

MEMBANGUN INSTRUMEN EVALUASI YANG BERKUALITAS MELALUI ANALISIS TINGKAT KESUKARAN, DAYA BEDA, DAN EFEKTIVITAS PENGECOH SOAL

Laila Febriani¹, Riri Safitri², Cinta Hamidah³, Juli Anisah⁴, Sholatul Ramadhana⁵,
Ahmad Husein⁶, Fitrah Al Fatih⁷

^{1,2,3,4,5,6,7}Pendidikan Guru Madrasah Ibtidaiyah

Sekolah Tinggi Agama Islam Negeri Mandailing Natal

lailafebriani2003@gmail.com, ririsafitri589@gmail.com,
cintahamidah92@gmail.com, Julianisah23@gmail.com, hsolah376@gmail.com,
huseinahmad9897@gmail.com, fitrahalfatih310104@gmail.com

ABSTRACT

This study aims to examine the importance of difficulty level, discrimination index, and distractor effectiveness analysis in developing high-quality educational evaluation instruments. The quality of test items plays a crucial role in ensuring that evaluation results accurately and objectively measure students' abilities. This study employed a library research method by utilizing various references, including books, scientific journals, and other relevant literature related to item analysis. Data were collected through reviewing, analyzing, and synthesizing theories concerning test item quality. The findings indicate that the difficulty level serves to classify items based on their degree of ease or difficulty, the discrimination index measures an item's ability to distinguish between high-achieving and low-achieving students, while distractor effectiveness evaluates the functionality of incorrect answer options in multiple-choice tests. These three aspects are interconnected and serve as key indicators in determining the quality of an evaluation instrument. Test items with balanced difficulty levels, strong discrimination power, and effective distractors provide more accurate information regarding students' learning outcomes and competencies. Therefore, item analysis should be conducted continuously to improve the validity, reliability, and overall quality of educational assessment instruments. The implementation of systematic item analysis contributes significantly to the development of more effective, objective, and accurate evaluation practices in educational settings.

Keywords: *Item Analysis; Difficulty Level; Distractor Effectiveness*

ABSTRAK

Penelitian ini bertujuan untuk mengkaji pentingnya analisis tingkat kesukaran, daya beda, dan efektivitas pengecoh dalam membangun instrumen evaluasi pembelajaran yang berkualitas. Kualitas butir soal merupakan faktor penting yang menentukan ketepatan hasil evaluasi dalam mengukur kemampuan peserta didik secara objektif dan akurat. Metode yang digunakan dalam penelitian ini adalah studi kepustakaan (library research) dengan memanfaatkan berbagai sumber referensi berupa buku, jurnal ilmiah, dan literatur yang relevan dengan topik pembahasan. Data diperoleh melalui proses pengkajian, analisis, dan sintesis berbagai teori yang berkaitan dengan analisis butir soal. Hasil kajian menunjukkan bahwa tingkat kesukaran berfungsi untuk mengidentifikasi kategori soal berdasarkan tingkat

kemudahan atau kesulitannya, daya beda digunakan untuk mengetahui kemampuan soal dalam membedakan peserta didik yang memiliki tingkat penguasaan materi tinggi dan rendah, sedangkan efektivitas pengecoh berperan dalam menilai keberfungsian alternatif jawaban yang salah dalam soal pilihan ganda. Ketiga aspek tersebut saling berkaitan dan menjadi indikator utama dalam menentukan kualitas suatu instrumen evaluasi. Instrumen yang memiliki tingkat kesukaran yang proporsional, daya beda yang baik, serta pengecoh yang efektif akan menghasilkan informasi yang lebih akurat mengenai kemampuan peserta didik. Oleh karena itu, analisis butir soal perlu dilakukan secara berkelanjutan untuk meningkatkan validitas, reliabilitas, dan kualitas evaluasi pembelajaran

Kata Kunci: Analisis Butir Soal; Tingkat Kesukaran; Efektivitas Pengecoh

A. Pendahuluan

Evaluasi pembelajaran digunakan dalam dunia pendidikan untuk mengetahui sejauh mana proses belajar mengajar dapat mencapai tujuan yang telah ditetapkan. Dari kegiatan evaluasi, guru dapat melihat kemampuan peserta didik sekaligus mengetahui efektivitas pembelajaran yang sudah dilakukan. Di sekolah, bentuk evaluasi yang paling sering digunakan adalah tes atau soal. Mutu soal yang digunakan sangat memengaruhi hasil penilaian kemampuan siswa. Karena itu, penyusunan soal perlu dilakukan dengan baik agar instrumen yang digunakan mampu mengukur kemampuan peserta didik secara tepat, objektif, dan akurat. Karena itu, diperlukan suatu proses untuk mengkaji kualitas soal yang telah dibuat sehingga dapat diketahui apakah soal tersebut layak digunakan

sebagai instrumen evaluasi pembelajaran. Proses pengkajian tersebut disebut analisis butir soal yang dilakukan untuk menilai mutu soal berdasarkan beberapa indikator tertentu dilakukan untuk mengetahui kualitas soal melalui beberapa indikator, seperti tingkat kesukaran, daya beda, dan efektivitas pengecoh. Analisis butir soal merupakan suatu prosedur yang digunakan untuk menilai mutu soal agar dapat diketahui apakah soal tersebut layak digunakan sebagai instrumen evaluasi dalam mengukur kemampuan peserta didik secara objektif dan tepat buat bodi note nya untuk dapus ini (Lapenia & Hidayati, 2023). Berdasarkan pendapat tersebut, dapat disimpulkan bahwa analisis butir soal merupakan kegiatan untuk menilai kualitas dan kelayakan soal melalui indikator tertentu agar dapat digunakan

sebagai alat evaluasi yang objektif dan tepat.

Pembahasan tentang pengecoh, tingkat kesukaran, dan daya beda soal jadi dasar penting dalam membuat alat evaluasi pembelajaran. Soal yang terlalu susah atau terlalu gampang biasanya kurang pas untuk mengukur kemampuan peserta didik dengan tepat. Begitu juga kalau daya beda soal rendah, soal tersebut jadi sulit membedakan siswa yang paham materi dan yang belum paham. Selain itu, pengecoh yang tidak bekerja dengan baik juga membuat soal kurang efektif untuk mengukur kemampuan siswa. Karena itu, analisis butir soal perlu dilakukan sebelum soal dipakai dalam kegiatan evaluasi pembelajaran.

Kenyataan yang terjadi di lapangan menunjukkan bahwa masih banyak guru yang menyusun soal tanpa melakukan analisis kualitas butir soal terlebih dahulu. Sebagian besar guru hanya berfokus pada penyusunan soal dan hasil nilai peserta didik tanpa memperhatikan kualitas soal yang digunakan. Akibatnya, masih ditemukan soal dengan tingkat kesukaran yang tidak seimbang, daya beda yang rendah, serta pengecoh yang tidak berfungsi

secara optimal. Kondisi tersebut menyebabkan hasil evaluasi belum sepenuhnya mampu menggambarkan kemampuan peserta didik yang sebenarnya.

Penelitian mengenai pengecoh, tingkat kesukaran, dan daya beda soal telah banyak dilakukan, namun sebagian besar penelitian hanya berfokus pada hasil analisis secara umum dan belum membahas secara mendalam keterkaitan ketiga aspek tersebut dalam menentukan kualitas soal. Selain itu, masih terdapat guru yang belum memahami pentingnya analisis butir soal dalam proses evaluasi pembelajaran. Hal ini menunjukkan adanya kesenjangan antara konsep ideal penyusunan soal dengan praktik yang terjadi di lapangan.

Penelitian ini difokuskan pada kualitas butir soal sebagai instrumen penilaian pembelajaran dengan sasaran pembahasan yaitu guru dan peserta didik dalam kegiatan evaluasi pendidikan. Pada kesempatan ini penulis menyajikan pembahasan mengenai pengecoh, tingkat kesukaran, dan daya beda soal agar pendidik dapat lebih memahami konsep analisis butir soal serta mampu menyusun instrumen evaluasi

yang lebih baik, efektif, dan berkualitas.

B. Metode Penelitian

Artikel ini disusun dengan menggunakan metode studi kepustakaan yang memanfaatkan berbagai sumber referensi, seperti buku, jurnal ilmiah, dan literatur lain yang memiliki keterkaitan dengan topik pembahasan. Proses penulisan dilakukan melalui pengkajian secara mendalam dan menyeluruh terhadap materi yang dibahas, kemudian dilanjutkan dengan pengolahan serta pengintegrasian berbagai informasi dari sumber-sumber tersebut. Selanjutnya, hasil kajian disusun kembali dalam bentuk uraian yang sistematis agar lebih jelas, mudah dipahami, dan mampu membangun pemahaman konseptual mengenai tingkat kesukaran, daya beda, serta efektivitas pengecoh pada butir soal. Melalui artikel ini diharapkan pembaca, khususnya mahasiswa pendidikan dan guru, dapat memperoleh pemahaman yang lebih baik terkait analisis butir soal. Selain itu, hasil pembahasan dalam artikel ini juga dapat dijadikan sebagai bahan rujukan dan sumber belajar yang dapat diterapkan dalam praktik

evaluasi pembelajaran di lingkungan pendidikan.

C. Hasil Penelitian dan Pembahasan

Secara sederhana, tingkat kesukaran dapat diartikan sebagai indikator yang digunakan untuk mengetahui sejauh mana suatu soal dianggap mudah atau sulit oleh peserta didik ketika dikerjakan. Menurut Nita et al (2024) tingkat kesukaran butir soal merupakan ukuran yang digunakan untuk mengetahui sejauh mana suatu soal dianggap mudah atau sulit bagi peserta didik. Analisis tingkat kesukaran dilakukan untuk melihat apakah suatu butir soal dapat dijawab oleh sebagian besar siswa atau hanya oleh sedikit peserta didik sehingga soal yang baik seharusnya memiliki tingkat kesukaran yang seimbang, tidak terlalu mudah dan tidak terlalu sulit.

Sejalan dengan pendapat tersebut, Pradita et al (2023) menjelaskan bahwa tingkat kesukaran merupakan salah satu komponen penting dalam analisis butir soal yang dilakukan setelah soal diujicobakan kepada siswa. Melalui analisis tersebut, guru dapat mengetahui proporsi peserta didik yang mampu

menjawab soal dengan benar. Sementara itu, Magdalena et al (2021) menyebutkan bahwa tingkat kesukaran berkaitan dengan peluang peserta didik dalam menjawab soal secara benar pada tingkat kemampuan tertentu. Semakin banyak siswa yang menjawab benar, maka soal tersebut tergolong mudah, sedangkan apabila hanya sedikit siswa yang mampu menjawab dengan benar maka soal tersebut termasuk sukar. Pendapat tersebut juga diperkuat oleh Basuki et al (2024) yang menyatakan bahwa tingkat kesukaran adalah bilangan yang menunjukkan mudah atau sulitnya suatu soal dengan indeks kesukaran berkisar antara 0,00 sampai 1,00. Jika nilai indeks mendekati 0, maka soal tergolong semakin sukar, sedangkan apabila mendekati 1, maka soal semakin mudah.

Dari teori-teori tersebut, Penulis memaknai tingkat kesukaran butir soal sebagai ukuran yang digunakan untuk melihat sejauh mana suatu soal dapat dikerjakan oleh peserta didik. Tingkat kesukaran suatu soal dapat diketahui melalui jumlah siswa yang mampu menjawab dengan benar. Apabila mayoritas siswa dapat menjawab soal dengan tepat, maka soal tersebut

termasuk mudah, sedangkan jika hanya sedikit siswa yang mampu menjawabnya maka soal tersebut tergolong sulit. Oleh sebab itu, dalam penyusunan instrumen evaluasi diperlukan soal dengan tingkat kesukaran yang proporsional, yakni tidak terlalu mudah maupun terlalu sukar, sehingga hasil penilaian dapat menggambarkan kemampuan peserta didik secara lebih akurat. Secara sederhana, kondisi tersebut dapat diibaratkan seperti tantangan dalam sebuah perlombaan. Tantangan yang terlalu ringan membuat hampir semua peserta dapat melaluinya dengan mudah, sedangkan tantangan yang terlalu berat menyebabkan banyak peserta tidak mampu menyelesaikannya. Namun, apabila tingkat tantangan dibuat sesuai, maka kemampuan masing-masing peserta akan tampak dengan lebih jelas.

Pengukuran tingkat kesukaran dilakukan untuk mengetahui apakah butir soal yang disusun telah memenuhi kriteria kualitas dalam proses evaluasi pembelajaran. Melalui analisis tingkat kesukaran, guru dapat melihat perbandingan jumlah soal yang tergolong mudah, sedang, dan sulit sehingga hasilnya dapat dijadikan sebagai acuan dalam memperbaiki

kualitas instrumen evaluasi yang digunakan (Lestari et al., 2023). Selain itu, analisis tingkat kesukaran juga bertujuan untuk mengetahui kelayakan suatu butir soal dalam mengukur kemampuan peserta didik agar hasil evaluasi dapat memberikan pengukuran yang sesuai dengan tingkat kemampuan siswa. Dengan mengetahui tingkat kesukaran soal, guru dapat menentukan apakah suatu soal masih layak digunakan, perlu direvisi, atau harus diganti sehingga instrumen evaluasi menjadi lebih efektif, objektif, serta mampu mengukur kemampuan peserta didik secara lebih akurat dan proporsional (Dewi et al., 2018).

Penentuan tingkat kesukaran soal dilakukan melalui beberapa tahapan yang diawali dengan mengumpulkan data hasil tes peserta didik. Setelah itu, dilakukan perhitungan terhadap jumlah siswa yang mampu menjawab benar pada setiap butir soal. Tahapan tersebut bertujuan untuk mengetahui proporsi peserta didik yang dapat menyelesaikan soal dengan tepat yang kemudian dinyatakan dalam bentuk indeks kesukaran. Melalui hasil analisis tersebut, guru dapat mengidentifikasi kategori soal, apakah

termasuk mudah, sedang, atau sukar (Zulpan et al., 2025). Perhitungan tingkat kesukaran dilakukan menggunakan rumus sederhana sebagai berikut:

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

Keterangan:

P = indeks tingkat kesukaran

B

= jumlah siswa yang menjawab benar

N = jumlah seluruh peserta tes

Arikunto (2018) menyatakan bahwa agar hasil perhitungan indeks tingkat kesukaran dapat diinterpretasikan secara lebih tepat, maka setiap butir soal perlu diklasifikasikan ke dalam beberapa kategori tertentu. Secara umum, pengelompokan tingkat kesukaran tersebut terdiri atas beberapa klasifikasi yang digunakan untuk menggambarkan karakteristik masing-masing butir soal.

Tabel 1 indeks tingkat kesukaran

No	Indeks Kesukaran (P)	Kategori Soal
1	0,00 – 0,30	Sukar
2	0,31 – 0,70	Sedang
3	0,71 – 1,00	Mudah

Dari penjelasan tersebut, dapat dipahami bahwa proses penentuan tingkat kesukaran diawali dengan perhitungan menggunakan formula yang sudah ditentukan. Hasil dari

perhitungan tersebut kemudian diberi makna melalui indeks yang berlandaskan teori yang sesuai, lalu dikelompokkan untuk menentukan apakah suatu butir soal berada pada kategori mudah, sedang, atau sulit. Pada tahap akhir, dilakukan evaluasi terhadap hasil analisis tingkat kesukaran hingga seluruh rangkaian penilaian butir soal dinyatakan tuntas.

Dalam analisis butir soal, daya beda digunakan sebagai indikator untuk menilai kemampuan suatu soal dalam memisahkan peserta didik berdasarkan tingkat penguasaan materi yang dimiliki. Indikator ini menunjukkan apakah suatu butir soal mampu membedakan antara peserta didik yang memiliki kemampuan akademik lebih tinggi dan peserta didik yang memiliki kemampuan relatif lebih rendah. Solichin (2017) menjelaskan bahwa daya beda merupakan kemampuan suatu butir soal untuk membedakan peserta didik berdasarkan tingkat kemampuannya. Butir soal yang memiliki daya beda tinggi umumnya dapat dijawab dengan benar oleh peserta didik yang berkemampuan tinggi, sedangkan peserta didik yang berkemampuan rendah cenderung mengalami

kesulitan dalam menjawab soal tersebut secara tepat.

Sejalan dengan pendapat tersebut, Rachmawati & Pradana (2024) menyatakan bahwa daya beda merupakan indeks yang menunjukkan kemampuan suatu butir soal dalam membedakan peserta didik berdasarkan tingkat penguasaan materi yang dimiliki. Melalui analisis daya beda, guru dapat mengetahui sejauh mana suatu soal mampu mengukur perbedaan kemampuan peserta didik secara objektif sehingga kualitas instrumen evaluasi dapat ditentukan dengan lebih akurat. Selain itu, Trianingsih (2023) mengemukakan bahwa daya beda berfungsi untuk mengidentifikasi kualitas suatu butir soal serta menentukan kelayakannya untuk digunakan dalam kegiatan evaluasi pembelajaran. Dengan demikian, analisis daya beda tidak hanya membantu dalam membedakan peserta didik berdasarkan tingkat penguasaan materi, tetapi juga menjadi dasar dalam menilai kualitas dan efektivitas suatu butir soal sebagai instrumen evaluasi.

Jika ditinjau dari sudut pandang penulis, daya beda dapat dipahami sebagai kemampuan suatu butir soal

dalam menghasilkan informasi yang menunjukkan adanya perbedaan tingkat penguasaan materi di antara peserta didik. Keberadaan daya beda memungkinkan suatu soal berfungsi sebagai alat seleksi yang dapat memperlihatkan siapa saja peserta didik yang telah memahami materi dengan baik dan siapa yang masih memerlukan penguatan pembelajaran. Dalam konteks pembelajaran, daya beda dapat dianalogikan seperti seorang guru yang memberikan pertanyaan pemecahan masalah setelah materi selesai diajarkan. Peserta didik yang benar-benar memahami konsep yang telah dipelajari umumnya mampu menemukan jawaban yang tepat, sedangkan peserta didik yang belum memahami materi akan mengalami kesulitan dalam menyelesaikannya. Apabila pertanyaan tersebut mampu menunjukkan perbedaan yang jelas antara kedua kelompok peserta didik tersebut, maka pertanyaan tersebut memiliki daya beda yang baik. Sebaliknya, jika hampir seluruh peserta didik dapat menjawab dengan benar atau justru hampir semuanya menjawab salah, maka kemampuan soal dalam membedakan tingkat

penguasaan materi menjadi kurang optimal.

Tujuan dilakukannya analisis daya beda adalah untuk mengetahui kemampuan suatu butir soal dalam menunjukkan perbedaan tingkat penguasaan materi yang dimiliki peserta didik. Melalui hasil analisis tersebut, guru dapat melihat apakah soal mampu membedakan peserta didik yang telah menguasai materi dengan peserta didik yang belum menguasainya. Ratni (2023) menjelaskan bahwa analisis daya beda bertujuan untuk mengidentifikasi kualitas suatu butir soal berdasarkan kemampuannya dalam membedakan peserta didik yang memiliki tingkat kemampuan yang berbeda. Melalui hasil analisis tersebut, guru dapat menentukan apakah suatu butir soal masih layak digunakan dalam kegiatan evaluasi pembelajaran atau memerlukan perbaikan agar kualitas instrumen penilaian menjadi lebih optimal. Oleh karena itu, daya beda menjadi salah satu indikator penting dalam analisis butir soal karena dapat memberikan informasi mengenai kualitas, efektivitas, dan kelayakan suatu soal untuk digunakan sebagai instrumen evaluasi pembelajaran.

Proses analisis daya beda diawali dengan pengelompokan peserta tes berdasarkan hasil yang diperoleh agar perbedaan kemampuan antar kelompok dapat diidentifikasi secara lebih jelas. Menurut Konijo et al. (2025), pembentukan kelompok dalam analisis daya beda disesuaikan dengan jumlah peserta tes. Jika jumlah peserta kurang dari 100 orang, maka peserta dibagi menjadi dua kelompok yang sama besar, yaitu kelompok atas dan kelompok bawah. Namun, apabila jumlah peserta mencapai 100 orang atau lebih, kelompok atas ditentukan dari 27% peserta dengan skor tertinggi dan kelompok bawah dari 27% peserta dengan skor terendah, sedangkan peserta yang berada pada kelompok tengah tidak diikutsertakan dalam analisis. Sebelum proses pengelompokan dilakukan, skor peserta harus diurutkan dari yang tertinggi hingga yang terendah. Selanjutnya, perhitungan daya beda dapat dilakukan dengan menggunakan rumus yang sebagai berikut:

$$D = \frac{B_A - B_B}{J_A}$$

Keterangan:

D = indeks daya beda

B_A= jumlah peserta kelompok atas yang menjawab benar.

B_B= jumlah peserta kelompok bawah yang menjawab benar

J_A= Jumlah peserta dalam kelompok atas

Setelah nilai daya beda setiap butir soal diperoleh melalui perhitungan menggunakan rumus yang telah ditetapkan, langkah selanjutnya adalah melakukan interpretasi terhadap nilai tersebut berdasarkan kriteria daya beda yang berlaku sebagai berikut :

Tabel. 2 Kriteria Interpretasi Daya Beda Butir Soal

Indeks Daya Beda	Interpretasi
≥ 0,40	Sangat Baik
0,30–0,39	Cukup Baik
0,20–0,29	Perlu Revisi
≤ 0,19	Jelek

(Ebel & Frisbie.,1991)

Berdasarkan uraian tersebut, dapat dipahami bahwa proses analisis daya beda dilakukan melalui beberapa tahapan yang saling berkaitan. Tahap awal dimulai dengan mengurutkan skor peserta didik dari nilai tertinggi hingga nilai terendah. Setelah itu, peserta dikelompokkan ke dalam kelompok atas dan kelompok bawah sesuai dengan ketentuan yang didasarkan pada jumlah peserta tes.

Tahap berikutnya adalah menghitung nilai daya beda menggunakan rumus yang telah ditetapkan. Hasil perhitungan tersebut kemudian diinterpretasikan berdasarkan kriteria indeks daya beda yang digunakan untuk mengetahui tingkat kualitas setiap butir soal. Melalui proses tersebut, dapat ditentukan kategori masing-masing butir soal sehingga diperoleh dasar yang tepat dalam menetapkan kelayakan soal untuk digunakan dalam kegiatan evaluasi pembelajaran.

Dalam konteks penilaian instrumen tes berbentuk pilihan ganda, pengecoh (distractor) dipahami sebagai opsi jawaban selain kunci yang benar, yang dirancang sedemikian rupa agar tetap terlihat logis dan meyakinkan sehingga memiliki kemungkinan untuk dipilih oleh peserta didik yang belum menguasai materi secara memadai. Dengan kata lain, pengecoh dirancang untuk “menjebak” peserta didik yang memiliki penguasaan konsep rendah atau belum memahami materi secara menyeluruh. Amalia & Widayati (2012) menjelaskan bahwa efektivitas penggunaan pengecoh merupakan salah satu indikator kualitas butir soal karena dapat

menunjukkan apakah pilihan jawaban yang disediakan berfungsi dengan baik atau tidak. Sejalan dengan pendapat tersebut, Fitriani (2021) menyatakan bahwa efektivitas pengecoh perlu dianalisis untuk mengetahui kemampuan suatu soal dalam membedakan peserta didik yang memahami materi dengan yang belum memahami materi. Selain itu, Pidekso et.al (2024) menegaskan bahwa pengecoh merupakan bagian penting dalam karakteristik butir soal karena berpengaruh terhadap kualitas evaluasi pembelajaran, terutama dalam aspek daya pembeda dan validitas soal.

Secara umum, pengecoh (distractor) dapat dimaknai sebagai alternatif jawaban yang tidak benar, tetapi disusun sedemikian rupa sehingga mampu mengidentifikasi perbedaan antara peserta didik yang telah menguasai materi dan peserta didik yang belum menguasainya. Oleh karena itu, efektivitas pengecoh menjadi salah satu aspek penting dalam menentukan mutu suatu butir soal karena berkaitan dengan kemampuan soal dalam membedakan tingkat penguasaan peserta didik, mendukung validitas instrumen, serta

meningkatkan akurasi hasil evaluasi pembelajaran.

Suatu pengecoh dinilai berfungsi dengan baik apabila mampu menarik perhatian peserta tes yang belum memahami konsep atau belum menguasai materi yang diujikan. Keberhasilan pengecoh dalam menjalankan fungsinya ditunjukkan oleh kemampuannya menciptakan keraguan pada peserta didik yang memiliki penguasaan materi yang rendah, sehingga mereka terdorong untuk memilih alternatif jawaban tersebut dan menganggapnya sebagai jawaban yang benar (Hariati et al., 2026).

Keberfungsian suatu pengecoh dapat dilihat dari jumlah peserta tes yang memilih alternatif jawaban tersebut. Semakin tinggi jumlah pemilih pada suatu pengecoh, semakin menunjukkan bahwa pengecoh tersebut bekerja secara efektif. Sebaliknya, apabila tidak ada peserta yang memilihnya, maka pengecoh tersebut dapat dikategorikan tidak berfungsi (Arbiatin & Mulabbiyah, 2020). Kategori efektivitas pengecoh yang rendah mengindikasikan bahwa alternatif jawaban pengecoh belum berfungsi secara efektif. Keadaan ini

menunjukkan bahwa pengecoh tidak cukup menarik bagi peserta didik yang belum menguasai konsep maupun materi yang diujikan (Qadir et al., 2024).

Berdasarkan uraian tersebut, dapat dipahami bahwa pengecoh soal merupakan alternatif jawaban yang tidak benar dalam suatu butir soal. Keberfungsian pengecoh dapat diketahui dari adanya peserta didik yang memilih alternatif jawaban tersebut. Sebaliknya, apabila suatu pilihan jawaban tidak dipilih oleh peserta didik sama sekali, maka pengecoh tersebut dapat dinyatakan tidak berfungsi. Untuk mengetahui apakah suatu pengecoh berfungsi dengan baik atau tidak, diperlukan perhitungan yang mengacu pada kriteria efektivitas pengecoh sebagai berikut:

$$IP = \frac{P(O)}{JS} \times 100\%$$

Keterangan:

IP = indeks daya pembeda

P (O) = jumlah memilih setiap pilihan jawaban

Js = jumlah seluruh siswa

Agar penerapan rumus P(O) lebih mudah dipahami, setiap alternatif jawaban perlu dihitung secara terpisah. Sebagai ilustrasi,

terdapat 100 peserta didik yang menjawab soal nomor 4 dengan distribusi pilihan jawaban yang berbeda-beda.

Tabel. 3 Contoh perhitungan P(O)

Pilihan Jawaban	Jumlah Siswa
A	8
B	12
C	20
D	60
Total	100

Tabel.4 Hasil Perhitungan indeks pengecoh

Pilihan Jawaban	Perhitungan	Persentase
A	$8/100 = 0,08$	8%
B	$12/100 = 0,12$	12%
C	$20/100 = 0,2$	20%

Tahapan berikutnya setelah memperoleh nilai efektivitas pengecoh adalah melakukan penelaahan terhadap hasil perhitungan yang telah diperoleh. Nilai tersebut selanjutnya dicocokkan dengan kriteria yang berlaku untuk menilai tingkat keberfungsian masing-masing pengecoh. Melalui analisis tersebut, dapat ditentukan apakah pengecoh telah mampu menjalankan perannya secara efektif dalam menarik peserta didik yang belum menguasai materi, atau masih memerlukan perbaikan karena belum menunjukkan fungsi yang optimal.

Tabel. 5 Kriteria Efektivitas

Pengecoh soal

Indeks Pengecoh (IP)	Kriteria
> 5%	Efektif / Berfungsi
< 5%	Tidak Efektif / Tidak Berfungsi

(Mardapi, 2008).

Pengecoh dalam soal pilihan ganda dirancang untuk mengarahkan peserta tes yang belum memahami materi agar tidak langsung memilih jawaban yang benar. Setiap butir soal umumnya dilengkapi dengan beberapa alternatif jawaban yang berperan sebagai pengecoh, yang berfungsi untuk menguji ketepatan penguasaan materi peserta didik. Pada umumnya, satu butir soal memiliki tiga pilihan pengecoh selain satu jawaban yang benar. Suatu pengecoh dinilai berfungsi dengan baik apabila dipilih oleh sedikitnya 5% peserta tes. Sebaliknya, apabila pengecoh tidak menunjukkan efektivitas yang memadai, tindakan yang dapat dilakukan adalah mengganti atau memperbaiki alternatif jawaban tersebut, kemudian mengujikannya kembali untuk mengetahui tingkat keberfungsian (Radja et al., 2023).

Berdasarkan uraian tersebut, dapat dipahami bahwa langkah awal dalam analisis pengecoh adalah

mengidentifikasi alternatif jawaban yang salah pada setiap butir soal. Selanjutnya, setiap pengecoh dianalisis melalui perhitungan tertentu untuk memperoleh nilai yang menunjukkan tingkat keberfungsian. Hasil perhitungan tersebut kemudian dijelaskan dengan mengacu pada indeks atau kriteria yang telah ditetapkan. Setelah proses penafsiran dilakukan, langkah berikutnya adalah menentukan tindak lanjut yang sesuai terhadap butir soal berdasarkan hasil analisis yang diperoleh.

D. Kesimpulan

Analisis butir soal merupakan bagian penting dalam kegiatan penilaian pembelajaran yang bertujuan untuk memastikan bahwa alat evaluasi dapat mengukur kemampuan peserta didik secara tepat, objektif, dan efektif. Dalam proses ini terdapat tiga hal utama yang perlu diperhatikan, yaitu tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Ketiga hal tersebut saling berhubungan dalam menentukan kualitas sebuah soal, sehingga hasil penilaian dapat menggambarkan kemampuan peserta didik dengan lebih akurat.

Tingkat kesukaran digunakan untuk melihat apakah suatu soal tergolong mudah, sedang, atau sulit. Sementara itu, daya pembeda berfungsi untuk mengetahui kemampuan soal dalam membedakan peserta didik yang memiliki pemahaman tinggi dan rendah terhadap materi. Soal yang baik adalah soal yang memiliki tingkat kesukaran yang seimbang serta daya pembeda yang baik, sehingga dapat memberikan informasi yang jelas tentang hasil belajar peserta didik. Dari hasil analisis ini, guru dapat menilai apakah soal tersebut layak digunakan atau perlu diperbaiki.

Selain itu, efektivitas pengecoh juga sangat penting dalam soal pilihan ganda. Pengecoh yang baik dapat menarik peserta didik yang belum memahami materi dengan baik, sehingga membantu soal dalam membedakan tingkat pemahaman siswa. Oleh karena itu, analisis tingkat kesukaran, daya pembeda, dan efektivitas pengecoh perlu dilakukan secara terus-menerus agar soal yang digunakan semakin baik, valid, dan mampu mendukung proses pembelajaran yang lebih efektif.

DAFTAR PUSTAKA

- Amalia, A. N., & Widayati, A. (2012). *Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di Kota Yogyakarta tahun 2012. Jurnal Pendidikan Akuntansi Indonesia*, 10(1).
<https://doi.org/10.21831/jpai.v10i1.919>
- Arbiatin, E., & Mulabbiyah, M. (2020). Analisis kelayakan butir soal tes penilaian akhir semester mata pelajaran matematika kelas VI di SDN 19 Ampenan tahun pelajaran 2019/2020. *El Midad*, 12(2), 146-171.
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan* (Edisi 3). Jakarta: Bumi Aksara.
- Basuki, R., dkk. (2024). Analisis butir soal ujian akhir semester ganjil Madrasah Aliyah Walisongo Tapung Kampar. *INTIFA (Journal of Education and Language)*, 1(2), 209–222.
- Dewi, S. S., Hariastuti, R. M., & Utami, A. U. (2018). Analisis tingkat kesukaran dan daya pembeda soal olimpiade matematika (OMI) tingkat SMP tahun 2018. *Transformasi: Jurnal Pendidikan Matematika dan Matematika*, 3(1).
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement*.
- Fitriani, N. (2021). *Analisis tingkat kesukaran, daya pembeda, dan efektivitas pengecoh soal pelatihan kewaspadaan kegawatdaruratan maternal dan neonatal. Paedagoria: Jurnal Kajian, Penelitian dan Pengembangan Kependidikan*, 12(2).
<https://journal.ummat.ac.id/index.php/paedagoria/article/view/4956/0>
- Hariati, B. S., Hakim, L., Pratiwi, V., Zakia, A., Azizah, R. C. A., Maheswari, A. D., ... & Rahmawati, T. D. (2026). Analisis Kualitas Butir Soal Pilihan Ganda Elemen Dokumen Berbasis Digital Menggunakan ANATES. *Jurnal Jendela Pendidikan*, 6(01), 46-58.
- Konijo, N. A. P. A., Machmud, T., & Rahmi, E. (2025). Analisis butir esai 10-item pada kelas xi sma muhammadiyah batudaa 2025/2026. *JP2M (Jurnal Pendidikan dan Pembelajaran Matematika)*, 11(2), 1546-1558.
- Lapenia, M., & Hidayati, H. (2023). Analisis Butir Soal Ujicoba Posttest untuk Mengukur Hasil Belajar dan Pemahaman Siswa Pada Materi Fluida. *Jurnal Pendidikan Tambusai*, 7 (3), 22935–22943.
<https://doi.org/10.31004/jptam.v7i3.10231>
- Lestari, M., Halini, & Indriani, T. (2023). Analisis tingkat kesukaran soal persamaan dan pertidaksamaan nilai mutlak melalui pendekatan teori tes klasik. *Indonesian Journal of Learning and Instructional Innovation*, 1(2), 28–33.
- Magdalena, I., Anggraini, I. A., & Khoiriah, S. (2021). Analisis daya pembeda dan taraf kesukaran pada soal bilangan Romawi kelas 4 SDN Tobat 1 Balaraja. *Nusantara*, 3(1), 151-158.

- Mardapi, D. (2008). Teknik penyusunan instrumen tes dan nontes. Mitra Cendekia Pres
- Nita, N. K. A. N. T., Wirani, I. A. S., & Rai, I. B. (2024). TINGKAT KESUKARAN DAN DAYA BEDA BUTIR SOAL SUMATIF BAHASA BALI KELAS X SMA NEGERI 1 MENGWI TAHUN PELAJARAN 2022/2023. *Jurnal Pendidikan Bahasa Bali Undiksha*, 11(2), 105-115.
- Pidekso, G. W. S., Ramadhan, G. P. S., Sampurno, M. R. A., Hafizh, M. N., Pratiwi, V., & Hakim, L. (2024). Analisis karakteristik butir soal elemen komputer akuntansi fase F menggunakan software Anates. *Jurnal Pendidikan Akuntansi Indonesia*, 23(2). <https://doi.org/10.21831/jpai.v23i2.91595>
- Pradita, E., Megawanti, P., & Yulianingsih, Y. (2023). Analisis tingkat kesukaran, daya pembeda, dan fungsi distraktor pts matematika smpn Jakarta. *Himpunan: Jurnal Ilmiah Mahasiswa Pendidikan Matematika*, 3(1), 109-118.
- Qadir, A., Huda, N., & Hermina, D. (2024). Analisis butir tes: tingkat kesukaran, daya pembeda dan efektivitas pengecoh. *Al-Furqan: Jurnal Agama, Sosial, Dan Budaya*, 3(3), 1450-1467.
- Rachmawati, D., & Bayu Pradana, A. (2025). Analisis butir soal mata pelajaran ekonomi: Validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. *Jurnal Pendidikan Ekonomi (JUPE)*, 13(3), 273–284. <https://doi.org/10.26740/jupe.v13n3.p273-284>
- Radja, S. P., Bano, V. O., & Ina, A. T. (2023). Analisis kualitas butir soal hasil belajar peserta didik berdasarkan tingkat kesukaran, daya beda dan efektivitas pengecoh di SMAN 1 Pandawai. *Jurnal Edusavana*, 1(1), 30-41.
- Ratni, R. (2023). Analisis Tingkat Kesukaran, Daya Pembeda dan Efektivitas Pengecoh Soal Ulangan Bahasa Indonesia di Sekolah Dasar. *Prosa: Jurnal Penelitian Pendidikan Guru Sekolah Dasar*, 1(1), 117-123.
- Saputri, HAS, Zulhijrah, Nabila Joti Larasati, & Shaleh. (2023). ANALISIS INSTRUMEN PENILAI : VALIDITAS, RELIABILITAS, TINGKAT KESUKARAN DAN DAYA BEDA BUTIR SOAL. *Didaktik : Jurnal Ilmiah PGSD STKIP Subang*, 9 (5), 2986 - 2995. <https://doi.org/10.36989/didaktik.v9i5.2268>
- Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirasat: Jurnal Manajemen Dan Pendidikan Islam*, 2 (2), 192–213. <https://doi.org/10.26594/dirasat.v2i2.879>
- Trianingsih, L. (2023). Analisis kualitas tes dan butir soal integral pada evaluasi formatif matematika teknik. *Indonesian Journal of Civil Engineering Education*, 9(2)