

An analysis of test question quality of the final semester examination in biology subject for class XI Of SMAN 2 Tana Tidung

Analisis kualitas butir soal ujian akhir semester pada mata pelajaran biologi kelas XI Di SMAN 2 Tana Tidung

¹Ibrahim, ¹Siti Musfiah, ¹Alfian Mucti

¹Program Studi Pendidikan Biologi, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Borneo Tarakan, Tarakan, Indonesia

Email*: sitimasfiah1998@gmail.com

Abstract: *The present research aimed to determine the test question quality of the Final Semester Examination in Biology Subject for Class XI of SMAN 2 Tana Tidung in terms of validity, reliability, level of difficulty, discrimination power, and the effectiveness of the distractors in each question. This research was a quantitative descriptive research with a population of all test questions of the final Semester Examination at SMAN 2 Tana Tidung for the 2020/2021 and 2021/2022 academic years. The sampling technique used was purposive sampling, in this case, the test questions of the final semester examination for class XI were taken as the sample. The data were collected using documentation and interviews. The data were then analyzed using Anates Version 4.0 software for multiple choice and essay test questions and using SPSS for matching test questions. Based on the results of data analysis, the test questions of the odd semester in the 2020/2021 academic year were obtained 6 (30%) qualified questions, 6 (30%) poor quality questions, and 8 (40%) questions which did not meet the quality criteria. The test questions of the even semester in the 2020/2021 academic year were attained 14 (56%) qualified questions, 4 (16%) poor quality questions, and 7 (28%) questions which had no quality. The test questions of the odd semester in the 2021/2022 academic year were found 6 (30%) qualified questions, 10 (50%) poor quality questions, and 4 (20%) questions which were not qualified. The test questions of the even semester in the 2021/2022 academic year on matching test questions were gained 5 (62,5%) statement items which got the quality criteria, 1 (12,5%) poor quality item, and 2 (25%) items which were not qualified; whereas in essay test questions were found 3 (100%) items which met the quality criteria.*

Keywords: *Analysis of Test Questions Quality, Examination Test Questions, and Biology Subject.*

Pendahuluan

Menurut Undang-Undang No. 20 Tahun 2003, menyatakan bahwa pendidikan merupakan suatu usaha sadar dan terencana melalui proses pembelajaran yang digunakan untuk mengembangkan segala potensi yang dimiliki peserta didik.

Kualitas pendidikan menjadi perhatian bagi setiap negara. Melalui pendidikan yang berkualitas sumber daya manusia dapat ditingkakan, sehingga diharapkan tingkat kesejahteraan masyarakat menjadi meningkat. Pendidikan dapat dikatakan berkualitas dilihat dari kemampuan lulusan suatu jenjang pendidikan. Kemampuan lulusan terdiri dari aspek pengetahuan, perilaku, dan keterampilan. Suatu dunia pendidikan mempunyai hubungan yang sangat erat antara pendidik dan Peserta didik, sehingga terjadi suatu proses pembelajaran di sekolah. Oleh karena itu, dibutuhkan pendidik yang profesional dalam mendidik peserta didik yang berkualitas dan unggul (Sukmela & Nofrion, 2018).

Keberhasilan peserta didik dapat dilihat melalui proses pembelajaran. Proses belajar mengajar di sekolah tidak bersifat mekanisme saja, akan tetapi memiliki visi, misi, serta tujuan sebagai target yang dicapai. Oleh karena itu, untuk mencapai target tersebut perlu dilakukan proses evaluasi. Kegiatan evaluasi dilakukan oleh pendidik. Pendidik harus memperoleh informasi terkait sejauh mana tingkat kemampuan dan perubahan peserta didik melalui kegiatan evaluasi. Evaluasi berarti proses yang secara sistematis dilakukan untuk mengetahui efisiensi dan tingkat keberhasilan pada program yang bersangkutan (Subali, 2012).

Adapun kegiatan evaluasi memiliki hubungan dengan kegiatan pengukuran. Pengukuran itu sendiri merupakan salah satu tahapan sangat penting dalam suatu proses penelitian ilmiah. Melalui pengukuran akan dihasilkan data penelitian, data penelitian tersebut dapat ditafsirkan, disimpulkan dan diimplikasikan. Karena pencapaian terkait perkembangan peserta didik perlu diukur terlebih dahulu baik secara individu maupun secara kelompok. Hal ini dapat disadari oleh pendidik karena pada masing-masing peserta didik mempunyai kemampuan yang bervariasi. Setiap peserta didik ada yang memiliki kemampuan dalam menangkap materi dengan cepat atau lambat. Sehingga pendidik dapat mengevaluasi terkait perkembangan kemampuan peserta didik dengan mengetahui apa yang mereka kerjakan dari awal hingga akhir. Pencapaian belajar dapat dievaluasi melalui kegiatan pengukuran. Kegiatan pengukuran pada bagian evaluasi digunakan untuk menghasilkan data, baik data kuantitatif maupun kualitatif (Jumaeda, 2016).

Kegiatan pengukuran memerlukan sebuah alat ukur yang biasa disebut dengan instrumen. Instrumen merupakan alat ukur yang digunakan untuk mengukur dengan cara mengumpulkan data. Instrumen yang digunakan dalam pendidikan datanya berupa tes atau non tes. Tes merupakan cara untuk mengadakan penilaian dalam bentuk tugas atau serangkaian tugas yang harus dikerjakan oleh peserta didik baik individu maupun kelompok, sehingga menghasilkan nilai terkait tingkah laku atau prestasi peserta didik. Adapun karakteristik instrumen yang baik sebagai alat evaluasi adalah memenuhi syarat utamanya harus validitas dan reliabilitas. Oleh karena itu, alat evaluasi yang baik dapat dilihat dari berbagai segi, yaitu validitas, reliabilitas, objektivitas, praktikalbitas, daya pembeda, taraf kesukaran, efektivitas option, dan efisiensi (Hadi dkk, 2019).

Berdasarkan uraian di atas, penulis tertarik untuk melakukan penelitian yang berjudul “Analisis Kualitas Butir Soal Ujian Akhir Semester Pada Mata Pelajaran Biologi Kelas XI di SMAN 2 Tana Tidung”.

Metode Penelitian

Jenis penelitian yang digunakan berupa penelitian deskriptif kuantitatif dengan pendekatan analisis kuantitatif. Adapun populasi dari penelitian ini adalah seluruh soal Ujian Akhir Semester di SMAN 2 Tana Tidung Tahun Ajaran 2020/2021 dan 2021/2022. Teknik sampling yang digunakan yaitu *Purposive sampling*. Sampel yang digunakan berupa soal Ujian Akhir Semester kelas XI di SMAN 2 Tana Tidung Instrumen yang digunakan berupa dokumentasi. Dokumentasi tersebut berupa naskah soal, lembar jawaban peserta didik, dan kunci jawaban yang berasal dari sekolah. Teknik pengumpulan data berupa dokumentasi dan wawancara. Teknik analisis data pada tiap butir-butir soal UAS semester pada mata pelajaran biologi kelas XI di SMAN 2 Tana Tidung dengan mencari validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektifitas pengecoh. Masing- masing kriteria tersebut akan dihitung menggunakan *Software Anates Versi 4.0* untuk soal dalam bentuk Pilihan Ganda (PG) dan uraian/essay, sedangkan *Statistical Product and Service Solution (SPSS) Versi 20* digunakan untuk analisis butir soal menjodohkan.

Hasil Penelitian

Analisis Soal Secara Kuantitatif

a. Analisis Validitas

Analisis validitas digunakan untuk menentukan tingkat kualitas butir soal. Analisis tingkat validitas butir soal dapat dilakukan dengan cara mengkorelasikan antara skor butir soal yang diperoleh peserta didik dengan skor total. Adapun interpretasi hasil dari analisis validitas butir soal disajikan pada tabel 1.

Tabel 1. Interpretasi Validitas Butir Soal

Soal UAS	Semester	Bentuk Soal	Interpretasi Validitas				
			Sangat Rendah	Rendah	Sedang	Tinggi	Sangat Tinggi
2020/2021	Ganjil	PG	1,4,5,6,8 ,9,10,18	2,7,12,1 4,19,20	3,11,13, 15,16	17	-
		Σ	8	6	5	1	-
	Persentase	40%	30%	25%	5%	-	
	Genap	PG	1,4,7,8,1 0,16,21	3,11,12, 19	5,6,9,13 ,14,15,2 0,23,24, 25	2,17,18, 22	-
Σ		7	4	10	4	-	

		Persentase	28%	16%	40%	16%	-
		PG	5,6,14,17	1,3,7,10,13,15,16,18	2,4,8,12,19	9,11,20	-
	Ganjil	Σ	4	8	5	3	-
		Persentase	20%	40%	25%	15%	-
2021/2022		Menjodohkan	5,8	1	-	3,4,6,7	2
		Σ	2	1	-	4	1
	Genap	Persentase	25%	12,5%	-	50%	12,5%
		Essay	-	-	-	-	1,2,3
		Σ	-	-	-	-	3
		Persentase	-	-	-	-	100%

Berdasarkan Tabel 1 interpretasi validitas butir soal menunjukkan bahwa dari butir soal Ujian Akhir Semester tahun ajaran 2021/2022 semester genap pada soal bentuk essay dan 1 butir pernyataan pada soal menjodohkan yang termasuk dalam kategori “Sangat Tinggi”.

b. Data Reliabilitas Tes

Setelah dilakukan perhitungan menggunakan *Software Anates Versi 4.0* pada soal PG dan essay dan SPSS untuk soal menjodohkan, diperoleh koefisien reliabilitas tes disajikan pada tabel 2.

Tabel 2. Analisis Koefisien Reliabilitas Tes

Soal UAS	Semester	Bentuk Soal	Rata-rata	Simpang Baku	Korelasi XY	Reliabilitas Tes	Interpretasi Reliabilitas
2020/2021	Ganjil	PG	7,07	2,88	0,44	0,61	Tinggi
	Genap	PG	18,41	3,13	0,43	0,60	Tinggi
2021/2022	Ganjil	PG	10,09	3,12	0,48	0,65	Tinggi
	Genap	Menjodohkan	4,70	1,79	-	0,67	Tinggi
		Essay	49,52	21,58	0,77	0,87	Sangat Tinggi

Berdasarkan Tabel 2 terlihat bahwa koefisien reliabilitas tes dari butir soal Ujian Akhir Semester Tahun ajaran 2020/2021 dan 2021/2022 memiliki koefisien reliabilitas tes $\geq 0,60$.

c. Tingkat Kesukaran

Hasil perhitungan dengan menggunakan *Software Anates Versi 4.0* untuk soal PG dan essay, sedangkan SPSS untuk soal menjodohkan, maka diperoleh data tingkat kesukaran untuk setiap butir soal termasuk kategori sukar, sedang, dan mudah. Data tingkat kesukaran butir soal yang dikategorikan disajikan pada tabel 3.

Tabel 3. Interpretasi Hasil Analisis Tingkat Kesukaran

Soal UAS	Semester	Bentuk Soal	Interpretasi Tingkat Kesukaran Butir Soal		
			Sukar	Sedang	Mudah
2020 /2021	Ganjil	PG	2,4,5,6,7,9,1 0,18	1,3,8,11,13,1 4,15,16,17,1 9,20	12
		Σ	8	11	1
		Persentase	40%	55%	5%
	Genap	PG	6,16	5,7,10,13,14, 20	1,2,3,4,8,9,1 1,12,15,17,1 8,19,21,22,2 3,24,25
		Σ	2	6	17
		Persentase	8%	24%	68%
2021 /2022	Ganjil	PG	4,5,6,17,19	2,3,8,9,10,11 ,12,13,16,18, 20	1,7,14,15
		Σ	5	11	4
		Persentase	25%	44%	16%
	Genap	Menjodohkan	5,8	4,6,7	1,2,3
		Σ	2	3	3
		Persentase	25%	37,5%	37,5%
		Essay	-	1,2,3	-
		Σ	-	3	-
		Persentase	-	100%	-

Berdasarkan Tabel 3 menunjukkan bahwa pada semester ganjil 2020/2021 terdapat 8 butir soal dengan kategori “sukar”, 11 butir soal dengan kategori “sedang”, dan 1 butir soal “mudah”. Terdapat 2 butir soal dengan kategori “sukar”, 6 butir soal “sedang”, dan 17 butir soal dengan kategori “mudah” pada soal Ujian akhir Semester Genap 2020/2021. Sedangkan pada soal Ujian Akhir Semester Ganjil 2021/2022, terdapat 5 butir soal dengan kategori “sukar”, 11 butir soal dengan kategori “sedang”, dan 4 butir soal dengan kategori “mudah”. Pada soal

tahun ajaran 2021/2022 semester genap, pada soal menjodohkan terdapat 1 butir pernyataan dengan kategori “sukar”, 3 butir pernyataan dengan kategori “sedang”, 3 butir pernyataan dengan kategori “mudah”, sedangkan 3 butir soal bentuk essay yang termasuk kategori “sedang”.

d. Analisis Daya Pembeda Soal Ujian Akhir Semester Pada Mata Pelajaran Biologi Kelas XI SMAN 2 Tana Tidung

Hasil dari perhitungan soal PG dan essay dengan menggunakan *Software Anates* Versi 4.0 dan SPSS yang digunakan untuk perhitungan soal menjodohkan pada butir soal Ujian Akhir Semester, maka diperoleh daya pembeda pada setiap butir soal termasuk kategori jelek sekali, jelek, cukup, baik, dan baik sekali. Kategori data daya pembeda pada butir soal disajikan pada tabel 4.

Tabel 4. Interpretasi Hasil Analisis Daya Pembeda

Soal UAS	Semester	Bentuk Soal	Interpretasi Daya Pembeda Butir Soal				
			Jelek Sekali	Jelek	Cukup	Baik	Baik Sekali
2020/2021	Ganjil	PG	8	1,4,5,6,10,18	7,9,14	2,11,1 2,13,1 6,19,20	3,15,17
		Σ	1	6	3	7	3
		Persentase	5%	30%	15%	35%	15%
	Genap	PG	7,8	1,3,10,18,21	9,11,14,25	2,4,13 ,19,20 ,22,24	5,6,
		Σ	2	8	6	7	2
		Persentase	8%	32%	24%	28%	8%
2021	Ganjil	PG	5,6,17	1,14	4,10,15,16,18	8,9,17,21 2,13,19	9,20
		Σ	3	2	5	8	2
		Persentase	15%	10%	25%	40%	10%
/2022	Menjodohkan		5	8	1	3,4,6,7	2
		Σ	1	1	1	4	1
	Genap	Persentase	12,5%	12,5%	12,5%	50%	12,5%
		Essay	-	-	-	2	1,3
		Σ	-	-	-	1	2
Persentase	-	-	-	25%	75%		

Berdasarkan Tabel 4 menunjukkan bahwa soal UAS 2020/2021 semester ganjil butir soal paling banyak pada kategori “baik” sebanyak 7 butir soal, semester genap pada soal UAS 2020/2021 butir soal paling banyak terdapat pada kategori “jelek” sebanyak 8 butir soal, serta di tahun ajaran 2021/2022 semester ganjil

jumlah butir soal yang terbanyak termasuk pada kategori “baik” sebanyak 8 butir soal. Soal tahun ajaran 2021/2022 semester genap, terdapat 4 butir pernyataan dalam soal menjodohkan dengan kategori “baik”, dan terdapat 2 butir soal essay yang termasuk kategori “sangat baik”.

e. Analisis Efektivitas Pengecoh Pada Soal Tes

Hasil dari perhitungan menggunakan *Software Anates Versi 4.0* pada butir soal Ujian Akhir Semester, maka diperoleh data efektivitas pengecoh pada soal ujian tahun ajaran 2020/2021 semester ganjil dan genap, sedangkan pada tahun ajaran 2021/2022 hanya untuk semester ganjil. Data tersebut disajikan pada tabel sebagai berikut:

Tabel 5. Hasil Rekapitulasi Analisis Indeks Efektivitas Pengecoh Butir Soal

Tahun	Semester	Indeks Efektivitas Pengecoh					
		SB	B	KB	BR	SBR	
2020/2021	Ganjil	1,2,3,5,6 ,7,9,10,1 2,13,15, 16,17	1,2,3,4,7 ,8,9,10,1 1,12,13, 14,15,16 ,17,18,1 9,20	2,3,4,5, 6,8,9,1 1,14,17 ,18,19, 20	6,7,9,1 1,13,14 ,20	6,8,13,14 ,17	
		Σ	13	18	13	7	5
		Persentase	65%	90%	65%	35%	25%
2020/2021	Genap	3,6,9,20, 23,24	1,2,4,10, 11,13,14 ,19,20,2 2,24	5,7,10, 16	1,2,3,4, 5,6,7,8, 9,10,11 ,12,13, 14,15,1 6,17,18 ,19,21, 22,23,2 4,25	3,4,6,7,1 1,12,13,1 4,16,18,2 1,22,23,2 4,25	
		Σ	6	11	4	24	15
		Persentase	24%	44%	16%	96%	60%
2021/2022	Genap	2,3,4,5,6 ,7,8,10,1 1,12,15, 16,17,19	2,3,5,6,9 ,10,11,1 2,13,18, 19,20	2,4,8,9, 10,16,1 7,20	1,2,3,4, 7,8,9,1 0,12,14 ,15,16, 19,20	4,7,15,19 ,20	
		Σ	14	12	8	14	5
		Persentase	70%	60%	40%	70%	25%

Berdasarkan Tabel 5 hasil rekapitulasi analisis efektivitas pengecoh pada butir soal UAS semester ganjil 2020/2021 menunjukkan bahwa sebanyak 13 butir soal (65%) dengan indeks pengecoh sangat baik, 18 (90%) butir soal dengan indeks Baik, 13 (65%) butir soal dengan indeks pengecoh kurang baik, 7 (35%) butir soal dengan indeks pengecoh buruk, dan 5 (25%) butir soal dengan indeks pengecoh

sangat buruk. Sedangkan pada semester genap terdapat 6 (24%) butir soal dengan indeks pengecoh sangat baik, 11 (44%) butir soal dengan indeks pengecoh baik, 4 (16%) butir soal dengan indeks pengecoh kurang baik, 24 (96%) butir soal dengan indeks pengecoh buruk, dan 15 (60%) butir soal dengan indeks pengecoh sangat buruk. Soal UAS semester ganjil 2021/2022 terdapat 14 (70%) butir soal dengan indeks pengecoh sangat baik, 12 (60%) butir soal dengan indeks pengecoh baik, 8 (40%) dengan indeks pengecoh kurang baik, 14 (70%) butir soal dengan indeks pengecoh buruk, serta 5 (25%) butir soal dengan indeks pengecoh sangat buruk.

Pembahasan

Analisis butir soal digunakan untuk mengetahui kualitas soal melalui analisis butir soal, maka akan diperoleh penyebab peserta didik belum mencapai ketuntasan. Peserta didik belum mencapai ketuntasan bisa saja disebabkan oleh soal yang terlalu sulit dan bukan karena belum menguasai materi. Menurut Purwanto (2012), menyatakan bahwa validitas memiliki hubungan dengan kemampuan untuk mengukur secara tepat sesuatu yang diinginkan diukur. Validitas digunakan untuk menentukan kehandalan suatu butir soal. Penentuan dalam validitas butir soal menggunakan korelasi *product moment* dengan cara mengkorelasikan skor yang diperoleh peserta didik dengan skor total. Hasil analisis validitas butir soal pada soal tahun ajaran 2020/2021 semester ganjil terdapat 8 butir soal (40%) dengan kategori “sangat rendah”, 6 butir soal (30%) dengan kategori “rendah”, 5 butir soal (25%) dengan kategori “sedang”, 1 butir soal (5%) dengan kategori “tinggi”, dan tidak terdapat butir soal yang termasuk dalam kategori “sangat tinggi”. Soal tahun ajaran 2020/2021 semester genap terdapat 7 butir soal (28%) dengan kategori “sangat rendah”, 4 butir soal (16%) dengan kategori “rendah”, 10 butir soal (40%) dengan kategori “sedang”, 4 butir soal (16%) dengan kategori “tinggi”, dan tidak ada butir soal yang termasuk dalam kategori “sangat tinggi”. Soal tahun ajaran 2021/2022 semester ganjil terdapat 4 butir soal (20%) dengan kategori “sangat rendah”, 8 butir soal (40%) dengan kategori “rendah”, 5 butir soal (25%) dengan kategori “sedang”, 3 butir soal (15%) dengan kategori “tinggi”, serta tidak ada butir soal yang termasuk dalam kategori “sangat tinggi”. Pada soal tahun ajaran 2021/2022 semester genap, pada soal menjodohkan terdapat 2 butir pernyataan (25%) dengan kategori “sangat rendah”, 1 butir pernyataan (12,5%) dengan kategori “rendah”, 4 butir pernyataan (50%) dengan kategori “tinggi”, 1 butir pernyataan (12,5%) dengan kategori “sangat tinggi”, serta tidak ada butir pernyataan yang termasuk kategori “sedang. Sedangkan pada soal essay terdapat 3 butir soal (100%) dengan kategori “sangat tinggi”.

Soal yang valid yaitu soal tersebut dapat mengukur kompetensi yang diharapkan, sedangkan soal tidak valid yaitu soal tersebut tidak dapat mengukur kompetensi yang diharapkan. Suatu teknik evaluasi dapat dikatakan valid apabila teknik evaluasi atau tes tersebut dapat sepenuhnya mengukur kemampuan sesuai yang diharapkan. Maka dari itu, alat tes tersebut dapat digunakan untuk mengukur kemampuan peserta didik dalam mencapai kompetensi yang diharapkan (Arikunto, 2006). Berdasarkan uraian di atas, dapat disimpulkan bahwa pada soal Ujian Akhir Semester masih perlu diperbaiki dilihat berdasarkan hasil validitas soal pilihan ganda tahun ajaran 2020/2021 semester ganjil menunjukkan bahwa diperoleh 7 butir soal yang valid yaitu nomor 3,11,13,14,15,16,17 dan 13 butir soal tidak valid yaitu nomor 1,2,4,5,6,7,8,9,10,12,18,19,20. Pada soal tahun ajaran 2020/2021 semester genap diperoleh 14 butir soal valid yaitu nomor 2,5,6,9,13,14,15,17,18,20,22,23,24,25 dan 11 butir soal tidak valid yaitu nomor

1,3,4,7,8,10,11,12,16,19,21. Pada soal tahun ajaran 2021/2022 semester ganjil diperoleh 6 butir soal valid yaitu nomor 2,9,11,12,19,20 dan 14 butir soal tidak valid yaitu nomor 1,3,4,5,6,7,8,10,13,14,15,16,17,18. Soal tahun ajaran 2021/2022 semester genap, pada soal menjodohkan diperoleh 5 butir pernyataan valid yaitu nomor 2,3,4,6,7 dan 3 butir pernyataan tidak valid yaitu nomor 1,5,8, sedangkan pada soal essay diperoleh bahwa 3 butir soal valid. Menurut Arikunto (2013), soal tersebut dapat dikatakan valid apabila koefisien korelasinya $> r_{tabel}$ dengan taraf signifikansi sebesar 5% (0,05).

Reliabilitas berarti sejauh mana hasil pengukuran suatu tes dapat dipercaya. Reliabilitas suatu tes mengacu pada derajat stabilitas, konsistensi, daya prediksi, dan akurasi. Pengukuran yang mempunyai reliabilitas yang tinggi merupakan pengukuran yang menghasilkan data reliabel (Sanaky dkk, 2021). Hasil analisis reliabilitas tes diperoleh data pada soal tahun ajaran 2020/2021 semester ganjil nilai koefisien reliabilitas 0,61 termasuk dalam kategori “tinggi”, pada soal tahun 2020/2021 semester genap nilai koefisien reliabilitas 0,60 termasuk dalam kategori “tinggi”, pada soal tahun ajaran 2021/2022 semester ganjil nilai koefisien reliabilitas 0,65 termasuk dalam kategori “tinggi”. Soal tahun ajaran 2021/2022 semester genap diperoleh nilai koefisien reliabilitas pada soal menjodohkan 0,67 kategori “tinggi”, sedangkan pada soal esay koefisien reliabilitas sebesar 0,87 dengan kategori “sangat tinggi”. Menurut Arikunto bahwa koefisien reliabilitas dikatakan “tinggi” apabila koefisien reliabilitas berada pada rentang nilai $0,60 < r_{11} \leq 0,80$ dan rentang nilai $0,80 < r_{11} \leq 1,00$ dikatakan “sangat tinggi. Berdasarkan uraian di atas dapat disimpulkan bahwa pada soal Ujian Akhir Semester dikatakan reliabel, karena memiliki reliabilitas “tinggi” dan “sangat tinggi”. Maka soal tersebut mempunyai keterandalan yang tinggi. Keterandalan yang dimaksud yaitu meliputi ketepatan atau kecermatan hasil pengukuran dan kestabilan dari hasil pengukuran. Sehingga apabila dilakukan pengujian beberapa kali pada soal tes, maka akan memberikan hasil yang relatif sama (Nuswawati, 2010).

Perhitungan tingkat kesukaran butir soal merupakan pengukuran yang digunakan untuk mengukur seberapa besar derajat kesukaran suatu butir soal. Apabila suatu soal mempunyai tingkat kesukaran yang seimbang, maka soal tersebut dapat dikatakan baik. Menurut Arikunto (2006), menyatakan bahwa penentuan kelayakan soal perlu memperhatikan tujuan penggunaan soal, apabila soal tersebut digunakan untuk memperoleh pencapaian hasil belajar peserta didik maka harus menggunakan soal yang tidak terlalu mudah dan tidak terlalu sukar (Lumbanraja, 2016). Berdasarkan Hasil analisis tingkat kesukaran soal Ujian Akhir Semester pada mata pelajaran biologi kelas XI di SMAN 2 Tana Tidung diperoleh data bahwa pada soal tahun ajaran 2020/2021 semester ganjil terdapat 8 butir soal (40%) dengan kategori “sukar”, 11 butir soal (55%) dengan kategori “sedang”, dan 1 butir soal (5%) dengan kategori “mudah”. Soal tahun ajaran 2020/2021 semester genap terdapat 2 butir soal (8%) dengan kategori “sukar”, 6 butir soal (24%) dengan kategori “sedang”, dan 17 butir soal (68%) dengan kategori “mudah”. Kemudian pada soal tahun ajaran 2021/2022 semester ganjil terdapat 5 butir soal (25%) dengan kategori “sukar”, 11 butir soal (44%) dengan kategori “sedang”, dan 4 butir soal (16%) dengan kategori “mudah”. Soal tahun ajaran 2021/2022 semester genap, diperoleh bahwa pada soal menjodohkan terdapat 2 butir pernyataan (25%) kategori “sukar”, 3 butir pernyataan (37,5%) kategori “sedang”, dan 3 butir pernyataan (37,5%) dengan kategori “mudah”, sedangkan pada soal essay terdapat 3 butir soal (100%) dengan kategori “sedang”.

Dari uraian diatas dapat disimpulkan bahwa tingkat kesukaran butir soal Ujian Akhir Semester pada mata pelajaran biologi kelas XI di SMAN 2 Tana Tidung tahun ajaran 2020/2021 semester genap belum tergolong baik, sedangkan soal tahun ajaran 2020/2021 semester ganjil dan soal tahun ajaran 2021/2022 semester ganjil sudah tergolong baik. Adapun persentase soal dengan kategori “sedang” pada tahun ajaran 2020/2021 semester ganjil mencapai 55% yaitu soal nomor 1,3,8,11,13,14,15,16,17,19, dan 20, soal tahun ajaran 2020/2021 semester genap mencapai 24% yaitu soal nomor 5,7,10,13,14, dan 20, soal tahun ajaran 2021/2022 semester ganjil mencapai 44% yaitu soal nomor 2,3,8,9,10,11,12,13,16,18, dan 20, soal tahun ajaran 2021/2022 semester genap pada soal menjodohkan mencapai 37,5% yaitu pernyataan nomor 4,6,7 dan pada soal essay mencapai 100%. Menurut Sudjana (2004), soal yang diberikan kepada peserta didik sebaiknya memiliki keseimbangan antara sukar : sedang : mudah dengan perbandingan 3 : 4 : 3 atau 2 : 5 : 3. Dilihat dari perbandingan tersebut bahwa Soal dapat dikatakan seimbang apabila soal yang berkategori sedang memiliki perbandingan yang tinggi dibanding dengan soal yang berkategori mudah dan sukar. Ditinjau berdasarkan perbandingan, dapat diketahui bahwa pada soal tahun ajaran 2020/2021 semester ganjil memiliki perbandingan 8 : 11 : 1, sedangkan pada semester genap memiliki perbandingan 2 : 6 : 17. Soal tahun ajaran 2021/2022 semester ganjil memiliki perbandingan 5 : 11 : 4, sedangkan pada soal menjodohkan semester genap memiliki perbandingan 2 : 3 : 3 dan soal essay pada semester genap memiliki perbandingan 0 : 3 : 0. Hasil perbandingan ini menghasilkan proporsi yang tidak seimbang pada soal tahun ajaran 2020/2021 pada semester genap dan soal tahun ajaran 2021/2022 semester genap pada soal menjodohkan.

Daya pembeda soal merupakan kemampuan suatu soal untuk membedakan antara peserta didik yang memiliki kemampuan akademik rendah dengan peserta didik yang memiliki kemampuan akademik tinggi (Asri, 2012). Hasil dari analisis daya pembeda soal Ujian Akhir Semester pada mata pelajaran biologi kelas XI di SMAN 2 Tana Tidung diperoleh data bahwa pada soal tahun ajaran 2020/2021 semester ganjil terdapat 1 butir soal (5%) dengan kategori “sangat jelek”, 6 butir soal (30%) dengan kategori “jelek”, 3 butir soal (15%) dengan kategori “cukup”, 7 butir soal (35%) dengan kategori “baik”, dan 3 butir soal (15%) dengan kategori “baik sekali”. Soal tahun ajaran 2020/2021 semester genap terdapat 2 butir soal (8%) dengan kategori “sangat jelek”, 8 butir soal (32%) dengan kategori “jelek”, 6 butir soal (24%) dengan kategori “cukup”, 7 butir soal (28%) dengan kategori “baik”, serta 2 butir soal (8%) dengan kategori “baik sekali”. Kemudian pada soal tahun ajaran 2021/2022 semester ganjil terdapat 3 butir soal (15%) dengan kategori “sangat jelek”, 2 butir soal (10%) dengan kategori “jelek”, 5 butir soal (25%) dengan kategori “cukup”, 8 butir soal (40%) dengan kategori “baik”, dan 2 butir soal (10%) dengan kategori “baik sekali”. Soal tahun ajaran 2021/2022 semester genap, pada soal menjodohkan terdapat 1 butir pernyataan (12,5%) dengan kategori “sangat jelek”, 1 butir pernyataan (12,5%) dengan kategori “jelek”, 1 butir pernyataan (12,5%) dengan kategori “cukup”, 4 butir pernyataan (50%) dengan kategori “baik”, serta 1 butir pernyataan (12,5%) dengan kategori “baik sekali”, sedangkan pada soal essay terdapat 1 butir soal (25%) dengan kategori “baik” dan 2 butir soal (75%) dengan kategori “baik sekali”.

Berdasarkan uraian di atas dapat disimpulkan bahwa daya pembeda pada soal tahun ajaran 2020/2021 semester ganjil daya pembeda “cukup” yaitu nomor 7,9,14, kategori “baik” yaitu nomor 2,11,12,13,16,19,20, dan kategori “baik sekali” yaitu nomor 3,15,17. Soal pada tahun ajaran

2020/2021 semester genap daya pembeda “cukup” yaitu nomor 9,11,14,17,23,25 kategori “baik” yaitu nomor 2,4,13,19,20,22,24, dan kategori “baik sekali” yaitu nomor 5,6. Daya pembeda pada soal tahun ajaran 2021/2022 semester ganjil dengan kategori “cukup” yaitu nomor 4,10,15,16,18, kategori “baik” yaitu nomor 2,3,7,8,11,12,13,19, dan kategori “baik sekali” yaitu nomor 9,20. Soal tahun ajaran 2021/2022 semester genap, daya pembeda dengan kategori “cukup” yaitu nomor 1, kategori “baik” yaitu nomor 3,4,6,7, dan kategori “baik sekali” yaitu nomor 2 pada soal menjodohkan, sedangkan pada soal essay daya pembeda dengan kategori “baik” yaitu nomor 2 dan kategori “baik sekali” yaitu nomor 1,3. Menurut Nuswowati (2010), daya pembeda suatu soal berarti kemampuan suatu soal untuk membedakan antara pesertadidik yang pintar dengan peserta didik yang kurang pintar. Soal yang dapat dijawab dengan benar oleh peserta didik yang pintar maupun kurang pintar, maka soal tersebut dikatakan tidak baik karena tidak memiliki daya pembeda. Sehingga alat tes yang baik itu memiliki daya pembeda yang cukup. Karena daya pembeda yang cukup ini mampu membedakan antara peserta didik yang pintar dengan peserta didik yang kurang pintar.

Menurut Sudijono (2012), menyatakan bahwa tes objektif dalam bentuk pilihan ganda tersebut untuk setiap butir soal yang dikeluarkan dalam tes hasil belajar telah dilengkapi dengan beberapa pilihan jawaban. Jumlah pilihan jawaban berkisar antara 3 sampai 5, dari salah satu pilihan jawaban tersebut merupakan jawaban yang benar (kunci jawaban), sedangkan sisanya merupakan jawaban salah. Jawaban-jawaban yang salah itulah yang sering disebut dengan pengecoh. Berdasarkan uraian di atas analisis efektivitas pengecoh pada butir soal Ujian Akhir Semester pada mata pelajaran biologi di SMAN 2 Tana Tidung mempunyai 5 kategori yaitu sangat baik, baik, kurang baik, buruk, dan sangat buruk. Hasil analisis menggunakan *Software Anates Versi 4.0* diperoleh data indeks pengecoh yang masih sangat buruk (pengecoh tidak berfungsi) pada soal tahun ajaran 2020/2021 semester ganjil sebanyak 5 (25%) butir soal yaitu 6C, 8E, 13D, 14B, dan 17E. Soal tahun ajaran 2020/2021 semester genap sebanyak 15 (60%) butir soal yaitu 3E, 4C, 6C, 7D, 11D, 12A, 13B, 14D, 16A, 18E, 21A, 22B, 23B, 24E, dan 25B. Kemudian pada soal tahun ajaran 2021/2022 semester ganjil indeks pengecoh yang sangat buruk sebanyak 5 (25%) butir soal yaitu 4C, 7B, 15D, 19C, dan 20B. Sehingga dapat disimpulkan bahwa pengecoh tersebut berfungsi dengan baik.

Berdasarkan hasil analisis efektivitas pengecoh pada butir soal Ujian Akhir Semester pada mata pelajaran biologi di SMAN 2 Tana Tidung mempunyai 5 kategori yaitu sangat baik, baik, kurang baik, buruk, dan sangat buruk (tidak berfungsi). Hasil penelitian pada kategori sangat buruk diperoleh persentase sebesar 25% pada soal tahun ajaran 2020/2021 semester ganjil, 60% pada soal tahun ajaran 2020/2021 semester genap, dan 25% pada soal tahun ajaran 2021/2022 semester ganjil. Menurut Oktanin & Sukirno (2015), efektivitas pengecoh dapat berfungsi dengan sangat baik apabila keempat pengecoh berfungsi, efektivitas pengecoh dapat berfungsi dengan baik apabila terdapat tiga pengecoh yang berfungsi, efektivitas pengecoh dikatakan cukup baik apabila terdapat dua jawaban pengecoh yang berfungsi, efektivitas pengecoh dikatakan kurang baik apabila terdapat 1 jawaban pengecoh yang berfungsi, serta efektivitas pengecoh dikatakan tidak baik apabila tidak semua jawaban pengecoh yang berfungsi.

Dengan merujuk pada sumber-sumber yang telah diuraikan sebelumnya, soal yang baik adalah soal yang memenuhi persyaratan validitas, reliabilitas, tingkat kesukaran yang sedang, daya

pembeda yang baik, dan efektivitas pengecoh yang baik. Hasil analisis soal yang memenuhi persyaratan tersebut terdapat 6 (30%) butir soal yang berkualitas (dipakai) yaitu nomor 3,11,13,15,16,17, terdapat 6 (30%) butir soal kurang berkualitas (revisi) yaitu nomor 2,7,12,14,19,20, dan terdapat 8 (40%) butir soal yang di eliminasi (tidak berkualitas) yaitu nomor 1,4,5,6,8,9,10,18 pada soal tahun 2020/2021 semester ganjil. Pada soal tahun 2020/2021 semester genap diperoleh 14 (56%) butir soal yang dipakai yaitu nomor 2,5,6,9,13,14,15,17,18,20,22,23,24,25, terdapat 4 (16%) butir soal di revisi yaitu nomor 3,11,12,19, dan terdapat 7 (28%) butir soal yang di eliminasi yaitu nomor 1,4,7,8,10,16,21. Pada soal tahun 2021/2022 semester ganjil diperoleh 6 (30%) butir soal yang dipakai yaitu nomor 2,9,11,12,19,20, terdapat 10 (50%) butir soal di revisi yaitu nomor 1,3,4,7,8,10,13,15,16,18, dan terdapat 4 (20%) butir soal di eliminasi yaitu nomor 5,6,14,17. Sedangkan pada soal tahun 2021/2022 semester genap pada soal menjodohkan diperoleh 5 (62,5%) butir pernyataan yang digunakan yaitu nomor 2,3,6,7, terdapat 1 (12,5%) butir pernyataan di revisi yaitu nomor 1, dan terdapat 2 (25%) butir pernyataan di eliminasi yaitu nomor 5,8, dan pada soal essay terdapat 3 (100%) butir soal yang dipakai. Menurut Sugiyono, kualitas butir soal dapat digunakan (di simpan di bank soal) apabila memenuhi 4 kriteria yaitu validitas, tingkat kesukaran, daya pembeda, efektivitas pengecoh. Kualitas butir soal dapat di revisi (perbaikan) apabila memenuhi 3 atau 2 dari keempat kriteria yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Sedangkan soal yang dapat di eliminasi apabila hanya memenuhi 1 atau tidak sama sekali dari keempat kriteria (validitas, tingkat kesukaran, daya pembeda, efektivitas pengecoh), serta syarat untuk tes secara keseluruhan harus reliabel.

Simpulan

Berdasarkan hasil analisis butir soal dan pembahasan yang telah diuraikan sebelumnya, maka dapat ditarik kesimpulan bahwa kualitas butir soal ujian akhir semester pada mata pelajaran biologi kelas XI di SMAN 2 Tana Tidung tahun ajaran 2020/2021 terdapat 43% soal yang berkualitas, 23% kurang berkualitas, dan 34% tidak berkualitas. Pada tahun ajaran 2021/2022 terdapat 64,2% soal yang berkualitas, 20,8% kurang berkualitas, dan 15% tidak berkualitas.

Ucapan Terima Kasih

Peneliti mengucapkan terimah kasih kepada Tuhan Yang Maha Esa, kedua orang tua, bapak/ibu dosen pembimbing, dan teman-teman saya yang telah mendukung dan membantu dalam menyelesaikan penelitian ini.

Daftar rujukan

Arikunto, Suharsimi. (2006). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara

Arikunto, Suharsimi. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara

Asri, A. F., & Burhan, A. (2014). Analisis tingkat kesukaran, daya pembeda dan fungsi distraktor soal ujian semester ganjil mata pelajaran produktif di smk negeri 1 indralaya utara tahun pelajaran 2012/2013. *Jurnal Pendidikan Teknik Mesin, 1*(2).

Hadi, K., Dazrullisa, D., Manurung, B., & Hasruddin, H. (2019). Analisis Daya Beda Soal Tes Berpikir Tingkat Tinggi Berbasis Kearifan Lokal Pada Materi Keanekaragaman Hayati

Kelas X SMA. *Bionatural: Jurnal Ilmiah Pendidikan Biologi*, 6(2).

- Jumaeda, S. T. (2016). Komparasi bentuk tes pilihan ganda dengan tes menjodohkan (matching test) ditinjau dari tingkat kesukaran pada mata pelajaran bahasa Arab di MAN Waimital Kabupaten Seram bagian barat. *al-Iltizam: Jurnal Pendidikan Agama Islam*, 1(1), 31-47.
- Lumbanraja, L. H. (2017). *Analisis Tingkat Kesukaran Dan Daya Pembeda Pada Butir Tes Soal Ujian Tengah Semester Bahasa Indonesia Kelas XII SMA Negeri 7 Medan Tahun Pembelajaran 2016/2017* (Doctoral dissertation, UNIMED).
- Nuswowati, M., Binadja, A., & Ifada, K. E. N. (2010). Pengaruh validitas dan reliabilitas butir soal ulangan akhir semester bidang studi kimia terhadap pencapaian kompetensi. *Jurnal inovasi pendidikan kimia*, 4(1).
- Oktanin, W. S., & Sukirno, S. (2015). Analisis Butir Soal Ujian Akhir Mata Pelajaran Ekonomi Akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 13(1).
- Purwanto. 2012. *Instrumen Penelitian Sosial dan Pendidikan*. Yogyakarta: Pustaka Belajar.
- Sanaky, M. M. (2021). Analisis Faktor-Faktor Keterlambatan Pada Proyek Pembangunan Gedung Asrama Man 1 Tulehu Maluku Tengah. *Jurnal Simetrik*, 11(1), 432-439.
- Subali, Bambang. (2012). *Prinsip Asesment & Evaluasi Pembelajaran*. Yogyakarta: UNY Press
- Sudijono, Anas. (2012). *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada
- Sudjana, Nana. (2004). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya Offset
- Sugiyono. (2010). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sukmela, Yelit Putri, and Nofrion. 2018. *Analisis Butir Soal Paket Ujian Akhir Semester Genap Geografi Tahun Pelajaran 2016/2017 SMA Negeri Se Kota Pekanbaru*. Universitas Negeri Padang: Jurnal Buana, Volume 2 No.2
- Undang-Undang RI. 2003. Undang-Undang Republik Indonesia, Nomor 20, Tahun 2003, tentang Sistem Pendidikan Nasional