

ANALISIS KUALITAS BUTIR SOAL HASIL BELAJAR PESERTA DIDIK BERDASARKAN TINGKAT KESUKARAN, DAYA BEDA DAN EFEKTIVITAS PENGECOH DI SMAN 1 PANDAWAI

Sofiani Putri Radja¹, Vidriana Oktoviana Bano^{2*}, Anita Tamu Ina³

^{1,2,3} Pendidikan Biologi, FKIP, Universitas Kristen Wira Wacana Sumba

Email penulis: sofianiputriradja@gmail.com, vidri.bano@unkriswina.ac.id
anitamuina@unkriswina.ac.id

Abstrak

Tujuan penelitian ini menganalisis kualitas butir soal hasil belajar peserta didik berdasarkan tingkat kesukaran, daya beda dan efektivitas pengecoh pada soal Ujian Akhir Semester Ganjil Mata Pelajaran IPA tahun ajaran 2021/2022 dengan soal yang berjumlah 30 soal. Sampel penelitian terdiri dari 72 lembar jawaban siswa kelas X IPA 1 dan X IPA 2 menggunakan metode *Purposive sampling*. Penelitian ini menggunakan aplikasi SPSS versi 22 dan Microsoft excel dengan pendekatan kuantitatif. Hasil penelitian menunjukkan bahwa di kelas X IPA 1 6,6% butir soal tergolong sukar, 50% butir soal tergolong sedang dan 43,4% tergolong mudah. Hasil penelitian tingkat kesukaran di kelas X IPA 2, diperoleh 43,4% butir soal tergolong sukar, 23,3% butir soal tergolong sedang dan 33,3% butir soal tergolong mudah. Untuk analisis daya beda di kelas X IPA 1, diperoleh 3,3% butir soal sudah memiliki daya beda baik, 6,7% memiliki daya beda cukup baik, 13,3% butir soal memiliki daya beda kurang baik serta perlu direvisi dan 76,7% butir soal berdaya beda tidak baik perlu direvisi atau dibuang. Tingkat daya beda di kelas X IPA 2, diperoleh 23,3% butir soal sudah memiliki daya beda baik, 23,3% memiliki daya beda cukup baik, 10% butir soal memiliki daya beda kurang baik dan 43,3% butir soal berdaya beda tidak baik. Analisis efektivitas pengecoh butir soal di kelas X IPA 1 diketahui 80% dari total option pengecoh pada butir soal tersebut termasuk baik dan sisanya tidak baik dan perlu direvisi, sedangkan untuk kelas X IPA 2 diketahui 81% dari total pengecoh termasuk kategori baik, dan sisanya perlu direvisi agar bisa digunakan.

Kata kunci: Analisis butir soal, tingkat kesukaran, daya beda, efektivitas pengecoh

Abstract

This study aims to measure the comparison of student learning outcomes based on the level of difficulty, differentiability and effectiveness of deceivers in the Odd Semester Final Examination questions for Science Subjects for the 2021/2022 school year with 30 questions. The research sample consisted of 72 answer sheets for class X science 1 and X science 2 students using the Purposive sampling method. This study used SPSS version 22 application and Microsoft Excel with a quantitative approach. The results showed that in class X science 1, 6.6% of the questions were difficult, 50% of the questions were classified as moderate and 43.4% were classified as easy. The results of the research on the level of difficulty in class X science 2, obtained 43.4% of question items classified as difficult, 23.3% of question items classified as medium and 33.3% of question items classified as easy. For the analysis of different power in class X IPA 1, 3.3% of the question items already have good differential power, 6.7% have a fairly good difference power, 13.3% of the question items have poor different power and need to be revised and 76.7% of the different power question items do not either need to be revised or discarded. The level of different power in the X IPA 2, obtained 23.3% of the question items already have good different power, 23.3% have a fairly good difference power, 10% of the question items have a bad difference power and 43.3% of the questions have a bad difference power. Analysis of the effectiveness of the question item deceivers in class X IPA 1 is known that 80% of the total deceptive options in the question item are good and the rest are not good and need to be revised,

while for class X IPA 2 it is known that 81% of the total deceptors belong to the good category, and the rest need to be revised so that they can be used.

Keywords: Item analysis, level of difficulty, discriminating power, effectiveness of distractors

PENDAHULUAN

Tes Hasil Belajar (THB) merupakan salah satu cara untuk melihat seberapa baik siswa dalam belajar (Purwanto, 2008). Ini menggunakan pertanyaan untuk mengetahui apa yang diketahui siswa. Hasil tes menunjukkan seberapa bagus tes tersebut (Widoyoko, 2018). Soal-soal dalam tes harus dapat diandalkan dan valid. Reliable berarti kita dapat mempercayai tes tersebut, dan valid berarti tes tersebut mengukur apa yang seharusnya. Hasil tes hanya baik jika sesuai dengan apa yang sebenarnya terjadi (Sugiyono, 2016).

Ada beberapa hal berbeda yang dilihat guru ketika mereka menguji siswa. Mereka ingin mengetahui seberapa sulit atau mudahnya soal-soal tersebut, seberapa baik mereka memisahkan siswa yang mengetahui materi dari yang tidak, dan seberapa baik jawaban yang salah dalam mengelabui siswa (Asep, 2013).

1. Tingkat Kesukaran

Tingkat kesulitan adalah angka yang menunjukkan sulit atau mudahnya suatu soal. Jika semua orang bisa melakukannya dengan benar, itu terlalu mudah, dan jika tidak ada yang bisa melakukannya dengan benar, itu terlalu sulit (Arikunto, 2013b). Tingkat kesukaran berarti seberapa sulit atau mudahnya suatu butir soal bagi sekelompok peserta didik atau dengan kata lain, tingkat kesukaran adalah tingkat mudah atau tidaknya suatu soal yang diberikan pada sekelompok peserta didik (Azis, 2016). Guru menginginkan soal-soal yang tingkat kesulitannya tepat, tidak terlalu sulit dan tidak terlalu mudah (Sudijono, 2009).

Asumsi yang digunakan untuk memperoleh kualitas soal yang baik, di samping memenuhi validitas dan reliabilitas adalah adanya keseimbangan dari tingkat kesukaran dari soal itu sendiri. Keseimbangan dan kriteria yang dimaksud adalah adanya soal yang termasuk mudah (30%), sedang (40%) dan sukar (30%) (Sudjana, 2014). Tingkat kesukaran soal adalah pengukuran seberapa besar derajat kesukaran suatu soal. Menganalisis tingkat kesukaran soal artinya mengkaji soal-soal mana yang termasuk mudah, sedang, dan sukar (Arifin, 2016).

Analisis tingkat kesukaran soal berarti mengkaji soal-soal mana yang termasuk mudah, sedang, dan sukar. Angka tingkat kesukaran soal dinyatakan berkisar 0,00-1,00. Artinya, angka indeks paling rendah adalah 0,00 dan paling tinggi adalah 1,00. Angka indeks kesukaran sebesar 0,00 ($P=0,00$) merupakan kategori item yang terlalu sukar, sebab seluruh peserta tes tidak dapat menjawab item dengan benar. Sebaliknya, apabila angka indeks kesukaran item itu adalah 1,00 ($P=1,00$) merupakan kategori item yang terlalu mudah, sebab seluruh peserta tes dapat menjawab dengan benar (Kadir, 2015).

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. Rumus untuk mencari tingkat kesukaran menurut (Widoyoko, 2018) :

$$p = \frac{\sum b}{N}$$

Keterangan:

P : Tingkat Kesukaran

 $\sum b$: Jumlah peserta didik yang menjawab benar

N : Jumlah seluruh peserta tes

Tingkat kesukaran secara keseluruhan dalam suatu naskah soal dihitung dengan menjumlahkan tingkat kesulitan semua butir soal, kemudian dibagi dengan jumlah butir soal.

Rumusnya :

$$P_s = \frac{\sum Pb}{N}$$

Keterangan:

Ps : Tingkat kesukaran naskah soal

 $\sum Pb$: Jumlah tingkat kesukaran butir soal

N : Jumlah butir soal

Berikut pembagian kategori/kriteria tingkat kesukaran dan hubungannya dengan kualitas butir soal, yaitu:

Indeks Kesukaran	Kategori Butir Soal
0,00 – 0,30	Sukar, butir soal kurang baik, perlu direvisi
0,31 – 0,70	Sedang, butir soal cukup baik, digunakan
0,71 – 1,00	Mudah, butir soal baik, direvisi

2. Tingkat Daya Beda

Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah atau daya pembeda ialah kemampuan suatu instrumen soal untuk mengklasifikasikan siswa berdasarkan tingkat kemampuannya. Makin tinggi nilai sebuah soal, makin baik soal tersebut membedakan siswa yang berkemampuan tinggi dari siswa yang berkemampuan rendah (Arikunto, 2013a).

Untuk mencari daya beda soal menggunakan persamaan (Widoyoko, 2018)

$$D = \frac{Ba - Bb}{\frac{1}{2}N}$$

Keterangan:

D : Daya beda

Ba : Jumlah jawaban benar kelompok atas

Bb : Jumlah jawaban benar kelompok bawah

N : Jumlah peserta tes kelompok bawah dan kelompok atas

Berikut pembagian kategori/kriteria tingkat daya beda dan hubungannya dengan kualitas butir soal yaitu :

Index Daya Beda	Kualitas Butir Soal
-1,00 – 0,00	Tidak baik, buang atau diganti
0,20 – 0,29	Kurang baik, perlu pembahasan dan revisi
0,30 - 0,39	Cukup baik, dapat digunakan dengan revisi
0,40 – 1,00	Baik, digunakan tanpa revisi

3. Efektivitas Pengecoh

Efektivitas pengecoh adalah kemampuan pengecoh untuk mengecoh peserta didik yang kurang cakap memilih alternatif jawaban. *Distractor* berfungsi mengecoh peserta tes bila peserta tes tidak memilih jawaban yang benar (Putri, 2020). Analisis efektivitas pengecoh (*Distractor*) adalah jawaban pengecoh yang memiliki perbedaan tipis dengan jawaban benar, sehingga sering menjebak siswa dalam menjawab soal tes yang diujikan (Setiawaty, 2017). Pengecoh akan efektif dengan baik apabila dipilih sekurang-kurang 5% dari jumlah peserta tes (Uno, H.B, 2018).

Opsi atau pilihan alternatif dalam setiap butir tes pilihan ganda bisa terdiri dari tiga sampai lima buah dan kemungkinan salah satunya adalah pilihan atau jawaban yang benar (kunci jawaban) sedangkan sisanya merupakan jawaban yang salah. Jawaban-jawaban yang salah itulah yang disebut pengecoh. Jadi, pengecoh adalah pilihan alternatif yang bisa mengecoh jawaban yang sebenarnya.

Untuk menjelaskan cara perhitungan tingkat efektivitas pengecoh tiap butir soal, dapat dilihat pada ilustrasi berikut yang menyajikan hasil uji coba instrumen hasil belajar yang terdiri dari 7 butir tes objektif yang terdiri dari 4 pilihan dan diujicobakan terhadap 40 orang peserta tes. Mengingat banyak peserta tes 40 orang, pengecoh dikatakan baik jika dipilih oleh minimal 5% dari 40 orang, atau minimal dipilih oleh 2 orang peserta tes.

Rumus untuk menganalisis tingkat efektivitas pengecoh dapat menggunakan rumus (Wardoyo W & Suprpto E., 2014):

Keterangan :

$$IPc = \frac{nPC}{N - nB/Alt - 1} \times 100\%$$

IPc = Indeks Pengecoh

nPC = Jumlah Peserta Didik Yang Memilih Pengecoh

N = Jumlah Peserta Yang Ikut Tes

nB = Jumlah Peserta Yang Menjawab Benar Pada Setiap Soal

1 = Bilangan Tetap

Jika semua peserta didik menjawab benar pada butir soal tertentu, maka IP=0 yang berarti soal tersebut jelek dan pengecohnya tidak berfungsi. Menurut Depdikbud (1997) untuk menilai pengecoh dari masing-masing butir soal dapat dikategorikan :

Kategori Pengecoh	Nilai Buttir Soal
≥0,025	Baik
<0,025	Revisi
0,000	Tolak/ Tidak baik

Berdasarkan hasil wawancara tanggal 23/02/2022 dengan guru mata pelajaran biologi kelas X IPA SMA N 1 Pandawai (YTM), mengatakan bahwa pada hasil ujian akhir semester ganjil tahun ajaran 2021/2022 dari 72 peserta didik dan KKM 68, terdapat 42 (58%) peserta didik tuntas dan 30 (42%) siswa yang tidak tuntas. Salah satu penyebab banyaknya peserta didik yang tidak tuntas adalah` karena peserta didik kurang memahami materi pembelajaran yang diajarkankan guru.

Sistem Belajar Dari Rumah (BDR) dan *shift-shiftan* adalah sistem pembelajaran yang digunakan saat itu karena dalam masa pandemi covid-19 mengakibatkan waktu pembelajaran di kelas sangat singkat, guru memaparkan materi pembelajaran hanya point-point besarnya saja sehingga peserta didik kurang memahami materi pembelajaran yang diajarkan guru. Guru juga belum optimal dalam membuat kisi-kisi soal yang didasarkan pada Taksonomi Bloom (TB), sehingga banyak siswa merasa soal yang diberikan guru terlalu sukar untuk dikerjakan. Selain itu guru belum optimal dalam menganalisis butir soal UAS untuk mengetahui tingkat kesukaran soal, daya beda dan efektivitas pengecoh soal tersebut. Inilah yang mengakibatkan banyak siswa tidak tuntas pada UAS ganjil tahun 2021/2022.

Berdasarkan permasalahan yang diuraikan di atas, maka Peneliti memandang penting melakukan analisis butir soal untuk mengetahui kelayakan soal yang diberikan guru supaya dapat memberikan hasil yang maksimal dalam mengukur tingkat pemahaman peserta didik. Penelitian terdahulu yang relevan dengan penelitian ini adalah penelitian yang dilakukan oleh (Vidriana & Marambaawang, 2022) yang berjudul "Analisis Kriteria Butir Soal Ujian Sekolah Mata Pelajaran IPA di SMP Negeri 1 Waingapu"., dengan Hasil analisis pada tingkat kesukarannya menunjukkan 15% tergolong sukar, 60% tergolong sedang dan 25% tergolong mudah. Untuk tingkat daya beda, 70% memiliki daya beda yang baik dan 30% tidak memiliki daya beda. Dan pada analisis efektivitas pengecoh, sekitar 96% termasuk baik sedangkan sisanya termasuk dalam option tidak baik, sehingga perlu direvisi. Perbedaan antara penelitian yang terdahulu dengan penelitian ini adalah pada penelitian terdahulu tidak dilakukan perbandingan hasil analisis, sedangkan pada penelitian ini dilakukan perbandingan hasil analisis antara kelas X IPA 1 dan kelas X IPA 2. Tujuan dari penelitian ini adalah selain untuk mengetahui kualitas butirsoal, membandingkan tingkat kemampuan siswa setiap kelas serta hasil penelitian ini juga dapat dijadikan masukan dan motivasi dalam meningkatkan kualitas soal dan kualitas pendidikan itu sendiri.

METODE PENELITIAN

Dalam penelitian ini menggunakan jenis penelitian kualitatif dengan pendekatan kuantitatif. Penelitian ini dilakukan di SMAN 1 Pandawai yang beralamat di Jl. Cendana Kawangu, Kawangu, Kecamatan Pandawai, Kab. Sumba Timur, Provinsi Nusa Tenggara Timur. Penelitian ini dilakukan bulan April-Mei tahun 2022. Variabel dalam penelitian ini adalah tingkat kesukaran, analisis daya beda dan efektivitas pengecoh butir soal. Jumlah populasi dalam penelitian ini 828 orang. Sampel dalam penelitian adalah hasil belajar peserta didik dalam lembar jawaban Penilaian Akhir Semester (PAS) kelas X IPA 1 dan X IPA 2 yang masing-masing berjumlah 36 orang (total 72 orang). Data yang telah terkumpul kemudian dilakukan analisis kriteria butir soal dengan menggunakan aplikasi SPSS versi 22 dan *Microsoft Excel*. Aplikasi SPSS digunakan untuk menganalisis Tingkat Validitas dan Reabilitas, dan *Microsoft Excel* digunakan untuk menghitung tingkat kesukaran, daya beda, dan efektivitas pengecoh.

HASIL & PEMBAHASAN

Pada bagian ini akan diuraikan hasil dan pembahasan analisis kriteria butir soal untuk tingkat kesukaran soal (1), daya pembeda soal (2), efektivitas pengecoh/distraktor (3), tingkat validitas soal (3), dan efektivitas pengecoh (4).

A. HASIL

1. Uji Validitas

Tabel 1. Hasil uji validitas soal di kelas X IPA 1

Butir Soal	rHitung	rTabel	Keterangan
1	0,347	0,231	Valid
2	0,566	0,231	Valid
3	-0,012	0,231	Tidak Valid
4	0,090	0,231	Tidak Valid
5	0,233	0,231	Valid
6	0,535	0,231	Valid
7	0,514	0,231	Valid
8	0,307	0,231	Valid
9	0,169	0,231	Tidak Valid
10	0,351	0,231	Valid
11	0,352	0,231	Valid
12	0,389	0,231	Valid
13	0,290	0,231	Valid
14	0,542	0,231	Valid
15	0,555	0,231	Valid
16	0,406	0,231	Valid
17	0,225	0,231	Tidak Valid
18	0,95	0,231	Tidak Valid
19	0,382	0,231	Valid
20	0,016	0,231	Tidak Valid
21	0,233	0,231	Valid
22	-0,044	0,231	Tidak Valid
23	0,158	0,231	Tidak Valid
24	0,502	0,231	Valid
25	0,128	0,231	Tidak Valid
26	0,304	0,231	Valid
27	0,365	0,231	Valid
28	0,108	0,231	Tidak Valid
29	0,132	0,231	Tidak Valid
30	0,250	0,231	Tidak Valid

Tabel 2. Hasil uji validitas soal di kelas X IPA 2

Butir Soal	rHitung	rTabel	Ket
1	-0,160	0,231	Tidak Valid
2	0,366	0,231	Valid
3	0,061	0,231	Tidak Valid
4	0,029	0,231	Tidak Valid
5	0,371	0,231	Valid
6	0,683	0,231	Valid
7	0,032	0,231	Tidak Valid
8	0,083	0,231	Tidak Valid
9	0,509	0,231	Valid
10	0,335	0,231	Valid
11	0,611	0,231	Valid
12	0,206	0,231	Tidak Valid
13	0,210	0,231	Tidak Valid
14	0,495	0,231	Valid
15	0,289	0,231	Valid
16	0,360	0,231	Valid
17	-0,029	0,231	Tidak Valid
18	0,446	0,231	Valid
19	0,175	0,231	Tidak Valid
20	0,293	0,231	Valid
21	0,317	0,231	Valid
22	0,529	0,231	Valid
23	0,094	0,231	Tidak Valid
24	0,359	0,231	Valid
25	0,327	0,231	Valid
26	-0,201	0,231	Tidak Valid
27	0,457	0,231	Valid
28	0,418	0,231	Valid
29	-0,094	0,231	Tidak Valid
30	0,502	0,231	Valid

Tabel 3. Persentase uji validitas soal di kelas X IPA 1

No	Ketentuan Validitas	Butir Soal	Jml	Persentase
1	rhitung > r tabel Valid	1,2,5,6,7,8,10,11,12,13,14,15,16,19,21,24,26,27	18	60%
2	rhitung < r tabel Tidak Valid	3,4,9,17,18,20,22,23,25,28,29,30	12	40%
Jumlah			30	100%

Tabel 4. Persentase uji validitas soal di kelas X IPA 2

No	Ketentuan Validitas	Butir Soal	Jml	Persentase
1	rhitung > r tabel Valid	2,5,6,9,10,11,14,15,16,18,20,21,22,24,25,27,28,30	18	60%
2	rhitung < r tabel Tidak Valid	1,3,4,7,8,12,13,17,19,23,26,29	12	40%
Jumlah			30	100%

Berdasarkan tabel 3 di atas dapat dilihat bahwa dari 30 soal pilihan ganda yang diberikan kepada 72 peserta tes, terdapat 18 (60%) soal termasuk kategori **valid** karena $r_{hitung} > r_{tabel}$ (0,231) dan 12 (40%) soal termasuk kategori tidak valid, karena $r_{hitung} < r_{tabel}$ (0,231). Tabel 4 di atas dapat dilihat bahwa dari 30 soal pilihan ganda yang diberikan kepada 72 peserta tes, terdapat 18 (60%) soal termasuk kategori **valid** karena $r_{hitung} > r_{tabel}$ (0,231) dan 12 (40%) soal termasuk kategori tidak valid, karena $r_{hitung} < r_{tabel}$ (0,231).

2. Uji Reabilitas

Tabel 5. Hasil uji reliabilitas soal untuk kelas X IPA 1

Cronbach's Alpha	N of Items
,603	30

Tabel 6. Hasil uji reliabilitas soal untuk kelas X IPA 2

Cronbach's Alpha	N of Items
,564	30

3. TINGKAT KESUKARAN

Tabel 7. Persentase tingkat kesukaran soal kelas X IPA 1

No	Indeks Kesukaran	Butir Soal	Jml	Persentase
1	0,00-0,30 Sukar	11,24	2	6,6%
2	0,31-0,70 Sedang	2,6,9,10,13, 14,17,18,19, 21,22,25,26, 29,30	15	50%
3	0,71-1,00 Mudah	1,3,4,5,7,8, 12,15,16,20, 23,27,28	13	43,4%
Jumlah			30	100%

Tabel 8. Persentase tingkat kesukaran soal kelas X IPA 2

No	Indeks Kesukaran	Butir Soal	Jml	Presentase
1	Indeks Kesukaran	2,6,9,11, 13,14,17, 18,21,24, 25,29,30	13	43,4%
2	0,00-0,30 Sukar	5,8,10,15,16, 19,20	7	23,3%
3	0,31-0,70 Sedang	1,3,4,7,12, 22,23,26, 27,28	10	33,3%
Jumlah			30	100%

Untuk soal dengan tingkat kesukaran yang baik, perlu diperhatikan keseimbangan dan kriteria soal. Keseimbangan dan kriteria soal yang dimaksud adalah adanya soal yang termasuk mudah (30%), sedang (40%) dan sukar (30%) (Sudjana, 2013). Pada tabel 7 dapat dilihat bahwa dari 30 butir soal dengan 72 peserta tes diperoleh 6,6% (2) soal tergolong **sukar**, 50% (15) soal tergolong **sedang** dan 43,4% (13) tergolong **mudah**. Pada tabel 8 dapat dilihat bahwa dari 30 butir soal dengan 72 peserta tes diperoleh 643,4% (13) soal tergolong **sukar**, 23,3% (7) soal tergolong **sedang** dan 33,3% (10) tergolong **mudah**.

4. TINGKAT DAYA BEDA

Tabel 9 Persentase tingkat daya beda butir soal kelas X IPA 1

No.	Indeks Daya Beda	Butir Soal	Jml	Persentase
1	-1-0,00 Tidak Baik	1,3,4,5,7,8, 9,10,12,15, 16,17,18,19, 20,22,23,25, 26,27,28,29,30	23	76,7%
2	0,20-0,29	6,11,13,21	4	13,3%

Tabel 10 Persentase tingkat daya beda butir soal kelas X IPA 2

No.	Indeks Daya Beda	Butir Soal	Jml	Persentase
1	-1-0,00 Tidak Baik	1,3,4,7,8,13, 15,17,19,20, 23,26,29	13	43,4%
2	0,20-0,29 Kurang Baik	16,21,24	3	10%

	Kurang Baik			
3	0,30-0,39 Cukup Baik	2,24	2	6,7%
4	0,40-1,00 Baik	14	1	3,3%
Jumlah			30	100%

3	0,30-0,39 Cukup Baik	2,12,14, 25,27,28, 30	7	23,3%
4	0,40-1,00 Baik	5,6,9,10,11, 18,22	7	23,3%
Jumlah			30	100%

Berdasarkan tabel 9 di atas dapat dilihat bahwa terdapat 3,3% (1) butir soal tergolong **baik**, dapat digunakan tanpa revisi, 6,7% (2) butir soal termasuk kategori **cukup baik**, 13,3% (4) butir soal termasuk kategori **kurang baik**, dan 76,7% (23) butir soal termasuk kategori **tidak baik**. Sedangkan tabel 10 di atas dapat dilihat bahwa terdapat 23,3% (7) butir soal termasuk kategori **baik**, dapat digunakan tanpa revisi, 23,3% (7) butir soal termasuk kategori **cukup baik**, 10% (3) butir soal termasuk kategori **kurang baik**, dan 43,4% (13) butir soal termasuk kategori **tidak baik**.

5. TINGKAT EFEKTIVITAS PENGECOH

Tabel 11 Hasil Analisis tingkat efektivitas pengecoh butir soal kelas X IPA 1

Butir Soal	Kunci Jawaban	Jumlah Siswa	Jumlah Yang Memilih Benar	Jumlah Yang Memilih Salah	Efektivitas Pengecoh		
					<2% Tidak Baik	<5% Revisi	≥5% Baik
1	D	36	29	7	-	B,C	A, E
2	C	36	12	24	-	-	A,B, D,E
3	C	36	29	7	-	A	B,D,E
4	C	36	32	4	B	A	D,E
5	A	36	32	4	D, E	-	B,C
6	A	36	12	24	-	-	B,C, D,E
7	B	36	28	8	-	C	A, D,E
8	B	36	29	7	A	-	C,D,E
9	A	36	12	24	-	-	B,C, D,E
10	A	36	23	13	B	-	C,D,E
11	A	36	10	26	-	-	B,C, D,E
12	C	36	31	5	A, D	-	A,E
13	B	36	11	25	-	-	A,C,D ,E
14	C	36	15	21	-	-	A,B,D ,E
15	E	36	29	7	A, D	-	B,C
16	A	36	28	8	D	-	B,C,E
17	D	36	11	25	-	-	A,B, C,E

Tabel 12. Hasil analisis tingkat efektivitas pengecoh butir soal kelas X IPA 2

Butir Soal	Kunci Jawaban	Jumlah Siswa	Jumlah Yang Memilih Benar	Jumlah Yang Memilih Salah	Efektivitas Pengecoh		
					<2% Tidak Baik	<5% Revisi	≥5% Baik
1	D	36	29	7	-	C	A,B,E
2	C	36	12	24	-	-	A,B, D,E
3	C	36	29	7	B	E	A,D
4	C	36	32	4	E	-	A,D
5	A	36	32	4	-	-	B,C
6	A	36	12	24	-	-	B,C, D,E
7	B	36	28	8	A	-	C D, E
8	B	36	29	7	-	-	C,D,E
9	A	36	12	24	-	-	B,C,D,E
10	A	36	23	13	-	D	C,E
11	A	36	10	26	-	-	B,C,D,E
12	C	36	31	5	A, B	D	E
13	B	36	11	25	-	-	A,C,D ,E
14	C	36	15	21	-	-	A,B,D ,E
15	E	36	29	7	A	-	B,C
16	A	36	28	8	-	-	B,C,E
17	D	36	11	25	-	-	A,B,C ,E
18	B	36	11	25	-	-	A,C,D ,E

18	B	36	11	25	-	-	A,C, D,E
19	A	36	23	13	-	D	B,C,E
20	B	36	26	10	-	-	A,C, D,E
21	B	36	12	24	-	-	A,C, D,E
22	A	36	20	16	-	D	B,C,E
23	A	36	28	8	-	D	B,C,E
24	E	36	12	24	-	-	A,B, C,D
25	D	36	11	25	-	C	A,B,E
26	A	36	19	17	-	-	B,C, D,E
27	D	36	30	6	-	A, B	C, E
28	A	36	31	5	E	B,C	D,
29	A	36	13	23	-	-	B,C, D,E
30	B	36	14	22	-	-	A,C, D,E
TOTAL					11	13	96

19	A	36	23	13	B	-	C,D,E
20	B	36	26	10	-	C	A,D,E
21	B	36	12	24	D	-	A,C,E
22	A	36	20	16	-	B	C,D,E
23	A	36	28	8	B, D	E	C
24	E	36	12	24	-	-	A,B,C ,D
25	D	36	11	25	B	-	A,C,E
26	A	36	19	17	B, D	-	C,E
27	D	36	30	6	-	C	A,B,E
28	A	36	31	5	B	D	C
29	A	36	13	23	-	-	B,C,D ,E
30	B	36	14	22	-	-	A,C,D ,E
TOTAL					14	9	97

Pada tabel 11 dapat dilihat bahwa terdapat 96 opsi jawaban pengecoh tergolong **baik**, karena memiliki indeks pengecoh $\geq 5\%$. 13 opsi jawaban pengecoh tergolong **direvisi**, karena memiliki indeks pengecoh $< 5\%$, sedangkan 11 opsi jawaban pengecoh lainnya tergolong **tidak baik**, karena memiliki indeks pengecoh $< 2\%$. Sedangkan Pada tabel 12 dapat dilihat bahwa terdapat 97 opsi jawaban pengecoh tergolong **baik**, karena memiliki indeks pengecoh $\geq 5\%$. 19 opsi jawaban pengecoh tergolong **direvisi**, karena memiliki indeks pengecoh $< 5\%$, sedangkan 14 opsi jawaban pengecoh lainnya tergolong **tidak baik**, karena memiliki indeks pengecoh $< 2\%$.

B. PEMBAHASAN

1. Uji Validitas dan Reliabilitas

Berdasarkan hasil analisis uji validitas pada 30 butir soal yang disajikan ke kelas X IPA 1 dan X IPA 2 (table 1, 2, 3 dan 4), diketahui bahwa kedua kelas memperoleh hasil yang sama, yaitu 60% (18) soal termasuk kategori valid karena r hitung $>$ r tabel (0,231) dan 40% (12) soal dinyatakan tidak valid karena r hitung $<$ r tabel (0,231). Suatu Tes Hasil Belajar (THB) dinyatakan valid r hitung $>$ r tabel, sebaliknya suatu THB dinyatakan tidak valid dan bila r hitung $\leq r$ tabel (Purwanto, 2011). Hampir sebagian soal UAS tidak valid, ini dapat disebabkan oleh tingkat kesukaran soal yang tidak seimbang, dan tidak memiliki daya beda yang baik, serta peserta didik yang kurang memahami materi pembelajaran. Ini didukung oleh teori yang dikemukakan oleh (Sukardi, 2013), yang menyatakan bahwa ada beberapa faktor yang mempengaruhi validitas hasil tes, yaitu:

- 1) Faktor yang berasal dari dalam tes (internal), di antaranya :
 - a. Makna tes yang disusun tidak jelas.
 - b. Kata-kata yang digunakan dalam soal terlalu sulit dipahami.
 - c. Item-item tes dikonstruksi (disusun) dengan jelek.

- d. Tingkat kesulitan soal, terlalu mudah/ terlalu sulit.
 - e. Waktu yang dialokasi tidak tepat/ waktu guru dalam memberi materi pembelajaran terlalu singkat.
 - f. Jumlah item tes terlalu sedikit, hanya memuat beberapa materi pembelajaran.
 - g. Jawaban masing-masing ide evaluasi bisa diprediksi peserta didik.
- 2) Faktor yang berasal dari administrasi dan skor, contohnya:
- a. Waktu mengerjakan soal tidak cukup.
 - b. Adanya kecurangan dalam mengerjakan soal, sehingga sulit membedakan peserta didik yang belajar dengan peserta didik yang curang.
 - c. Pemberian petunjuk oleh pengawas ujian tidak kepada semua peserta tes.
 - d. Teknik pemberian skor pada tes lisan yang tidak konsisten.
 - e. Adanya orang lain yang membantu menjawab item tes yang diberikan.
 - f. Faktor-faktor yang berasal dari jawaban peserta didik, contohnya :
 - g. Dalam menjawab item tes (soal), peserta didik tegang atau tertekan.
 - h. Peserta didik menjadi tidak berkonsentrasi karena ruangan tiba-tiba gaduh.

Untuk hasil penelitian uji reliabilitas di kelas X IPA 1 (table 5 dan 6), terlihat bahwa soal tes dinyatakan reliabel, karena diperoleh nilai Cronbach's Alpha hitungnya $0,603 > r_{tabel}$ ($0,231$) atau lebih mendekati angka 1, demikian pula untuk hasil penelitian uji reliabilitas di kelas X IPA 2 (table 4) diperoleh nilai Cronbach's Alpha hitungnya $0,564 > r_{tabel}$ ($0,231$) atau lebih mendekati angka 1, sehingga dinyatakan reliabel. Ketentuan suatu instrumen dikatakan reliabel jika $r_{hitung} > r_{table}$ ($r_{hitung} > 0,231$) maka skor hasil pengukuran THB berkorelasi signifikan atau reliabel, tetep jika $r_{hitung} < r_{table}$ ($r_{hitung} < 0,231$), maka THB dinyatakan tidak berkorelasi atau tidak reliabel. Signifikansi korelasi menunjukkan adanya konsistensi sehingga THB telah dapat dikatakan reliabel (Purwanto, 2011). Ini didukung oleh teori yang dikemukakan oleh (Purwanto, 2009), yang menyatakan bahwa ada beberapa faktor yang dapat mempengaruhi reliabilitas hasil tes, yaitu :

- 1) Tingkat Validitas soal.
- 2) Luas atau banyak tidaknya sampel yang diambil.
- 3) Perbedaan bakat dan kemampuan peserta tes.
- 4) Suasana dan kondisi saat berlangsungnya tes.

2. Tingkat Kesukaran Soal

Berdasarkan hasil penelitian tingkat kesukaran soal di kelas X IPA 1 (Tabel 7 dan 8), terdapat 6,6% (2) soal dinyatakan **sukar**, 50% (15) soal dinyatakan **sedang**, dan 43,4% (13) soal dinyatakan **mudah**, serta nilai rata-rata tingkat kesukaran 1 adalah 0,54 atau memiliki tingkat kesukaran **sedang**. Sedangkan hasil penelitian tingkat kesukaran di kelas X IPA 2 (tabel 9 dan 10), dari 30 soal yang disajikan terdapat 43,4% (13) soal tergolong **sukar**, 23,3% (7) soal tergolong **sedang**, dan 33,3% (10) soal tergolong **mudah**, serta nilai rata-rata tingkat kesukarannya adalah 0,25 atau masuk kategori **sukar**. Hasil penelitian di kelas X IPA 1 yang berkategori sedang berbeda dengan hasil penelitian di kelas X IPA 2 yang berkategori sulit. Ini dapat disebabkan oleh tingkat kesukaran soal yang tidak seimbang. Ini didukung oleh teori yang berbunyi soal yang baik adalah soal yang memiliki tingkat kesukaran seimbang, perbandingan antara soal mudah, sedang dan sukar bisa dibuat 3-4-3, yang artinya 30% soal mudah, 40% soal sedang dan 30% soal sukar (Sujana, 2010).

3. Tingkat Daya Beda Soal Pilihan Ganda

Berdasarkan hasil penelitian tingkat daya beda, disetiap kelas peserta tes dikelompokkan menjadi 2 kelompok, yaitu kelompok atas dan kelompok bawah. Dari 30 soal yang diberikan kepada kedua kelompok di kelas X IPA 1 (table 9 dan 10), diperoleh hasil penelitian 1 (3,3%) soal tergolong **baik**, 2 (6,7%) soal tergolong **cukup baik**, 4 (13,3%) soal tergolong **kurang baik**, dan 23 (76,7%) soal tergolong **tidak baik**. Sedangkan hasil penelitian di kelas X IPA 2 (tabel 13 dan 14), diperoleh 7 (23,3%) soal tergolong **baik**, 7 (23,3%) soal tergolong **cukup baik**, 3 (10%) soal tergolong **kurang baik** dan 13 (43,4%) soal tergolong **tidak baik**. Nilai rata-rata daya beda soal di kelas X IPA 1 adalah 0,05 sehingga tergolong berdaya beda **tidak baik**. Sedangkan nilai rata-rata daya beda soal di kelas X IPA 2 adalah 0,24 atau tergolong **kurang baik**. Oleh sebab itu, soal yang berdaya beda tidak baik, harus dibuang atau diganti dan soal yang berdaya beda kurang baik harus direvisi agar bisa digunakan lagi. Ini didukung oleh teori menyatakan soal yang berdaya beda baik (digunakan) adalah soal yang memiliki index pembeda 0,40-1, cukup baik, dapat digunakan dengan revisi adalah soal yang memiliki index pembeda 0,30-0,39, kurang baik, perlu pembahasan /revisi adalah soal yang memiliki index pembeda 0,20-0,29, sedangkan jika index pembeda -1,00-0,00, maka soal harus dibuang atau diganti (Dali S., 1992).

4. Tingkat Efektif Pengecoh Soal Pilihan Ganda

Berdasarkan hasil penelitian tingkat efektivitas pengecoh di kelas X IPA 1 (tabel 11) terlihat bahwa terdapat 96 opsi jawaban pengecoh tergolong baik, karena memiliki indeks pengecoh $\geq 5\%$, 13 opsi jawaban pengecoh tergolong direvisi, karena memiliki indeks pengecoh $< 5\%$, sedangkan 11 opsi jawaban pengecoh lainnya tergolong tidak baik, karena memiliki indeks pengecoh $< 2\%$. Sedangkan untuk hasil penelitian tingkat pengecoh di kelas X IPA 2 (tabel 12), diperoleh 97 opsi jawaban tergolong baik, 9 opsi jawaban tergolong kurang baik dan direvisi dan 14 soal tergolong tidak baik. Pengecoh yang tidak baik adalah pengecoh hanya dipilih oleh 1 atau 2 peserta tes, dan bahkan tidak satupun peserta tes yang memilihnya. Nilai (Arikunto, 2013) efektivitas pengecoh ini dilihat berdasarkan ketentuan dari (Uno, H.B, 2018) yang berbunyi "pengecoh akan efektif dengan baik apabila dipilih sekurang-kurang 5% dari jumlah peserta tes".

KESIMPULAN & SARAN

Berdasarkan hasil analisis soal tingkat kesukaran, daya beda, efektivitas pengecoh pada butir soal ujian akhir semester ganjil mata pelajaran biologi kelas X IPA 1 dan X IPA 2 di SMAN 1 Pandawai, diperoleh kesimpulan sebagai berikut :

- 1) Pada tingkat kesukaran, berdasarkan rata-rata nilai tingkat kesukaran mencapai 0,54 (tergolong **sedang**) untuk kelas X IPA 1 tetapi 0,25 (tergolong **sukar**) untuk kelas X IPA 2, maka dapat disimpulkan bahwa guru perlu merevisi soal yang tergolong sukar, sedang dan mudah, agar bisa digunakan untuk mengukur tingkat kemampuan siswa, karena tingkat kemampuan siswa yang berbeda dan tidak ada keseimbangan soal untuk mencapai kriteria soal yang berkualitas baik.
- 2) Pada tingkat daya beda, berdasarkan nilai rata-rata tingkat daya beda untuk kelas X IPA 1 berjumlah 0,05 tergolong tidak baik dan 0,24 tergolong kurang baik untuk kelas X IPA 2. Guru juga perlu merevisi soal yang tergolong tidak baik dan kurang baik jika ingin digunakan lagi.
- 3) Pada tingkat efektivitas pengecoh di kelas X IPA 1, terdapat 13 opsi pengecoh **kurang baik**, harus direvisi, karena memiliki indeks pengecoh $< 5\%$, dan 11 opsi pengecoh tergolong **tidak**

baik digunakan karena memiliki indeks pengecoh >2%. Selanjutnya tingkat efektivitas pengecoh di kelas X IPA 2, terdapat 9 opsi pengecoh **kurang baik**, harus direvisi, dan 14 opsi pengecoh tergolong **tidak baik**, harus dibuang atau diganti. Jika guru ingin memakai opsi yang tergolong kurang baik dan tidak baik, maka guru perlu merevisinya agar dapat digunakan sebagai alat ukur kemampuan siswa.

DAFTAR PUSTAKA

- Abdul Kadir. (2015). Menyusun dan Menganalisis Tes Hasil Belajar. *Jurnal Al-Ta'dib*, 8(2), 81.
- Arifin, Z. (2016). *Evaluasi Program Teori dan Praktek dalam Konteks Pendidikan dan Nonpendidikan*. Bandung: PT Remaja Rosdakarya.
- Arikunto, S. (2013b). *Dasar-dasar Evaluasi Pendidikan*. Jakarta :Bumi Aksara.
- Asep, A. H. &. (2013). *Evaluasi Pembelajaran*. Yogyakarta: Multi Pressindo
- Azis. (2016). Analisis Tes Buatan Guru Bidang Studi Matematika Kelas V SD I Katobengke. *ISSN:2008-2157, 6(2)*, 18.
- Bano, V. O.,& Marambaawang N.D. (2022). Analisis Kriteria Butir Soal Ujian Sekolah Mata Pelajaran IPA di SMA Negeri 1 Waingapu. *Jurnal Ideas*.
- Departemen Pendidikan Dan Kebudayaan. (1997). *Manual Item And Test Analysis (Iteman)*. Badan Penelitian dan Pengembangan Pendidikan dan Kebudayaan: Pusat Penelitian dan Pengembangan Sistem Pengujian.
- Naga Dali S. (1992). *Pengantar Teori Sekor Pada Pengukuran Pendidikan*. Jakarta: Besbats.
- Nana Sudjana. (2013). *Penilaian Hasil Proses Belajar Mengajar*. Jakarta : Raja Grafindo Persada.
- Nana Sujana. (2010). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Ngalim Purwanto. (2009). *Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya.
- Purwanto. (2011). *Evaluasi Hasil Belajar, Komponen dan Model-Model Evaluasi Pendidikan*. Yogyakarta : Pustaka Pelajar.
- Purwanto, M. N. (2008). *Prinsip-prinsip dan Teknik Evaluasi Pembelajaran*. Bandung: PT. Remaja Rosdakarya.
- Resty Putri A, & R. Y. (2020). *Analisis Distraktor Butir Soal*. Bunpou Shokyuu Kouhan Semester 2 Tahun 2018/2019. *Prodi Pendidikan Bahasa Jepang Unnes*. Kiryoku.
- Sudijono, A. (2009). *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo.
- Sudjana, A. (2014). *Pengantar Evaluasi Pendidikan*. Raja Grafindo Persada.
- Sudjana, Nana. (2010). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Sugiyono. (2016). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sukardi. (2009). *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta: PT. Bumi Akasara
- Sukardi. (2013). *Metode Penelitian Pendidikan Tindakan Kelas Implementasi dan Pengembangannya*. Jakarta: Bumi Aksara
- Setiawaty. (2017). Index Tingkat Kesulitan (ITK) dan Keberfungsian Distraktor Soal Pilihan Ganda UAS Genap Mata Pelajaran Bahasa Indonesia.
- Uno, H. B., & Koni, S. (2018). *Assessment Pembelajaran*. PT Bumi.
- Wardoyo, W., & Suprpto, E. (2014). Rancang Bangun Program Analisis Butir Soal Pilihan Ganda Sebagai Pendukung Proses Evaluasi Pembelajaran. *Jurnal Teknik Elektro*, 6(2). <https://doi.org/10.15294/jte.v6i2.3589>
- Widoyoko, E. P. (2018). *Penilaian Hasil Pembelajaran di Sekolah*. Yogyakarta: Pustaka Belajar. Hal. 307, 242-244.