



Kualitas Soal: Tingkat Kesukaran Daya Pembeda Pengecoh Soal

Zulpan¹, Ahmad Syafi², Dite Umbara Alfansuri³, Salsabila⁴, Santri Wardiah Hasibuan⁵, Nur Azizah⁶, Diana Langgeng Mustikawati⁷

^{1,4,5,6}PGMI STAIN Mandailing Natal, ²Universitas Negeri Semarang, ³Universitas Musamus, ⁷Universitas Ivet Semarang

*Correspondensi: Zulpan

Email: zulpan200990@gmail.com



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

Abstrak: Artikel ini bertujuan memberikan penjelasan konseptual yang sistematis dan mudah dipahami mengenai tiga komponen utama analisis butir soal, yaitu tingkat kesukaran, daya pembeda, dan pengecoh. Penulisan artikel ini menggunakan metode kajian pustaka dengan menelaah berbagai sumber relevan seperti buku dan artikel ilmiah, kemudian menyintesis teori-teori yang ada untuk membangun pemahaman konseptual yang terstruktur. Hasil kajian ini menghasilkan rancangan konsep yang sederhana namun komprehensif dalam memahami serta menerapkan analisis tingkat kesukaran, daya pembeda, dan pengecoh soal. Artikel ini diharapkan dapat menjadi referensi dan memperluas wawasan mahasiswa keguruan maupun guru dalam menganalisis butir soal, sekaligus mendorong penerapannya dalam proses pembelajaran. Tingkat kesukaran diukur berdasarkan proporsi peserta yang menjawab benar dibandingkan dengan jumlah seluruh peserta tes, yang kemudian diklasifikasikan ke dalam kategori mudah, sedang, atau sulit. Soal yang berkualitas bukan semata-mata yang mudah atau sulit, melainkan yang mampu membedakan tingkat pemahaman peserta didik. Sementara itu, pengecoh soal merupakan alternatif jawaban yang salah, yang efektivitasnya dapat dianalisis melalui perhitungan indeks tertentu untuk menentukan kelayakan setiap pilihan jawaban sebelum dilakukan tindak lanjut terhadap revisi butir soal.

Kata Kunci: Tingkat Kesukaran, Daya Beda, Pengecoh Soal

Abstrack: This article aims to provide a systematic and comprehensible conceptual explanation of the three main components of item analysis: difficulty level, discriminating power, and distractors. The study employs a literature review method by examining various relevant sources, including books and scientific articles, and synthesizing existing theories to develop a well-structured conceptual framework. The findings result in a simple yet comprehensive conceptual model for understanding and applying item analysis based on difficulty level, discrimination index, and distractor effectiveness. This article is expected to serve as a reference and enhance the understanding of pre-service teachers and educators in analyzing test items, as well as encourage their application in the learning process. The difficulty level is measured by the proportion of students who answer correctly compared to the total number of participants, which is then categorized as easy, moderate, or difficult. A good test item is not merely one that is easy or difficult, but one that can effectively differentiate students based on their level of understanding. Meanwhile, distractors refer to incorrect answer options whose effectiveness can be evaluated through specific index calculations to determine their appropriateness before revising the item.

Keywords: Item Difficulty, Discrimination Index, Distractor Effectiveness

Pendahuluan

Pembelajaran secara umum dikatakan sebuah aktivitas interaksi guru dengan muridnya untuk mencapai tujuan yang diinginkan. Melihat keberhasilan siswa tentu guru melakukan proses penilaian terhadap murid sesuai dengan pembelajaran yang telah dilakukan. Sebelum melakukan penilaian, guru mempersiapkan tes dari berbagai bentuk yang mudah dipahami siswa. Penggunaan tes atau alat evaluasi untuk mengukur kemampuan tertentu yang berkaitan dengan indikator tujuan pembelajaran. (Komarudin, dkk, 2024). Umumnya guru di sekolah lebih sering menggunakan soal bentuk pilihan ganda dan essay. Soal yang disusun oleh guru diujikan pada saat pembelajaran dengan memberikan soal kepada murid untuk dijawab, setelah murid selesai menjawab, kemudian guru melakukan koreksi, seterusnya memberikan skor, dan memberikan penilaian.

Pada umumnya guru-guru masih sering mengabaikan terkait dengan kualitas soal yang telah diberikan kepada murid, sangat jarang guru melakukan analisis soal sebelum dan sesudah soal dikerjakan. Paparan pada artikel Tutut (2015) menyatakan soal yang telah diujikan kepada murid tidak melakukan tahapan analisis butir soal, sehingga belum diketahui kualitas butir soal yang disusun. Seterusnya Anita, dkk (2018) menyatakan bahwa guru belum melakukan analisis tes terhadap soal yang digunakan dikarenakan guru tidak memiliki waktu yang cukup untuk melakukan proses analisis tes. diperkuat oleh Fauziah, dkk, (2020) menyatakan hasil ujian siswa belum dijadikan sebagai gambaran apakah soal tersebut sukar, mudah atau sedang sehingga tidak dapat membedakan siswa yang paham dan kurang paham terhadap materi pembelajaran.

Idealnya soal yang telah diujikan hendaklah melakukan analisis tes karena bertujuan: 1) Memberikan pengetahuan tentang kelebihan dan kekurangan suatu lembar soal; 2) Memberikan informasi lengkap mengenai spesifikasi butir soal; 3) Mengetahui permasalahan yang terdapat pada butir soal; 4) Penilaian kumpulan bank soal; 5) Referensi untuk menyusun item soal (Zohrabi, 2013). Sejalan dengan pendapat Liesfi (2016) menyatakan analisis soal bertujuan untuk mengidentifikasi soal-soal yang baik, kurang baik dan soal yang tidak baik. Aspek lain yang mendukung baiknya kualitas perangkat soal adalah aspek tingkat kesukaran butir soal dan daya pembeda soal (Wahyuningsih, 2015). Akhmadi (2021) menegaskan jika efektifitas pengecoh buruk maka indeks daya pembeda akan rendah dan tingkat kesukaran soal tersebut juga rendah. Begitu sebaliknya, jika efektifitas pengecoh suatu jawaban pada butir soal itu baik maka indeks daya pembeda akan baik dan tingkat kesukaran soal menjadi tinggi. Paparan di atas menjelaskan secara umum serta fenomena yang telah terjadi di lapangan terkait dengan analisis soal mengenai tingkat kesukaran, daya pembeda dan pengecoh soal. Hal tersebut membuat penulis tertarik membuat konsep yang lebih mudah dipahami tentang tingkat kesukaran, daya pembeda, pengecoh soal. Artikel diharapkan dapat memberikan pemahaman serta kontribusi sebagai rujukan dan sumber belajar bagi mahasiswa keguruan dan guru untuk diimplementasikan pada pembelajaran.

Metode

Artikel ini ditulis menggunakan metode kajian pustaka yang bersumber dari buku dan artikel ilmiah serta referensi lain yang relevan. Artikel ini dilakukan memahami secara mengakar dan mendalam terhadap topik-topik yang diangkat kemudian dilakukan sintesis, seterusnya mengonstruksi kalimat demi kalimat untuk mudah dipahami serta menciptakan konsep. Hasil pada tulisan ini yakni konstruksi konsep yang lebih mudah dipahami terhadap materi tingkat kesukaran, daya pembeda, serta pengecoh soal. Pada akhirnya artikel dapat menjadi rujukan serta sumber belajar

kepada mahasiswa keguruan dan guru, selanjutnya dapat diimplementasikan pada pembelajaran nyata.

Hasil dan Pembahasan

Secara bahasa umum terkait istilah tingkat kesukaran sebagian berpandangan bahwa tingkat kesukaran itu diartikan soal-soal (butir soal) dianggap susah untuk dikerjakan. Tingkat kesukaran butir soal diukur oleh persentase siswa yang menjawab soal dengan benar (Kustati & Amelia, 2024). Kemudian Aziz (2016) menyatakan tingkat kesukaran berarti seberapa sulit atau mudahnya suatu butir soal bagi sekelompok peserta didik. Azwar (2006) menyatakan bahwa tingkat kesukaran butir soal adalah proporsi antara banyaknya peserta tes yang menjawab butir soal benar. Tingkat kesukaran butir soal merupakan proporsi antara siswa yang dapat menjawab benar suatu butir soal dengan jumlah seluruh peserta tes (Rahayu & Djazari, 2016). Seterusnya Anastasi & Susan (2007) tingkat kesukaran soal berkaitan dengan persentase peserta yang menjawab soal dengan benar, semakin mudah butir soal, semakin besarlah persentasenya.

Penulis memberikan pemahaman dari hasil teori-teori bahwa tingkat kesukaran soal yakni butir-butir soal siswa menjawab benar dan salah, seterusnya dilakukan akumulasi proporsi menjawab benar dan salah sehingga mendapat tingkat kesukaran butir soal. Analogi sederhana satu kelas berjumlah 30 siswa diberikan tes pilihan ganda jumlah 10 butir soal. Pada butir soal nomor 5 terdapat 25 siswa menjawab benar maka kita dapat menyatakan bahwa soal nomor 5 mudah. Seterusnya pada butir soal nomor 2 terdapat 7 orang yang menjawab benar maka kita dapat mengatakan soal nomor 2 sulit.

Tujuan utama dari pengukuran tingkat kesukaran adalah untuk mengetahui sejauh mana sebuah soal dapat dijawab oleh peserta didik, sehingga dapat dijadikan dasar menentukan apakah soal tersebut perlu diperbaiki, dipertahankan, atau dibuang. Melalui analisis tingkat kesukaran, guru dapat menyusun komposisi soal seimbang antara soal mudah, sedang, dan sulit setiap ujian, agar pencapaian hasil belajar peserta didik dapat diukur secara adil dan proporsional. Mengetahui tingkat kesukaran, pendidik juga dapat menyesuaikan tingkat tantangan soal sesuai kompetensi pada kurikulum pembelajaran (Arikunto, 2012). Analisis tingkat kesukaran juga bertujuan untuk meningkatkan efektivitas instrumen penilaian sebagai alat ukur kemampuan kognitif siswa. Soal terlalu mudah tidak mampu membedakan antara siswa kompeten dan tidak, sementara soal terlalu sulit dapat menurunkan motivasi dan tidak merepresentasikan pencapaian pembelajaran yang sebenarnya. Oleh karena itu, pengukuran tingkat kesukaran menjadi bagian penting untuk proses validasi soal, serta menjadi dasar pengambilan keputusan perbaikan instrumen evaluasi pada ujian berikutnya. (Fitriyani, dkk, 2020).

Tahapan dalam menentukan tingkat kesukaran soal dimulai dari pengumpulan data hasil tes peserta didik, kemudian dilanjutkan menghitung jumlah peserta menjawab benar tiap butir soal. Proses ini dilakukan untuk mengetahui proporsi siswa memiliki kemampuan menyelesaikan soal benar dan dinyatakan dengan indeks kesukaran. Berdasarkan hasil analisis ini, guru dapat mengetahui apakah soal tergolong mudah, sedang, atau sukar. Tingkat kesukaran dihitung melalui formula sederhana sebagai berikut:

$P = \frac{B}{JS}$	Pengertian Huruf; P : tingkat kesukaran
--------------------	--

	<p>B : banyaknya siswa yang menjawab soal benar</p> <p>JS : jumlah seluruh siswa</p>
--	--

Tingkat kesukaran soal memiliki kriteria mudah, sedang, dan sukar (Ratnawulan, 2014). Menurut Matondang, dkk (2019) bahwa perbandingan antara soal mudah-sedang-sukar bisa dibuat 3-4-3. Artinya, 30% soal kategori mudah, 40% soal kategori sedang, dan 30% soal kategori sukar. Perbandingan lain yang termasuk. Penafsiran tingkat kesukaran butir diartikan semakin kecil indeks yang diperoleh, maka soal tersebut sulit, sebaliknya makin besar indeks yang diperoleh, makin soal tersebut mudah. Indeks tingkat kesukaran sebagai berikut:

Tabel 1 Indeks tingkat kesukaran

Indeks Kesukaran (P)	Kriteria
0,00 - 0,30	Soal Sukar
0,31 - 0,70	Soal Sedang
0,71- 1,00	Soal Mudah

Selain indeks dari indeks di atas tersebut ada juga menggunakan lima skala yakni;

Tabel 2. Indeks lima skala

Indeks Kesukaran (P)	Kriteria
0,00 - 0,20	Soal Sukar Sekali
0,21 - 0,40	Soal Sukar
0,41- 0,60	Soal Sedang
0,61 - 0,80	Soal Mudah
0,81- 1,00	Soal Mudah Sekali

(Wayan, 2011).

Berdasarkan paparan tersebut dapat dipahami bahwa hal yang mendasar kita lakukan untuk melihat tingkat kesukaran adalah melakukan perhitungan dengan menggunakan formula/ rumus yang telah ditentukan seterusnya melakukan interpretasi menggunakan indeks berdasarkan teori yang relevan, kemudian membuat kesimpulan apakah butir soal mudah, sedang maupun susah dan akhir menganalisis tingkat kesukaran soal sudah selesai dikerjakan.

Daya pembeda soal merupakan salah satu elemen penting untuk menganalisis kualitas butir soal untuk membedakan antara peserta tes yang memiliki kemampuan tinggi dan rendah. Soal berkualitas baik seharusnya mampu menyaring siswa yang benar-benar memahami materi dari mereka yang belum menguasainya. Menurut Dali S. Naga (2002) daya pembeda dapat ditentukan melalui skornya untuk membedakan peserta tes dari kelompok berkemampuan tinggi dan kelompok berkemampuan rendah. Fatimah, L. U., & Alfath, K. (2019) juga berpendapat bahwa kemampuan antara butir soal dapat membedakan antara peserta didik menguasai materi dan peserta didik belum menguasai materi. Sedangkan menurut Arifin (2012) menyatakan bahwa daya pembeda adalah sejauh mana suatu butir soal mampu membedakan peserta didik sudah menguasai materi dengan peserta didik belum atau kurang menguasai materi berdasarkan kriteria tertentu. Hanifah Nani (2017) juga mengemukakan pendapat bahwa daya pembeda soal adalah kemampuan soal untuk membedakan kelompok peserta tes berkemampuan tinggi dan

kelompok peserta tes berkemampuan rendah. Senada dengan itu, Pramono & Kusri (2018) serta Al-Bahrum menjelaskan bahwa daya pembeda merupakan kemampuan sebuah butir soal untuk menunjukkan perbedaan antara mahasiswa yang berkemampuan tinggi dan mahasiswa berkemampuan rendah.

Dari paparan penulis dapat disimpulkan bahwa daya pembeda adalah salah satu hal terpenting dalam proses menganalisis butir soal yang berkualitas baik dengan kemampuan butir soal untuk membedakan peserta didik berkemampuan tinggi dan rendah. Soal dengan kualitas baik akan lebih banyak dijawab oleh peserta didik yang memang menguasai materi yang diajarkan. Analogi sederhananya adalah ketika guru mengujikan soal tes kepada peserta didik dengan jumlah 10 soal. Soal tersebut dapat dijawab oleh kelompok siswa berkemampuan tinggi dan berkemampuan rendah. Maka bisa dikatakan daya pembeda soal tersebut tidak berfungsi dan tidak berkualitas baik. Karena soal yang memiliki daya pembeda yang baik tidak akan terjawab oleh kelompok siswa yang memiliki kemampuan rendah.

Tujuan daya pembeda soal ini adalah dapat menentukan bahwa soal tes yang diujikan benar - benar mengukur sejauh mana tingkat kemampuan peserta didik dalam memahami materi yang diajarkan sesuai aspek yang diukur. Jadi, tidak selamanya soal berkualitas baik itu sulit atau mudah, tetapi juga soal yang dapat menyaring peserta didik berdasarkan tingkat pemahamannya. Suryabrata (1999) berpendapat bahwa tujuan utama dari analisis daya beda adalah untuk menentukan apakah suatu butir soal mampu membedakan kelompok peserta tes berdasarkan aspek yang diukur, sesuai dengan perbedaan nyata yang ada dalam kelompok tersebut. Oleh karena itu, daya pembeda menjadi unsur yang sangat penting dalam penyusunan soal, karena dapat memastikan bahwa proses penilaian berjalan dengan adil, tepat sasaran, dan benar-benar mencerminkan kemampuan peserta didik yang sesungguhnya.

Tahapan-tahapan yang harus dilakukan, salah satunya dengan membagi kelompok peserta tes berdasarkan jumlahnya. Syamsudin (2012) menjelaskan bahwa jika jumlah peserta kurang dari 100 orang (kelompok kecil), maka peserta dibagi menjadi dua kelompok yang sama besar: kelompok atas dan kelompok bawah. Sedangkan untuk jumlah peserta 100 orang atau lebih (kelompok besar), digunakan persentase 27% peserta dengan skor tertinggi sebagai kelompok atas, dan 27% peserta dengan skor terendah sebagai kelompok bawah. Peserta di bagian tengah tidak diperhitungkan dalam analisis ini. Perlu diperhatikan sebelum membagi kelompok atas dan bawah terlebih dulu urutkan hasil skor anak dari yang tertinggi hingga sampai skor terendah, seterusnya gunakan metode pembagi kelompok berdasarkan jumlah peserta dan seterusnya dapat dihitung dengan menggunakan rumus berikut:

$DB = \frac{KA-KB}{JP}$	Pengertian Huruf; DP = Daya Pembeda KA = Kelompok Atas KB = Kelompok Bawah JP = Jumlah Pembagi
---	--

Setelah melakukan perhitungan dengan menggunakan rumus, maka dapat dilakukan penafsiran melalui indeks antara lain:

Tabel 3. Penafsiran Indeks

Indeks Daya Pembeda (DB)	Kriteria
< 0,19	Tidak Baik
0,20 - 0,29	Cukup Perlu Sedikit Revisi
0,30- 0,39	Baik
0,40 – 1,00	Sangat Baik

(Mansyur & Rasyid, 2007)

Dapat menggunakan indeks dengan lima skala antara lain:

Tabel 5. Indeks lima skala

Indeks Daya Pembeda (DB)	Kriteria
< 0,00	Tidak Baik
0,00 - 0,20	Kurang Baik
0,21- 0,40	Cukup
0,41 – 0,70	Baik
0,71 – 1,00	Sangat Baik

(Zalukhu dkk, 2023)

Berdasarkan paparan tersebut dapat dipahami bahwa hal yang mendasar kita lakukan untuk melihat daya pembeda melakukan pengurutan skor perolehan dari jumlah siswa dari paling tinggi ke rendah, seterusnya membagi kelompok atas dan bawah sesuai dengan teori jumlah peserta yang digunakan, kemudian menghitung dengan rumus yang sudah ditentukan, setelah mendapatkan hasil baru dilakukan penafsiran melalui indeks daya pembeda, dan akhirnya mendapatkan keputusan soal dengan kategori berdasarkan indeks yang digunakan.

Pengecoh Soal

Bahasa Inggris dari pengecoh adalah (*distractor*) ini biasanya ditemukan pada soal pilihan ganda pada opsi pilihan jawaban memiliki satu jawaban benar dan sisa pilihannya disebut pengecoh. Sejalan dengan Leonardus et. al. (2024) menyatakan efektivitas pengecoh adalah suatu pola yang menunjukkan bagaimana peserta tes memilih jawaban berdasarkan pilihan tersedia pada setiap butir soal. Ini mencerminkan sejauh mana pengecoh berfungsi mengarahkan peserta tes untuk memilih jawaban salah namun masuk akal. Basuki & Hariyanto (2017) menyatakan pengecoh merupakan jawaban sifatnya itu menyesatkan atau mengecoh jawaban siswa. Seterusnya Setiawaty (2017) berpendapat bahwa pengecoh merupakan jawaban pengecoh yang memiliki perbedaan tipis dengan jawaban benar, sehingga sering menjebak siswa dalam menjawab soal tes.

Berdasarkan teori-teori tersebut penulis memberikan pandangan bahwa pengecoh itu pilihan jawaban yang salah. Pengecoh dirancang untuk menyesatkan siswa dan siswi, sehingga mereka mungkin memilih jawaban salah. Ini menunjukkan bahwa pengecoh memiliki tujuan untuk menguji kemampuan siswa dalam memilih jawaban yang benar nantinya, pengecoh sering kali mirip dengan jawaban benar, sehingga bisa membuat siswa bingung dan terjebak saat menjawab soal tes. Ini menunjukkan bahwa pengecoh dirancang secara cermat untuk meningkatkan kesulitan soal tetapi jika peserta tes tidak hanya sekedar menghafal tapi juga

mencermati sekaligus memahami setiap tes pasti peserta tes tidak ada kesulitan untuk mengetahui mana jawaban yang benar dan mana pengecoh soal dalam tes tersebut.

Menurut Arikunto (2008) mengatakan sebuah pengecoh dikatakan berfungsi apabila mempunyai daya tarik besar bagi peserta-peserta tes yang kurang memahami konsep atau kurang menguasai bahan materi. pengecoh dikatakan telah dapat menjalankan fungsinya dengan baik apabila pengecoh tersebut mempunyai daya tarik sedemikian rupa sehingga peserta didik yang kurang menguasai konsep merasa bimbang dan pada akhirnya mereka terkecoh untuk memilih pengecoh sebagai jawaban benar (Rahayu & Djazari, 2016) dan pengecoh yang tidak efektif karena tidak ada peserta memilihnya harus dibuang atau direvisi (Putri et al., 2020). Semakin banyak yang memilih pengecoh maka dapat dikatakan pengecoh tersebut berfungsi, sebaliknya jika pengecoh tidak dipilih maka pengecoh tersebut tidak berfungsi (Eni Arbiatin dan Mulabbiyah, 2020). Butir soal yang efektivitas pengecohnya tidak baik dan sangat tidak baik mengindikasikan bahwa pengecoh belum berfungsi dengan baik, pengecoh tersebut tidak memiliki daya tarik bagi peserta didik yang kurang menguasai konsep atau materi (Rahayu & Djazari, 2016).

Paparan di atas penulis memberikan pandangan bahwa pengecoh soal adalah pilihan jawaban salah, seterusnya efektivitas pengecoh dapat dilihat ketika butir soal salah ada memilihnya maka pengecohnya berfungsi, apabila tidak ada memilih salah satu pilihan jawaban maka pengecoh soal tidak berfungsi.

Contoh sederhana:

Soal dengan opsi pilihan

A	B	C	D
---	---	---	---

Pada kasus ini kunci jawaban A, maka B,C,D adalah pengecoh. Apabila pilihan B,C,D ada yang menjawab maka pengecoh pada pilihan B,C,D berfungsi.

Apabila B tidak ada yang menjawab maka pengecoh pada pilihan B tidak berfungsi.

Maka Pengecoh dapat dihitung dengan menggunakan perhitungan antara lain:

$IP = \frac{P(O)}{JS} \times 100\%$	Pengertian Huruf;
	IP : Indeks daya Pembeda
	P (O) : Jumlah memilih setiap pilihan Jawaban
	JS : jumlah seluruh siswa

Untuk memudahkan pemahaman maksud rumus P (O) yakni ada setiap pilihan jawaban harus dihitung contoh 100 siswa menjawab soal nomor 1 dengan jawaban antara lain:

Pilihan	A	B	C	D
Jawaban Siswa	4	16	70	10

Kasus ini kunci Jawaban C

Tabel 6. Perhitungan Pengecoh

Pilihan	Perhitungan	Pengkali 100%
A	$4/100 = 0,04$	4 %
B	$16/100 = 0,16$	16 %
D	$10/100 = 0,10$	10 %

Menghitung pengecoh soal sudah selesai maka kita perlu membuat keputusan berdasarkan indeks antara lain:

Indeks Pengecoh	Kriteria
$\geq 5\%$	Efektif/ Berfungsi
$5\% <$	Tidak Efektif

(Muchlizani, dkk., 2023)

Sejalan dengan Putri, dkk (2020) berpendapat bahwa pengecoh berfungsi untuk mengecoh peserta tes memilih jawaban benar. Soal pilihan ganda terdapat alternatif jawaban yang berfungsi untuk mengecoh jawaban yang benar. Terdapat tiga jawaban pengecoh yang terpasang pada setiap butir jawaban soal. Sehingga dikatakan bahwa sebuah pengecoh dapat dikatakan baik jika setidaknya 5% dari peserta tes memilih opsi tersebut. Seterusnya apabila pengecoh tidak efektif tindak lanjut yakni; membuang soal tersebut, atau memperbaiki pilihan pengecoh yang tidak efektif dan diujikan kembali untuk melihat hasilnya. Berdasarkan paparan tersebut dapat dipahami bahwa hal yang mendasar kita lakukan untuk melihat pengecoh soal adalah pilihan jawaban salah, seterusnya dilakukan perhitungan setiap pilihan jawaban salah tersebut, setelah mendapatkan hasil perhitungan maka dapat ditafsirkan berdasarkan indeks yang ditetapkan, baru selanjutnya dilakukan tindak lanjut terhadap soal.

Simpulan

Tingkat kesukaran soal merupakan salah satu indikator penting dalam menilai kualitas butir soal dalam evaluasi pembelajaran. Tingkat kesukaran diukur melalui proporsi siswa yang menjawab benar terhadap jumlah seluruh peserta, yang selanjutnya diinterpretasikan menggunakan kriteria tertentu untuk menentukan kategori soal: mudah, sedang, atau sulit. hal mendasar dilakukan untuk melihat tingkat kesukaran adalah melakukan perhitungan seterusnya interpretasi menggunakan indeks berdasarkan teori yang relevan, kemudian membuat kesimpulan terhadap kriteria butir soal mudah, sedang maupun susah dan akhir. Tujuan utama dari analisis ini adalah untuk memastikan bahwa soal-soal yang digunakan mampu mengukur kemampuan siswa secara adil, proporsional, dan sesuai dengan target dan tujuan pembelajaran.

Daya pembeda adalah kapasitas, potensi sebuah soal dalam hal membedakan dan menentukan peserta didik yang memiliki kognitif atau pengetahuan di atas rata-rata dengan peserta didik yang memiliki tingkat kemampuan di bawah rata-rata yang bertujuan untuk menilai, menetapkan dan menentukan sejauh mana suatu butir soal mampu membedakan antara peserta didik yang memiliki tingkat kemampuan di atas rata-rata dengan peserta didik yang memiliki tingkat kemampuan di bawah rata-rata sesuai aspek yang diukur. Hal yang mendasar dilakukan untuk melihat daya pembeda melakukan pengurutan skor perolehan dari jumlah siswa dari paling tinggi ke rendah, seterusnya membagi kelompok atas dan bawah sesuai dengan teori

jumlah peserta yang digunakan, kemudian melakukan perhitungan, setelah mendapatkan hasil perhitungan, selanjutnya dilakukan penafsiran melalui indeks daya pembeda, dan akhirnya mendapatkan keputusan soal dengan kategori berdasarkan indeks yang digunakan. Jadi, soal yang baik bukan hanya soal sulit atau mudah, akan tetapi juga soal bisa menyaring peserta berdasarkan tingkat pemahamannya.

Pengecoh soal merupakan pilihan jawaban salah, seterusnya efektivitas pengecoh dapat dilihat ketika butir soal salah ada memilihnya maka pengecohnya berfungsi, apabila tidak ada memilih salah satu pilihan jawaban maka pengecoh soal tidak berfungsi. hal mendasar dilakukan untuk melihat pengecoh soal adalah pilihan jawaban salah, seterusnya dilakukan perhitungan setiap pilihan jawaban salah tersebut, setelah mendapatkan hasil perhitungan maka dapat ditafsirkan berdasarkan indeks yang ditetapkan, baru selanjutnya dilakukan tindak lanjut terhadap soal.

Daftar Pustaka

- Amelia, M. A. (2017). Analisis Soal Tes Hasil Belajar High Order Thinking Skills (HOTs) Matematika Materi Pecahan Untuk Kelas 5 Sekolah Dasar. *Jurnal Penelitian*, 20(2), <http://e-journal.usd.ac.id/index.php/JP/article/viewFile/869/685>
- Anastasi, Anne; Urbina, Susan. (2007.) Tes Psikologi. Jakarta: Prenhallindo
- Anita, A., Tyowati, S., & Zulfadrial, Z. (2018). Analisis kualitas butir soal fisika kelas x sekolah menengah atas. *Edukasi: Jurnal Pendidikan*, 16(1), 35-47.
- Arifin, Z. (2017). Kriteria instrumen dalam suatu penelitian. *Jurnal Theorems (the original research of mathematics)*, 2(1), 28-36.
- Arikunto, S. (2015). Dasar-dasar Penilaian Pendidikan; ke-4). PT Bumi Aksara.
- Arikunto, S. (2008). Dasar-Dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara.
- Azis. (2016). Analisis Tes Buatan Guru Bidang Studi Matematika Kelas V SD I Katobengke. ISSN:2008-2157, 6(2), 18.
- Azwar, Saifuddin. (2006). Tes Prestasi. Yogyakarta: Pustaka Pelajar.
- Azwar, Saifuddin. (2006). Reliabilitas, Validitas, Interpretasi dan Komputasi. Yogyakarta: Liberty.
- Basuki, I., & Hariyanto. (2017). Asesmen Pembelajaran; ke-4). PT Remaja Rosdakarya.
- Crocker, L. M., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart and Winston.
- Eni Arbiatin dan Mulabbiyah. (2020). Penilaian Akhir Semester Mata Pelajaran Matematika. 12(2), 146–171.
- Fatimah, L. U., & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda dan fungsi distraktor. *AL-MANAR: Jurnal Komunikasi dan Pendidikan Islam*, 8(2), 37-64.
- Fauziah, I. R. N., Syihabudin, S., & Sopian, A. (2020). Analisis Kualitas Tes Bahasa Arab Berbasis Higher Order Thinking Skill (Hots). *لساننا (LISANUNA): Jurnal Ilmu Bahasa Arab dan Pembelajarannya*, 10(1), 45-54.
- Hanifah, N. (2017). Perbandingan tingkat kesukaran, daya pembeda butir soal dan reliabilitas tes bentuk pilihan ganda biasa dan pilihan ganda asosiasi mata pelajaran ekonomi. *Sosio e-KONS*, 6(1).

- Hopkins, Charles D dan Antes, Richard L. (1999). *Classroom Measurement and Evaluation*. Illionis, F.E. Peacock.
- Joseph, M., & Lumban Gaol, P. (2019). Analisis efektivitas pengecoh soal UAS mapel Matematika kelas IV SD GMIH 4 Tobelo. Widyasari Press. <https://widyasari-press.com/analisis-efektivitas-pengecoh-soal-uas/>
- Leonardus Kaka, Vidriana Oktoviana Bano & Yohana Njoeroemana. 2024. Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates di SMPN 2 Kanatang. *Jurnal Inovasi Pendidikan*, 4(9), 1441-1449, h
- Kementerian Pendidikan dan Kebudayaan. (2018). *Buku Tematik Terpadu Kurikulum 2013 Revisi untuk SD/MI*. Jakarta: Kemendikbud.
- Komarudin, U., Wulandari, M. N., Ramadhani, F. A., & Marfu'ah, S. (2024). Pengembangan Instrumen Penelitian. *YPAD Penerbit*.
- Kurniawan, T. K. (2015). Analisis butir soal ulangan akhir semester gasal mata pelajaran IPS Sekolah Dasar. *Journal of Elementary Education*, 4(1), 1-6.
- Kustati, M., & Amelia, R. (2024). Pelajaran Pai Di Sma Negeri 10 (Telaah Terhadap Reliabilitas , Daya Kesukaran Menggunakan Software Anates). 7, 6954–6961.
- Mochammad Noor Akhmadi. (2021). Analisis Butir Soal Evaluasi Tema 1 Kelas 4 Sdn Plumbungan Menggunakan Program Anates. *Ed-Humanistics: Jurnal Ilmu Pendidikan*, 6(1), 799–806.