
**PENGEMBANGAN ASESMEN BIOLOGI BERBASIS
KETERAMPILAN BERPIKIR KRITIS SISWA
KELAS XI SEMESTER GENAP DI SMA
MUHAMMADIYAH 1 PALEMBANG**

Aseptianova¹
Sulton Nawawi²
Maulidiyah Yuliandina³

^{1,2,3} Program Studi Pendidikan Biologi, FKIP, Universitas Muhammadiyah Palembang
E-mail: Aseptianova@gmail.com

Abstract: This study aims to determine the results of the development and the appropriateness of Biological Assessment based on the critical thinking skills. This research is a research and a development. The development used in this study was a development model of the McIntire adaptation. The validity was done by the 2 expert validators; 1) material experts; and 2) critical thinking skills evaluation experts. The results of this study was the Biology assessment based on critical thinking skills in the form of description / essay with the material from even semester students grade XI and produce the final product questions from the first and second trials amounted to 34 items. The appropriateness of Biology assessment based on critical thinking skills by the expert validators showed that the results is sufficiently high and highly qualified. Moreover, and based on the results of the empirical validity of the items the questions are declared valid and reliable. In the test of the difficulty of the questions in the first trial and the second trial the average question was classified as difficult and moderate, while the distinguishing test for the 1st and 2nd trials was categorized sufficient.

Kata Kunci: Asesmen biologi, keterampilan berpikir kritis.

PENDAHULUAN

Pendidikan merupakan aspek terpenting di dalam kehidupan manusia. Pendidikan sebagai wadah untuk membentuk karakter seseorang, potensi diri dan pola pikir seseorang yang dapat ditingkatkan melalui proses pembelajaran. Pendidikan dapat diperoleh dan dikembangkan melalui sekolah, keluarga, dan masyarakat, sehingga menghasilkan karakter siswa yang berkualitas.

Pendidikan yang berkualitas mengupayakan siswa untuk memiliki kecerdasaan otak, kecerdasan hati, serta membekali keterampilan-keterampilannya (Widodo, 2015). Sebab dengan adanya pendidikan yang

berkualitas dapat menciptakan manusia yang cerdas sehingga mampu memiliki keterampilan berpikir kritis di era globalisasi. Salah satunya dengan mengembangkan potensi diri, kreativitas, keterampilan, serta kecakapan siswa di dalam dunia pendidikan. Pemerintah Indonesia sekarang ini sedang melakukan perbaikan paradigma pendidikan, sehingga dapat menciptakan lulusan yang mempunyai keterampilan dan berpikir tingkat tinggi, tetapi hakikatnya kualitas pendidikan di Indonesia masih belum maksimal.

Pernyataan tersebut dapat diketahui dengan pencapaian Indonesia dalam mengikuti dua

ajang yakni *Program for International Student Assessment* (PISA) dan *Trends in International Mathematics and Science Study* (TIMSS). hasil PISA, Pada tahun 2015 negara Indonesia menduduki peringkat ke 63 dari 72 negara dan berdasarkan data hasil TIMSS, pada tahun 2015, Indonesia menduduki peringkat ke 38 dari 39 negara. Hasil dua data tersebut menunjukkan bahwa prestasi pelajar Indonesia masih jauh tertinggal dari negara-negara lainnya, sehingga menyebabkan kualitas pendidikan di Indonesia masih kategori rendah. Penyebab kualitas pendidikan di Indonesia masih kurang adalah penggunaan kurikulum di Indonesia dan kemampuan seorang guru atau pendidik.

Pada Kurikulum 2013, khususnya dalam proses pembelajaran di tuntut agar siswa turut aktif dan menjadi pusat dalam proses pembelajaran, serta dituntut untuk berpikir, menganalisis, mengevaluasi, dan menyimpulkan sendiri apa yang menjadi permasalahan di setiap materi yang dipelajari guna untuk memecahkan masalah tersebut, guru hanya membimbing, mendorong, dan memberikan fasilitas bagi seorang siswa agar tercapai suatu tujuan pembelajaran.

Selain itu, tugas pendidik adalah mampu mengembangkan keterampilan-keterampilan yang dimiliki oleh siswa, salah satunya adalah keterampilan berpikir kritis. Pada keterampilan berpikir kritis, guru harus mempunyai asesmen berpikir kritis, karena Kurikulum 2013 guru dituntut untuk mengembangkan asesmen atau penilaian yang berbasis keterampilan berpikir kritis dan melatih keterampilan berpikir kritis siswa dengan memberikan soal-soal

berbasis keterampilan berpikir kritis berupa asesmen. Asesmen adalah sebagai alat penilaian untuk siswa di dalam proses pembelajaran, sebab dengan adanya asesmen ini, seorang guru bisa mengukur bagaimana perkembangan nilai siswa dalam proses pembelajaran di dalam kelas.

Berdasarkan persentase penguasaan materi soal dari pamer ujian nasional tingkat SMA/MA tahun 2017/2018, pembelajaran Biologi pada materi kelas XI khususnya materi semester genap yaitu sistem pencernaan 67,74%, sistem saraf 52,69%, sistem pernapasan dan sistem ekskresi 65,59%, gangguan/penyakit yang dialami 66,67%. Oleh karena itu, materi kelas XI semester genap dipergunakan dalam penelitian ini dan juga nilai tersebut masih dibawah nilai KKM yaitu 70 sedangkan berdasarkan analisis kebutuhan melalui hasil wawancara dan pemberian kuesioner bahwa guru Biologi kelas XI masih mengalami kesulitan dalam Kurikulum 2013 khususnya bagian penilaian, dan juga guru Biologi kelas XI tersebut belum pernah membuat soal-soal yang berpikir kritis, hanya menerapkan soal-soal dalam tingkatan ranah C1–C3.

Pada penelitian ini pengembangan soal-soal *essay* digunakan agar siswa tidak asal menjawab setidaknya siswa masih perlu berpikir dalam menjawab pertanyaan-pertanyaan yang diberikan oleh gurunya, karena dengan soal *essay*, siswa bisa menuangkan pikiran ke dalam lembar jawaban dan itu membuat siswa untuk melatih berpikir dalam menjawab soal.

METODE

Penelitian yang digunakan adalah Penelitian dan Pengembangan (*Research and Development*), adapun yang dikembangkan berupa asesmen (penilaian) Biologi yang berbasis

keterampilan berpikir kritis. Prosedur yang digunakan oleh peneliti dalam mengembangkan asesmen Biologi yaitu sesuai dengan model pengembangan dari McIntire dalam Mulyatiningsih (2011), adapun model tersebut terdiri atas 10 tahapan sebagai berikut:

1. Menjelasan kompetensi yang dapat ditunjukkan oleh peserta tes dan tujuan tes.
2. Mengembangkan rancangan tes.
3. Menyusun butir-butir soal.
4. Menulis petunjuk pelaksanaan tes.
5. Melakukan uji coba pada tes yang telah disusun.
6. Merevisi butir tes.
7. Melakukan analisis butir tes.
8. Validasi soal tes.
9. Setelah validasi, menetapkan norma acuan.
10. Melengkapi tes manual.

Subjek dalam penelitian ini adalah guru Biologi kelas XI, dosen validasi, dan siswa kelas XII SMA Muhammadiyah 1 Palembang. Penelitian ini menggunakan instrumen berupa lembar kuesioner, lembar wawancara, dan lembar observasi yang akan diisi oleh guru Biologi pada saat pengambilan data awal, kemudian lembar validasi yang akan diisi oleh dosen validasi ahli berpikir kritis dan dosen ahli materi, dan juga tes berupa soal essai berbasis keterampilan berpikir kritis yang akan dijawab oleh siswa SMA kelas XII. Kemudian teknik analisis data menggunakan analisis kualitatif dan analisis kuantitatif.

Perhitungan validasi dosen ahli ini menggunakan formula statistik sebagai berikut Azwar (2017):

$$V = \frac{\sum s}{n(c - 1)}$$

keterangan:

s = $r - lo$

lo = Angka rating yang terendah

c = Angka rating yang tertinggi

r = Angka rating yang diberikan oleh seorang penilai

Tabel 1. Kriteria Validitas

No	Hasil Validitas	Kriteria Validitas
1	$0,81 < V \leq 1,00$	Sangat Tinggi
2	$0,61 < V \leq 0,80$	Tinggi
3	$0,41 < V \leq 0,60$	Cukup
4	$0,21 < V \leq 0,40$	Rendah
5	$0,01 < V \leq 0,20$	Sangat Rendah

(Azwar dalam Madliya, Abdurachman, & Hartono, 2017)

Uji prasyarat instrumen merupakan salah satu cara untuk menentukan butir soal atau kelayakan soal berpikir kritis. Uji prasyarat instrumen meliputi: validitas soal, reliabilitas soal, tingkat kesukaran dan uji daya beda.

a. Validitas

Pada analisis validitas item soal menggunakan program SPSS 16. Pada uji validitas dalam menilai apakah item soal tersebut valid atau tidak, melalui uji signifikansi, yaitu uji-t dengan menggunakan derajat kebebasan $n-2$ dengan n adalah jumlah subyek. Dengan menggunakan taraf signifikansi tertentu akan diketahui nilai t-teoritisnya melalui tabel t-student yang banyak terdapat dilampirkan berbagai buku statistik maupun metode penelitian (Aritonang, 2008).

b. Reliabilitas

Reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Pada uji reliabilitas dihitung menggunakan statistik *Alpha Cronbach's* dengan menggunakan SPSS 16. Nilai reliabilitas dibandingkan dengan Tabel 2 nilai *Alpha Cronbach's* pada Tabel 2.

Tabel 2. Nilai Alpha Cronbach's

Tingkat Reliabilitas	Kriteria
<0,5	Buruk
0,5–0,6	Jelek
0,6–0,7	Cukup
0,7–0,8	Bagus
>0,8	Bagus sekali

(Sumintono & Widhiarso, 2014)

Untuk menghitung tingkat kesukaran tes bentuk uraian menurut Arifin (2016). Skor yang diperoleh kemudian diinterpretasikan pada Tabel 3.

$$\text{Rata-rata} = \frac{\text{jumlah skor peserta didik tiap soal}}{\text{jumlah peserta didik}}$$

$$\text{Tingkat kesukaran} = \frac{\text{Rata-Rata}}{\text{skor maksimum tiap soal}}$$

Tabel 3. Kriteria Tingkat Kesukaran

Skor	Keterangan
0,00–0,30	Sukar
0,31–0,70	Sedang
0,71–1,00	Mudah

(Arifin, 2016)

c. Uji Daya Beda

Menurut Arifin (2012: 146), untuk soal bentuk uraian, teknik yang digunakan untuk menghitung daya pembeda yaitu:

$\bar{X}_{KA} - \bar{X}_{KB}$	Skor Maks
Keterangan:	
DP	= daya pembeda
\bar{X}_{KA}	= rata-rata dari kelompok atas
\bar{X}_{KB}	= rata-rata dari kelompok bawah
Skor Maks	= skor maksimum

Tabel 4. Kriteria Daya Pembeda

Skor	Keterangan
0,00–0,19	Jelek (<i>Poor</i>)
0,20–0,39	Cukup (<i>Satisfactory</i>)
0,40–0,69	Baik (<i>Good</i>)
0,70–1,00	Baik Sekali (<i>Excellent</i>)
Tanda Negatif (-)	Tidak Baik

(Arikunto, 2018)

HASIL DAN PEMBAHASAN

A. Hasil Validasi Dosen Ahli

Pengembangan instrumen soal divalidasi ke ahli materi dan KBK (Keterampilan Berpikir Kritis) berupa kisi-kisi atau matriks penilaian, rubrik penilaian, naskah soal, dan lembar penilaian untuk mengetahui kelayakan soal. Hasil validasi ke dosen para ahli dibagi menjadi tiga yaitu valid tanpa revisi, valid dengan revisi, dan tidak valid. Ketika soal yang telah divalidasi dosen ahli, maka soal tersebut terdapat beberapa saran dan masukan untuk memperoleh produk soal yang baik. Berikut ini merupakan hasil validasi dari dosen para ahli meliputi saran dan masukan dari dosen ahli dapat dilihat pada Tabel 5.

Tabel 5. Saran dan Masukan dari Dosen Ahli

No.	Saran dan Masukan dari Dosen Para Ahli	
	Ahli KBK	Ahli Materi
1.	Terlalu banyak soal yang diuji cobakan	Bahasa tidak komunikatif
2.	Sesuaikan soal dengan indikator	Terjadinya miskonsepsi
3.	Pertanyaan tidak menggunakan kata-kata perintah atau kata Tanya	Pertanyaan tidak menggunakan kata-kata perintah atau kata Tanya
4.	Jawaban tidak tersirat pada stimulus	Secara umum matriks penilaian sudah memenuhi skala penilaian

Berdasarkan Tabel 5. Menunjukkan bahwa hasil validasi oleh dosen para ahli yaitu validator ahli KBK (keterampilan berpikir kritis), dan validator ahli materi kelas XI semester genap, bahwa hasil validasi oleh para ahli menunjukkan dari 40 soal *essay* yang divalidasi oleh dosen ahli KBK ada 33 item soal yang valid tanpa revisi, dan 7 item soal yang valid dengan revisi, serta ada beberapa saran dan masukan dari dosen ahli KBK sehingga soal yang valid dengan revisi diperbaiki agar soal tersebut menjadi baik. sedangkan menurut ahli materi ada 36 item soal yang valid tanpa revisi, dan 4 item soal yang valid dengan revisi. Hasil dari validasi dosen ahli terdapat saran dan masukan agar soal tersebut menjadi lebih baik, saran dan masukan dari dosen ahli antara lain menurut ahli evaluasi keterampilan berpikir kritis berupa soal tersebut terlalu banyak untuk diuji cobakan, sesuaikan soal dengan indikator berpikir kritis, pertanyaan tidak menggunakan kata-kata perintah atau kata tanya, dan ada

jawaban tidak tersirat pada stimulus, hal ini sejalan dengan pendapat Septiana (2016), menyatakan bahwa dari segi konstruksi masih ada beberapa soal yang perlu direvisi karena tidak sesuai dengan beberapa aspek penelaahan. Sama halnya pendapat TFT (2016), menyatakan bahwa komentar dan saran yang didapat terletak pada pemberian ilustrasi gambar yang sesuai, menambahkan gambar yang dibutuhkan oleh soal, penggunaan bahasa dan kalimat tiap soal yang perlu diperjelas dan kesesuaian indikator dengan butir soal. sedangkan menurut ahli materi adalah bahasa tidak komunikatif, terjadinya miskonsepsi, pertanyaan tidak menggunakan kata-kata perintah atau kata tanya, tetapi secara umum matriks penilaian sudah memenuhi skala penilaian. Setelah itu, penilaian validasi produk soal oleh dosen para ahli maka dapat dilihat dari rekapitulasi hasil validasi oleh dosen para ahli pada Tabel 6.

Tabel 6. Rekapitulasi Hasil Validasi Dosen Para Ahli

No.	Validator	Nomor Butir Soal		
		Valid tanpa Revisi	Valid dengan Revisi	Tidak Valid
1.	Ahli Soal KBK	1, 4, 5, 6, 7, 8, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40	2, 3, 9, 10, 11, 27, 28	–
	Jumlah	33 Butir Soal	7 Butir Soal	–
2.	Ahli Materi	1, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40	2, 10, 27, 28	–
	Jumlah	36 Butir Soal	4 Butir Soal	–

Tabel 7. Hasil Validasi Pengembangan Soal Berpikir Kritis Berdasarkan Statistik Aiken

No	Keterangan	No. Item	Jumlah
1.	Sangat Tinggi	1, 4, 5, 6, 7, 8, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40	33
2.	Tinggi	3, 9, 11	3
3.	Cukup	2, 10, 27, 28	4
	Jumlah item		40

Berdasarkan Tabel 6. Menunjukkan bahwa dari 40 soal yang telah divalidasi oleh dosen ahli keterampilan berpikir kritis mendapatkan 33 butir soal yang valid tanpa revisi dan 7 butir soal yang valid dengan revisi, sedangkan dosen ahli materi mendapatkan dari 40 soal ada 36 butir soal yang valid tanpa revisi dan 4 soal yang valid dengan revisi. Rekapitulasi hasil uji statistik Aiken menggunakan *Microsoft Excel* yang dapat dilihat pada Tabel 7.

Selanjutnya berdasarkan Tabel 7. Menunjukkan hasil validasi dosen para ahli dianalisis menggunakan statistik Aiken pada *Microsoft Excel* yang menunjukkan bahwa soal-soal tersebut dalam kategori sangat tinggi. Pada keseluruhan 40 item soal yang di analisis diperoleh hasil data berupa 33 item soal dengan kriteria sangat tinggi dengan nilai statistik 1.00, 3 item soal dengan kriteria tinggi dengan nilai 0.75, dan 4 item soal dengan kriteria cukup dengan nilai 0.50. ini berarti secara keseluruhan soal dalam kriteria sangat tinggi dengan nilai 1.00 dan menunjukkan bahwa secara keseluruhan soal yang telah dianalisis menggunakan statistik Aiken telah valid sesuai dengan yang ingin diukur dan soal-soal tersebut memiliki validitas yang tinggi, sehingga layak untuk diuji cobakan ke siswa, sebab

suatu kevalidan tersebut mengacu pada kelayakan suatu soal agar dapat digunakan dalam proses pembelajaran. Sependapat Azwar (2014), pengukuran dikatakan validitas yang tinggi apabila menghasilkan data yang secara akurat memberikan gambaran mengenai variabel yang diukur seperti yang dikehendaki oleh tujuan pengukuran tersebut. Sama halnya dengan pendapat Retnawati dalam Arifin (2017), instrumen pengukuran dibuktikan valid jika ahli menyakini bahwa instrumen tersebut mengukur penguasaan kemampuan yang didefinisikan dalam domain ataupun juga konstruk psikologi yang diukur. Hasil Uji Coba ke-1 (Melakukan Uji coba pada tes yang telah disusun)

Penelitian pengembangan asesmen Biologi ini menggunakan soal berbentuk *essay/ uraian* yang akan diuji cobakan ke siswa kelas XII dengan jumlah 36 soal berbasis berpikir kritis. Setelah diuji cobakan ke siswa, maka hasil dari uji coba tersebut dilakukan analisis yakni uji validitas, reliabilitas dengan bantuan *SPSS*, uji tingkat kesukaran, dan uji daya pembeda dengan menggunakan *Microsoft Excel*. Adapun hasil dari uji validitas untuk uji coba pertama dapat dilihat pada tingkat signifikan 0.01 atau 0.05 (1 tailed) dan dalam bentuk Tabel 8.

Tabel 8. Hasil Uji Validitas Pada Uji Coba ke-1

Nomor Soal Valid	Jumlah Soal Valid	Nomor Soal Tidak Valid	Jumlah Soal Tidak Valid
1, 2, 4, 5, 8, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36	30	3, 6, 7, 11, 14, 20	6

Berdasarkan Tabel 8. Menunjukkan hasil uji coba pertama kelas XII dengan sampel dua kelas diperoleh 30 soal yang valid dan 6 yang tidak valid dari materi kelas XI semester genap. Soal yang tidak valid disebabkan karena uji daya pembeda terkategori jelek karena uji daya pembeda merupakan uji butir soal yang dapat membedakan siswa berkemampuan yang tinggi dengan siswa yang berkemampuan yang rendah. Sehingga soal-soal yang tidak valid dengan kategori jelek tersebut tidak dapat membedakan kelompok siswa yang berkemampuan tinggi dengan kelompok siswa berkemampuan yang rendah. Sependapat dengan Suwarto dalam Pardimin (2007), yang menyatakan bahwa butir soal merupakan kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa berkemampuan rendah berdasarkan kriteria tertentu. Sependapat dengan Nawawi & Wijayanti (2018), menyatakan bahwa Butir soal dapat dikatakan valid jika dapat membedakan siswa yang mampu dan tidak mampu. Jika butir soal tidak

dapat membedakan siswa yang mampu dan tidak mampu maka butir soal perlu direvisi atau dibuang.

Sama halnya dengan pendapat Pardimin (2007), menyatakan bahwa besarnya indeks daya pembeda (DB) antara -1,00 sampai dengan 1,00. Semakin kecil indeks daya pembeda menunjukkan bahwa butir soal semakin jelek, untuk membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah, begitu juga sebaliknya jika diperoleh indeks daya pembeda semakin besar maka butir soal semakin baik untuk membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Selain itu, soal yang tidak valid yang berjumlah 6 soal tersebut berarti soal tidak dapat mengukur kompetensi yang diharapkan, sehingga soal yang tidak valid perlu direvisi atau diperbaiki agar menjadi soal yang baik. Selanjutnya data dianalisis nilai reliabilitasnya, uji reliabilitas ini merupakan sejauhmana hasil suatu proses pengukuran dapat konsisten atau dapat dipercaya dari pengembangan ini, hasil reliabilitas dapat dilihat pada Tabel 9.

Tabel 9. Hasil Uji Reliabilitas Pada Saat Uji Coba ke-1

Nilai Alpha Cronbach's	Jumlah soal	Kriteria
0,841	36	Bagus Sekali

Berdasarkan Tabel 9. bahwa butir soal keterampilan berpikir kritis

yang diuji cobakan di SMA Muhammadiyah 1 Palembang. Secara

keseluruhan item yang dikembangkan yaitu reliabel dengan reliabel nilai alpha sekitar 0,841 dengan kategori bagus sekali. Setelah itu, data di uji tingkat kesukaran untuk mengetahui butir soal tersebut apakah terlalu sukar atau terlalu mudah, karena soal yang baik adalah soal yang tidak terlalu sukar dan tidak terlalu mudah. Pada uji coba ke-1 nilai rata-rata reliabilitasnya adalah 0,841 dengan kategori bagus sekali. Ini berarti soal-soal tersebut dapat dipercaya dalam mengukur keterampilan berpikir kritis dan juga soal-soal tersebut konsisten dalam penilaian serta memiliki keterandalan yang sangat tinggi, sehingga soal tersebut memiliki ketepatan dalam hasil pengukuran dan

kestabilan. Sejalan dengan pendapat Reynold dalam Arifin (2017), menyatakan bahwa reliabilitas mengacu pada kekonsistenan atau kestabilan hasil penilaian. Sama halnya pendapat Syahriandi (2015), menyatakan bahwa reliabilitas berarti keterandalan, artinya suatu tes memiliki keterandalan, jika tes tersebut memperoleh hasil yang sama ketika digunakan untuk mengukur berulang-ulang dan keterandalan yang dimaksud dalam hal ini meliputi ketepatan hasil pengukuran dan keajegan atau kestabilan dari hasil pengukuran. Hasil uji tingkat kesukaran dengan menggunakan *Microsoft Excel* yang dapat dilihat pada Tabel 10.

Tabel 10. Hasil Uji Tingkat Kesukaran Soal Berbasis Keterampilan Berpikir Kritis Pada Uji Coba ke-1

Tingkat Kesukaran	Nomor Soal	Jumlah Soal
Sukar	10, 15, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36	21
Sedang	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 16, 18	15
Mudah	-	-

Berdasarkan Tabel 10. bahwa ada beberapa kategori tingkat kesukaran pada tahap ujicoba yang dibagi ke dalam tiga kategori yaitu soal yang tergolong mudah, sedang, dan sukar. Butir soal yang tergolong mudah pada ujicoba pertama tidak ada. Butir soal yang tergolong kategori sedang ada 15 item soal yaitu butir soal nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 16, dan 18 dan butir soal yang tergolong sukar ada 21 item soal yaitu butir soal nomor 10, 15, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, dan 36.

Pada uji tingkat kesukaran ke-1 dari 36 soal di dapatlah soal ada

yang sukar dan ada yang sedang. Soal yang sukar berjumlah 21 butir soal, dan soal yang sedang berjumlah 15 butir soal, sehingga soal tersebut masih dalam proporsi soal yang sukar karena menurut Sukiman dalam Amri (2018: 61), bahwa untuk tes sejenis formatif, proporsi tingkat kesukarannya adalah 25% kategori mudah, 50% kategori sedang, dan 25% kategori sukar. Tetapi secara keseluruhan soal masih dalam kategori sukar dengan nilai 0,289 sehingga secara keseluruhan soal, masih sangat sulit dimengerti oleh siswa dan siswa masih kesulitan dalam menginterpretasikan soal-soal tersebut dan ini yang menyebabkan

masih ada soal yang belum valid. Sejalan dengan pendapat Subali (2009), yaitu item soal yang tergolong sukar dapat digunakan sebagai penanda bahwa siswa belum menguasai indikator yang diukur pada item tersebut. Demikian juga dengan item soal yang tergolong mudah digunakan sebagai penanda bahwa siswa telah menguasai apa yang hendak diukur pada indikator. Selain itu, soal yang sukar sulit bagi siswa untuk memahami makna dari soal tersebut, oleh karena itu,

sebaiknya soal yang baik itu adalah soal yang tidak terlalu sukar ataupun soal yang terlalu mudah, dengan kata lain soal yang baik dalam kategori sedang.

Selanjutnya uji daya pembeda, uji imi dilakukan untuk mengetahui kemampuan suatu soal agar dapat membedakan antara siswa yang berkemampuan tinggi dengan siswa berkemampuan rendah. Hasil uji daya pembeda yang dapat dilihat pada Tabel 11.

Tabel 11. Hasil Uji Daya Beda Pada Uji Coba ke-1

Tingkat Daya Pembeda	Nomor Soal	Jumlah Soal
Baik	17	1
Cukup	4, 5, 9, 10, 13, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 36	23
Jelek	1, 2, 3, 6, 7, 8, 11, 12, 14, 20, 31, 35	12

Berdasarkan Tabel 11. bahwa ada beberapa kategori daya pembeda pada tahap ujicoba yang dibagi ke dalam lima kategori yaitu jelek, cukup, baik, baik sekali, dan tidak baik. Butir soal yang memiliki kategori daya beda jelek memiliki daya pembeda pada rentang 0,00–0,20. Butir soal yang memiliki kategori tersebut yaitu nomor 1, 2, 3, 6, 7, 8, 11, 12, 14, 20, 31, dan 35. Daya pembeda jelek berarti kemampuan butir-butir tersebut masih kurang dalam membedakan kemampuan siswa berkemampuan tinggi dengan siswa berkemampuan rendah. Butir soal yang memiliki kategori daya pembeda cukup memiliki nilai daya pembeda pada rentang 0,21–0,40. Butir soal yang cukup artinya kemampuan butir soal tersebut cukup bisa membedakan kemampuan siswa berkemampuan tinggi dengan siswa berkemampuan rendah. Butir soal yang memiliki daya pembeda cukup adalah butir soal

nomor 4, 5, 9, 10, 13, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, dan 36. Butir soal yang memiliki kategori daya pembeda baik memiliki nilai daya pembeda pada rentang 0,41–0,70. Butir soal yang baik artinya kemampuan butir-butir tersebut bisa membedakan kemampuan siswa berkemampuan tinggi dengan siswa berkemampuan rendah. Butir soal yang memiliki daya pembeda baik adalah butir soal nomor 17. Butir soal yang memiliki daya pembeda baik sekali memiliki nilai daya pembeda pada rentang 0,71–1,00. Butir soal yang memiliki kategori tersebut tidak ada. Butir soal yang memiliki daya pembeda tidak baik memiliki nilai daya pembeda pada rentang nilai dengan tanda negatif. Butir soal yang memiliki kategori tersebut tidak ada.

B. Uji Coba Ke-2 (Validasi Soal Tes)

Pada uji coba kedua ini menggunakan dua kelas yaitu XII IPA 1 dan XII IPA 2 sebagai sampel

penelitian pengembangan soal berpikir kritis. Setelah dilakukan uji coba kedua diperoleh hasil jawaban siswa yang akan di analisis uji validitas

menggunakan SPSS. 16. Hasil validitas dari hasil uji coba kedua yang dapat dilihat pada Tabel 12.

Tabel 12. Hasil Uji Validitas Uji Coba ke-2

Nomor Soal Valid	Jumlah Soal Valid	Nomor Soal Tidak Valid	Jumlah Soal Tidak Valid
1, 2, 3, 4, 6, 7, 8, 10, 11,	34	5, 9	2
12, 13, 14, 15, 16, 17,			
18, 19, 20, 21, 22, 23,			
24, 25, 26, 27, 28, 29,			
30, 31, 32, 33, 34, 35, 36			

Berdasarkan Tabel 12. analisis uji validitas pada uji coba soal kedua ini menunjukkan bahwa dari 36 item soal, 34 soal dinyatakan valid dan 2 soal dinyatakan tidak valid. Berdasarkan uji coba pertama dan kedua ini menunjukkan soal yang konsisten valid berjumlah 28 soal yang valid dan hanya 8 yang tidak konsisten valid. Pada ujicoba pertama nomor soal 3, 6, 7, 11, 14, dan 20 tidak valid tetapi pada uji

coba kedua menjadi valid sedangkan pada ujicoba pertama nomor soal 5, dan 9 valid tetapi pada ujicoba kedua menjadi tidak valid sehingga ada 8 soal yang tidak konsisten.

Selanjutnya uji reliabilitas pada uji coba kedua dengan menggunakan sampel dua kelas pada kelas XII yaitu XII IPA 1 dan XII IPA 2, hasil uji reliabilitas untuk uji coba kedua dapat dilihat pada Tabel 13.

Tabel 13. Uji Reliabilitas Pada Uji Coba ke-2

Nilai Alpha Cronbach's	Jumlah soal	Kriteria
0,836	36	Bagus Sekali

Tabel 14. Uji Tingkat Kesukaran Pada Uji Coba ke-2

Tingkat Kesukaran	Nomor Soal	Jumlah Soal
Sukar	25, 28, 31, 32, 33, 34, 35, 36	8
Sedang	2, 3, 4, 5, 7, 8, 11, 12, 14, 15, 16, 19, 20, 21, 22, 23, 24, 26, 27, 29, 30	21
Mudah	1, 6, 9, 10, 13, 17, 18	7

Berdasarkan hasil penelitian di SMA Muhammadiyah 1 Palembang pada uji coba ke dua secara keseluruhan bahwa nilainya reliabel yaitu dengan nilai alpha sekitar 0,836 dengan kategori bagus sekali. Selanjutnya dilakukan uji tingkat kesukaran pada uji coba kedua yang dapat dilihat pada Tabel 14.

Berdasarkan hasil Tabel 14. bahwa ada beberapa kategori tingkat kesukaran pada tahap ujicoba yang dibagi ke dalam tiga kategori yaitu soal yang tergolong mudah, sedang, dan sukar. Kategori mudah memiliki nilai tingkat kesukaran 0,71–1,00. Butir soal yang tergolong kategori mudah adalah butir soal nomor 1, 6, 9, 10, 13, 17, dan 18. Kategori sedang memiliki nilai tingkat kesukaran 0,31–0,70. Butir soal yang tergolong sedang adalah butir soal

nomor 2, 3, 4, 5, 7, 8, 11, 12, 14, 15, 16, 19, 20, 21, 22, 23, 24, 26, 27, 29, dan 30. Kategori sukar memiliki nilai tingkat kesukaran 0,00–0,30. Butir soal yang tergolong kategori sukar adalah butir soal nomor 25, 28, 31, 32, 33, 34, 35, dan 36. Sesuai kriteria kualitas instrumen tes, terdapat butir soal instrumen tes kemampuan berpikir kritis yang dinyatakan memiliki tingkat kesukaran tidak baik apabila terlalu mudah dan terlalu sukar. Semakin banyak yang menjawab suatu soal, maka soal tersebut semakin dikategorikan mudah, sebaliknya semakin sedikit suatu soal dijawab, maka soal tersebut semakin dikategorikan sukar. Kemudian analisis uji daya pembeda pada uji coba kedua yang dapat dilihat pada Tabel 15.

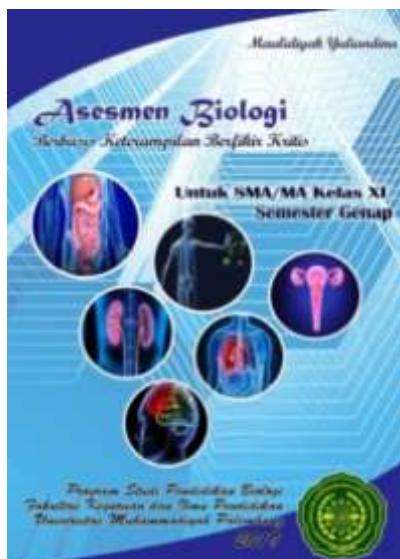
Tabel 15. Hasil Uji Daya Beda Pada Uji Coba ke-2

Tingkat Daya Pembeda	Nomor Soal	Jumlah Soal
Baik	15, 25, 26, 27, 29, 30	6
Cukup	3, 10, 11, 17, 20, 21, 22, 23, 24, 28, 33, 34, 35, 36	14
Jelek	1, 2, 4, 5, 6, 7, 8, 9, 12, 13, 14, 16, 18, 19, 31, 32	16

Berdasarkan Tabel 15. diatas bahwa ada beberapa kategori daya pembeda pada tahap ujicoba yang dibagi ke dalam lima kategori yaitu jelek, cukup, baik, baik sekali, dan tidak baik. Butir soal yang memiliki kategori tersebut yaitu nomor 1, 2, 4, 5, 6, 7, 8, 9, 12, 13, 14, 16, 18, 19, 31, dan 32. Butir soal yang memiliki daya pembeda cukup adalah butir soal nomor 3, 10, 11, 17, 20, 21, 22, 23, 24, 28, 33, 34, 35, dan 36. Butir soal yang memiliki daya pembeda baik adalah butir soal nomor 15, 25, 26, 27, 29, dan 30. Butir soal yang memiliki kategori tersebut tidak ada.

Selanjutnya hasil produk akhir dalam pengembangan asesmen Biologi

berbasis keterampilan berpikir kritis siswa kelas XI semester genap di SMA Muhammadiyah 1 palembang bahwa dari 36 soal yang telah di uji cobakan pada uji coba ke-1 dan uji coba ke-2 menunjukkan hasil akhir soal yang akan dijadikan produk akhir adalah berjumlah 34 soal dari 36 soal pada materi kelas XI semester genap. Hasil produk akhir ini berupa cover soal, kata pengantar, daftar isi, sejarah pengembangan asesmen, latar belakang, petunjuk penggunaan tes, soal-soal serta daftar pustaka. Identitas produk ini dapat dijelaskan dari Tabel 16.



Gambar 1. Cover Depan Produk Soal Berpikir Kritis

Tabel 16. Identitas Produk Tes Keterampilan Berpikir Kritis

No	Desain	Keterangan
1.	Bentuk fisik	Bahan cetak
2.	Judul	Pengembangan Asesmen Biologi Berbasis Keterampilan Berpikir Kritis
3.	Ukuran kertas	A4
4.	Tebal Halaman	56 halaman + cover
5.	Bahasa	Indonesia
6.	Jenis huruf	Times new roman dengan ukuran huruf 12

KESIMPULAN

Hasil pengembangan asesmen Biologi berbasis keterampilan berpikir kritis berbentuk uraian/ *essay* dengan materi kelas XI semester genap dengan melakukan uji coba sebanyak dua kali diperoleh hasil yaitu pada uji coba ke-1 dan uji coba ke-2 diperoleh 34 soal yang valid untuk dijadikan produk akhir dalam pengembangan asesmen berbasis keterampilan berpikir kritis.

Kelayakan hasil pengembangan asesmen Biologi yang dikembangkan berdasarkan validator ahli dan validasi empiris butir soal yaitu Hasil validator

dosen ahli terkualifikasi cukup dan sangat tinggi dan berdasarkan validasi empiris butir soal dinyatakan valid dan reliabel. Pada uji kesukaran soal pada uji coba ke-1 rata-rata soal tergolong sukar dan sedang. Pada uji daya pembeda untuk uji coba ke-1 dan ke-2 terkategori cukup sehingga dapat digunakan untuk mengukur keterampilan berpikir kritis siswa.

SARAN

Berdasarkan hasil dari penelitian ini, dapat dikemukakan bahwa dalam beberapa saran yaitu: mengembangkan dan mengoptimalkan keterampilan berpikir kritis disarankan kepada guru agar siswa lebih dibiasakan dalam mengerjakan soal-soal yang mengasah keterampilan berpikir kritis siswa. Pada peneliti selanjutnya sebaiknya jika menggunakan soal uraian atau *Essay* jangan terlalu banyak jumlah soalnya, sebaiknya sekitar 1-10 soal uraian saja. Pada peneliti selanjutnya dipikirkan lagi waktu dalam mengerjakan soal uraian.

DAFTAR RUJUKAN

- Amri, S. 2018. Pengembangan Tes Sumatif Mata Pelajaran Biologi Semester Ganjil Kelas IX. *Jurnal Biotek*, 6(1): 53-64.
- Arifin, Z. 2012. *Evaluasi Pembelajaran*. Jakarta Pusat: Subdit Kelembagaan Direktorat Pendidikan Tinggi Islam.
- Arifin, Z. 2016. *Evaluasi Pembelajaran: Prinsip, Teknik, dan Prosedur*. Bandung: PT. Remaja Rosdakarya.
- Arifin, Z. 2017. Kriteria Instrumen dalam Suatu Penelitian. *Jurnal Theorems (The Original Research of Mathematics)*, 2(1): 28-36.
- Arikunto, S. 2018. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

- Aritonang, L. R. 2008. Validitas dan Reliabilitas Butir Instrumen. *Jurnal Pendidikan Universitas Tarumanagara Akademika*, 10(2): 159-180.
- Azwar, S. 2014. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Belajar.
- Azwar, S. 2017. *Metode Penelitian Psikologi*. Yogyakarta: Pustaka Pelajar.
- Madliya, S., F. Abdurachman, & Hartono. 2017. Pengembangan Instrumen Penilaian Keterampilan Proses Sains Dasar Mata Pelajaran Kimia Pada Kompetisi Dasar Kelarutan dan Hasil Kali Kelarutan di SMA. *Jurnal Prosiding Seminar Nasional Pendidikan IPA*, Halaman 327-337.
- Mulyatiningsih, E. 2011. *Riset Terapan Bidang Pendidikan dan Teknik*. Yogyakarta: UNY Press.
- Nawawi, S., & T. F. Wijayanti. 2018. Pengembangan Asesmen Biologi Berbasis Keterampilan Berpikir Kritis Terintegrasi Nilai Islam. *Jurnal Inovasi Pendidikan IPA*, 4(2): 136-148.
- Pardimin, W. S. 2007. Analisis Butir Soal Tes Pemecahan Masalah Matematika. *Jurnal Wacana Akademika*, 1(1): 69-76.
- Septiana, N. 2016. Analisis Butir Soal Ulangaan Akhir Semester (UAS) Biologi Tahun Pelajaran 2015/2016 Kelas X dan XI Pada MAN Sampit. *Jurnal Edusains*, 4(2): 115-121.
- Subali, B. 2009. Pengembangan Tes Pengukur Keterampilan Proses Sains Pola Divergen Mata Pelajaran Biologi SMA. *Jurnal Jurdik Biologi*, Halaman: 581-593.
- Sumintono, B., & W. Widhiarso. 2014. *Aplikasi Model RASCH untuk Penelitian Ilmu-Ilmu Sosial*. Cimahi: Trim Komunikata Publishing House.
- Syahriandi. 2015. Validitas dan Reliabilitas Butir Soal Ujian Semester Bahasa Indonesia. 3(1): 67-78.
- TFT, N. A. 2016. Pengembangan Instrumen Asesmen Penguasaan Konsep Tes Testlet Pada Materi Suhu dan Kalor. *Jurnal Pendidikan*, 1(6): 1197-1203.
- Widodo, H. 2015. Potret Pendidikan di Indonesia dan Kesiapannya dalam Menghadapi Masyarakat Ekonomi Asia (MEA). *Jurnal Cendikia*, 13 (2): 293-307.