



Pengembangan asesmen biologi berbasis keterampilan berpikir kritis terintegrasi nilai Islam

Sulton Nawawi *, Tutik Fitri Wijayanti

Program Studi Pendidikan Biologi, Universitas Muhammadiyah Palembang.
Jalan Jenderal Ahmad Yani, 13 Ulu, Seberang Ulu II, Kota Palembang, 30116, Indonesia

* Coresponding Author. E-mail: sulton_nawawi@um-palembang.ac.id

Received: 23 November 2017; Revised: 14 September 2018; Accepted: 18 September 2018

Abstrak

Penelitian ini bertujuan untuk mengetahui karakteristik dan kelayakan asesmen berbasis berpikir kritis terintegrasi nilai Islam. Penelitian ini merupakan penelitian dan pengembangan. Pengembangan yang dilakukan menggunakan model prosedural dengan mengadaptasi model penelitian dan pengembangan Mcintire dan Miller. Validasi ahli dilakukan oleh 3 validator ahli: (1) ahli materi; (2) ahli evaluasi pendidikan; dan (3) ahli nilai Islam. Uji coba pertama dilakukan pada siswa kelas XI SMA Nadhatul Ulama dan SMA Muhammadiyah 4 Palembang sebanyak 70 siswa. Uji validasi tes dilakukan pada siswa Kelas X SMA Daarul Aitam, SMA Nadhatul Ulama, SMA Azhariyah, SMA Muhammadiyah 3 Palembang, SMA Muhammadiyah 4 Palembang, SMA Muhammadiyah 9 Palembang, SMA YP Nurfauzan, dan SMA Ashanadiyah. Hasil penelitian ini berupa asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam berbentuk soal pilihan ganda yang di dalamnya terdapat materi biologi kelas X semester 1 dan 2 yang dintegrasikan dengan indikator berpikir kritis serta nilai Islam. Kelayakan asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam berdasarkan validator ahli hasilnya berkualifikasi baik sampai dengan sangat baik serta berdasarkan hasil validitas empirik butir soal dinyatakan valid dan reliabel sehingga dapat digunakan untuk mengukur keterampilan berpikir kritis terintegrasi nilai Islam pada siswa.

Kata Kunci: asesmen, berpikir kritis, biologi, nilai Islam.

Development of biological assessment based critical thinking skills integrated with Islamic Values

Abstract

This study aims to determine the characteristics and feasibility of assessment-based integrated critical thinking of Islamic values. This research is research and development. The development was carried out using a procedural model by adapting the research and development model of McIntire and Miller. Expert validation is carried out by 3 expert validators: (1) material experts; (2) education evaluation experts; and (3) Islamic values experts. The first trial was conducted on students of class XI Nadhatul Ulama High School and Muhammadiyah 4 Palembang High School as many as 70 students. Test validation tests were carried out on students of Class X Daarul Aitam High School, Nadhatul Ulama High School, Azhariyah High School, Muhammadiyah 3 Palembang High School, Muhammadiyah 4 Palembang High School, Muhammadiyah 9 Palembang High School, YP Nurfauzan High School, and Ashanadiyah High School. The results of this study in the form of a biological assessment based on integrated critical thinking in Islamic values in the form of multiple choice questions in which there are biological materials of class X semester 1 & 2 that are integrated with indicators of critical thinking and Islamic values. Feasibility of biology assessment based on integrated critical thinking Islamic values based on validator experts results in good quality to very good and based on the results of empirical validity items are declared valid and reliable so that they can be used to measure integrated critical thinking skills of Islamic values in students.

Keywords: *assessment, critical thinking, biology, the value of Islam.*

How to Cite: Nawawi, S., & Wijayanti, T. (2018). Pengembangan asesmen biologi berbasis keterampilan berpikir kritis terintegrasi nilai Islam. *Jurnal Inovasi Pendidikan IPA*, 4(2), 136-148. doi:<https://doi.org/10.21831/jipi.v4i2.21265>



<https://doi.org/10.21831/jipi.v4i2.21265>

PENDAHULUAN

Kemajuan teknologi di abad ke-21, menuntut setiap orang untuk memiliki keterampilan yang begitu kompleks, seperti mampu mengembangkan atau mempertahankan keterampilan dan kecakapan dalam bekerja untuk mengatasi kompetisi di berbagai negara (Tremblay, Lalancette, & Roseveare, 2012). Untuk mengatasi hal tersebut, tentunya setiap orang dituntut untuk memiliki keterampilan berpikir tingkat tinggi. Salah satu aspek yang terdapat di berpikir tingkat tinggi adalah berpikir kritis.

Berpikir kritis merupakan cara berpikir tentang subjek, konten, atau masalah yang mampu meningkatkan kualitas pemikirannya dengan terampil. Berpikir kritis juga digunakan untuk memilih dan menganalisis antara yang benar dan yang salah (Wood, 2002). Menurut Kementerian Pendidikan dan Kebudayaan (2016) pada kehidupan abad 21 siswa hendaknya dapat memenuhi tuntutan hidup yang salah satunya adalah berpikir kritis. Berpikir kritis penting dilatihkan bagi siswa melalui kegiatan pembelajaran agar siswa memiliki bekal dalam memutuskan gagasan di kehidupan mendatang.

Memilih dan memutuskan gagasan di abad 21 tidak hanya harus memiliki keterampilan berpikir kritis, tetapi juga harus memiliki landasan etika dan norma. Etika dan norma selain diatur dalam suatu negara, juga diatur oleh Agama. Di dalam agama Islam pada khususnya, memiliki nilai-nilai Islam yang bersumber dari Al-Quran dan Hadist yang di dalamnya banyak pedoman hidup bagi manusia (Pimpinan Pusat Muhammadiyah, 2000). Pengambilan keputusan hidup yang berlandaskan nilai Islam akan menjadi lebih baik, mengingat di abad 21 banyak terdapat pengaruh positif dan negatif dalam kehidupan (Dacholfany, 2015).

Mengingat pentingnya keterampilan berpikir kritis yang harus dilandasi nilai Islam di kehidupan abad 21, siswa muslim hendaknya mengembangkan dan meningkatkan kemampuannya dalam menghadapi tantangan yang semakin kompleks. Hal ini dapat dilatihkan melalui kegiatan pembelajaran, khususnya di pembelajaran Biologi. Membelajarkan biologi yang terintegrasi nilai Islam akan membuat siswa semakin menyadari akan keagungan Allah SWT, menambah iman dan ketaqwaan.

Berdasarkan hasil observasi awal di delapan sekolah berbasis Islam wilayah seberang Ulu Palembang, masih jarang mengintegrasikan nilai Islam ke dalam materi pembelajaran bio-

logi. Hal ini wajar terjadi, karena pada dasarnya di dalam kurikulum belum mengatur tentang materi yang diintegrasikan dengan nilai-nilai agama secara langsung. Begitu pula dengan soal UAS ganjil yang digunakan pada 8 SMA berbasis Islam di wilayah Seberang Ulu Palembang masih belum memaksimalkan aspek berpikir kritis yaitu interpretasi sebesar 7,02%; analisis sebesar 1,89%; evaluasi 0%; inferensi 5,52%; penjelasan sebesar 11,37%; dan pengaturan diri sebesar 1,75%, serta belum mengintegrasikan nilai Islam (0%). Hal ini terjadi karena: (1) guru menganggap soal yang berbasis nilai Islam difokuskan pada mata pelajaran Agama, (2) salah satu guru mengaku baru terfikirkan mengenai soal yang terintegrasi nilai Islam meskipun dalam proses pembelajaran sudah mulai mengintegrasikan nilai Islam, (3) membuat soal disesuaikan dengan kemampuan siswa, (3) lebih ditekankan pada proses pembelajaran (khususnya aspek evaluasi) namun tidak diberikan saat ujian, (4) banyak mengambil soal dari buku-buku paket yang digunakan, dan (5) lebih mengutamakan aspek kognitif.

Bila melihat aspek kognitif (terdapat 6 aspek yaitu mengingat, mengetahui, mengaplikasi, menganalisis, mengevaluasi, dan mencipta) dan melihat aspek berpikir kritis (terdapat aspek eksplanasi, interpretasi, analisis, evaluasi, inferensi, dan regulasi diri), keduanya saling berkaitan. Siswa yang memiliki kemampuan di hasil belajar kognitif, belum tentu dapat memaksimalkan kemampuan berpikir kritisnya, namun jika siswa memiliki kemampuan berpikir kritis yang bagus akan memiliki kemampuan hasil belajar yang bagus pula. Pernyataan ini diperkuat oleh Surayya, Subagia, & Tika (2014) yang menyatakan bahwa “semakin tinggi keterampilan berpikir kritis yang dimiliki oleh siswa akan memperkuat model pembelajaran yang diterapkan dalam rangka meningkatkan hasil belajar.”

Itu sebabnya diperlukan asesmen yang mampu mengukur kemampuan berpikir kritis terintegrasi nilai Al-Islam. (Wiliam, 2013) menjelaskan bahwasannya asesmen merupakan jembatan antara mengajar dengan pembelajaran. Dalam kegiatan mengajar, terdapat indikator yang akan digunakan sebagai acuan dalam memberikan materi pembelajaran. Setiap siswa yang belajar, belum tentu akan paham dengan semua materi yang diperoleh meskipun telah belajar berjam-jam di kelas. Tujuan asesmen inilah yang akan digunakan sebagai alat untuk mengungkap proses dan kemajuan belajar.

Ketika hasil asesmen siswa masih tergolong rendah, maka di sinilah letak kegunaan asesmen untuk mengetahui apa yang sudah diperoleh siswa selama belajar. Hal ini terjadi karena asesmen dapat memberikan umpan balik secara berkesinambungan untuk memperbaiki mutu pembelajaran (Syaban, 2006; Wulan, 2007).

METODE

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development*). Produk yang akan dikembangkan adalah asesmen berbasis berpikir kritis terintegrasi nilai Islam. Pengembangan yang dilakukan menggunakan model prosedural dengan mengadaptasi model penelitian dan pengembangan (McIntire & Miller, 2007).

Terdapat 10 langkah pengembangan tes, langkah-langkah pengembangan asesmen berbasis berpikir kritis terintegrasi nilai Islam sebagai berikut: Langkah pertama yaitu mendefinisikan kompetensi, peserta tes dan tujuan tes (*Defining the test universe, audience and purpose*). Kompetensi yang digunakan meliputi Kompetensi Dasar (KD) materi Biologi kelas X semester 1 dan semester 2 dari Permendikbud No. 24 Tahun 2016. Ada 11 Kompetensi Dasar yang digunakan pada penelitian ini, dimulai dari KD 3.1 sampai KD 3.11. KD digunakan sebagai penentuan materi yang digunakan untuk membuat asesmen yang dikaitkan dengan indikator keterampilan berpikir kritis serta diintegrasikan dengan nilai Islam.

Peserta tes ini adalah siswa kelas XI SMA Nadhatul Ulama dan SMA Muhammadiyah 4 Palembang untuk tahap *Conduct Piloting Test*. Kelas X SMA SMA Daarul Aitam, SMA Nadhatul Ulama, SMA Azhariyah, SMA Muhammadiyah 3 Palembang, SMA Muhammadiyah 4 Palembang, SMA Muhammadiyah 9 Palembang, SMA YP Nurfauzan, dan SMA Ashanadiyah untuk tahap *Validation the Test*. Tujuan tes ini adalah untuk mengukur dan mengetahui tingkat keterampilan berpikir kritis dan nilai Islam siswa pada sekolah wilayah Seberang Ulu kota Palembang.

Langkah kedua yaitu mengembangkan rencana uji (*Developing a test plan*). Meliputi pembuatan kisi-kisi, format pertanyaan atau jawaban, bentuk penyelenggaraan dan cara penskorannya.

Kisi-kisi dibuat dari Kompetensi Dasar (KD) materi Biologi kelas X semester 1 dan semester 2 dari Permendikbud No. 24 Tahun 2016. Materi dikaitkan dengan indikator kete-

rampilan berpikir kritis serta diintegrasikan dengan nilai Islam. Enam indikator keterampilan berpikir kritis yang digunakan yaitu interpretasi, analisis, evaluasi, kesimpulan, penjelasan, dan pengaturan diri (Facione, 2013). Nilai Islam yang digunakan bersumber dari Al-Qur'an dan Hadist.

Instrumen test berupa soal pilihan ganda pada dengan lima pilihan jawaban (A, B, C, D, E). Cara penskoran tes ini adalah:

$$\text{Skor} = \left(\frac{B}{N}\right) \times 100$$

Keterangan:

B: Banyaknya butir soal yang dijawab benar.

N: Banyaknya butir soal (Mardapi, 2008).

Setelah didapatkan hasil dengan menggunakan rumus tersebut, selanjutnya menentukan kriteria penilaian dengan menggunakan skala huruf pada Tabel 1.

Tabel 1. Kriteria Penilaian

Nilai	Huruf	Keterangan
80-100	A	Baik sekali
66-79	B	Baik
56-65	C	Cukup
40-55	D	Kurang
30-39	E	Gagal

Sumber: (Arikunto, 2016)

Langkah ketiga yaitu menulis item tes (*Composing the test item*). Soal yang akan dikembangkan berjumlah 60 butir soal pilihan ganda. Tahap selanjutnya yaitu validasi ahli (*expert judgment*), validasi dilakukan oleh 3 validator ahli; (1) ahli materi; (2) ahli evaluasi pendidikan; dan (3) ahli nilai Islam. Validasi digunakan untuk menguji aspek materi, evaluasinya dan nilai Islam. Instrumen lembar validasi yang digunakan berupa angket, yang mencakup aspek materi/substansi, aspek konstruksi, aspek bahasa, aspek berpikir kritis dan aspek nilai Islam. Pedoman penskoran dapat dilihat pada Tabel 2.

Tabel 2. Pedoman penskoran pada angket: (Riduwan, 2009)

No.	Analisis Kuantitatif	Skor
1.	Sangat setuju	4
2.	Setuju	3
3.	Tidak setuju	2
4.	Sangat tidak setuju	1

Langkah keempat yaitu Menulis instruksi administrasi (*Writing the administration instructions*). Pada langkah ini disusun petunjuk penyelenggaraan tes yang terdiri dari petunjuk

untuk penyelenggara, pengawas ujian dan petunjuk untuk peserta tes itu sendiri. Petunjuk penyelenggaraan yang digunakan dalam penelitian ini berisikan kondisi yang harus disiapkan pada saat tes dilaksanakan.

Langkah kelima yaitu Melakukan uji coba (*Conduct piloting test*). Pada langkah ini perangkat tes yang telah disusun kemudian diuji coba untuk memperoleh data empiris yang berguna pada pengujian kualitas butir tes. Uji coba ini dilakukan kepada siswa kelas XI SMA Nadhatul Ulama dan SMA Muhammadiyah 4 Palembang sebanyak 70 siswa. Uji coba dilakukan dengan memberikan produk berupa 60 soal pilihan ganda.

Langkah keenam yaitu Analisis item (*Item analysis*). Setelah melakukan ujicoba maka langkah selanjutnya menganalisis butir soal, validitas, dan reliabilitas menggunakan model Rasch berbantuan program Winstep. Sumintono & Widhiarso (2015) menjelaskan butir soal valid ditentukan menggunakan nilai *outfit mean square* (MNSQ), *outfit Z-standard* (ZSTD), dan *point measure correlation* (Pt Mean Score) pada *output Item Fit Order*, yang memenuhi 2-3 kriteria tersebut. Ketentuan yang digunakan adalah sebagai berikut:

- a. Nilai Outfit MNSQ yang diterima:
 $0,5 < \text{MNSQ} < 1,5$.
- b. Nilai Outfit ZSTD yang diterima:
 $-2,0 < \text{ZSTD} < +2,0$.
- c. Nilai Pt Mean Corr yang diterima:
 $0,4 < \text{Pt Mean Corr} < 0,85$.

Begitu pula untuk melihat kesesuaian pola respon siswa dapat dilihat menggunakan nilai logit pada *output Person Fit Order* (ketentuan yang digunakan sama seperti melihat kesesuaian butir soal/butir soal yang valid), dan lebih lanjut secara detail dapat dilihat melalui *Scalogram* mengenai persebaran jawaban setiap siswa/responden.

Untuk melihat tingkat kesulitan butir soal dapat dilihat pada *output Item Measure* menggunakan nilai logit, sedangkan untuk melihat tingkat abilitas siswa dapat dilihat melalui *output Person Measure* atau dapat juga keduanya (tingkat kesukaran butir soal dan abilitas siswa) dilihat melalui *Variable Map*. Nilai *mean item measure* selalu menunjukkan hasil 0,00 sedangkan untuk nilai *mean person measure* dapat bernilai negatif ataupun positif. Jika nilai *mean person measure* negatif menandakan kecenderungan abilitas siswa lebih kecil daripada tingkat kesulitan soal, begitu pula sebaliknya.

Pengelompokan soal dan responden dapat diketahui melalui nilai *separation*. Semakin besar nilai *separation*, maka kualitas instrumen dalam hal keseluruhan responden dan butir soal semakin bagus, karena dapat digunakan untuk mengidentifikasi kelompok soal dan kelompok responden. Persamaan yang digunakan untuk melihat pengelompokan secara lebih teliti dapat dilihat melalui perhitungan pemisahan strata (H):

$$H = \frac{[(4 \times \text{separation}) + 1]}{3}$$

Dalam model Rasch juga terdapat nilai Alpha Cronbach yang digunakan untuk mengukur reliabilitas, yaitu interaksi antara responden dengan butir soal secara keseluruhan. Kriteria yang digunakan adalah: < 0,5: buruk, 0,5-0,6: jelek, 0,6-0,7: cukup, 0,7-0,8: bagus, dan > 0,8: bagus sekali.

Selain itu terdapat pula nilai *Person Reliability* untuk melihat konsistensi jawaban siswa dan *Item Reliability* untuk melihat kualitas butir soal. Ketentuan yang digunakan pada *Person Reliability* dan *Item Reliability* adalah sebagai berikut: < 0,67: lemah, 0,67-0,80: cukup, 0,8-0,90: bagus, 0,91-0,94: bagus sekali, dan > 0,94: istimewa.

Data lain yang dapat digunakan adalah nilai ideal pada INFIT MNSQ dan OUTFIT MNSQ mendekati angka 1,00. Nilai ideal pada INFIT ZSTD dan OUTFIT ZSTD mendekati angka 0,00. Semua kriteria ini berlaku untuk tabel responden maupun butir soal yang tercantum pada *output Summary Statistics*.

Langkah ketujuh yaitu merevisi tes (*Revising the test*). Hasil analisis butir untuk merevisi butir yang kurang baik. Pengambilan keputusan terhadap butir-butir yang perlu direvisi dilakukan dengan menggunakan beberapa pertimbangan hasil analisis model Rasch. Soal yang dianggap valid akan diambil, sedangkan soal yang tidak valid akan dibuang. Selanjutnya mengecek kembali soal-soal yang sudah valid terkait kalimat-kalimat dan bahasa yang masih belum tepat.

Langkah kedelapan yaitu validasi tes (*Validation the test*). Pada tahap validasi tes ini dilakukan pada siswa Kelas X SMA Daarul Aitam, SMA Nadhatul Ulama, SMA Azhariyah, SMA Muhammadiyah 3 Palembang, SMA Muhammadiyah 4 Palembang, SMA Muhammadiyah 9 Palembang, SMA YP Nurfauzan, dan SMA Ashanadiyah. Hasil yang diperoleh dianalisis kembali menggunakan mo-

del Rasch berbantuan program Winstep, untuk menghitung validitas dan reliabilitas.

Langkah kesembilan yaitu mengembangkan norma (*Developing norms*). Pengembang tes dapat membuat dan menetapkan norma acuan dari distribusi skor tes untuk menginterpretasikan posisi skor tes individu dibandingkan dengan skor tes peserta tes yang lain. Selain itu, pengembang tes juga dituntut untuk menetapkan skor potong yaitu batas skor kelulusan yang digunakan untuk menetapkan keputusan seseorang termasuk dalam kategori kelompok peserta yang lolos atau gagal.

Langkah kesepuluh yaitu melengkapi tes manual (*Complete test manual*). Akhir dari kegiatan pengembangan tes adalah menyusun buku petunjuk penggunaan tes (manual tes). Isi buku petunjuk menjelaskan latar belakang pembuatan tes, sejarah proses pengembangan, hasil-hasil studi validasi, deskripsi target sasaran yang sesuai, petunjuk penyelenggaraan, cara penyekoran tes, dan informasi tentang cara menginterpretasikan skor individu.

Teknik pengumpulan data yang digunakan yaitu menggunakan wawancara, observasi, angket (kuesioner), dan tes tertulis. Wawancara merupakan metode pengumpulan data yang menghendaki komunikasi langsung antara Penyidik dengan subjek atau responden. Pada penelitian ini, wawancara dilakukan dengan melakukan tanya jawab kepada guru dan siswa pada saat analisis kebutuhan awal, dimana jawaban tersebut akan dicatat secara lengkap agar diperoleh data yang benar.

Observasi merupakan metode pengumpulan data yang menggunakan pengamatan terhadap objek penelitian. Observasi yang dilakukan merupakan observasi langsung pada saat analisis kebutuhan awal, dimana observasi dilakukan terhadap proses yang terjadi dalam situasi yang sebenarnya dan langsung diamati tanpa adanya instrumen pengamatan.

Kuesioner atau angket merupakan teknik pengumpulan data dengan memberikan seperangkat pernyataan tertulis kepada responden untuk ditanggapi. Pada penelitian ini, angket digunakan pada saat analisis kebutuhan awal, diberikan kepada guru dan siswa.

Tes tertulis merupakan alat atau prosedur yang digunakan untuk mengetahui atau meng-

ukur sesuatu dalam suasana, dengan cara dan aturan-aturan yang sudah ditentukan. Pada penelitian ini, tes yang digunakan berupa soal pilihan ganda materi biologi kelas X semester 1 dan 2 dengan 5 pilihan jawaban yaitu A, B, C, D dan E.

Teknik analisis data yang digunakan yaitu teknik persentase deskriptif dan menggunakan Model Rasch. Teknik persentase deskriptif digunakan untuk menganalisis hasil observasi awal pada kajian soal dan digunakan untuk menganalisis hasil validator ahli (*expert judgment*). Persentase jawaban dari setiap pernyataan pada instrumen *validasi* dan angket dengan menggunakan rumus sebagai berikut:

$$\text{Persentase Jawaban Responden} = \frac{\text{jumlah skor yang diperoleh}}{\text{jumlah skor tertinggi}} \times 100\%$$

Persentase kelayakan yang didapatkan kemudian diinterpretasikan ke dalam kategori berdasarkan Tabel 3.

Tabel 3. Pengambilan Keputusan Angket Validasi

Tingkat Pencapaian	Kualifikasi	Keterangan
90%-100%	Sangat baik	Tidak perlu direvisi
75%-89%	Baik	Tidak perlu direvisi
65%-74%	Cukup	Direvisi
55%-64%	Kurang baik	Direvisi
0%-54%	Kurang cukup	Direvisi

Sumber: (Suwastono, 2011)

Model Rasch berbantuan program Winstep digunakan untuk menganalisis data hasil *conduct piloting test* dan *validation test* untuk menganalisis validitas empiris pada butir soal/instrumen.

HASIL DAN PEMBAHASAN

Hasil penelitian pengembangan yang telah dilaksanakan berupa asesmen berbasis berpikir kritis terintegrasi nilai Islam. Hal-hal yang direncanakan pada tahap awal setelah menentukan kompetensi, peserta tes dan tujuan tes, yaitu ini membuat kisi-kisi, format pertanyaan atau jawaban, bentuk penyelenggaraan dan cara penyekorannya. Contoh tampilan kisi-kisi soal asesmen berbasis berpikir kritis terintegrasi nilai Islam dapat dilihat pada Tabel 4.

Tabel 4. Contoh Kisi-Kisi Soal Asesmen Berbasis Berpikir Kritis Terintegrasi Nilai Islam.

Kompetensi Dasar	Materi	Indikator Berpikir Kritis	Terintegrasi Nilai Islam	Nomor Soal	Kunci Jawaban
3.1 Memahami melalui penerapan tentang ruang lingkup Biologi (permasalahan pada berbagai objek Biologi dan tingkat organisasi kehidupan), metode ilmiah dan prinsip keselamatan kerja berdasarkan pengamatan dan percobaan	Ruang Lingkup Biologi <ul style="list-style-type: none"> • Permasalahan Biologi pada berbagai objek Biologi, dan tingkat organisasi kehidupan • Cabang-cabang ilmu dalam Biologi dan kaitannya dengan pengembangan karir di masa depan • Manfaat mempelajari Biologi bagi diri sendiri dan lingkungan, serta masa depan peradaban bangsa • Metode Ilmiah • Keselamatan Kerja 	Interpretasi	√	2	A
		Analisis	√	3	C
		Kesimpulan	√	1	B
		Pengaturan Diri	-	22	D
3.2 Menganalisis data hasil observasi tentang berbagai tingkat keanekaragaman hayati (gen, jenis dan ekosistem) di Indonesia serta ancaman dan pelestariannya	Keanekaragaman Hayati <ul style="list-style-type: none"> • Konsep keanekaragaman gen, jenis, ekosistem • Keanekaragaman hayati Indonesia, flora dan fauna, serta penyebarannya berdasarkan Garis Wallace dan Garis Weber • Keunikan hutan hujan tropis Indonesia • Pemanfaatan keanekaragaman hayati Indonesia • Upaya pelestarian keanekaragaman hayati Indonesia 	Interpretasi	-	19	C
		Analisis	√	14	A
		Evaluasi	-	15	C
		Pengaturan Diri	-	18	A

Tabel 5. Hasil Validasi Ahli Nilai Islam

Aspek yang Dinilai	Persentase Skor Rata-rata (%)	Kualifikasi	Kategori
Aspek Materi/substansi	92,75	Sangat Baik	Tidak perlu direvisi
Aspek Konstruksi	86,36	Baik	Tidak perlu direvisi
Aspek Bahasa	100	Sangat Baik	Tidak perlu direvisi
Aspek berpikir kritis	83,33	Baik	Tidak perlu direvisi
Aspek Nilai Islam	100	Sangat Baik	Tidak perlu direvisi
Rata-rata	92,48	Sangat Baik	Tidak perlu direvisi

Tabel 6. Hasil Validasi Ahli Materi

Aspek yang Dinilai	Persentase Skor Rata-rata (%)	Kualifikasi	Kategori
Aspek Materi/substansi	87,5	Baik	Tidak perlu direvisi
Aspek Konstruksi	79,54	Baik	Tidak perlu direvisi
Aspek Bahasa	95,00	Sangat Baik	Tidak perlu direvisi
Aspek berpikir kritis	100	Sangat Baik	Tidak perlu direvisi
Aspek Nilai Islam	83,33	Baik	Tidak perlu direvisi
Rata-rata	89,07	Baik	Tidak perlu direvisi

Tabel 7. Hasil Validasi Ahli Evaluasi Pendidikan

Aspek yang Dinilai	Persentase Skor Rata-rata (%)	Kualifikasi	Kategori
Aspek Materi/substansi	87,50	Baik	Tidak perlu direvisi
Aspek Konstruksi	79,54	Baik	Tidak perlu direvisi
Aspek Bahasa	95,00	Sangat Baik	Tidak perlu direvisi
Aspek berpikir kritis	75,00	Baik	Tidak perlu direvisi
Aspek Nilai Islam	100	Sangat Baik	Tidak perlu direvisi
Rata-rata	87,40	Baik	Tidak perlu direvisi

Tabel 8. Rekapitulasi Hasil Perhitungan Rasch tentang Soal Valid dan Soal Tidak Valid pada *Item Fit Order*

Nomor Soal Valid	Jumlah Soal Valid	Nomor Soal Tidak Valid	Jumlah Soal Tidak Valid
1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 24, 25, 26, 28, 30, 31, 32, 33, 34, 35, 37, 39, 40, 42, 43, 44, 46, 48, 49, 50, 52, 54, 55, 57, 58, 59, 60.	45	4, 7, 17, 19, 23, 27, 29, 36, 38, 41, 45, 47, 51, 53, 56.	15

Hasil validasi ahli materi, ahli evaluasi pendidikan dan ahli nilai Islam asesmen berbasis berpikir kritis terintegrasi nilai Islam dapat dilihat pada Tabel 5, Tabel 6 dan Tabel 7.

Hasil penilaian ahli nilai Islam, ahli materi dan ahli evaluasi pendidikan tentang asesmen berbasis berpikir kritis terintegrasi nilai Islam berkualifikasi sangat baik dan baik sehingga layak digunakan. Namun ada beberapa saran dan masukan untuk diperbaiki seperti tata tulis, ejaan, penulisan simbol dan penulisan nama ilmiah, pilihan jawaban “semua benar” hendaknya diganti, beberapa kata/kalimat dalam soal masih harus diperbaiki untuk mempertegas maksud pertanyaan. Semua saran dan masukan sudah diperbaiki oleh peneliti dan dilanjutkan membuat petunjuk penyelenggaraan tes yang terdiri dari petunjuk untuk penyelenggara, pengawas ujian dan petunjuk untuk peserta tes itu sendiri. Setelah itu dilanjutkan ke tahap uji coba (*conduct piloting test*).

Uji coba dilakukan pada siswa kelas XI di SMA Nadhatul Ulama dan SMA Muhammadiyah 4. Alasan dilakukannya pada siswa kelas XI adalah karena siswa sudah mendapatkan semua pelajaran materi kelas X, sedangkan kelas X saat penelitian berlangsung belum mendapatkan semua materi.

Berdasarkan hasil analisis Rasch pada *Item Fit Order* diperoleh butir soal valid yang sesuai adalah 45 soal dan soal yang tidak valid diperoleh 15 soal. Secara lebih detail tercantum pada Tabel 8.

Berdasarkan hasil pada Tabel 9 dapat dijelaskan bahwa instrumen soal yang dikembangkan memiliki reliabilitas *Alpha Cronbach* sebesar 0,78 yang artinya interaksi antara siswa dengan butir-butir soal secara keseluruhan

adalah bagus. Untuk nilai *person reliability* dan *item reliability*, secara berurutan sebesar 0,76 dan 0,93 dapat diartikan bahwa konsistensi jawaban siswa terkategori cukup dan kualitas butir-butir soal dalam instrumen yang dikembangkan aspek reliabilitasnya bagus.

Tabel 9. Nilai Uji Reliabilitas

Jenis Uji Reliabilitas	Nilai Hasil	Keterangan
<i>Alpha Cronbach</i>	0,78	Bagus
<i>Item Reliability</i>	0,93	Bagus
<i>Person Reliability</i>	0,76	Cukup

Tabel 10. Kualitas Butir Soal dalam Instrumen yang Dikembangkan

	Nilai
INFIT MNSQ	0,99
OUTFIT MNSQ	1,12
INFIT ZSTD	-0,1
OUTFIT ZSTD	0,1
<i>Mean Item Measure</i>	0,00
<i>Separation Item</i>	3,64
<i>Nilai Pemisahan Strata Item</i>	5

Berdasarkan Tabel 10 dapat dijelaskan bahwasanya nilai INFIT dan OUTFIT MNSQ yang diperoleh secara berurutan sebesar 0,99 dan 1,12 memiliki arti bahwa instrumen terkategori bagus, karena nilai mendekati nilai ideal yaitu 1,00. Nilai INFIT dan OUTFIT ZSTD secara berurutan adalah -0,1 dan 0,1 yang artinya instrumen memiliki kualitas bagus, karena mendekati nilai ideal 0,0.

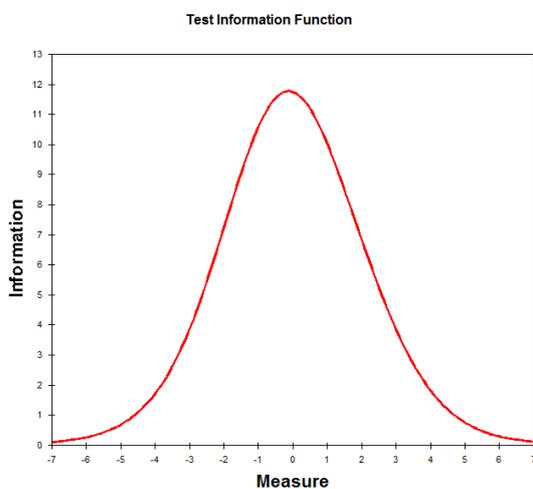
Nilai rata-rata *item measure* diperoleh 0,00 yang pada dasarnya nilai rata-rata *item measure* selalu diperoleh 0,00. Untuk nilai *separation* diperoleh 3,64 yang dapat digunakan untuk melihat pengelompokan instrumen (*pemisahan strata*) yaitu 5. Hal ini menyatakan bahwa instrumen terdiri dari 5 kelompok butir soal

sehingga instrumen dapat dikategorikan sangat bagus. Lima kelompok butir soal ini terdiri dari kelompok butir soal sangat sukar, sukar, sedang, mudah, dan sangat mudah (lebih rinci lihat Tabel 11). Soal yang paling sulit bagi siswa adalah soal nomor 53 dengan nilai logit tertinggi yaitu 2,86. Soal yang paling mudah dijawab bagi siswa adalah soal nomor 58 dengan nilai logit terendah yaitu -2,35.

Tabel 11. Persebaran Soal pada Tingkat Kesulitan

Tingkat Kesulitan	Nomor Soal
Sangat Sukar	53, 4, 27, dan 45.
Sukar	29, 51, 1, 47, 39, 22, dan 35.
Sedang	23, 38, 21, 56, 26, 34, 5, 12, 28, 36, 42, 6, 15, 40, 52, 7, 46, 57, 2, 31, 24, 37, 43, 54, 41, 48, dan 50.
Mudah	11, 17, 19, 8, 20, 25, 30, 33, 9, 3, 32, 44, 49, 13, 14, 16, 55, dan 59.
Sangat Mudah	18, 10, 60, dan 58.

Sumintono & Widhiarso (2015) menjelaskan bahwa setiap pengukuran selalu menghasilkan informasi fokus pengukuran. Pada penelitian kali ini, informasi fokus pengukuran dapat dilihat pada Gambar 1.



Gambar 1. Informasi Fokus Pengukuran Keterampilan Berpikir Kritis Siswa

Sumbu X menunjukkan tentang kemampuan siswa, sedangkan sumbu Y menunjukkan besaran fungsi informasi yang diperoleh. Dapat disimpulkan dari Gambar 2 bahwa 60 butir soal yang diberikan kepada 70 siswa di kelas XI menunjukkan butir soal cocok untuk mengetahui tingkat kemampuan siswa yang sedang saja. Dapat diartikan bahwa instrumen soal berpikir kritis disertai nilai Islam menghasilkan informasi yang optimal ketika diberikan kepada

siswa yang memiliki keterampilan berpikir kritis pada level sedang.

Selanjutnya melihat pola jawaban siswa pada masing-masing butir soal. Abidin (2016) menjelaskan, pola jawaban soal perlu dicek dengan cara menghitung jumlah siswa yang memilih opsi tertentu untuk suatu soal. Hasil perhitungan ini akan menunjukkan apakah opsi pengecoh dan opsi kunci jawaban dapat berfungsi dengan optimal. Hal ini bertujuan agar soal yang dikerjakan siswa dapat berfungsi secara optimal, sehingga mendapatkan hasil yang lebih baik. Suatu pengecoh dapat dikatakan berfungsi apabila dipilih sedikitnya 2,5% dari semua siswa yang mengerjakan. Meskipun pola jawaban soal termasuk dalam teori tes klasik, peneliti tetap memasukkan dalam pembahasan karena dianggap perlu untuk mengecek apakah butir soal masih memiliki kalimat yang membingungkan dan mengecek pola jawaban setiap butir soal yang dianggap valid. Dengan adanya pertimbangan tersebut, maka diperoleh hasil sebagai berikut yang tertera pada Tabel 12.

Tabel 12. Sebaran Pola Jawaban Siswa

	Nomor Soal
Terdapat opsi pengecoh dipilih lebih banyak dibanding kunci jawaban	1, 11, 12, 15, 21, 22, 28, 31, 34, 35, 39, 40, 42, 46, 52.
Terdapat opsi pengecoh banyak dipilih meskipun kunci jawaban lebih banyak dipilih	6, 24, 26, 30, 37, 48, 50, 57.
Terdapat opsi pengecoh yang kurang berfungsi	3, 8, 14, 59.

Setelah melakukan ujicoba terbatas dan hasil analisis model Rasch maka terdapat keputusan dan hasil revisi, yaitu: (1) Membuang butir soal yang tidak valid, butir soal valid akan diambil dan digunakan pada tahapan selanjutnya *validation test* sebanyak 45 soal; (2) Merevisi pola jawaban soal, merevisi pola jawaban soal terkait susunan kalimat dan bahasa yang masih belum pas dan membingungkan bagi siswa.

Langkah selanjutnya melakukan ujicoba soal kembali untuk melihat kestabilan validitas dan reliabilitas soal. Ujicoba dilakukan pada 8 sekolah SMA berbasis Islam yang ada di wilayah Seberang Ulu Palembang. Sekolah yang digunakan antara lain: SMA Daarul Aitam, SMA NU, SMA Azhariyah, SMA Muhammadiyah 3 Palembang, SMA Muhammadiyah 4 Palembang, SMA Muhammadiyah 9 Palembang, SMA YP Nurfauzan, dan SMA Ashanadiyah. Siswa yang digunakan adalah siswa kelas X yang sudah mendapatkan semua materi Biologi kelas

X, sehingga instrumen diujicobakan di akhir semester genap.

Berdasarkan hasil analisis yang kedua, diperoleh soal valid 43 butir dan soal yang tidak valid 2 butir yaitu nomor 32 dan 43.

Tabel 13. Rekapitulasi Hasil Perhitungan Rasch tentang Soal Valid dan Soal Tidak Valid pada Item Fit Order

Nomor Soal Valid	Jumlah Soal Valid	Nomor Soal Tidak Valid	Jumlah Soal Tidak Valid
1, 3, 4, 6, 7, 8, 9, 10, 12, 13, 16, 17, 19, 20, 21, 23, 24, 25, 26, 29, 33, 34, 35, 36, 37, 38, 39, 41, 42, 2, 5, 11, 14, 15, 18, 22, 27, 28, 30, 31, 40, 44, dan 45.	43	32 dan 43	2

Selain melihat validitas, instrumen juga harus reliabel yang artinya soal memiliki tingkat keajegan dalam mengukur keterampilan siswa. Hasil uji pada pemodelan Rasch diperoleh pada Tabel 14.

Tabel 14. Nilai Uji Reliabilitas

Jenis Uji Reliabilitas	Nilai Hasil	Keterangan
<i>Alpha Cronbach</i>	0,68	Cukup
<i>Item Reliability</i>	0,93	Bagus
<i>Person Reliability</i>	0,64	Lemah

Berdasarkan hasil pada Tabel 14 dapat dijelaskan bahwa instrumen soal yang dikembangkan memiliki nilai alpha cronbach sebesar 0,68 yang artinya interaksi antara siswa dengan butir-butir soal secara keseluruhan adalah cukup. Untuk nilai *person reliability* dan *item reliability*, secara berurutan sebesar 0,64 dan 0,93 dapat diartikan bahwa konsistensi jawaban siswa terkategori lemah namun kualitas butir-butir soal dalam instrumen yang dikembangkan aspek reliabilitasnya bagus.

Tabel 15. Kualitas Butir Soal dalam Instrumen yang Dikembangkan

	Nilai
INFIT MNSQ	1,00
OUTFIT MNSQ	1,02
INFIT ZSTD	-0,1
OUTFIT ZSTD	0,0
<i>Mean Item Measure Separation Item</i>	0,00
<i>Separation Item</i>	3,63
Nilai Pemisahan Strata Item	5

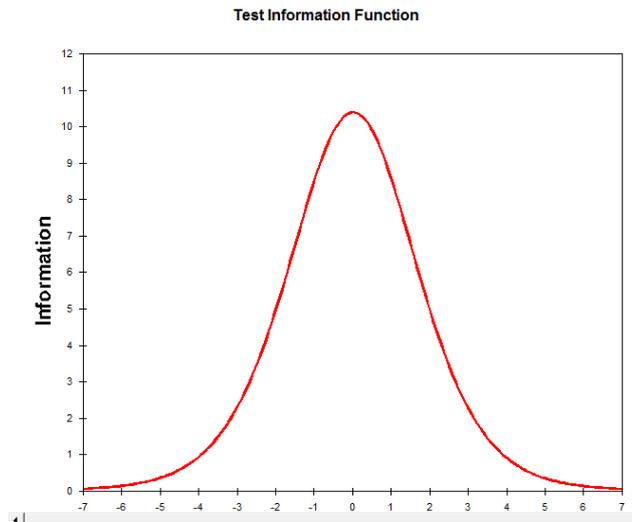
Berdasarkan Tabel 15 dapat dijelaskan bahwasannya Nilai INFIT dan OUTFIT MNSQ yang diperoleh secara berurutan sebesar 1,00 dan 1,02 memiliki arti bahwa instrumen terkategori bagus, karena nilai mendekati nilai ideal yaitu 1,00. Nilai INFIT dan OUTFIT ZSTD secara berurutan adalah -0,1 dan 0,0 yang artinya instrumen memiliki kualitas bagus, karena mendekati nilai ideal 0,0. Jika dibandingkan dengan hasil ujicoba soal pertama, maka nilai MNSQ dan ZSTD kali ini lebih baik.

Nilai rata-rata *item measure* diperoleh 0,00 yang pada dasarnya nilai rata-rata *item measure* selalu diperoleh 0,00. Untuk nilai separation diperoleh 3,63 yang dapat digunakan untuk melihat pengelompokan instrumen (pemisahan strata) yaitu 5. Hal ini menyatakan bahwa instrumen terdiri dari 5 kelompok butir soal sehingga instrumen dapat dikategorikan sangat bagus. Lima kelompok butir soal ini terdiri dari kelompok butir soal sangat sukar, sukar, sedang, mudah, dan sangat mudah. Melihat dari nilai *logit*, soal yang paling sulit bagi siswa adalah soal nomor 17 dan 28 dengan nilai *logit* yang sama yaitu 1,10. Soal yang paling mudah dijawab bagi siswa adalah soal nomor 12 dengan nilai *logit* terendah yaitu -1,48. Dilihat dari hasil perhitungan Strata pada Tabel 18 terdapat 5 kelompok butir soal, lebih lengkap dapat dilihat pada Tabel 16.

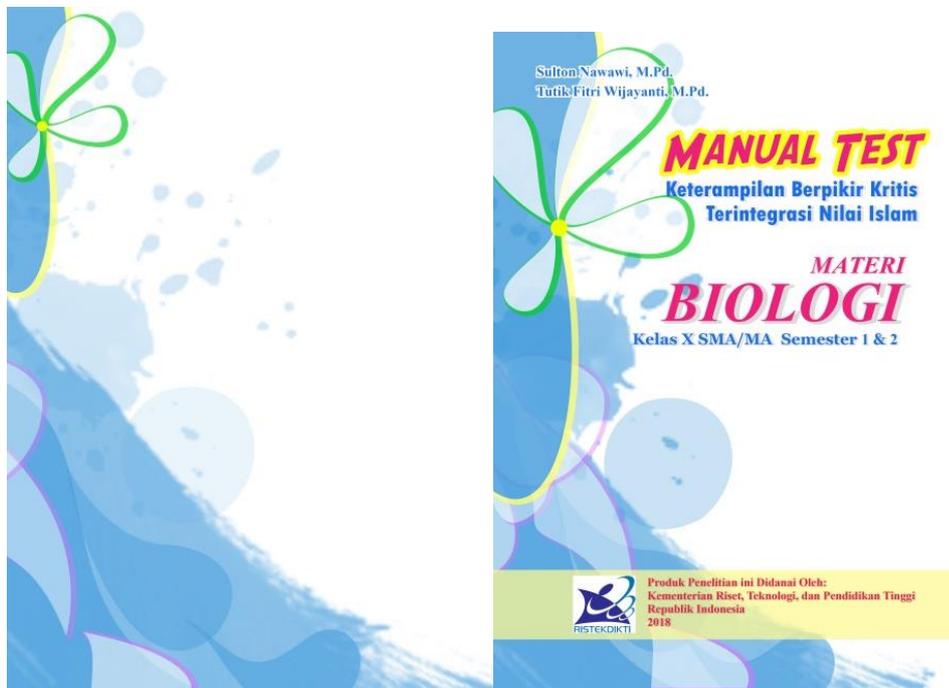
Tabel 16. Persebaran Soal pada Tingkat Kesukaran

Tingkat Kesulitan	Nomor Soal
Sangat Sukar	17 dan 28
Sukar	2, 31, 40, 22, 25, 9, 24, 21, 32, 16, 35, dan 36.
Sedang	38, 4, 14, 20, 27, dan 44.
Mudah	13, 23, 36, 30, 33, 39, 41, 42, 5, 6, 1, 18, 43, 3, 7, 37, 19, 34, 45, 11, 29, dan 10.
Sangat Mudah	15, 8, dan 12.

Pada tahapan kali ini, informasi fokus pengukuran dapat dilihat pada Gambar 2, dari 45 butir soal yang diberikan kepada 252 siswa di kelas X menunjukkan butir soal cocok untuk mengetahui tingkat kemampuan siswa yang sedang saja. Dapat diartikan bahwa instrumen soal berpikir kritis disertai nilai Islam menghasilkan informasi yang optimal ketika diberikan kepada siswa yang memiliki keterampilan berpikir kritis pada level sedang.



Gambar 2. Informasi Fokus Pengukuran Keterampilan Berpikir Kritis Siswa



Gambar 3. Cover Produk Instrumen Tes Biologi Berbasis Keterampilan Berpikir Kritis Terintegrasi Nilai Islam

Setelah validasi lengkap, maka selanjutnya membuat norma acuan dari distribusi skor tes. Perolehan skor yang diperoleh dapat digunakan untuk melihat tingkat keterampilan berpikir kritis terintegrasi nilai Islam pada siswa. Norma acuan bisa dilihat pada Tabel 17.

Akhir dari kegiatan pengembangan tes adalah menyusun buku petunjuk penggunaan tes (*manual test*). Isi buku petunjuk menjelaskan latar belakang pembuatan tes, sejarah proses pengembangan, hasil-hasil studi validasi, deskripsi target sasaran yang sesuai, petunjuk penyelenggaraan, cara penyekoran tes, dan infor-

masi tentang cara menginterpretasikan skor individu.

Tabel 17. Norma Acuan Penguasaan Tingkat Keterampilan Berpikir Kritis Terintegrasi Nilai Islam

Nilai	Huruf	Keterangan
80-100	A	Baik sekali
66-79	B	Baik
56-65	C	Cukup
40-55	D	Kurang
30-39	E	Gagal

Sumber: (Arikunto, 2016)

Identitas produk dari Instrumen Tes Biologi Berbasis Keterampilan Berpikir Kritis Terintegrasi Nilai Islam dapat dilihat pada Tabel 18.

Tabel 18. Instrumen Tes Biologi Berbasis Keterampilan Berpikir Kritis Terintegrasi Nilai Islam

No.	Desain	Keterangan
1.	Bentuk fisik	Bahan cetak
2.	Judul	Instrumen Tes Biologi Berbasis Keterampilan Berpikir Kritis Terintegrasi Nilai Islam
3.	Ukuran kertas	A4
4.	Tebal halaman	31 halaman + cover
5.	Bahasa	Indonesia
6.	Jenis huruf	Cambria dengan ukuran huruf 11

Karakteristik Soal Biologi Berbasis Berpikir Kritis Terintegrasi Nilai Islam

Asesmen Biologi yang dikembangkan merupakan soal Biologi SMA/MA yang menggunakan materi kelas X pada semester ganjil dan genap dengan mengacu pada kurikulum 2013. Asesmen diintegrasikan menggunakan aspek berpikir kritis Facione (2013) yaitu aspek interpretasi, analisis, evaluasi, inferensi, penjelasan, dan pengaturan diri. Asesmen juga disusun dengan mengaitkan permasalahan kehidupan sehari-hari dan fakta-fakta yang ada di lingkungan, serta diintegrasikan nilai Islam, meskipun tidak semua butir soal diintegrasikan nilai Islam karena keterbatasan dari peneliti. Nilai Islam yang digunakan merupakan nilai-nilai yang bersumber dari Alquran dan Hadits.

Asesmen yang dikembangkan sudah sesuai dengan komponen urutan penyusunan asesmen. Langkah awal pengembangan asesmen dimulai dengan analisis kebutuhan awal untuk menentukan kompetensi, peserta tes dan tujuan tes, langkah selanjutnya yaitu mengembangkan rencana uji yang meliputi membuat kisi-kisi soal, membuat format pertanyaan, cara penyelenggaraan dan cara penyekorannya, langkah selanjutnya yaitu validasi ahli (*expert judgment*), memberikan draft awal soal untuk kepada ahli untuk divalidasi, ada tiga ahli yaitu ahli materi, ahli evaluasi pendidikan dan ahli nilai Islam, setelah validasi ahli langkah selanjutnya yaitu menulis instruksi administrasi, berupa petunjuk penyelenggara, petunjuk pengawas ujian dan petunjuk peserta didik, langkah selanjutnya yaitu melakukan uji coba soal pertama kepada peserta didik, setelah uji coba pertama,

dilakukan analisis item butir soal untuk mengetahui validitas dan reliabilitas soal, selanjutnya soal direvisi dan dilakukan uji coba soal kepada peserta didik yang kedua untuk melihat validitas dan reliabilitas soal, setelah selesai dikembangkan norma acuan dan melengkapi manual tes. Hal ini sesuai dengan pendapat Subali (2010) untuk mendapatkan asesmen yang baik dalam pengembangan harus mengikuti langkah-langkah yang benar, meliputi perencanaan asesmen yang memuat maksud dan tujuan asesmen yaitu penyusunan kisi-kisi, penyusunan instrumen/alat ukur, penelaahan (*review*) untuk menilai kualitas alat ukur/instrumen secara kualitatif, yakni sebelum digunakan, uji coba alat ukur, untuk menyelidiki kesahihan dan keandalan secara empiris, pelaksanaan pengukuran, asesmen yang merupakan interpretasi hasil pengukuran dan pemanfaatan hasil asesmen.

Asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam menggunakan indikator berpikir kritis dan nilai Islam, disajikan melalui pemberian pertanyaan yang terorganisasi dan sistematis dalam menilai suatu topik, sehingga mengantarkan siswa pada kesimpulan yang mandiri dan dapat dipercaya. Pertanyaan-pertanyaan yang terorganisasi dan sistematis memungkinkan siswa untuk terlibat dalam kegiatan yang membantunya untuk mendapatkan pemahaman yang mendalam. Pertanyaan-pertanyaan yang dihadirkan harus dikemukakan sesuai dengan urutan untuk membantu siswa meneliti dan memahami setiap masalah, isu, proyek atau keputusan (Johnson, 2010). Hal ini senada dengan pendapat Rustaman (2005) bahwa dengan menggunakan pertanyaan yang efektif berarti mendorong siswa untuk berpikir dan bernalar. Begitu pula dengan pendapat Nasution (2011) bahwa bertanya merupakan stimulus yang mendorong anak untuk berpikir dan belajar, dengan bertanya, siswa akan memperoleh pengetahuan. Asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam membuat siswa aktif dalam pembelajaran, sebab siswa berpikir bagaimana memecahkan masalah yang dihadapi, dalam proses pemecahan masalah terjadi suatu proses berpikir tingkat tinggi yaitu berpikir kritis (Johnson et al., 2009).

Asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam, berarti asesmen tersebut diselipkan nilai-nilai Islam. Nilai Islam berasal dari Al-Qur'an dan hadits, yang diintegrasikan pada butir soal, selanjutnya peserta didik akan menjawab, menelaah, menganalisis, dan mengevaluasi soal tersebut. Implementasi nilai Islam

diharapkan dapat membuat siswa memahami karakteristik Islam, juga melahirkan siswa yang menguasai ilmu pengetahuan dan teknologi. Disamping itu juga ada harapan dapat mengubah sikap ke arah yang lebih baik setelah benar-benar terjun ke masyarakat memberikan uswan hasanah.

Kelayakan Soal Biologi Berbasis Berpikir Kritis Terintegrasi Nilai Islam

Kelayakan soal biologi berbasis berpikir kritis terintegrasi nilai Islam diuji melalui tahapan *expert judgment validity* yang dilakukan oleh validator materi, validator evaluasi pendidikan, validator nilai Islam yang didalamnya terdapat aspek materi/substansi, aspek konstruksi, aspek bahasa, aspek berpikir kritis dan aspek nilai Islam. Berdasarkan hasil *expert judgment validity* diperoleh hasil validasi nilai Islam memperoleh nilai rata-rata 92,48 dengan kategori sangat baik, validasi materi memperoleh nilai rata-rata 89,07 dengan kategori baik, validasi evaluasi pendidikan memperoleh nilai rata-rata 87,40 dengan kategori baik. Penilaian Ahli Nilai Islam tidak ada revisi, untuk Ahli Materi dan Ahli Evaluasi ada revisi berupa kritik, saran atau tanggapan mengenai kekurangan instrumen soal, seperti perbaikan tata tulis, ejaan, penulisan simbol dan penulisan nama ilmiah, *option* jawaban pada beberapa butir soal disederhanakan, serta pilihan jawaban “semua benar” hendaknya diganti. Semua kritik, saran atau tanggapan sudah direvisi.

Kelayakan juga dapat dilihat pada uji *empiric validity* yang dianalisis menggunakan model Rasch berbantuan program Winstep. Berdasarkan hasil *empiric validity* diperoleh hasil soal yang valid adalah 43 butir soal dari jumlah awal 60 butir soal. Sumintono, B & Widhiarso (2015) menjelaskan bahwa pada model Rasch, validitas ditunjukkan oleh pola respon siswa (*person*) kepada setiap butir soal (*item*). Pemodelan Rasch dapat mengetahui adanya responden yang memang tidak sesuai dilibatkan dalam pengumpulan data dan dapat menjelaskan alasan responden perlu dikeluarkan karena tidak sesuai dengan model yang ada. Hal ini dapat dilihat dari banyaknya siswa yang tidak sesuai dengan model karena konsistensi jawaban siswa lemah. Banyaknya siswa yang menebak jawaban dan kurang teliti ketika mengerjakan soal dapat menyebabkan pola respon siswa menjadi kurang bagus, sehingga interaksi jawaban siswa dengan respon butir soal terkategori

cukup yang dibuktikan dari hasil perhitungan *alpha cronbach*.

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. Hal ini dapat dilihat dari *Variable Map* dan hasil pengolahan nilai *logit* pada *item fit*, sehingga diperoleh jumlah soal valid 43 butir soal. Nilai *item reliability* pada instrumen yang dikembangkan diperoleh 0,93 menandakan kualitas butir soal dalam instrumen memiliki aspek reliabilitas bagus. Selain itu soal yang dikembangkan terbagi menjadi 5 strata/tingkatan. Hal ini menyatakan bahwa instrumen terdiri dari 5 kelompok butir soal sehingga instrumen dapat dikategorikan sangat bagus. Lima kelompok butir soal ini terdiri dari kelompok butir soal sangat sukar, sukar, sedang, mudah, dan sangat mudah.

Dapat disimpulkan bahwa soal biologi berbasis berpikir kritis terintegrasi nilai Islam dapat dinyatakan valid secara empiris maupun *expert judgment* dan reliabel sehingga dapat digunakan untuk mengukur keterampilan berpikir kritis terintegrasi nilai Islam pada siswa.

SIMPULAN

Karakteristik asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam menggunakan aspek berpikir kritis yang dikembangkan oleh Facione dan diintegrasikan dengan nilai Islam, nilai Islam bersumber dari Alquran dan Hadits. Asesmen menggunakan materi Biologi kelas X semester ganjil dan genap, dibuat berdasarkan permasalahan kehidupan sehari-hari dan fakta yang ada di lingkungan. Asesmen biologi berbasis berpikir kritis terintegrasi nilai Islam berdasarkan validator ahli hasilnya berkualifikasi baik sampai dengan sangat baik serta berdasarkan hasil validitas empirik butir soal dinyatakan valid dan reliabel sehingga dapat digunakan untuk mengukur keterampilan berpikir kritis terintegrasi nilai Islam pada siswa.

DAFTAR PUSTAKA

- Abidin, Y. (2016). *Revitalisasi penilaian pembelajaran dalam konteks pendidikan multiliterasi abad ke-21*. Bandung: Refika Aditama.
- Arikunto, S. (2016). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.
- Dacholfany, M. I. (2015). Reformasi pendidikan islam dalam menghadapi era globalisasi:

- sebuah tantangan dan harapan. *Akademika*, 20(1), 173–194.
- Facione, P. A. (2013). *Critical thinking, what it is and why it counts*. California: California Academic Press.
- Johnson, E. . (2010). *CTL: Contextual teaching and learning (menjadikan belajar-mengajar lebih mengasyikan dan bermakna)*. Bandung: Kaifa Learning.
- Johnson, L. F., Smith, R. S., Smythe J. T., & Varon R. K. (2009). *Challenge-based learning: an approach for our time. new media consortium*. Retrieved from [http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED505102&site=ehost-live](http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED505102&lang=es&scope=site%0Ahttp://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED505102&site=ehost-live)
- Kementrian Pendidikan dan Kebudayaan. (2016). *Silabus mata pelajaran sekolah menengah atas/madrasah aliyah (SMA/MA): mata pelajaran biologi*. Retrieved from <http://www.syaiflash.com/rpp2016/>
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan no-tes*. Yogyakarta: Mitra Cendikia Offset.
- McIntire, S., & Miller, L. (2007). *Foundations of psychological testing: a practical approach*. Boston: McGraw-Hil.
- Nasution, S. (2011). *Berbagai Pendekatan dalam proses belajar mengajar*. Jakarta: Bumi Aksar.
- Pimpinan Pusat Muhammadiyah. (2000). *Pedoman kehidupan islami warga muhammadiyah*. Retrieved from http://tasikmalaya-kota.muhammadiyah.or.id/muhfile/tasikmalaya-kota/download/PEDOMAN_KEHIDUPAN_ISLAMIS_WARGA_MUHAMMADIYAH-1.pdf.
- Riduwan. (2009). *Skala Pengukuran variabel-variabel penelitian*. Jakarta: Alfabeta.
- Rustaman, N. (2005). *Strategi belajar mengajar biologi*. Bandung: Universitas Pendidikan Indonesia.
- Subali, B. (2010). *Penilaian, evaluasi, dan remedial pembelajaran biologi*. Yogyakarta: UNY Press.
- Sumintono, B & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada asesmen pendidikan*. Cimahi: Trim Komunikata.
- Surayya, L., Subagia, I. W., & Tika, I. N. (2014). Pengaruh model pembelajaran think pair share terhadap hasil belajar ipa ditinjau dari keterampilan berpikir kritis siswa. *E-Journal Program Pascasarjana Universitas Pendidikan Ganesha*, 4, 1–11.
- Suwastono, A. (2011). *Pengembangan pembelajaran e-learning berbasis moodle pada mata kuliah penginderaan jauh*. Universitas Negeri Malang, Malang. Retrieved from <http://mulok.library.um.ac.id/index3.php/50112.html>
- Syaban, M. (2006). Asesmen proses. *EDUCARE*, 2(1), 43–51. Retrieved from <http://jurnal.fkip.unla.ac.id/index.php/educare/article/view/20>
- Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *The assessment of higher education learning outcomes* (vol. 1). OECD. Retrieved from <http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf>
- Wiliam, D. (2013). Assessment: the bridge between teaching and learning. *Voices from the Middle*, 21(2), 15–20. Retrieved from <http://cccc.ncte.org/library/NCTEFiles/Resources/Journals/VM/0212-dec2013/VM0212Assessment.pdf>
- Wood, R. (2002). *Critical thinking*. Retrieved from <http://www.robinwood.com/Democracy/GeneralEssays/CriticalThinking.pdf>
- Wulan, A. R. (2007). *Pengertian dan esensi konsep evaluasi, asesmen, tes, dan pengukuran*. Retrieved from http://file.upi.edu/Direktori/SPS/PRODI.PENDIDIKAN_IPA/197404171999032-ANA_RATNAWULAN/pengertian_asesmen.pdf