabila mampu menarik perhatian peserta didik yang kurang menguasai konsep, menimbulkan keraguan, dan akhirnya dipilih sebagai jawaban oleh peserta didik tersebut.

Tabel 6 Kriteria Penilaian Pengecoh

Pengecoh yang berfungsi	Kriteria
3	Sangat Baik
2	Baik
1	Kurang Baik
0	Tidak Baik

Sumber: Auliya et al. (2025)

Tabel 7 Penilaian Pengecoh

Kategori	Pengecoh yang berfungsi	Jumlah Soal	Persentase Pengecoh
Sangat Baik	3	5	25%
Baik	2	4	20%
Kurang Baik	1	6	30%
Tidak Baik	0	5	25%

Efektivitas pengecoh yang baik ditandai dengan kemampuannya mengalihkan perhatian peserta didik dari jawaban yang benar. Suatu distraktor dikatakan efektif apabila dipilih minimal 5% oleh peserta didik, hal ini akan menunjukkan bahwa pilihan tersebut cukup meyakinkan serta mampu membedakan tingkat pemahaman peserta didik. Hal ini diperkuat oleh pendapat Sudijono

(2015), yang menyatakan bahwa distraktor dianggap berfungsi dengan baik apabila sekurang-kurangnya dipilih oleh 5% dari keseluruhan peserta tes.

Berdasarkan hasil analisis, dari keseluruhan 20 butir soal yang dianalisis terdapat sebanyak 5 butir soal (25%) berada pada kategori Sangat Baik, hal ini menunjukkan bahwa pengecoh pada soal-soal tersebut mampu berfungsi secara optimal karena dipilih oleh peserta didik dan dapat membedakan kemampuan siswa. Selanjutnya, 4 butir soal (20%) termasuk kategori Baik, artinya distraktor masih berfungsi dan dapat menilai kemampuan siswa. Sementara itu, jumlah terbesar terdapat pada kategori kurang baik, yaitu 6 butir soal (30%), yang mengindikasikan bahwa sebagian besar distraktor kurang menarik bagi peserta didik atau tidak berfungsi sebagaimana mestinya. Adapun 5 butir soal (25%) masuk dalam kategori tidak baik, Hal ini menunjukkan bahwa distraktor pada soal-soal tersebut tidak dipilih sama sekali atau sangat lemah dalam mengecoh peserta didik. Dari hasil analisis Distraktor tersebut dapat disimpulkan bahwa butir soal tersebut

sudah baik dan berfungsi secara optimal.

E. Kesimpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa kualitas butir soal pilihan ganda yang disusun oleh mahasiswa Pendidikan Ekonomi melalui simulasi tugas proyek secara umum sudah cukup baik, terutama dari aspek validitas (75% valid) dan reliabilitas yang tinggi (*Cronbach's Alpha* = 0,802). Hal ini menunjukkan bahwa sebagian besar instrumen telah mampu mengukur kompetensi secara tepat dan konsisten. Namun demikian, masih terdapat beberapa kelemahan pada aspek tingkat kesukaran yang didominasi oleh kategori mudah, serta pada daya pembeda dan efektivitas pengecoh yang belum sepenuhnya optimal.

Oleh karena itu, diperlukan perbaikan dalam penyusunan soal, khususnya dalam menyeimbangkan tingkat kesukaran, meningkatkan kemampuan soal dalam membedakan peserta didik, serta merancang distraktor yang lebih berfungsi. Untuk penelitian selanjutnya, disarankan menggunakan jumlah sampel yang lebih luas dan mengembangkan analisis dengan pendekatan yang

lebih komprehensif agar diperoleh gambaran kualitas instrumen yang lebih mendalam.

DAFTAR PUSTAKA

- Ambarwati, Y. F., and I. Ismiyati. 2021. "Analisis Butir Soal Pilihan Ganda Ulangan Akhir Semester Genap Mata Pelajaran Kearsipan." *Measurement In Educational Research (Meter)* 1(2):64–75.
- Arifin, Z. 2017. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Arikunto, Suharsimi. 2018. *Dasar-Dasar Evaluasi Pendidikan*. Ed Ketiga. Jakarta: Bumi Aksara.
- Ashayidatina, L., and Alber. 2023. "Analisis Butir Soal Ujian Semester Ganjil Bahasa Indonesia Kelas X SMK Negeri 1 Mempura." *Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 8:131–37.
doi:<https://dx.doi.org/10.26737/jp-bsi.v8i2.5039>.
- Auliya, Annisa Syufa, Nurul Hikmah, Ratna Yulis Tyaningsih, and Muhammad Turmuzi. 2025. "Analisis Kualitas Butir Soal Asesmen Tengah Semester Genap Mata Pelajaran Matematika Kurikulum Merdeka Kelas VII SMP Negeri 2 Kuripan Tahun Ajaran 2024/2025." *Pendas : Jurnal Ilmiah Pendidikan Dasar* 10(2):210–24.
- Basri, Khairil, Baidowi Baidowi, Junaidi Junaidi, and Muhammad Turmuzi. 2021. "Analisis Butir Soal Ulangan Semester Ganjil Mata Pelajaran Matematika Kelas VIII SMP." *Griya Journal of Mathematics Education and Application* 1(4):682–94.
doi:10.29303/griya.v1i4.107.
- Elizanoviyanti, Fitrialin, and Asep Saepurokhman. 2025. "Analisis Tingkat Validitas Dan Reliabilitas Butir Soal Penilaian Akhir Semester Mata Pelajaran Bahasa Indonesia Kelas XI SMA Negeri Situraja." *Literat: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 4(2):168–76.
- Firdaus, D., and Y. Sri Rahayu. 2020. "Pengembangan Instrumen Tes Miskonsepsi Menggunakan Four-Tier Multiple-Choice Test Submateri Katabolisme Karbohidrat Kelas XII SMA." 9(2).
- Fiska, J. M., Y. Hidayati, N. Qomaria, and W. P. Hadi. 2021. "Analisis Butir Soal Ulangan Harian Ipa Menggunakan Software Anates Pada Pendekatan Teori Tes Klasik." *Natural Science Education Research* 4(1):65–76.
doi:<https://doi.org/10.21107/nser.v4i1.8133>.
- Ghozali, I. 2016. *Aplikasi Multivariate Dengan Program IBM SPSS 23.0*. Edisi 8. Semarang: Badan Penerbit Universitas Diponegoro.
- Iskandar, A., and M. Rizal. 2018. "Analisis Kualitas Soal Di Perguruan Tinggi Berbasis Aplikasi TAP." *Jurnal Penelitian Dan Evaluasi Pendidikan* 22(1):12–23.
- Khasanah, I., A. Fuady, and Sunismi. 2023. "Analisis Soal Ulangan Harian Matematika Bentuk Pilihan Ganda." *MATHEMA JOURNAL* 5(2):110–25.
doi:<https://doi.org/10.33365/jm.v5i2.2818>.

- Nafs, H., N. Sridana, N. Hikmah, and H. Soeprianto. 2023. "Analisis Kualitas Butir Soal Ulangan Akhir Semester Genap Mata Pelajaran Matematika Kelas VII SMPN 6 Mataram Tahun Ajaran 2022/2023." *Jurnal Ilmiah Profesi Pendidikan* 2324–31.
- Radja, Sofiani Putri, Vidriana Oktoviana Bano, and Anita Tamu Ina. 2023. "Analisis Kualitas Butir Soal Hasil Belajar Peserta Didik Berdasarkan Tingkat Kesukaran, Daya Beda Dan Efektivitas Pengecoh Di SMAN 1 Pandawai." *Jurnal Edusavana* 30–41.
- Ramadhan, Wandri, Fildza Malahati, Kharisma Romadhon, and Syahrul Ramadhan. 2023. "Analisis Butir Soal Tipe Multiple Choice Questions Pada Penilaian Harian Sekolah Dasar." *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan Dan Pembelajaran* 10(2):93–105.
doi:10.21093/twt.v10i2.6155.
- Rosadi, Mohammad Munib, Basuki, and Fajar Satriya Hadi. 2022. "Pengaruh Pembelajaran Menggunakan Trainer Kopling Terhadap Hasil Belajar Mahasiswa Teknik Mesin Universitas Hasyim Asy'ari." *Jurnal Pendidikan Dan Konseling* 4(5):1163–71.
- Saputri, Hera Aoriliana, Zulhijrah, Nabila Joti Larasati, and Shaleh. 2023. "Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkatkesukaran Dan Daya Beda Butir Soal." *Didatik: Jurnal Ilmiah PGSD FKIP Universitas Mandiri* 09(05):2986–95.
- Soesana, S., H. Subakti, Karwanto, A. Fitri, S. Kuswandi, L. Sastri, I. Falani, N. Aswan, F. A. Hasibuan, and H. Lestari. 2023. *Metodologi Penelitian Kuantitatif*. Medan: Yayasan Kita Menulis.
- Sudijono, Anas. 2015. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Sudjana, Nana. 2017. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya.
- Sugiyono. 2017. *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, Dan R & D*. Bandung: Alfabeta.
- Waminton, R. 2015. *Evaluasi Hasil Belajar Matematika*. Yogyakarta: Media Akademi.
- Zahiroh, Ulfah. 2020. "Analisis Kualitas Butir Soal Pilihan Ganda Mata Pelajaran Kimia Pada Ujian Akhir Semester (UAS) Kelas XI MAN 2 Kepulauan Meranti." Universitas Islam Negeri Sultan Syarif Kasim Riau.