



## **Analisis Kualitas Butir Soal Pilihan Ganda pada Mata Pelajaran Konsentrasi Keahlian**

**<sup>1</sup>Nofiri Gea\*, <sup>1</sup>Paras Patikasari, <sup>1</sup>Deswita Maharani Padang, <sup>1</sup>Dina Ampera, <sup>1</sup>Untung Desy Purnamasari**

<sup>1</sup>Universitas Negeri Medan, Deli Serdang, Sumatera Utara, Indonesia

DOI: <https://doi.org/10.70115/semesta.v4i1.448>

<b>Article Info</b>	<b>Abstract</b>
<p><b>Article History</b> Received: February 20, 2026 Accepted: May 3, 2026 Published: May 20, 2026</p> <p><b>Keywords</b> Item analysis; multiple-choice test; difficulty level; discrimination index; distractor effectiveness; learning evaluation</p>	<p><i>This study aims to analyze the quality of multiple-choice test items. This research employed a descriptive quantitative approach involving 30 students of SMK Negeri 3 Tebing Tinggi as the research subjects. The object of the study consisted of 25 multiple-choice items. The data were collected through tests and documentation, then analyzed quantitatively to determine the quality of each test item. The results of the difficulty level analysis showed that 10 items were difficult, moderate items consisting of 13 items, and 2 easy items. The discrimination index analysis indicated that most items were categorized as very good, with 17 items in this category, while 2 items were categorized as good, 4 items as sufficient, and 2 items as poor. Meanwhile, the distractor effectiveness analysis showed that 17 items had very good distractors, while 8 items required revision due to ineffective answer options. Based on the findings, it can be concluded that item analysis is essential to ensure the quality of evaluation instruments, so that the scoring and assessment of students' learning outcomes can be conducted more objectively, accurately, and fairly.</i></p>
<b>Informasi Artikel</b>	<b>Abstrak</b>
<p><b>Kata kunci</b> Analisis butir soal; pilihan ganda; tingkat kesukaran; daya pembeda; efektivitas pengecoh; evaluasi pembelajaran</p> <p><b>Corresponding Author</b> Nofiri Gea Universitas Negeri Medan, Indonesia *E-mail: nofirig@gmail.com</p>	<p>Penelitian ini bertujuan untuk menganalisis kualitas butir soal pilihan ganda. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan subjek sebanyak 30 siswa SMK Negeri 3 Tebing Tinggi. Objek penelitian berupa 25 butir soal pilihan ganda. Data diperoleh melalui teknik tes dan dokumentasi, kemudian dianalisis secara kuantitatif. Hasil analisis tingkat kesukaran menunjukkan bahwa soal kategori sedang dengan jumlah 13 butir, diikuti 10 soal kategori sukar dan 2 soal kategori mudah. Analisis daya pembeda menunjukkan bahwa sebagian besar soal berada pada kategori sangat baik, yaitu sebanyak 17 soal, sedangkan 2 soal berkategori baik, 4 soal cukup, dan 2 soal berkategori jelek. Sementara itu, analisis efektivitas pengecoh menunjukkan bahwa 17 soal memiliki pengecoh yang sangat baik, sedangkan 8 soal masih memerlukan perbaikan karena terdapat pilihan jawaban yang tidak berfungsi secara efektif. Berdasarkan hasil penelitian, dapat disimpulkan bahwa analisis butir soal penting dilakukan untuk memastikan instrumen evaluasi memiliki kualitas yang baik sehingga proses penskoran dan penilaian hasil belajar siswa menjadi lebih objektif, akurat, dan adil.</p>
 <p>This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.</p>	

Copyright ©2026 Nofiri Gea, et. al.

## PENDAHULUAN

Pendidikan merupakan fondasi utama untuk pembangunan suatu negara, dan peran guru sangat penting untuk menghasilkan generasi penerus yang berkualitas. Dalam situasi seperti ini, kinerja guru menjadi komponen strategis penting dalam pengelolaan sumber daya manusia pendidikan. Kinerja guru mempengaruhi hasil belajar siswa selain menunjukkan seberapa efektif sistem pendidikan secara keseluruhan. Pendidikan berperan penting sebagai fondasi suatu bangsa karena berfungsi membekali peserta didik dengan pengetahuan, keterampilan, dan nilai-nilai karakter yang dibutuhkan untuk menghadapi suatu dinamika perkembangan zaman (Thinking, 2025)

Seorang guru dikatakan memiliki kinerja yang tinggi apabila mereka memahami apa artinya kinerja sesungguhnya dan dapat dibuktikan dengan melakukan apa yang mereka lakukan saat mengerjakan tugas mereka. Jika guru melakukan hal-hal berikut dalam pembelajaran yang baik: menguasai dan mengembangkan bahan pelajaran, disiplin dalam mengajar, kreatif dalam mengajar, bekerja sama dengan semua warga sekolah, memiliki contoh yang patut diteladani siswa, dan objektif dalam membimbing dan menilai siswa (Kusumaningrum et al., 2024). Pendidikan pada hakikatnya merupakan proses pembentukan manusia secara utuh yang mencakup aspek kognitif, afektif, dan psikomotorik (Creswell & Creswell, 2022).

Guru adalah pendidik profesional yang memiliki tugas, fungsi, dan peran penting dalam mencerdaskan kehidupan bangsa. Kinerja mengajar guru, di sisi lain, adalah kemampuan seorang guru untuk menyelesaikan tugas sesuai dengan tujuan yang telah ditetapkan. Kehadiran guru dalam proses pembelajaran di sekolah masih sangat penting. Pekerjaan ini tidak dapat diganti atau diambil alih oleh orang lain. Sebagai bagian penting dari proses pembelajaran, kinerja guru sangat penting dalam pengelolaan sumber daya manusia. Tidak hanya kemampuan mengajar, kinerja guru mencakup profesionalisme, komitmen, motivasi, dan kemampuan untuk beradaptasi dengan kemajuan pendidikan dan teknologi. Kinerja guru sebagai penggerak perubahan di ruang kelas sangat penting untuk menetapkan dan mencapai tujuan pendidikan nasional (Asiah, 2016).

Guru pasti ingin mengetahui hasil setelah melakukan proses pembelajaran. Evaluasi diri secara jujur, objektif, dan menyeluruh adalah salah satu cara untuk melakukannya. Hal ini dimaksudkan agar guru dapat segera mengetahui kesalahan dalam pelaksanaan pembelajaran dan berusaha memperbaikinya untuk pembelajaran berikutnya (Febriana, 2021). Salah satu jenis penilaian yang dapat dilakukan guru dalam pembelajaran adalah penilaian diagnostik, yang berfungsi untuk mengidentifikasi faktor-faktor Penyebab Kegagalan dan Pendukung Keberhasilan dalam Pembelajaran. Jika guru tidak tahu faktor-faktor ini, akan sulit bagi mereka untuk melakukan perbaikan untuk meningkatkan kualitas pembelajaran (Fatkhan, 2025).

Salah satu bentuk evaluasi yang sering digunakan dalam proses pembelajaran adalah tes tertulis, khususnya dalam bentuk soal pilihan ganda. Tes pilihan ganda digunakan karena dianggap praktis, efisien, dan mampu mengukur pemahaman siswa dalam jumlah yang banyak. Selain itu, bentuk soal ini juga mempermudah guru dalam proses menskor dan menilai hasil belajar siswa. Kegiatan menskor dan menilai merupakan bagian penting dalam evaluasi pembelajaran karena hasil skor yang diperoleh siswa akan menentukan tingkat pencapaian hasil belajar mereka (Juwita et al., 2026).

Menskor merupakan kegiatan memberikan angka atau nilai terhadap jawaban siswa berdasarkan pedoman penilaian yang telah ditentukan. Dalam tes objektif seperti pilihan ganda, proses penskoran biasanya dilakukan dengan memberikan skor pada jawaban benar dan tidak memberikan skor pada jawaban salah. Hasil penskoran tersebut kemudian digunakan sebagai dasar dalam kegiatan penilaian. Penilaian adalah proses menginterpretasikan hasil skor untuk mengetahui tingkat kemampuan atau prestasi belajar siswa. Oleh karena itu, kegiatan menskor dan menilai harus dilakukan secara tepat agar hasil evaluasi benar-benar mencerminkan kemampuan siswa yang sebenarnya (Juwita et al., 2026).

Dalam analisis butir soal terdapat beberapa aspek penting yang perlu diperhatikan, yaitu tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Tingkat kesukaran digunakan untuk mengetahui apakah suatu soal tergolong mudah, sedang, atau sukar. Tingkat kesukaran yang sesuai akan membantu guru dalam memberikan skor dan penilaian yang lebih objektif. Daya pembeda digunakan untuk mengetahui kemampuan soal dalam membedakan siswa yang memiliki kemampuan tinggi dan siswa yang memiliki kemampuan rendah. Sementara itu, efektivitas pengecoh digunakan untuk mengetahui apakah pilihan jawaban salah dapat berfungsi dengan baik dalam mengecoh siswa yang belum memahami materi (Sahra et al., 2026).

Kegiatan analisis butir soal memiliki hubungan yang sangat erat dengan proses menskor dan menilai hasil belajar siswa. Soal yang berkualitas baik akan menghasilkan skor yang lebih akurat sehingga penilaian terhadap kemampuan siswa juga menjadi lebih tepat. Sebaliknya, apabila soal yang digunakan memiliki kualitas rendah, maka hasil skor dan penilaian siswa dapat menjadi tidak objektif. Oleh karena itu, analisis butir soal perlu dilakukan agar instrumen evaluasi yang digunakan benar-benar mampu mengukur kemampuan siswa secara valid dan reliabel (Kadir, 2015). Keterampilan berpikir kritis merupakan salah satu kompetensi esensial yang harus dimiliki peserta didik dalam menghadapi tantangan pembelajaran abad ke-21. Kemampuan ini tidak hanya mencakup proses memahami informasi, tetapi juga melibatkan aktivitas kognitif tingkat tinggi seperti menganalisis, mengevaluasi, serta menarik kesimpulan secara logis dan sistematis berdasarkan bukti yang tersedia (Nisa & Chusni, 2026). Rapor pendidikan merupakan alat yang dapat dijadikan bahan evaluasi kualitas satuan pendidikan dan daerah. Rapor Pendidikan diperoleh melalui asesmen nasional dan survey karakter dan survey lingkungan belajar (Ardiansyah, 2023).

Berdasarkan permasalahan tersebut, penelitian ini dilakukan untuk menganalisis butir soal pilihan ganda dalam kaitannya dengan proses menskor dan menilai hasil belajar siswa. Analisis dilakukan untuk mengetahui kualitas soal berdasarkan tingkat kesukaran, daya pembeda, dan efektivitas pengecoh sehingga dapat diketahui apakah soal tersebut layak digunakan sebagai alat evaluasi pembelajaran. Hasil penelitian ini diharapkan dapat membantu guru dalam meningkatkan kualitas instrumen evaluasi agar proses penskoran dan penilaian hasil belajar siswa menjadi lebih efektif, objektif, dan akurat (Sahra et al., 2026).

## METODE

Penelitian ini menerapkan pendekatan kuantitatif deskriptif. Pendekatan kuantitatif digunakan karena data yang diolah berbentuk angka, yaitu hasil jawaban siswa pada tes yang telah diberikan. Sementara itu, pendekatan deskriptif digunakan untuk menjelaskan kualitas setiap butir soal berdasarkan hasil analisis tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Penelitian ini merupakan studi pendahuluan dengan pendekatan deskriptif kualitatif-kuantitatif, yang berfokus pada pemetaan profil kemampuan pemecahan masalah (Hanifah et al., 2026).

Analisis kualitas soal dilakukan dengan memperhatikan tiga aspek utama, yaitu tingkat kesukaran soal, kemampuan soal dalam membedakan siswa berkemampuan tinggi dan rendah, serta keberfungsian alternatif jawaban pengecoh. Subjek dalam penelitian ini adalah siswa SMK Negeri 3 Tebing Tinggi sebanyak 30 orang. Objek yang dikaji dalam penelitian ini adalah butir soal pilihan ganda yang digunakan dalam kegiatan evaluasi pembelajaran. Adapun jumlah soal yang dianalisis sebanyak 25 butir soal. Setiap butir soal ditelaah untuk mengetahui tingkat kesukarannya, kemampuan daya pembedanya, serta efektivitas pilihan jawaban pengecohnya.

Pengumpulan data dalam penelitian ini dilakukan melalui teknik tes dan dokumentasi. Teknik tes dilakukan dengan memberikan soal pilihan ganda kepada siswa SMK Negeri 3 Tebing Tinggi. Setelah tes selesai, jawaban siswa dikoreksi untuk mengetahui jumlah siswa yang menjawab benar dan salah pada masing-masing butir soal. Teknik dokumentasi dilakukan dengan mengumpulkan dokumen pendukung, seperti lembar jawaban siswa, kunci jawaban, jumlah peserta tes, serta tabel hasil analisis butir soal. Data tersebut kemudian digunakan sebagai dasar dalam menghitung tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Data dalam penelitian ini dianalisis secara kuantitatif melalui analisis butir soal. Analisis dilakukan dengan tiga langkah utama, yaitu analisis tingkat kesukaran, analisis daya pembeda, dan analisis efektivitas pengecoh.

## HASIL DAN PEMBAHASAN

### Tingkat Kesukaran

Tingkat kesukaran menyatakan sejauh mana soal tersebut mudah dan sulit bagi siswa. Semakin besar persentase siswa menjawab soal dengan benar maka semakin mudah soal, sebaliknya semakin kecil persentase siswa menjawab soal dengan benar maka semakin sulit soal (Setiyawan & Wijayanti, 2020). Soal disebut baik apabila soal tersebut tidak termasuk dalam kategori terlalu mudah atau terlalu sukar. Soal terlalu sukar akan membuat siswa menjadi putus asa dan patah semangat karena di luar kemampuannya, sedangkan soal terlalu mudah tidak dapat merangsang siswa untuk mempertinggi usaha memecahkannya (Saputri & Larasati, 2023).

Suatu tes tidak boleh terlalu mudah, dan juga tidak boleh terlalu sukar. Soal yang terlalu mudah dapat dijawab dengan benar oleh semua siswa bukanlah item soal yang baik. Begitu pula soal yang terlalu sukar tidak dapat dijawab oleh semua siswa bukanlah item soal yang baik. Angka yang dapat memberi petunjuk mengenai tingkat kesukaran soal dikenal dengan istilah *difficulty index* (angka indeks kesukaran soal) dalam dunia evaluasi hasil belajar (Widoyoko, 2014).

Tingkat kesukaran butir soal digunakan untuk menentukan perbandingan antara soal mudah, sedang dan sukar. Soal yang dibuat sebaiknya tidak terlalu mudah ataupun terlalu sukar. Jika terlalu mudah, membuat siswa tidak termotivasi untuk berusaha lebih keras memecahkan soal. Namun, jika terlalu sukar, akan membuat siswa putus asa atau tidak semangat karena tidak terjangkau oleh siswa. Menurut Sudjana, tingkat kesukaran soal dapat menggunakan perbandingan 1:1:1, 3:4:3, atau 3:5:2 (Ambarwati & Ismiyati, 2021). Berdasarkan hasil analisis data diperoleh hasil sebagai berikut.

Tabel 1. Tingkat Kesukaran Butir Soal

No Butir Soal	Indeks Tingkat Kesukaran ( <i>P</i> )	Kategori
1	0,97	Mudah
2	0,63	Sedang
3	0,67	Sedang
4	0,03	Sukar
5	0,13	Sukar
6	0,23	Sukar
7	0,33	Sedang
8	0,40	Sedang
9	0,63	Sedang
10	0,70	Sedang
11	0,63	Sedang
12	0,23	Sukar
13	0,50	Sedang
14	0,37	Sedang
15	0,30	Sukar
16	0,57	Sedang
17	0,23	Sukar
18	0,37	Sedang
19	0,87	Mudah
20	0,50	Sedang
21	0,33	Sedang
22	0,13	Sukar
23	0,27	Sukar
24	0,03	Sukar
25	0,37	Sedang

Berdasarkan hasil analisis tingkat kesukaran, perangkat tes di atas didominasi oleh soal kategori sedang sebanyak 13 butir. Selain itu, terdapat 10 butir soal berkategori sukar dan 2 butir soal berkategori mudah. Komposisi ini menunjukkan bahwa sebagian besar butir soal berada pada tingkat kesukaran yang cukup proporsional untuk mengukur kemampuan siswa.

### Daya Pembeda

Analisis daya pembeda artinya mengkaji soal-soal tes dari segi keangguapan tes tersebut dalam membedakan siswa yang termasuk kedalam katagori lemah/rendah dan katagori kuat/tinggi presentasinya. Daya pembeda butir soal memiliki manfaat yaitu untuk meningkatkan mutu setiap-butir soal melalui data empiriknya dan untuk mengetahui seberapa

jauh masing-masing butir soal dapat membedakan kemampuan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang diajarkan pendidik (Magdalena et al., 2020). Salah satu penyebab buruknya daya pembeda suatu soal karena soal tersebut kemungkinan terlalu sukar atau mudah, sehingga antara siswa yang tidak paham dan siswa yang sudah paham tentang materi menjadi tidak dapat dibedakan (Purniasari et al., 2021). Berdasarkan hasil analisis data diperoleh hasil sebagai berikut.

Tabel 2. Daya Pembeda Butir Soal

No Butir Soal	Indeks Daya Pembeda ( <i>D</i> )	Kategori
1	0,27	Cukup
2	0,60	Sangat Baik
3	0,67	Sangat Baik
4	0,00	Jelek
5	0,27	Cukup
6	0,53	Sangat Baik
7	0,47	Sangat Baik
8	0,60	Sangat Baik
9	0,53	Sangat Baik
10	0,53	Sangat Baik
11	0,60	Sangat Baik
12	0,60	Sangat Baik
13	0,60	Sangat Baik
14	0,67	Sangat Baik
15	0,40	Sangat Baik
16	0,73	Sangat Baik
17	0,07	Jelek
18	0,53	Sangat Baik
19	0,33	Baik
20	0,67	Sangat Baik
21	0,53	Sangat Baik
22	0,33	Baik
23	0,27	Cukup
24	0,20	Cukup
25	0,73	Sangat Baik

Berdasarkan hasil analisis daya pembeda terhadap 25 butir soal, sebagian besar soal berada pada kategori sangat baik, yaitu sebanyak 17 butir. Selain itu, terdapat 2 butir soal berkategori baik, 4 butir soal berkategori cukup, dan 2 butir soal berkategori jelek, yaitu butir nomor 4 dan 17. Dengan demikian, mayoritas soal telah mampu membedakan siswa berkemampuan tinggi dan rendah, meskipun beberapa butir masih perlu diperbaiki.

### Efektivitas Pengecoh

Efektivitas pengecoh atau fungsi distraktor adalah analisis yang dilakukan untuk mengetahui sejauh mana pilihan jawaban yang salah (pengecoh) dapat menarik perhatian peserta tes, khususnya bagi mereka yang belum menguasai materi. Dalam evaluasi hasil belajar, sebuah pengecoh dikatakan berfungsi secara efektif apabila dipilih oleh sekurang-

kurangnya 5% dari total peserta tes. Analisis ini sangat penting untuk memastikan bahwa setiap pilihan jawaban dalam soal pilihan ganda memiliki alasan logis dan tidak hanya sekadar pelengkap yang mudah ditebak keberadaannya.

Fungsi utama dari pengecoh yang efektif adalah untuk mendeteksi tingkat ketelitian dan penguasaan konsep siswa. Jika sebuah pengecoh tidak ada yang memilih sama sekali, maka pilihan tersebut dianggap tidak berfungsi dan perlu direvisi atau diganti karena tidak memberikan kontribusi dalam membedakan kemampuan siswa. Dengan mengevaluasi efektivitas pengecoh, pendidik dapat menciptakan perangkat tes yang lebih kompetitif dan objektif, di mana siswa benar-benar dituntut untuk memahami materi daripada sekadar mengandalkan teknik eliminasi jawaban yang tampak tidak masuk akal. Berdasarkan hasil analisis data diperoleh hasil sebagai berikut.

Tabel 3. Efektivitas Pengecoh Butir Soal

No Soal	Kunci Jawaban	A (%)	B (%)	C (%)	D (%)	E (%)	Status
1	D	0%	0%	3,3%	96,7%	0%	A, B, C, Tidak Efektif
2	B	6,7%	63,3%	3,3%	13,3%	13,3%	C Tidak Efektif
3	C	3,3%	6,7%	60%	26,7%	3,3%	A & E Tidak Efektif
4	B	10%	23,3%	26,7%	30%	10%	Sangat Baik
5	E	23,3%	10%	23,3%	26,7%	16,7%	Sangat Baik
6	E	13,3%	13,3%	6,7%	13,3%	53,3%	Sangat Baik
7	C	10%	13,3%	46,7%	16,7%	13,3%	Sangat Baik
8	B	6,7%	60%	3,3%	16,7%	13,3%	C Tidak Efektif
9	C	10%	30%	36,7%	10%	13,3%	Sangat Baik
10	E	10%	26,7%	16,7%	6,7%	40%	Sangat Baik
11	B	16,7%	53,3%	13,3%	0%	16,7%	D Tidak Efektif
12	A	46,7%	20%	13,3%	6,7%	13,3%	Sangat Baik
13	A	36,7%	13,3%	10%	13,3%	26,7%	Sangat Baik
14	D	36,7%	6,7%	10%	33,3%	13,3%	Sangat Baik
15	A	33,3%	20%	13,3%	13,3%	20%	Sangat Baik
16	E	16,7%	30%	10%	3,3%	40%	D Tidak Efektif
17	A	40%	23,3%	10%	23,3%	3,3%	E Tidak Efektif

18	C	13,3%	16,7%	36,7%	13,3%	20%	Sangat Baik
19	C	10%	10%	46,7%	10%	23,3%	Sangat Baik
20	A	33,3%	10%	20%	13,3%	23,3%	Sangat Baik
21	C	30%	10%	26,7%	13,3%	20%	Sangat Baik
22	E	23,3%	10%	13,3%	13,3%	40%	Sangat Baik
23	B	16,7%	36,7%	10%	13,3%	23,3%	Sangat Baik
24	D	33,3%	6,7%	10%	40%	10%	Sangat Baik
25	E	23,3%	20%	3,3%	10%	33,3%	Sangat Baik

Berdasarkan hasil analisis efektivitas pengecoh, sebanyak 17 butir soal memiliki pengecoh yang berfungsi dengan baik. Sementara itu, 8 butir soal masih memerlukan perbaikan karena terdapat satu atau lebih pilihan jawaban yang tidak dipilih atau sangat sedikit dipilih oleh siswa, antara lain pada butir nomor 1, 2, 3, 8, 11, 16, 17, dan 25. Kondisi tersebut menunjukkan bahwa beberapa alternatif jawaban perlu direvisi agar lebih logis dan mampu berfungsi sebagai pengecoh yang efektif.

Analisis efektivitas pengecoh atau distraktor dalam evaluasi hasil belajar secara umum mengklasifikasikan pilihan jawaban salah ke dalam tiga kategori utama berdasarkan pola pilihan peserta tes. Kategori tersebut adalah pengecoh yang **berfungsi**, **kurang berfungsi**, dan **tidak berfungsi**. Sebuah pengecoh dikatakan berfungsi apabila dipilih oleh sekurang-kurangnya 5% dari total peserta tes, serta lebih banyak dipilih oleh kelompok siswa berkemampuan rendah dibandingkan kelompok berkemampuan tinggi. Sebaliknya, pengecoh yang kurang berfungsi biasanya hanya dipilih oleh sedikit peserta (di bawah 5%), sedangkan pengecoh yang tidak berfungsi adalah pilihan yang sama sekali tidak dipilih oleh satu pun peserta tes. Dalam penyusunan tes yang berkualitas, pengecoh yang tidak berfungsi harus segera diganti atau direvisi total karena keberadaannya tidak memberikan kontribusi dalam mengevaluasi pemahaman siswa. Dengan mengkategorikan pengecoh ini, pendidik dapat memastikan bahwa setiap opsi dalam soal pilihan ganda benar-benar memiliki daya tarik yang logis bagi siswa yang belum menguasai materi secara mendalam.

## KESIMPULAN

Analisis butir soal yang mencakup tingkat kesukaran, daya pembeda, dan efektivitas pengecoh merupakan prosedur esensial untuk menjamin kualitas instrumen evaluasi hasil belajar. Berdasarkan pembahasan, dapat disimpulkan bahwa ketiga parameter ini saling berkaitan dalam menentukan validitas sebuah tes; soal yang ideal harus memiliki tingkat kesukaran yang proporsional (dominan sedang), daya pembeda yang positif dan signifikan untuk membedakan kemampuan siswa, serta pengecoh yang berfungsi secara efektif (dipilih minimal 5% peserta). Tanpa analisis yang komprehensif, pendidik tidak dapat memastikan apakah skor yang diperoleh siswa benar-benar merepresentasikan kompetensi yang sesungguhnya atau hanya hasil dari kebetulan dan konstruksi soal yang lemah.

Berdasarkan hasil analisis butir soal, disarankan bagi para pendidik untuk tidak hanya terpaku pada pemberian nilai akhir, tetapi juga meluangkan waktu untuk melakukan audit terhadap kualitas soal secara berkala menggunakan aplikasi analisis data atau perhitungan manual. Butir soal yang terbukti memiliki daya beda negatif atau pengecoh yang tidak berfungsi sebaiknya segera direvisi atau diganti sebelum dimasukkan ke dalam bank soal untuk ujian berikutnya. Selain itu, lembaga pendidikan perlu memberikan pelatihan berkelanjutan bagi guru mengenai teknik penyusunan soal yang baik agar proses evaluasi dapat berjalan lebih objektif, adil, dan mampu memberikan umpan balik yang akurat bagi perbaikan proses pembelajaran.

## DAFTAR PUSTAKA

- Asiah, S. (2016). Efektivitas Kinerja Guru. *TADBIR: Jurnal Manajemen Pendidikan Islam*, 4(2), 1–11.
- Febriana, R. (2021). *Evaluasi pembelajaran*. Bumi Aksara.
- Juwita, R., Hasania, Y. A., & Azis, A. (2026). Penyusunan Instrumen Evaluasi Pembelajaran: Analisis Pembuatan Soal Pilihan Ganda dan Esai/Uraian. *Al-Ilmiya: Jurnal Pendidikan Islam*, 1(4), 1837–1844.
- Kadir, A. (2015). Menyusun dan menganalisis tes hasil belajar. *Al-TA'DIB: Jurnal Kajian Ilmu Kependidikan*, 8(2), 70–81.
- Kusumaningrum, H., Chaerany, C., Kholisah, T. A., & Cahyani, R. (2024). Kinerja guru sebagai aspek strategis dalam pengelolaan sumber daya manusia pendidikan. *Jurnal Pendidikan Islam Muta'allimin*, 1(3), 103–114.
- Sahra, A. A., Rahman, A., Pendit, S. S. D., Azizah, A., & Rini, E. S. (2026). Improved Science Learning Outcomes through Technology Assisted Experimentation: Peningkatan Hasil Belajar IPAS melalui Eksperimen Berbantuan Teknologi. *Academia Open*, 11(1), 10–21070.
- Arikunto, S. (2021). *Dasar-Dasar Evaluasi Pendidikan Edisi 3*. Bumi Aksara. <https://books.google.co.id/books?id=j5EmEAAAQBAJ>
- Ikawati, D. S., Agustin, E. W., Hakim, L., & Pratiwi, V. (2024). Analisis butir soal pilihan ganda pada siswa SMK menggunakan Anates. *Edukatif: Jurnal Ilmu Pendidikan*, 6(6), 6773–6781.
- Trianingsih, L. (2023). Analisis Kualitas Tes dan Butir Soal Integral pada Evaluasi Formatif Matematika Teknik. *Indonesian Journal Of Civil Engineering Education*, 9(2), 44. <https://doi.org/10.20961/ijcee.v9i2.84711>
- Fitriani, N. (2021). Analisis tingkat kesukaran, daya pembeda, dan efektivitas pengecoh soal pelatihan kewaspadaan kegawatdaruratan maternal dan neonatal. *Paedagoria*, 12(2), 199–205.
- Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Al Ameer, A. Y., Yahia, A. I. O., Mohammed, O. A., & Adam, M. I. E. (2024). Item analysis: the impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(1), 1–7. <https://doi.org/10.1186/s12909-024-05433-y>

- Sirajuddin, S., Jufari, J., & Muh, N. (2021). Efisiensi Kajian Ilmu Administrasi. *Prodi P.ADP FF UNY*, 18(1), 1–14.
- Mulyani, H., Tanuatmodjo, H., & Iskandar, R. (2020). Quality analysis of teacher-made tests in financial accounting subject at vocational high schools. *Jurnal Pendidikan Vokasi*, 10(1), 1–9. <https://doi.org/10.21831/jpv.v10i1.29382>
- Setiyawan, R. A., & Wijayanti, P. S. (2020). Analisis kualitas instrumen untuk mengukur kemampuan pemecahan masalah siswa selama pembelajaran daring di masa pandemi. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(2), 130–139.
- Saputri, H. A. S., & Larasati, N. J. (2023). Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkat Kesukaran Dan Daya Beda Butir Soal. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 9(5), 2986–2995.
- Magdalena, I., Fauzi, H. N., & Putri, R. (2020). Pentingnya evaluasi dalam pembelajaran dan akibat memanipulasinya. *Bintang*, 2(2), 244–257.
- Widoyoko, E. P. (2014). Penilaian hasil pembelajaran di sekolah. *Yogyakarta: Pustaka Pelajar*, 1(2), 8.
- Ambarwati, Y. F., & Ismiyati, I. (2021). Analisis butir soal pilihan ganda ulangan akhir semester genap mata pelajaran kearsipan. *Measurement In Educational Research*, 1(2), 64–75.
- Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMA N 1 Kutowinangun Tahun Pelajaran 2019 / 2020 Menggunakan Model Iteman dan Rasch. *Jurnal Pendidikan Kimia*, 10(2). <https://doi.org/10.20961/jpkim.v10i2.48244>
- Slepkov, A. D., Van Bussel, M. L., Fitze, K. M., & Burr, W. S. (2021). A Baseline for Multiple-Choice Testing in the University Classroom. *SAGE Open*, 11(2). <https://doi.org/10.1177/21582440211016838>
- Ardiansyah, B. (2023). Profil Capaian Iklim Kebhinekaan Jenjang Sekolah Dasar di Provinsi NTB Berbasis Rapor Pendidikan 2023. *SEMESTA: Jurnal Ilmu Pendidikan Dan Pengajaran*, 1(3), 130–136.
- Nisa, W. A., & Chusni, M. M. (2026). Analisis Peningkatan Keterampilan Berpikir Kritis Peserta Didik pada Materi Suhu dan Kalor. *SEMESTA: Jurnal Ilmu Pendidikan Dan Pengajaran*, 4(1), 12–21.
- Thinking, C. (2025). *SEMESTA Journal of Science Education and Teaching The Influence of Problem Based Learning Model on Students* '. 8(1), 42–48.
- Creswell, J. W., & Creswell, J. D. (2022). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications. <https://books.google.co.id/books?id=Pr2VEAAAQBAJ>
- Hanifah, R. A., Chusni, M. M., Studi, P., Fisika, P., Islam, U., Sunan, N., & Djati, G. (2026). Analisis Kemampuan Pemecahan Masalah Fisika Peserta Didik pada Materi Fluida Dinamis: Studi Pendahuluan di SMA Mekar Arum 1. 4(1), 71–80.