

## Analisis Kualitas Butir Soal Ujian Tengah Semester (UTS) Ganjil Biologi Kelas XI SMAN 1 Pancung Soal

Annisa Anike Putri<sup>a,1\*</sup>, Zulyusri<sup>b,2</sup>, Violita<sup>c,3</sup>

<sup>a,b,c</sup> Program Studi Pendidikan Biologi, FMIPA UNP, Padang, Indonesia

<sup>1</sup> [annisaputri8465@gmail.com](mailto:annisaputri8465@gmail.com); <sup>2</sup> [zulyusri0808@mail.com](mailto:zulyusri0808@mail.com); <sup>3</sup> [violita@fmipa.unp.ac.id](mailto:violita@fmipa.unp.ac.id)

\* Corresponding author

### Informasi Artikel

#### Histori Artikel

Submission: 8/12/2021

Accepted: 26/7/2023

Published: 31/08/2023

#### Kata Kunci

Validitas;

Reabilitas;

Tingkat Kesukaran;

Daya Pembeda;

Distractor

### Abstrak

Penelitian ini bertujuan untuk mengetahui kualitas butir soal UTS Ganjil Biologi Kelas XI SMAN 1 Pancung Soal. Metode penelitian yang digunakan adalah penelitian deskriptif kuantitatif. Adapun sampel dalam penelitian ini yaitu siswa Biologi kelas XI SMAN 1 Pancung Soal tahun ajaran 2021/2022 dengan jumlah 16 siswa. Teknik pengumpulan data yang digunakan berupa dokumentasi. Instrumen penelitian ini meliputi 25 soal pilihan ganda, kunci jawaban, dan lembar jawaban hasil pekerjaan siswa. Hasil penelitian dianalisis validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh menggunakan bantuan *software* Anates 4.0.9. Hasil penelitian menunjukkan bahwa dari aspek validitasnya yaitu terdapat 1 (4%) butir soal dengan kategori tinggi, 4 (16%) butir soal dengan kategori cukup, 14 (56%) butir soal dengan kategori sangat rendah, 3 (12%) butir soal dengan kategori rendah serta juga ada 3 (12%) butir soal yang tidak terbaca tingkat kevalidannya. Aspek reliabilitasnya yaitu memiliki kriteria sangat rendah dengan harga  $r$  statistik sebesar -2,88. Aspek tingkat kesukarannya yaitu tidak diperoleh butir soal dengan kriteria mudah, hanya kriteria sedang dengan jumlah 9 (36%) butir soal dan kriteria sukar dengan jumlah 16 (64%) butir soal. Aspek daya pembedanya yaitu terdapat 5 (20%) dengan kriteria sangat baik (diterima), 6 (24%) butir soal dengan kriteria sedang (diterima dan revisi), 8 (31%) butir soal dengan kriteria jelek (direvisi), dan 5 (24%) butir soal dengan kriteria sangat jelek (harus dibuang/diganti). Kemudian aspek kualitas pengecoh (*distractor*) yaitu dari 25 butir soal hanya terdapat 3 (12%) butir soal yang semua distraktornya berfungsi dengan baik dan 22 (88%) butir soal yang distraktornya ada yang berfungsi dan ada yang tidak berfungsi sama sekali. Hasil analisis ini mengindikasikan bahwa secara keseluruhan soal tes tersebut harus direvisi kembali sehingga dapat digunakan lagi dalam tes hasil belajar yang akan datang.

©2023 The Author's

This is an open-access article under the [CC-BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/) license.



<https://doi.org/10.37058/metaedukasi.v4i2.4087>

## Pendahuluan

Kualitas pendidikan menjadi perhatian setiap negara. Melalui pendidikan kualitas sumber daya manusia dapat ditingkatkan, sehingga kesejahteraan masyarakat juga diharapkan akan meningkat. Kualitas pendidikan selalu dilihat pada kemampuan lulusan suatu jenjang pendidikan. Kemampuan lulusan terdiri dari aspek pengetahuan, keterampilan, dan sikap. Kemampuan tersebut dicapai melalui proses pembelajaran di sekolah. Pembelajaran hakikatnya adalah usaha sadar dari seorang guru untuk membelajarkan siswanya (mengarahkan interaksi siswa dengan sumber belajar lainnya) dalam rangka mencapai tujuan yang diharapkan. Oleh karena itu, dibutuhkan guru yang profesional dalam mendidik peserta didik yang unggul dan berkualitas (Sukmela & Nofrion, 2018).

Dalam rangka mewujudkan tujuan pendidikan dan untuk melihat keberhasilan siswa perlu adanya proses pembelajaran. Proses pembelajaran di sekolah tidak bersifat mekanisme saja, tapi memiliki visi, misi dan tujuan sebagai target yang harus dicapai. Untuk mencapai target tersebut perlu dilakukan

evaluasi sebagai tolak ukur tingkat pencapaian pembelajaran dari materi yang sudah dipelajari. Dengan demikian peran utama dari evaluasi adalah untuk menentukan tingkat ketercapaian tujuan pendidikan dan pembelajaran (Monica & Sudarman, 2019).

Menurut Magdalena et al., (2021), evaluasi hasil belajar dilakukan oleh pendidik untuk memantau proses, kemajuan, pencapaian dan perbaikan hasil belajar peserta didik secara berkesinambungan. Oleh karena itu seorang pendidik atau guru dituntut memiliki kemampuan dalam merencanakan, menyusun dan membuat alat penilaian (evaluasi). Selain itu Solichin (2017) juga mengungkapkan bahwa kegiatan evaluasi ini perlu dilakukan terutama untuk menciptakan kesempatan bagi para siswa untuk memperlihatkan prestasi mereka dalam kaitannya dengan tujuan yang telah ditentukan dalam kurikulum. Oleh karena itu evaluasi merupakan alat pemicu pengantar prestasi belajar siswa secara merata.

Menurut Susdelina et al., (2018), salah satu cara yang dapat ditempuh untuk mengukur kemampuan peserta didik dalam melaksanakan kegiatan evaluasi adalah dengan menggunakan instrumen tes atau non tes. Jenis instrumen yang umum digunakan untuk mengukur kemampuan peserta didik di sekolah adalah instrumen tes. Menurut Thoha (2001), tes adalah alat pengukuran berupa pertanyaan, perintah, dan petunjuk dalam bentuk soal-soal yang ditujukan kepada *testee* untuk mendapatkan respons sesuai dengan soal yang diberikan. Tes yang dijadikan bahan evaluasi sering kali tidak dilakukan analisis kualitas butir soal oleh guru, sehingga belum diketahui kualitas soal yang telah dibuat. Oleh karena itu, untuk mengetahui kualitas butir soal yang digunakan untuk tes perlu dilakukan analisis butir soal.

Analisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan kualitas soal yang telah ditulis. Menurut Ulum (2017), analisis kualitas tes bertujuan untuk mengidentifikasi soal-soal yang baik, kurang baik, dan soal yang dapat masuk dalam bank soal, direvisi, atau dibuang. Febriani & Saksono (2016) menyatakan bahwa tes yang baik harus dapat dipertanggungjawabkan dalam segi kelayakan, kesahihan (validitas), keterpercayaan (reliabilitas), ketafsiran, kebergunaan, dan efektivitas butir soal yang meliputi tingkat kesulitan dan daya pembeda soal yang baik. Hal ini selaras dengan Arifin (2013) bahwa analisis butir soal dapat dilakukan melalui uji validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh atau pola jawaban soal.

Salah satu bentuk tes yang digunakan guru di sekolah adalah soal Ujian Tengah Semester (UTS) yang termasuk dalam kategori tes sumatif. Menurut Widoyoko (2014), fungsi tes sumatif yaitu untuk mengetahui sejauh mana pencapaian kompetensi siswa dalam mata pelajaran tertentu. Hasil tes sumatif siswa selanjutnya dibandingkan dengan tujuan yang seharusnya dicapai atau kriteria ketuntasan minimal (KKM). Oleh karena itu, nilai tes sumatif dijadikan sebagai catatan kemajuan belajar siswa dan penentu seorang siswa dapat atau tidaknya menerima program berikutnya.

Berdasarkan uraian di atas, maka perlu dilakukan analisis butir soal secara kuantitatif untuk mengetahui kualitas soal tersebut. Kualitas butir soal UTS dapat dilihat dari hasil analisis validitas, reliabilitas, tingkat kesukaran, daya pembeda dan kualitas pengecoh (*distractor*) soal. Peneliti tertarik untuk melakukan penelitian dengan judul “ Analisis Kualitas Butir Soal Ujian Tengah Semester (UTS) Ganjil Biologi Kelas XI SMAN 1 Pancung Soal”.

## Metode

Metode penelitian yang digunakan adalah penelitian deskriptif kuantitatif. Adapun sampel dalam penelitian ini yaitu siswa Biologi kelas XI SMAN 1 Pancung Soal tahun ajaran 2021/2022 dengan jumlah 16 siswa. Teknik pengumpulan data yang digunakan berupa dokumentasi. Instrumen penelitian ini meliputi 25 soal pilihan ganda UTS ganjil, kunci jawaban, dan lembar jawaban hasil pekerjaan siswa.

Hasil penelitian dianalisis validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh menggunakan bantuan *software* Anates 4.0.9.

## Hasil dan Pembahasan

### Hasil

Berikut ini data hasil penelitian dapat dilihat dari aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda dan kualitas pengecoh soal yang dianalisis dengan menggunakan bantuan *software* Anates 4.0.9.

**Tabel 1.** Hasil analisis validitas, reliabilitas, daya pembeda dan tingkat kesukaran soal

Nomor Soal	Validitas		Reliabilitas	Daya pembeda		Tingkat kesukaran	
	Rxy	Kriteria		DP	Kriteria	TK	Kriteria
1	0,493	Cukup	-2,88	0,5	Sangat baik	0,125	Sukar
2	0,092	Sangat rendah	(Sangat rendah)	0	Jelek	0,0625	Sukar
3	-0,404	Sangat rendah		-0,25	Dibuang/diganti	0,125	Sukar
4	0,627	Tinggi		0,5	Sangat baik	0,1875	Sukar
5	0,019	Sangat rendah		0	Jelek	0,1875	Sukar
6	0,016	Sangat rendah		-0,25	Dibuang/diganti	0,6875	Sedang
7	0,314	Rendah		0,25	Sedang	0,125	Sukar
8	0,069	Sangat rendah		0,25	Sedang	0,25	Sukar
9	0,398	Rendah		0,75	Sangat baik	0,375	Sedang
10	0,135	Sangat rendah		0	Jelek	0,125	Sukar
11	-	-		0	Jelek	0	Sukar
12	0,153	Sangat rendah		0	Jelek	0,375	Sedang
13	0,475	Cukup		0,25	Sedang	0,1875	Sukar
14	-	-		0	Jelek	0	Sukar
15	0,337	Rendah		0,25	Sedang	0,0625	Sukar
16	0,496	Cukup		0,50	Sangat baik	0,3125	Sedang
17	-0,144	Sangat rendah		0,25	Sedang	0,3125	Sedang
18	-0,144	Sangat rendah		-0,25	Dibuang/diganti	0,3125	Sedang
19	-	-		0	Jelek	0	Sukar
20	-0,206	Sangat rendah		-0,25	Dibuang/diganti	0,25	Sukar
21	-0,404	Sangat rendah		-0,50	Dibuang/diganti	0,4375	Sedang
22	0,496	Cukup		0,50	Sangat baik	0,3125	Sedang
23	0,092	Sangat rendah		0,25	Sedang	0,0625	Sukar
24	0,092	Sangat rendah		0	Jelek	0,0625	Sukar
25	-0,112	Sangat rendah		-0,25	Dibuang/diganti	0,6875	Sedang

\*Sumber: dokumentasi pribadi

**Tabel 2.** Data tentang berfungsi atau tidaknya distraktor

Nomor Soal	Kunci Jawaban	Fungsi distraktor	
		Baik	Tidak Baik
1	E	A,B,D	C
2	E	A,B	C,D
3	E	A,C	B,D
4	B	D,E	A,C

5	E	A,C	B,D
6	C	-	A,B,D,E
7	B	D,E	A,C
8	B	C	A,D,E
9	C	B,D	A,E
10	E	A,B,C,D	-
11	E	C	A,B,D
12	B	A,D	C,E
13	A	B,C	D,E
14	B	A,C,E	D
15	D	A,C	B,E
16	E	B,C	A,D
17	E	A,B,C,D	-
18	A	C	B,D,E
19	E	A,B,C	D
20	D	A,B,C,E	-
21	C	D,E	A,B
22	E	A,B,C,D	-
23	B	D,E	A,C
24	D	A,C	B,E
25	A	-	B,C,D,E

<sup>a</sup>(annisa, 2021)

## Pembahasan

Berdasarkan analisis soal yang dilakukan dengan menggunakan Anates 4.0.9 maka dapat diketahui validitas, reabilitas, tingkat kesukaran, daya pembeda dan kualitas pengecoh soal.

### 1. Validitas

Menurut Arikunto (2010) validitas adalah suatu ukuran yang menunjukkan tingkatan-tingkatan kevalidan atau kesahihan suatu instrumen sehingga dapat mengukur apa yang hendak diukur. Hal ini didukung oleh Elviana (2020) bahwa suatu tes mempunyai validitas tinggi apabila tes tersebut dapat menjalankan fungsi ukurnya atau memberikan hasil ukur yang sesuai dengan maksud dan tujuan diadakan tes tersebut. Sebaliknya, suatu tes yang menghasilkan data tidak relevan dengan tujuan pengukuran dikatakan sebagai tes yang memiliki validitas rendah. Jika instrumen valid maka dapat dilihat dari kategori penafsiran mengenai indeks korelasi ( $r$ ) pada tabel 3.

**Tabel 3.** Kriteria validitas butir soal

Indeks korelasi	Keterangan
$0.80 \leq r_{xy} \leq 1.00$	Sangat tinggi
$0.60 \leq r_{xy} \leq 0.79$	Tinggi
$0.40 \leq r_{xy} \leq 0.59$	Cukup
$0.20 \leq r_{xy} \leq 0.39$	Rendah
$0.00 \leq r_{xy} \leq 0.19$	Sangat rendah

\*Sumber: Arif (2014)

Berdasarkan Tabel 1, diketahui bahwa dari 25 butir soal hanya terdapat 4% butir soal yang masuk dalam kategori tinggi yaitu soal nomor 4, 16% butir soal masuk dalam kategori cukup (1, 13, 16, 22), 56% butir soal masuk dalam kategori sangat rendah (2, 3, 5, 6, 8, 10, 12, 17, 18, 20, 21, 23, 24, 25), dan 12% butir soal masuk dalam kategori rendah (7, 9, 15), serta juga ada 12% butir soal yang tidak

terbaca tingkat kevalidannya (11, 14, dan 19) artinya soal tersebut harus diganti karena sama sekali tidak bisa digunakan.

Berdasarkan hasil di atas, maka dapat disimpulkan hanya 1 butir soal yang memiliki kriteria valid dan layak digunakan. Sedangkan 24 butir soal belum valid, oleh karena itu perlu dilakukan perbaikan terhadap soal-soal tersebut sehingga dapat digunakan kembali untuk tes yang akan datang. Menurut [Setyaningrum et al., \(2018: 95\)](#), butir soal dengan validitas tinggi menunjukkan soal itu secara tepat mampu mengukur kemampuan siswa. Sebaliknya, butir soal dengan validitas rendah mencerminkan perlu diadakannya perbaikan.

## 2. Reliabilitas

Menurut [Arikunto2 \(2013: 221\)](#), reliabilitas menunjukkan pada pengertian bahwa suatu instrumen dapat dipercaya untuk digunakan sebagai alat pengumpulan data karena instrumen tersebut sudah baik. Sedangkan [Purwati et al., \(2021: 50\)](#) dan [Balau et al., \(2021: 16\)](#) menyatakan bahwa reliabilitas artinya instrumen mampu menggambarkan sejauh mana konsistensi hasil ukur yang dilakukan berulang-ulang terhadap responden yang memiliki gejala sama dengan alat ukur yang sama. Untuk memberi interpretasi pengujian reliabilitas instrumen maka dapat dilihat pada Tabel 4.

**Tabel 4.** Kriteria Reliabilitas Butir Soal

Reabilitas soal	Interpretasi
0,81 – 1,00	Sangat tinggi
0,61 – 0,80	Tinggi
0,41 – 0,60	Cukup
0,21 – 0,40	Rendah
0,00 – 0,20	Sangat rendah

\*Sumber: [Marthunis et al., \(2015\)](#)

Berdasarkan hasil analisis yang dilakukan pada 25 butir soal menunjukkan bahwa diperoleh nilai reliabilitas data sebesar -2,88 dengan kriteria sangat rendah. Dapat disimpulkan bahwa soal-soal tersebut tidak memenuhi kriteria untuk digunakan dalam Ujian Tengah Semester (UTS) Ganjil dan perlu dilakukan revisi kembali.

Menurut [Agustiana et al., \(2018\)](#), rendahnya koefisien reliabilitas soal karena kesulitan tes. Karena salah satu faktor yang mempengaruhi reliabilitas yaitu kesulitan tes. Tes yang terlalu mudah atau terlalu sulit untuk siswa, cenderung menghasilkan reliabilitas rendah. Selain itu [Ebel dan Frisbie dalam Suzana \(2018\)](#) juga mengungkapkan bahwa tinggi rendahnya koefisien reliabilitas suatu tes dipengaruhi oleh beberapa faktor di antaranya yaitu: 1) banyaknya butir, 2) homogenitas materi tes, 3) homogenitas karakteristik butir, dan 4) variabilitas skor. Reliabilitas yang berhubungan dengan peserta didik dipengaruhi oleh faktor yaitu: 1) heterogenitas kelompok, 2) pengalaman peserta didik mengikuti tes, dan 3) motivasi peserta didik. Sedangkan faktor yang mempengaruhi reliabilitas yang berhubungan dengan administrasi adalah batas waktu dan kesempatan menyontek.

## 3. Tingkat kesukaran

Menurut [Sukardi \(2011\)](#) tingkat kesukaran adalah angka yang menunjukkan proporsi siswa yang menjawab betul dalam satu soal yang dilakukan dengan menggunakan tes objektif. [Daryanto \(2010\)](#) mengungkapkan bahwa soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya. Jadi dapat disimpulkan bahwa tingkat kesukaran adalah angka

yang menunjukkan bahwa apakah soal yang diujikan termasuk mudah, sedang atau sukar. Berikut ini kriteria indeks kesukaran butir soal pada Tabel 5.

**Tabel 5.** Kriteria Interpretasi tingkat kesukaran

Indeks kesukaran	Interpretasi
Kurang dari 0,3	Sukar
0,30 - 0,70	Sedang
Lebih dari 0,70	Mudah

\*Sumber: Marthunis et al. (2015)

Berdasarkan Tabel 1, diketahui bahwa dari 25 butir soal tidak terdapat butir soal dengan kriteria mudah, hanya kriteria sedang (6, 9, 12, 16, 17, 18, 21, 22, 25) dengan persentase 36% dan kriteria sukar (1, 2, 3, 4, 5, 7, 8, 10, 11, 13, 14, 15, 19, 20, 23, 24) dengan persentase 64%. Menurut Jusrianto et al., (2018), untuk butir soal yang termasuk kriteria sedang dapat dimasukkan dalam bank soal dan digunakan kembali sebagai alat evaluasi pada waktu yang akan datang. Sedangkan butir soal yang termasuk kriteria sukar perlu dilakukan tindak lanjut yaitu soal tersebut ditelusuri sehingga bisa diketahui faktor penyebab kegagalan siswa dalam menjawabnya. Kemudian dilakukan revisi untuk dapat diujikan kembali pada tes yang akan datang.

Menurut Nurfauziah & Windyariani (2020) menyatakan bahwa soal dengan kategori sukar, dapat diartikan sebagian besar siswa tidak dapat menjawab soal itu dengan baik. Kemudian didukung oleh Kemendiknas (2010), jika suatu butir soal termasuk kriteria sukar maka prediksinya adalah butir soal itu mungkin salah kunci jawaban, butir soal itu mempunyai 2 atau lebih jawaban benar, materi yang ditanyakan belum diajarkan atau belum tuntas pembelajarannya sehingga kompetensi minimum yang harus dikuasai siswa belum tercapai, materi yang diukur tidak cocok ditanyakan dengan menggunakan bentuk soal yang diberikan atau kalimat soal terlalu kompleks dan panjang.

#### 4. Daya pembeda soal

Daryanto (2010) menjelaskan bahwa daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang kurang pandai (berkemampuan rendah). Daya pembeda dapat diinterpretasikan seperti pada Tabel 6 di bawah ini.

**Tabel 6.** Indeks Daya Pembeda Butir Soal

Indeks Daya Pembeda	Interpretasi
$ID > 0,40$	Sangat baik
$0,30 \leq ID < 0,39$	Baik
$0,20 \leq ID < 0,29$	Sedang
$0,00 \leq ID < 0,19$	Jelek
$ID < 0.00$ (Negatif)	Dibuang/ Diganti

\*Sumber: Wahyuni dan Ibrahim (2012)

Berdasarkan Tabel 1 dapat diketahui bahwa dari 25 butir soal terdapat 20% butir soal dengan kriteria sangat baik (1, 4, 9, 16, 22), 24% butir soal dengan kriteria sedang (7, 8, 13, 15, 17, 23), dan 32% butir soal dengan kriteria jelek (2, 5, 10, 11, 12, 14, 19, 24) serta 24% butir soal harus dibuang/diganti (3, 6, 18, 20, 21, 25).

Menurut Sudijono (2009), untuk butir-butir soal yang sudah memiliki daya pembeda soal yang baik (*satisfactory*, *good*, dan *excellent*) hendaknya dimasukkan dalam buku bank soal tes hasil belajar.

Kemudian untuk butir-butir soal yang daya pembedanya masih rendah (kriteria sedang bahkan jelek), maka *tester* (guru) perlu menelusuri untuk kemudian diperbaiki sehingga dapat diajukan lagi dalam tes hasil belajar yang akan datang. Warju et al. (2020) mengungkapkan bahwa soal dengan kriteria rendah belum mampu membedakan siswa yang memiliki kemampuan tinggi dengan siswa yang memiliki kemampuan rendah.

Menurut Kemendiknas (2010), apabila suatu butir soal tidak dapat membedakan kedua kemampuan siswa, maka kemungkinan yang terjadi pada soal tersebut adalah kunci jawaban butir soal itu tidak tepat, butir soal itu memiliki 2 atau lebih kunci jawaban yang benar, kompetensi yang diukur tidak jelas, pengecoh tidak berfungsi, materi yang ditanyakan terlalu sulit sehingga banyak siswa yang menebak atau sebagian besar siswa yang memahami materi yang ditanyakan berpikir ada yang salah informasi dalam butir soalnya.

#### 5. Kualitas Pengecoh (*Distractor*)

*Option* atau pilihan jawaban dalam tes objektif dalam bentuk pilihan ganda berkisar tiga sampai lima. Salah satunya adalah pilihan yang merupakan kunci jawaban (jawaban benar) dan sisanya adalah jawaban salah. Jawaban salah inilah yang disebut dengan pengecoh atau distraktor Dewi (2021). Menurut Kurniawan (2015), analisis kualitas pengecoh merupakan kegiatan analisis yang bertujuan untuk mengetahui seberapa efektif suatu pengecoh dalam menjalankan fungsi ukurnya. Berdasarkan analisis terhadap perhitungan banyaknya *testee* yang memilih *option*/alternatif jawaban soal tersebut, maka dapat diperoleh informasi sebagaimana tertera pada Tabel 2 di atas.

Data tersebut menunjukkan bahwa dari 100 distraktor, ada 54 distraktor (54%) yang dapat dipakai atau telah berfungsi dengan baik dan 46 distraktor (46%) yang tidak berfungsi dengan baik. Selain itu, dapat juga diketahui bahwa dari 25 butir soal hanya terdapat 3 butir soal (12%) yang semua distraktornya berfungsi dengan baik dan 22 butir soal (88%) yang distraktornya ada yang berfungsi dan ada yang tidak berfungsi sama sekali.

Maksud dari tidak berfungsinya pengecoh bukan dikarenakan tidak ada yang memilih opsi dalam pelaksanaan tes, akan tetapi para *testee* lebih banyak tertarik pada satu opsi pengecoh dibandingkan dengan kunci jawaban dan opsi jawaban yang lainnya. Akibatnya jawaban para *testee* menumpuk pada satu opsi pengecoh, maka pengecoh yang lain tersebut dikatakan tidak berfungsi atau tidak menjalankan fungsinya dengan baik Zahiroh dan Ritonga (2021). Kemudian didukung oleh Iskandar & Rizal (2017) bahwa pengecoh yang tidak berfungsi akan memberikan dampak terhadap tingkat kesukaran butir soal. Karena jika terdapat satu atau dua pengecoh pada suatu butir soal yang tidak berfungsi maka indeks tingkat kesukaran butir soal akan menurun, karena peluang peserta tes untuk menjawab dengan benar semakin meningkat.

Menurut Sudijono (2009) menjelaskan bahwa distraktor yang sudah dapat menjalankan fungsinya dengan baik dapat dipakai lagi pada tes-tes yang akan datang, sedangkan distraktor yang belum dapat berfungsi dengan baik sebaiknya diperbaiki atau diganti dengan distraktor lain.

## Kesimpulan

Dari analisis kualitas butir soal mata pelajaran Biologi pada Ujian Tengah Semester (UTS) ganjil Biologi kelas XI SMAN 1 Pancung Soal dapat peneliti simpulkan bahwa dari aspek validitasnya yaitu terdapat 1 (4%) butir soal dengan kategori tinggi, 4 (16%) butir soal dengan kategori cukup, 14 (56%) butir soal dengan kategori sangat rendah, 3 (12%) butir soal dengan kategori rendah serta juga ada 3 (12%) butir soal yang tidak terbaca tingkat kevalidannya. Aspek reliabilitasnya yaitu memiliki kriteria sangat rendah dengan harga  $r$  statistik sebesar -2,88. Aspek tingkat kesukarannya yaitu tidak diperoleh

butir soal dengan kriteria mudah, hanya kriteria sedang dengan jumlah 9 (36%) butir soal dan kriteria sukar dengan jumlah 16 (64%) butir soal. Aspek daya pembedanya yaitu terdapat 5 (20%) dengan kriteria sangat baik (diterima), 6 (24%) butir soal dengan kriteria sedang (diterima dan revisi), 8 (31%) butir soal dengan kriteria jelek (direvisi), dan 5 (24%) butir soal dengan kriteria sangat jelek (harus dibuang/diganti). Kemudian aspek kualitas pengecoh (*distractor*) yaitu dari 25 butir soal hanya terdapat 3 (12%) butir soal yang semua distraktornya berfungsi dengan baik dan 22 (88%) butir soal yang distraktornya ada yang berfungsi dan ada yang tidak berfungsi sama sekali. Hasil analisis ini mengindikasikan bahwa secara keseluruhan soal tes tersebut harus direvisi kembali sehingga dapat digunakan lagi dalam tes hasil belajar yang akan datang.

## Referensi

- Agustiana, M., Mayrita, H., dan Muchti, A. 2019. "Analisis Butir Soal Ulangan Akhir Semester Mata Pelajaran Bahasa Indonesia Kelas XI." *Jurnal Ilmiah Bina Edukasi*, 11(01):26–35. [10.33557/jedukasi.v11i01.203](https://doi.org/10.33557/jedukasi.v11i01.203).
- Arif, M. 2015. "Penerapan Aplikasi Anates Bentuk Soal Pilihan Ganda." *Eduatic - Scientific Journal of Informatics Education*, 1(1):1–9. [10.21107/edutic.v1i1.398](https://doi.org/10.21107/edutic.v1i1.398)
- Arifin, Z. 2013. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Arikunto, S. 2010. *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Arikunto2, S. 2013. *Prosedur Penelitian Suatu Pendekatan Praktik Edisi Revisi*. Jakarta: PT. Rineka Cipta.
- Balau, M., Pesik, A., dan Damai, I., W. 2021. "Analisis Kualitas Butir Soal Buatan Guru Mata Pelajaran Matematika Kelas VIII SMP Negeri Satap Matabulu Kabupaten Bolaang Mongondow Timur." *MARISEKOLA: Jurnal Matematika Riset Edukasi Dan Kolaborasi*, 2(1):13–18. [10.53682/marisekola.v2i1.1112](https://doi.org/10.53682/marisekola.v2i1.1112).
- Daryanto. 2010. *Evaluasi Pendidikan*. Jakarta: Rineka Cipta.
- Dewi, S. 2021. "Kualitas Butir Soal Matematika pada Ujian Sekolah SDIT Al Fatih." *Faktor : Jurnal Ilmiah Kependidikan*, 8(2): 213-222.
- Elviana. 2020. "Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates." *Jurnal Mudarrisuna*, 10(2):58–74
- Febriani, Melinda, I., dan Saksono, L. 2016. "Analisis Butir Soal Ujian Akhir Semester (UAS) Bahasa Jerman Kelas X MIA 6 SMA Negeri 1 Maospati Tahun Pelajaran 2015/2016." *Jurnal Laterne*, V(2):1–7.
- Iskandar, A., dan Rizal, M. 2017. "Analisis Kualitas Soal di Perguruan Tinggi Berbasis Aplikasi TAP." *Jurnal Penelitian dan Evaluasi Pendidikan*, 21(2): 12-23.
- Jusrianto, Zahir, A., dan Megawati. 2018. "Analisis Kualitas Tes Ujian Akhir Semester Mata Kuliah Pengetahuan Komputer." *Jurnal Studi Guru Dan Pembelajaran*, 1(1):1–9. [10.30605/jsgp.1.1.2018.7](https://doi.org/10.30605/jsgp.1.1.2018.7).
- Kemendiknas. 2010. "Materi Bimbingan Teknis KTSP Dan Soal Terstandar 2010: Panduan Analisis Butir Soal." <http://smp3bonang.files.wordpress.com/2010/08/panduanpenulisan-butir-soal.pdf>. Online. Diakses tanggal 30 November 2021.
- Kurniawan, T. 2015. "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar (Analysis of Odd Semester Final Test Items in Elementary School of Social Studies Subjects)." *Journal of Elementary Education*, 4(1):1–6.
- Magdalena, I et al. 2021. "Analisis Validitas, Reliabilitas, Tingkat Kesulitan Dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan." *Bintang*, 3(2):198–214..



- Martunis M., Khaldun, I., dan Zulfadli. 2015. "Analisis Kualitas Butir Soal Ujian Semester Genap Mata Pelajaran Kimia Kelas X MAN Model Banda Aceh Tahun Pelajaran 2014 / 2015 Menggunakan Program Proanaltes." *Jurnal Ilmiah Mahasiswa Pendidikan Kimia*, 1(4):70–78.
- Monica, S., dan Sudarman, Y. 2019. "Analisis Butir Soal Ujian Tengah Semester Ganjil Seni Budaya Kelas VII Di SMPN 29 Sijunjung." *Journal of Chemical Information and Modeling* 53(9):1689–99
- Nurfauziah, W., dan Windyariani, S. 2020. "Analisis Butir Soal Keterampilan Proses Sains Pada Instrumen Uji Coba Materi Ekosistem." *Journal Of Biology Education*, 3(2):135. [10.21043/job.e.v3i2.7468](https://doi.org/10.21043/job.e.v3i2.7468).
- Purwati, et al. 2021. "Analisis Karakteristik Butir Soal Ujian Nasional Matematika SMP/MTS Berdasarkan Pendekatan Teori Tes Klasik." *Jurnal Sains Dan Edukasi Sains*, 4(2):46–51.
- Setyaningrum, P., M., P., Ramli, M., dan Rinanto, Y. 2018. "Analisis Kualitas Butir Soal Instrumen Assessment Diagnostic Untuk Mendeteksi Miskonsepsi Siswa SMA Pada Materi Virus." *Jurnal Bioedukatika*, 6(2):91. [10.26555/bioedukatika.v6i2.10925](https://doi.org/10.26555/bioedukatika.v6i2.10925).
- Solichin, M. 2017. "Analisis Daya Beda Soal Taraf Kesukaran, Butir Tes, Validitas Butir Tes, Interpretasi Hasil Tes Validitas Ramalan Dalam Evaluasi Pendidikan." *Jurnal Manajemen Dan Pendidikan Islam*, 2(2):192–213.
- Sudijono, A. 2009. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Sukardi. 2011. *Evaluasi Pendidikan: Prinsip Dan Operasionalnya*. Jakarta: PT Bumi Aksara.
- Sukmela, Y., P., dan Nofrion. 2018. "Analisis Butir Soal Paket A Ujian Akhir Semester Genap Geografi Tahun Pelajaran 2016/2017 SMA Negeri Se-Kota Pekanbaru." *Jurnal Buana*, 2(5):467–79.
- Susdelina, Perdana, S., A., dan Febrian. 2018. "Analisis Kualitas Instrumen Pengukuran Pemahaman Konsep Persamaan Kuadrat Melalui Teori Tes Klasik dan Rasch Model." *Jurnal Kiprah*, VI(1):41–48.
- Suzana, A. 2018. "Analisis Tingkat Kesukaran Dan Daya Beda Butir-Butir Soal Penilaian Akhir Tahun Matematika Kelas X Di SMA Negeri 1 Purbalingga." *MathGram Matematika*, 2(2):1–8.
- Thoha. 2001. *Teknik Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Ulum, M. 2017. "Analisis Butir Soal Ulangan Harian Pada Mata Pelajaran Ekonomi KD 3.1 Pendapatan Nasional Kelas XI IPS 1 di SMA Negeri 1 Gresik." *Jurnal Pendidikan Ekonomi*, 5(3):1–5.
- Wahyuni, S., dan Ibrahim, A., S. 2012. *Asesmen Pembelajaran Bahasa*. Bandung: PT. Refika Aditama.
- Warju, et al. 2020. "Analisis Kualitas Butir Soal Tipe Hots Pada Kompetensi Sistem Rem Di Sekolah Menengah Kejuruan." *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1):95. [10.23887/jptk-undiksha.v17i1.22914](https://doi.org/10.23887/jptk-undiksha.v17i1.22914).
- Widoyoko. 2014. *Penilaian Hasil Belajar Di Sekolah*. Yogyakarta: Pustaka Pelajar.
- Zahiroh, U., dan Ritonga, P., S. 2021. "Analisis Kualitas Butir Soal Pilihan Ganda Mata Pelajaran Kimia Pada Ujian Akhir Semester (UAS) Kelas XI MAN 2 Kepulauan Meranti." *Jedchem (Journal Education and Chemistry)*, 3(1):11–20. [10.36378/jedchem.v3i1.780](https://doi.org/10.36378/jedchem.v3i1.780)