

PENGEMBANGAN INSTRUMEN ASESMEN BERPIKIR KRITIS UNTUK SISWA SMP KELAS VII PADA MATERI INTERAKSI MAKHLUK HIDUP DENGAN LINGKUNGAN

Dharmawati, Sri Rahayu, Susriyati Mahanal
Pendidikan Dasar Pascasarjana-Universitas Negeri Malang
Jalan Semarang 5 Malang. E-mail: dharmawati@ymail.com

Abstract: This study aims to generate critical thinking assessment instrument for seven graders within organism interaction in an environment with an appropriate validity and reliability level. The design of this instrument is developed using stages suggested by Borg and Gall. The stages are researching and collecting information, planning, developing a preliminary product, conducting the limited examination, revising product from limited examination, field testing, revising product from field testing, and finalizing product. The result of content and construction validation shows that the level of feasibility is 88,35% and categorized as very feasible. While the degree of readability of assessment items is 93,51% and it is categorized as excellent. The coefficient of inter-rater reliability of assessment items is 0,951 and categorized as excellent. The instrument acquires reliability coefficient 0,792 for multiple choice and 0,753 for essay items. This study shows that the content, construction, and items in assessment are feasible.

Keywords: development, assessment instrument, critical thinking

Abstrak: Penelitian ini bertujuan untuk menghasilkan instrumen asesmen berpikir kritis untuk siswa SMP kelas VII pada materi interaksi makhluk hidup dengan lingkungan dengan tingkat validitas dan reliabilitas yang memadai. Rancangan penelitian menggunakan model pengembangan menurut Borg & Gall, yang meliputi langkah-langkah: penelitian dan pengumpulan informasi, perencanaan, pengembangan produk awal, uji coba terbatas, revisi produk uji coba terbatas, uji coba lapangan, revisi uji coba produk lapangan, dan penyempurnaan produk akhir. Berdasarkan hasil validasi isi dan konstruk diperoleh tingkat kelayakan produk sebesar 88,35% berada pada kriteria sangat layak. Tingkat keterbacaan soal asesmen sebesar 93,51% dengan kategori sangat baik. Koefisien *inter-rater reliability* pada soal asesmen bentuk penugasan sebesar 0,951 dengan kategori sangat baik. Instrumen tersebut mempunyai koefisien reliabilitas sebesar 0,792 (soal pilihan ganda) dan 0,753 (esai). Kesimpulan hasil penelitian menunjukkan bahwa berdasarkan validasi isi, konstruk, dan butir soal adalah layak untuk digunakan.

Kata kunci: pengembangan, instrumen asesmen, berpikir kritis

Tujuan utama keberhasilan pendidikan adalah peningkatan sumber daya manusia. Keberhasilan tersebut dipengaruhi oleh berbagai faktor, salah satunya adalah kemampuan guru dalam melakukan dan memanfaatkan asesmen, evaluasi proses, dan hasil belajar (Arifin, 2009:12). Kemampuan tersebut sangat diperlukan untuk mengetahui tercapai tidaknya tujuan pembelajaran yang sudah ditetapkan dalam kurikulum. Selain itu, kemampuan tersebut juga dapat digunakan untuk memperbaiki atau meningkatkan proses pembelajaran yang telah dilakukan guru (Budiman & Jailani, 2014:140). Pernyataan tersebut sesuai dengan Permendiknas Nomor 16 Tahun 2007 tentang Standar Kualifikasi Akademik dan Kompetensi Guru mata pelajaran, bahwa kompetensi guru mata pelajaran, antara lain adalah mengembangkan instrumen asesmen. Prinsip dan standar asesmen menekankan dua ide pokok, yaitu (1) asesmen harus meningkatkan hasil belajar peserta didik, (2) asesmen merupakan sebuah alat yang berharga untuk membuat keputusan pengajaran (Kemendiknas, 2007). Oleh karena itu, kegiatan asesmen sangat penting dalam pembelajaran karena dapat memberikan umpan balik yang konstruktif bagi guru maupun peserta didik, dapat memberikan motivasi kepada peserta didik untuk berprestasi lebih baik dan dapat memengaruhi perilaku belajar sesuai dengan asesmen yang dilakukan guru (Van de Walle, 2007:78).

Kualitas instrumen asesmen hasil belajar berpengaruh langsung dalam keakuratan status pencapaian hasil belajar peserta didik. Oleh karena itu, kedudukan instrumen asesmen hasil belajar sangat strategis dalam pengambilan keputusan guru dan sekolah terkait pencapaian hasil belajar peserta didik adalah kemampuan berpikir kritis (Hariyanto, 2014:7). Sesuai dengan Peraturan Menteri Pendidikan dan Kebudayaan No. 64 Tahun 2013 tentang Standar Isi Pendidikan Dasar dan Menengah untuk memenuhi kebutuhan masa depan dan menyongsong Generasi Emas Indonesia Tahun 2045, telah ditetapkan Standar Kompetensi

Lulusan yang berbasis pada Kompetensi Abad XXI, bahwa pembelajar pada abad ke-21 harus mampu mengembangkan keterampilan kompetitif yang berfokus pada pengembangan keterampilan berpikir tingkat tinggi (*higher order thinking skills*) yang mana salah satunya adalah berpikir kritis (*critical thinking*). Oleh karena itu, tugas guru adalah menciptakan peluang bagi siswa untuk mengembangkan dan meningkatkan keterampilan berpikir kritis, karena siswa dengan keterampilan berpikir kritis mampu memahami dunia di sekitarnya dan membuat keputusan yang baik, meningkatkan kinerja, dan meningkatkan motivasi belajar (Stobaugh, 2013).

Berpikir kritis adalah berpikir yang masuk akal dan reflektif yang terfokus untuk memutuskan apa yang harus dipercaya atau dilakukan (Ennis, 2011:1). Berpikir kritis dengan jelas menuntut interpretasi dan evaluasi terhadap observasi, komunikasi, dan sumber-sumber informasi lainnya. Berpikir kritis juga menuntut keterampilan dalam memikirkan asumsi-asumsi, dalam menarik implikasi-implikasi singkat, dalam mengajukan pertanyaan-pertanyaan yang relevan, dalam memikirkan dan memperdebatkan isu-isu secara terus menerus (Fisher, 2009:13).

Beberapa ahli menyatakan pendapatnya terkait indikator berpikir kritis, di antaranya menurut Chaffee (2012), indikator berpikir kritis terdiri atas (1) melakukan refleksi terhadap pengambilan keputusan dan bijaksana menganalisis situasi dalam pemecahan masalah, (2) mengevaluasi argumen, dan (3) menarik kesimpulan yang tepat. Scriven (1987) mengemukakan indikator berpikir kritis, antara lain memformulasikan pertanyaan dan masalah, mengumpulkan informasi yang relevan dalam pengambilan keputusan, memikirkan alternatif dan memberikan penyelesaian terhadap suatu masalah. Lebih lanjut, Ennis (1985) memaparkan bahwa tiga tingkatan terakhir dalam taksonomi Bloom yaitu menganalisis, mengevaluasi, dan mencipta, termasuk dalam kemampuan berpikir kritis. Namun, tingkatan taksonomi Bloom tersebut kurang dapat digunakan untuk mengukur kemampuan berpikir kritis. Dari beberapa definisi tentang berpikir kritis yang telah dikemukakan tersebut, pemaparan yang lebih lengkap yaitu menurut Ennis. Ennis (1985; 1996; 2011) memaparkan 12 indikator tentang kemampuan berpikir kritis, yaitu (1) memfokuskan pertanyaan; (2) menganalisis argumen; (3) bertanya dan menjawab pertanyaan yang membutuhkan penjelasan; (4) menilai kredibilitas suatu sumber; (5) mengobservasi dan mempertimbangkan hasil observasi; (6) melakukan deduksi dan menilai hasil deduksi; (7) melakukan induksi; (8) membuat dan mempertimbangkan nilai keputusan; (9) mendefinisikan istilah dan mempertimbangkan definisi menggunakan kriteria yang tepat; (10) mengidentifikasi asumsi; (11) memutuskan suatu tindakan; (12) berintegrasi dengan yang lain.

Perlunya peningkatan kemampuan berpikir kritis ditunjukkan melalui beberapa penelitian, seperti yang dilakukan oleh Higgins, Hall, Baumfield, dan Moseley (2005). Mereka menginvestigasi pengaruh kemampuan berpikir kritis pada fungsi kognitif, hasil belajar, dan sikap siswa. Mereka menyatakan bahwa siswa yang diminta untuk berpikir kritis, akan belajar lebih baik. Sedangkan studi yang dilakukan Wenglinsky (2004), menunjukkan data dari *National Assessment Educational Progress* (NAEP) dan *Trends in International Mathematics and Science Study* (TIMSS) bahwa dalam kategori matematika dan sains, jenis pertanyaan yang menekankan pada pemberian alasan diberikan skor yang lebih tinggi dalam semua tingkat yang diujikan. Newmann, Bryk, dan Nagaoka (2001), menunjukkan bahwa siswa yang memperoleh asesmen tentang mengorganisasikan pengetahuan dan pengalaman yang dimilikinya untuk menyelesaikan suatu permasalahan, mendapatkan nilai lebih tinggi dari rata-rata dalam *Iowa Tests of Basic Skill* (ITBS) dan *Illionis Goals Assessments Program* (IGAP).

Mullis, *et al.* (2012:40) menyatakan hasil prestasi TIMSS tahun 2011 menunjukkan skor pencapaian prestasi belajar peserta didik SMP yaitu 406 (skala 0—800) dengan skor rata-rata 500. Sementara itu, dalam *Program for International Student Assessment* (PISA), Indonesia berada di peringkat 5 ke bawah dari total 65 negara peserta dalam semua kategori. Data ini menunjukkan bahwa keadaan kemampuan siswa SMP Indonesia berada di bawah rata-rata. Hasil prestasi TIMSS dan PISA yang rendah tersebut tentunya disebabkan oleh beberapa faktor, diantaranya adalah masih rendahnya kemampuan siswa dalam menyelesaikan soal-soal dengan karakteristik yang sama dengan soal-soal TIMSS dan PISA yang menggunakan masalah kontekstual dan menuntut penalaran, kreativitas serta argumentasi dalam menyelesaikannya (Wardhani, 2011).

Studi pendahuluan yang dilakukan oleh penulis terhadap guru mata pelajaran IPA SMP melalui pemberian angket kepada 20 responden tentang ketersediaan instrumen asesmen berpikir kritis diperoleh informasi pengembangan instrumen asesmen yang dilakukan oleh guru IPA tidak mengacu pada kompetensi berpikir kritis (dokumen pribadi). Salah satu penyebabnya adalah guru tidak memiliki petunjuk atau pedoman dalam menyusun instrumen asesmen berpikir kritis. Instrumen asesmen yang digunakan guru untuk menilai hasil belajar siswa pada aspek kognitif biasanya diambil dari berbagai buku paket atau kumpulan soal. Menurut Budiman & Jaelani (2014), jenis pertanyaan yang diajukan atau tugas yang diberikan oleh guru sangat berpengaruh terhadap perkembangan keterampilan berpikir siswa, karena pertanyaan atau tugas tersebut bukan hanya untuk memfokuskan siswa pada kegiatan, tetapi juga untuk menggali potensi belajar mereka, serta memicu siswa untuk berpikir analitis, evaluatif, dan kreatif dapat melatih siswa dalam berpikir kritis.

Sesuai dengan uraian yang telah dikemukakan tersebut, dengan berpedoman pada pembelajaran abad 21 yaitu membuat *benchmark* asesmen sumatif secara global, dimana soal-soal pada PISA dan TIMSS adalah contoh yang baik, maka peneliti melakukan studi pendahuluan dengan menganalisis soal-soal PISA dan TIMSS khususnya yang sesuai dengan materi pembelajaran pada kelas VII SMP. Ditemukan bahwa soal-soal yang diujikan pada PISA dan TIMSS disesuaikan dengan kurikulum yang ada di Indonesia pada materi SMP kelas VII, yaitu tentang interaksi makhluk hidup dengan lingkungan, seperti

pemanasan global, pencemaran lingkungan, serta keseimbangan pada ekosistem. Oleh karena itu, peneliti mengembangkan suatu instrumen asesmen berpikir kritis untuk siswa SMP kelas VII pada materi interaksi makhluk hidup dengan lingkungan.

Berdasarkan latar belakang yang telah diuraikan, tujuan penelitian dan pengembangan ini yaitu menghasilkan instrumen asesmen berpikir kritis untuk siswa SMP kelas VII pada materi interaksi makhluk hidup dengan lingkungan yang valid dan reliabel. Penelitian ini mempunyai manfaat antara lain bagi guru sebagai alternatif alat ukur yang dapat digunakan untuk mengukur kemampuan berpikir kritis siswa berkaitan dengan materi interaksi makhluk hidup dengan lingkungan, bagi siswa sebagai sumber belajar untuk dapat melatih dan meningkatkan kemampuan berpikir kritis, serta bagi sekolah sebagai bahan rujukan dalam mengembangkan instrumen asesmen berpikir kritis pada materi pelajaran yang lain.

METODE

Penelitian ini merupakan penelitian pengembangan yang dirancang untuk menghasilkan produk berupa instrumen asesmen berpikir kritis pada mata pelajaran IPA untuk siswa SMP kelas VII pada materi interaksi makhluk hidup dengan lingkungan. Model pengembangan yang dipilih adalah model *Research and Development (R & D)* menurut Borg & Gall. Model *Research and Development (R & D)* menurut Borg dan Gall (1983:775) terdiri atas 8 langkah, yaitu (1) *research and information collecting* (penelitian dan pengumpulan informasi); (2) *planning* (perencanaan); (3) *develop preliminary form of product* (pengembangan draf produk); (4) *preliminary field testing* (uji coba lapangan awal); (5) *main product revision* (revisi hasil uji coba); (6) *main field testing* (uji coba lapangan); (7) *operational product revision* (penyempurnaan hasil uji coba lapangan); (8) *final product revision* (penyempurnaan produk akhir).

Pada tahap penelitian dan pengumpulan informasi sering disebut sebagai tahap analisis kebutuhan. Pada tahap ini dilakukan pengumpulan berbagai informasi yang dilakukan melalui studi literatur dan studi lapangan. Pada tahap perencanaan dilakukan penyusunan rencana penelitian, meliputi kemampuan yang diperlukan dalam pelaksanaan penelitian, rumusan tujuan yang hendak dicapai, desain penelitian, dan kemungkinan pengujian dalam lingkup yang terbatas dan luas.

Pada tahap pengembangan draf produk, kegiatan yang dilakukan pertama-tama yaitu penyusunan draf produk yang meliputi penyusunan kisi-kisi butir soal, butir-butir soal, dan rubrik asesmen berpikir kritis. Kemudian dilakukan penyusunan instrumen validasi untuk menetapkan kriteria-kriteria yang baik ditinjau dari tiga aspek, yaitu materi, konstruksi, dan bahasa dalam pengembangan instrumen asesmen berpikir kritis. Instrumen validasi disusun dalam bentuk skala Likert dengan rentang skor 1 hingga 5. Instrumen validasi juga dilengkapi dengan kolom komentar dan saran sehingga dapat digunakan oleh validator untuk memberikan komentar dan sarannya terhadap produk hasil pengembangan.

Validasi isi dan konstruk dilakukan oleh tiga orang dosen Universitas Negeri Malang dan satu orang guru IPA SMP. Setelah proses validasi, kemudian merevisi draf produk yang didasarkan dari hasil validasi ahli yang telah dilakukan. Komentar dan saran akan diperhatikan dalam kegiatan revisi. Revisi perlu dilakukan jika hasil validasi menunjukkan bahwa draf produk hasil pengembangan dinilai belum mencapai tingkat kelayakan yang diharapkan. Draft produk divalidasi kembali oleh ahli, setelah benar-benar dinyatakan layak oleh ahli maka produk dapat digunakan dalam uji coba skala terbatas.

Uji coba produk terbatas asesmen berpikir kritis pada materi interaksi makhluk hidup dengan lingkungan yang telah melalui proses validasi dan revisi selanjutnya diujicobakan terhadap 34 siswa pada salah satu sekolah negeri di Kabupaten Blitar. Pada tahap uji coba produk terbatas ini, siswa diberikan angket untuk mengetahui tanggapannya terhadap soal-soal asesmen berpikir kritis yang diberikan. Setelah dilakukan uji coba produk terbatas, kemudian dilakukan revisi berdasarkan hasil tanggapan siswa. Berdasarkan hasil tanggapan siswa, dilakukan penyempurnaan untuk memperbaiki kekurangan yang ada, kemudian asesmen berpikir kritis dapat digunakan dalam uji coba produk lapangan atau skala luas.

Uji coba produk skala luas dilakukan pada dua sekolah negeri di Kabupaten Blitar. Pada tahap ini, soal berpikir kritis diuji kualitas soalnya dengan menggunakan uji validitas, reliabilitas, daya beda soal, dan tingkat kesukaran soal. Revisi produk dilakukan berdasarkan hasil ujicoba produk lapangan. Produk direvisi kembali untuk melihat beberapa kekurangan dan segera diperbaiki sehingga dapat digunakan untuk penyempurnaan instrumen asesmen berpikir kritis yang telah dikembangkan.

Data, Instrumen, dan Teknik Pengumpulan Data

Data dalam penelitian ini meliputi data kualitatif dan kuantitatif. Data-data ini bertujuan untuk memberi gambaran mengenai kualitas produk yang dikembangkan. Data kualitatif diperoleh dari hasil validasi produk awal instrumen asesmen berpikir kritis dan pelaksanaan kegiatan pada uji coba produk skala terbatas, sedangkan data kuantitatif diperoleh dari uji coba produk lapangan. Instrumen penelitian diklasifikasikan menjadi dua macam yang masing-masing digunakan untuk memenuhi kriteria valid dan reliabel.

Instrumen untuk mengukur kevalidan digunakan lembar validasi yang dianalisis secara kualitatif. Soal tes valid atau layak digunakan berdasarkan penilaian dari validator. Kriteria validasi produk menggunakan lima pilihan yaitu sangat layak digunakan (nilai 5), layak (nilai 4), cukup layak (nilai 3), kurang layak (nilai 2), dan tidak layak (nilai 1). Soal tes berpikir kritis disusun berdasarkan 12 indikator berpikir kritis.

Teknik pengumpulan data yang digunakan peneliti adalah sebagai berikut: (1) menyusun instrumen yang akan digunakan dalam penelitian, seperti kisi-kisi soal asesmen berpikir kritis, kunci jawaban dan pedoman penyekoran, (2) menentukan validitas isi instrumen dengan *expert judgment* atau meminta beberapa ahli untuk memvalidasi instrumen yang telah dibuat, (3) melakukan revisi instrumen sesuai dengan saran validator, (4) melakukan uji coba instrumen penelitian, (5) menentukan validitas, reliabilitas, tingkat kesukaran, dan daya pembeda butir soal, (6) melakukan revisi instrumen berdasarkan analisis hasil uji coba.

Teknik Analisis Data

Analisis kualitatif soal asesmen berpikir kritis diperoleh dari hasil lembar validasi yang dilakukan dengan cara deskriptif kualitatif dan catatan lapangan saat pelaksanaan kegiatan pembelajaran pada uji coba produk terbatas.

Analisis Kuantitatif Soal Tes Berpikir Kritis

Analisis butir soal untuk mengetahui tingkat validitas soal pilihan ganda menggunakan koefisien korelasi biseral dan uji reliabilitasnya menggunakan KR-20. Sedangkan analisis untuk soal esai bagian pertama tingkat validitasnya menggunakan koefisien korelasi antar skor dan uji reliabilitasnya menggunakan Cronbach Alfa. Pada soal asesmen bagian kedua untuk mengetahui tingkat validitas dan reliabilitasnya menggunakan koefisien Kappa atau kesepakatan antar rater (*inter-rater reliability*).

HASIL

Deskripsi Hasil Pengembangan

Produk yang dihasilkan dalam penelitian dan pengembangan ini berupa instrumen asesmen berpikir kritis pada materi interaksi makhluk hidup dengan lingkungan. Instrumen asesmen ini terdiri atas bagian pendahuluan, bagian isi, dan bagian penutup. Bagian pendahuluan dari instrumen asesmen yang dikembangkan ini terdiri atas petunjuk penggunaan instrumen asesmen dan indikator-indikator berpikir kritis beserta penjelasannya. Bagian isi terdiri atas kisi-kisi instrumen asesmen, butir-butir soal berpikir kritis, rubrik penilaian, dan pedoman pengolahan skor. Kisi-kisi instrumen asesmen merupakan acuan untuk mengembangkan butir-butir soal yang ada dalam instrumen asesmen berpikir kritis. Butir-butir soal berpikir kritis dikembangkan dengan memerhatikan indikator berpikir kritis yang telah dijelaskan pada bagian awal instrumen asesmen. Butir-butir soal dalam instrumen asesmen ini dibagi menjadi dua bagian, yaitu bagian pertama dan kedua. Soal-soal bagian pertama dapat digunakan setelah materi interaksi makhluk hidup dengan lingkungan selesai diajarkan pada siswa. Sedangkan soal-soal yang ada pada bagian kedua dapat digunakan dalam proses pembelajaran. Adapun sebaran indikator berpikir kritis dalam soal-soal yang telah dikembangkan diuraikan pada Tabel 1. Rubrik penilaian terdiri atas kunci jawaban soal pilihan ganda dan esai yang dilengkapi dengan pedoman penyekoran. Pada bagian akhir disertakan penjelasan dari pedoman penyekoran yang digunakan dalam soal esai pada asesmen berpikir kritis. Pada bagian penutup berisi daftar rujukan yang digunakan dalam mengembangkan soal-soal asesmen.

Tabel 1. Sebaran Indikator Berpikir Kritis dalam Soal-soal Berpikir Kritis pada Materi Interaksi Makhluk Hidup dengan Lingkungan

No.	Indikator Berpikir Kritis	Jumlah Soal	
		Pilihan Ganda	Esai
1.	Memfokuskan pertanyaan	1 soal	-
2.	Menganalisis argumen	4 soal	6 soal
3.	Bertanya dan menjawab pertanyaan yang membutuhkan penjelasan	3 soal	-
4.	Menilai kredibilitas suatu sumber	-	5 soal
5.	Mengobservasi dan mempertimbangkan hasil observasi	1 soal	2 soal
6.	Melakukan deduksi dan menilai hasil deduksi	2 soal	-
7.	Melakukan induksi	2 soal	-
8.	Membuat dan mempertimbangkan nilai keputusan	3 soal	-
9.	Mendefinisikan istilah dan mempertimbangkan definisi menggunakan kriteria yang tepat	1 soal	-
10.	Mengidentifikasi asumsi	1 soal	3 soal
11.	Memutuskan suatu tindakan	1 soal	3 soal
12.	Berinteraksi dengan yang lain	1 soal	-



(Sumber: modifikasi PISA, 2015)

Gambar 1. Persentase runtuhnya koloni lebah

Gangguan Runtuhnya Koloni Lebah

Peneliti percaya bahwa ada beberapa penyebab gangguan runtuhnya koloni lebah. Salah satu penyebabnya adalah insektisida imidakloprid yang dapat menyebabkan lebah akan kehilangan arah ketika berada di luar sarang. Mereka menguji apakah paparan imidakloprid menyebabkan runtuhnya koloni. Mereka menambahkan insektisida pada berbagai sarang lebah selama tiga minggu dengan konsentrasi yang berbeda (takaran dalam mikrogram insektisida per kilogram makanan). Ada beberapa sarang tidak terpapar insektisida. Peneliti menemukan bahwa tidak ada koloni lebah yang langsung terjatuh setelah terpapar insektisida. Namun pada minggu ke 14, beberapa sarang telah ditinggalkan. Hasil pengamatan digambarkan dalam grafik berikut.

Manakah dari salah satu kesimpulan berikut yang sesuai dengan hasil yang ditunjukkan pada grafik tersebut?

- Koloni yang terpapar imidakloprid dengan konsentrasi lebih tinggi cenderung runtuh lebih cepat
- Koloni yang terpapar imidakloprid runtuh dalam waktu 10 minggu jumlahnya lebih banyak
- Paparan imidakloprid pada konsentrasi di bawah 20 µg/kg tidak membahayakan koloni
- Koloni yang terpapar imidakloprid tidak dapat bertahan hidup selama lebih dari 14 minggu

Tabel 2. Penggunaan Bahan Bakar Fosil

Sumber Bahan Bakar	Energi yang Dilepaskan (kJ Energi /g Bahan Bakar)	Karbon dioksida yang Dilepaskan (mg CO ₂ /kJ Energi yang Dihasilkan Oleh Bahan Bakar)
Minyak bumi	43,6	78
Etanol	27,3	59

(Sumber: modifikasi PISA, 2015)

Banyak pembangkit listrik menggunakan bahan bakar berbasis karbon dan memancarkan karbon dioksida (CO₂). CO₂ yang dilepaskan ke atmosfer memiliki dampak negatif pada iklim global. Insinyur telah menggunakan strategi yang berbeda untuk mengurangi jumlah CO₂ yang dilepaskan ke atmosfer. Salah satu strategi tersebut adalah untuk membakar bahan bakar hayati sebagai pengganti bahan bakar fosil. Bahan bakar fosil berasal dari makhluk hidup yang telah lama mati, sedangkan bahan bakar hayati berasal dari tanaman yang hidup dan baru saja mati. Meskipun bahan bakar hayati memiliki kelebihan bagi lingkungan, bahan bakar fosil masih banyak digunakan.

Tabel berikut membandingkan jumlah energi dan CO₂ yang dilepaskan pada saat minyak bumi dan etanol dibakar. Minyak bumi adalah bahan bakar fosil, sedangkan etanol adalah bahan bakar hayati. Berdasarkan tabel, kemukakan alasanmu mengapa seseorang lebih memilih menggunakan minyak bumi bukan etanol, meskipun harganya sama?

Salah satu butir soal esai bagian kedua pada asesmen berpikir kritis

Indikator Berpikir Kritis: Menilai Kredibilitas Suatu Sumber

Deskriptor: Kredibilitas berdasarkan sumber, penulis, tahun penulisan dan korelevanan dengan tema yang diberikan.

Hasil Uji Coba Produk

Validasi oleh ahli dilakukan untuk melihat isi produk awal. Validasi ini bertujuan untuk mendapatkan masukan, saran perbaikan, dan sekaligus penilaian terhadap produk awal sebelum dilakukan uji coba terbatas. Kegiatan validasi dilakukan dengan cara memberikan naskah produk awal kepada empat orang validator. Selanjutnya dilakukan analisis terhadap penilaian dari validator. Data hasil penilaian oleh validator terhadap produk awal disajikan pada Tabel 2.

Tabel 2. Data Hasil Validasi Ahli

NO.	Aspek Penilaian	Persentase (%) Rata-rata Tiap Aspek	Kriteria
1.	Petunjuk penggunaan	90	Sangat layak
2.	Tampilan dan tata letak	85	Sangat layak
3.	Keterterapan	75	Layak
4.	Penggunaan bahasa	91,67	Sangat layak
5.	Kisi-kisi	90	Sangat layak
6.	Isi	90,79	Sangat layak
7.	Rubrik penilaian dan kunci jawaban	96	Sangat layak
	Rata-rata aspek	88,35	Sangat layak

Berdasarkan hasil validasi yang dilakukan oleh empat orang validator diperoleh nilai persentase rata-rata untuk seluruh aspek penilaian produk sebesar 88,35% berada pada kriteria sangat layak, sehingga dapat disimpulkan bahwa produk instrumen asesmen berpikir kritis telah dapat digunakan. Akan tetapi, terdapat beberapa komentar dan saran yang diberikan oleh validator dan digunakan peneliti sebagai perbaikan instrumen asesmen berpikir kritis yang dikembangkan, di antaranya pada petunjuk penggunaan cara mengerjakan soal pilihan ganda, tampilan dan tata letak pada sampul, penggunaan bahasa asing yang lebih diperjelas, isi pada soal pilihan ganda dan esai yang lebih disesuaikan dengan konsep materi dan indikator berpikir kritis, serta pada rubrik penilaian dan kunci jawaban soal esai yang perlu diperbaiki.

Pada uji coba produk terbatas instrumen asesmen berpikir kritis ini, yang dilakukan analisis butir soal yaitu hanya soal asesmen bagian kedua, yaitu dengan menggunakan reliabilitas yang melibatkan rater atau penyekor, jadi posisi butir soal digantikan dengan posisi orang. Pada analisis reliabilitas antar rater (*inter-rater reliability*) yang diuji konsistensinya adalah raternya. Untuk soal asesmen bagian kedua ini melibatkan dua orang rater. Berikut disajikan hasil *inter-rater reliability* untuk masing-masing butir soal bagian kedua asesmen berpikir kritis untuk uji coba produk terbatas pada Tabel 3.

Tabel 3. Nilai *inter-rater reliability* masing-masing butir soal

No. Butir Soal	Nilai <i>Inter-rater reliability</i>	Kategori Kekuatan Kesepakatan
1	0,871	Sangat Baik
2	0,960	Sangat Baik
3a	1,000	Sangat Baik
3b	0,958	Sangat Baik
4	1,000	Sangat Baik
5	0,959	Sangat Baik
6	0,924	Sangat Baik
7	0,938	Sangat Baik
Rata-rata	0,951	Sangat Baik

(Kategori menurut Fleiss, 2003)

Berdasarkan hasil *inter-rater reliability* tersebut, maka dapat disimpulkan bahwa secara keseluruhan rata-rata koefisien kesepakatan antar rater adalah 0,951 dengan kategori sangat baik. Pada uji coba terbatas ini siswa juga diberikan angket untuk mengetahui respon/tanggapannya terhadap soal asesmen berpikir kritis yang telah diberikan. Persentase respon siswa terhadap keterbacaan soal asesmen berpikir kritis secara keseluruhan pada setiap item pertanyaan dirangkum pada Tabel 4.

Tabel 4. Hasil Respon Siswa terhadap Keterbacaan Instrumen Asesmen Bepikir Kritis

NO.	Item Pertanyaan	Persentase (%) Perolehan Siswa Tiap Item	Kategori
1.	Apakah bahasa yang digunakan pada soal asesmen mudah dipahami?	85,29	Sangat Baik
2.	Apakah gambar yang digunakan pada soal asesmen terlihat jelas dan menarik?	100	Sangat Baik
3.	Apakah struktur kalimat yang digunakan pada soal asesmen jelas?	94,12	Sangat Baik
4.	Apakah istilah yang terdapat pada soal asesmen mudah dipahami?	88,24	Sangat Baik
5.	Apakah isi soal asesmen sesuai dengan materi yang diajarkan oleh guru?	100	Sangat Baik
	Rata-rata	93,51	Sangat Baik

(Kategori menurut Riduwan, 2011)

Rata-rata hasil respon siswa terhadap keterbacaan soal asesmen berpikir kritis yang diberikan, yaitu sebesar 93,51% dengan kategori sangat baik. Meskipun demikian terdapat saran dan komentar dari siswa yang digunakan oleh peneliti untuk merevisi soal asesmen, yaitu pada soal esai nomor 6 bagian pertama bahwa struktur kalimat pada pertanyaan diperbaiki karena membingungkan.

Setelah melakukan revisi berdasarkan hasil uji coba produk terbatas, kemudian dilakukan uji coba produk skala luas. Uji coba produk skala luas dilakukan pada 75 siswa yang berasal dari dua sekolah di Kabupaten Blitar. Pada uji coba produk skala luas ini, soal asesmen yang diuji cobakan yaitu soal bagian pertama yang terdiri dari 20 nomor soal pilihan ganda dan 10 nomor soal esai. Analisis data uji coba produk lapangan menggunakan parameter teori tes klasik yang terdiri dari validitas, reliabilitas, daya beda, dan tingkat kesukaran butir soal, untuk mengetahui kualitas butir soal asesmen berpikir kritis secara empiris sebagai dasar untuk revisi produk pengembangan.

Pengujian validitas instrumen dilakukan dengan menggunakan metode Korelasi Poin Biserial untuk soal pilihan ganda dan Korelasi Product Moment untuk soal esai. Hasil uji validitas soal pilihan ganda dapat dilihat pada Tabel 5 dan hasil validasi untuk soal esai dapat dilihat pada Tabel 6.

Tabel 5. Hasil Uji Validitas Butir Soal Pilihan Ganda

Nomor Butir Soal	Kategori	Jumlah	Persentase (%)
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20.	Valid	17	90
17, 19	Tidak Valid	2	10

Berdasarkan Tabel 4 dapat diketahui bahwa hasil uji validitas instrumen yang terdiri dari 20 soal pilihan ganda diperoleh 17 soal atau 85% adalah valid, sedangkan yang tidak valid adalah 3 soal atau 15%.

Tabel 6. Hasil Uji Validitas Butir Soal Esai

Nomor Butir Soal	Kategori	Jumlah	Persentase (%)
1, 2, 3, 4, 5, 6, 7, 8, 11	Valid	9	81,82
9a, 9b	Tidak Valid	2	18,18

Hasil uji validitas instrumen yang terdiri dari 11 soal esai diperoleh 9 soal atau 81,82% adalah valid, sedangkan yang tidak valid adalah 2 soal atau 18,18%. Pengujian reliabilitas instrumen dilakukan dengan menggunakan metode KR-20 untuk soal pilihan ganda dan Cronbach's Alpha untuk soal esai yaitu dengan melihat kriteria Indeks Koefisien Reliabilitas. Hasil uji reliabilitas soal pilihan ganda dan esai dapat dilihat pada Tabel 7.

Tabel 7. Uji Reliabilitas Instrumen

Soal	Koefisien Reliabilitas	Kriteria
Pilihan Ganda	0.792	Tinggi
Esai	0.753	Tinggi

(Kategori menurut Frankel, 2012)

Hasil analisis daya pembeda menunjukkan bahwa dari 20 soal pilihan ganda yang diujikan, 5 soal atau 25% mempunyai daya pembeda yang baik sekali, 11 soal atau 55% mempunyai daya pembeda yang baik, 2 soal atau 10% mempunyai daya pembeda cukup, serta 2 soal atau 10% mempunyai daya pembeda jelek. Sementara itu, untuk soal esai dari 11 soal yang diujikan, terdapat 3 soal atau 27,3% mempunyai daya pembeda baik sekali, 5 soal atau 45,5% mempunyai daya pembeda baik, 2 soal atau 18,2% mempunyai daya pembeda cukup, serta 1 soal atau 9,1% mempunyai daya pembeda jelek. Hasil analisis tingkat kesukaran menunjukkan bahwa dari 20 soal pilihan ganda yang diujikan, 3 soal atau 15% adalah soal sukar, 15 soal atau 75% adalah soal sedang, serta 2 soal atau 10% adalah soal mudah. Sedangkan untuk soal esai dari 11 soal yang diujikan, terdapat 3 soal atau 27,3% adalah soal sukar, 6 soal atau 54,5% adalah soal sedang, serta 2 soal atau 18,2% adalah soal mudah.

Revisi Produk

Revisi produk dilakukan untuk memperoleh produk akhir yang memenuhi kriteria valid dan reliabel. Revisi dilakukan berdasarkan hasil penilaian dan analisis instrumen asesmen pada setiap tahap uji coba produk. Revisi produk pada penelitian ini terdiri atas revisi produk hasil validasi, revisi produk hasil uji coba terbatas, dan revisi produk hasil uji coba lapangan.

Berdasarkan hasil evaluasi setelah penilaian ahli, uji coba terbatas, dan uji coba lapangan, instrumen asesmen yang dikembangkan mengalami beberapa revisi. *Pertama*, revisi butir soal didasarkan masukan dan saran validator. Secara umum, masukan dan saran tersebut mengenai perbaikan pada petunjuk penggunaan cara mengerjakan soal pilihan ganda, tampilan dan tata letak pada sampul, penggunaan bahasa asing yang lebih diperjelas, isi pada soal pilihan ganda dan esai yang lebih disesuaikan dengan konsep materi dan indikator berpikir kritis, serta pada rubrik penilaian dan kunci jawaban soal esai yang perlu diperbaiki. *Kedua*, revisi butir soal berdasarkan hasil ujicoba terbatas yang diperoleh dari respon siswa terhadap soal asesmen. *Ketiga*, revisi butir soal berdasarkan hasil uji coba lapangan produk utama soal asesmen.

Berdasarkan hasil analisis data pada uji coba produk lapangan, maka perlu dilakukan beberapa revisi pada produk instrumen asesmen. Revisi yang dilakukan yaitu pada butir soal yang tidak valid, tidak perlu digunakan lagi dan dikeluarkan dari produk pengembangan. Butir soal yang dikeluarkan adalah soal pilihan ganda nomor 8, 17 dan 19, serta soal esai nomor 9a dan 9b.

Kajian Produk Akhir

Produk hasil pengembangan merupakan instrumen asesmen berpikir kritis untuk siswa SMP kelas VII pada materi interaksi makhluk hidup dengan lingkungan. Instrumen asesmen yang dihasilkan terdiri atas petunjuk penggunaan, indikator berpikir kritis beserta penjelasannya, kisi-kisi butir soal, butir-butir soal yaitu pilihan ganda dan esai, kunci jawaban dan pedoman penyekoran, serta rubrik penilaian. Instrumen asesmen berisi soal-soal tentang materi interaksi makhluk hidup dengan lingkungan, dikembangkan berdasarkan indikator berpikir kritis menurut Ennis (1985; 1996; 2011), yaitu (1) memfokuskan pertanyaan; (2) menganalisis argumen; (3) bertanya dan menjawab pertanyaan yang membutuhkan penjelasan; (4) menilai kredibilitas suatu sumber; (5) mengobservasi dan mempertimbangkan hasil observasi; (6) melakukan deduksi dan menilai hasil deduksi; (7) melakukan induksi; (8) membuat dan mempertimbangkan nilai keputusan; (9) mendefinisikan istilah dan mempertimbangkan definisi menggunakan kriteria yang tepat; (10) mengidentifikasi asumsi; (11) memutuskan suatu tindakan; (12) berinteraksi dengan yang lain.

Selain itu, kriteria suatu instrumen asesmen yang baik yaitu mempertimbangkan aspek validitas dan reliabilitas. Sebuah instrumen asesmen yang baik harus memiliki tingkat kestabilan dan konsistensi data (Creswell, 2015). Berdasarkan hasil validasi ahli, uji coba terbatas, uji coba lapangan, perbaikan-perbaikan, dan analisis data yang dilakukan dapat diketahui bahwa instrumen asesmen berpikir kritis yang dikembangkan telah memenuhi kriteria valid dan reliabel, memiliki kriteria tingkat keterbacaan yang sangat baik, serta kualitas butir soal yang baik. Soal instrumen asesmen terdiri atas dua bagian, yaitu bagian pertama terdiri atas 20 soal pilihan ganda dan 11 soal esai, bagian kedua terdiri dari 7 soal esai.

Validasi yang dilakukan meliputi validasi isi, validasi konstruk, dan validasi butir soal. Kevalidan instrumen asesmen berpikir kritis berdasarkan analisis hasil validasi isi dan konstruk yang dilakukan oleh empat orang ahli yang terdiri dari tiga dosen Universitas Negeri Malang dan satu orang guru IPA SMP, dinyatakan bahwa produk instrumen asesmen berpikir kritis yang dikembangkan memperoleh persentase 88,35% dan memenuhi kriteria sangat layak untuk digunakan.

Berdasarkan hasil analisis validasi butir soal diperoleh bahwa dari 20 soal pilihan ganda yang diujicobakan, terdapat 2 soal yang tidak valid, sedangkan dari 11 butir soal esai terdapat 2 soal yang tidak valid. Oleh karena itu, jumlah butir soal yang ada dalam instrumen asesmen berpikir kritis yang dihasilkan yaitu 18 soal pilihan ganda dan 9 soal esai.

Reliabilitas instrumen asesmen yang berupa soal tes berpikir kritis didasarkan pada analisis butir soal produk hasil pengembangan. Pada soal asesmen bagian kedua analisis reliabilitas butir soal menggunakan reliabilitas inter-rater (*inter-rater reliability*), diperoleh nilai koefisien sebesar 0,951. Analisis butir soal asesmen berpikir kritis bagian pertama diperoleh rata-rata koefisien reliabilitas yang tinggi pada masing-masing sekolah uji coba, baik untuk soal pilihan ganda maupun esai.

SIMPULAN DAN SARAN

Simpulan

Berdasarkan hasil penelitian dan pembahasan diperoleh simpulan sebagai berikut. *Pertama*, produk akhir dalam penelitian ini menghasilkan instrumen asesmen berpikir kritis untuk siswa SMP kelas VII pada materi interaksi makhluk hidup dengan lingkungan yang valid dan reliabel. *Kedua*, instrumen asesmen yang dihasilkan terdiri atas bagian pendahuluan, yaitu petunjuk penggunaan instrumen asesmen dan indikator-indikator berpikir kritis beserta penjelasannya. Bagian isi terdiri atas kisi-kisi instrumen asesmen, butir-butir soal berpikir kritis, rubrik penilaian, dan pedoman pengolahan skor, serta bagian penutup. Instrumen asesmen dikembangkan melalui langkah-langkah pengembangan, yaitu (a) penelitian dan pengumpulan informasi, (b) perencanaan, (c) pengembangan produk awal, (d) uji coba terbatas, (e) revisi produk uji coba terbatas, (f) uji coba lapangan, dan (g) revisi produk akhir. Kevalidan instrumen dibuktikan dengan hasil penilaian ahli yang menunjukkan bahwa instrumen sangat layak digunakan berdasarkan telaah aspek materi dan konstruksi. Butir soal pada instrumen asesmen juga telah memenuhi kriteria valid dan reliabel.

Saran

Berdasarkan hasil penelitian dan simpulan di atas, ada beberapa saran pemanfaatan produk akhir instrumen asesmen berpikir kritis, yaitu (1) siswa dapat menggunakan produk akhir instrumen asesmen berpikir kritis sebagai bahan latihan untuk melatih keterampilan berpikir kritis, (2) guru IPA SMP dapat menggunakan produk akhir instrumen asesmen berpikir kritis untuk mengukur penguasaan pengetahuan dan keterampilan berpikir kritis siswa, (3) produk akhir instrumen asesmen berpikir kritis hasil pengembangan dapat digunakan sebagai acuan dalam mengembangkan instrumen asesmen berpikir kritis pada materi pelajaran IPA lainnya.

DAFTAR RUJUKAN

- Borg, W. R. & Gall, M.D. 1983. *Educational Research: An Introduction Fourth Edition*. New York: Longman Inc.
- Budiman, A & Jailani. 2014. *Pengembangan Instrumen Asesmen Higher Order Thinking Skill (HOTS) pada Mata Pelajaran Matematika SMP Kelas VIII Semester 1*. Jurnal Riset Pendidikan Matematika Volume 1-Nomor 2, (online) (<http://journal.uny.ac.id/index.php/jrpm/article/viewFile/2671/2224>), diakses 5 Maret 2015.
- Ennis, R. H. 1985. *A Logical Basis for Measuring Critical Thinking Skills*, (online), (http://www.ascd.org/ASCD/pdf/journal/ed_lead/el_1985_ennis.pdf), diakses 20 Maret 2015.
- Ennis, R.H. 1996. *Critical Thinking Disposition: Their Nature and Assessability Informal Logic*, (Online). 18 (2&3):165—182, (<http://celt.uwindsor.ca>), diakses 7 April 2015.
- Ennis, R.H. 2011. *The Nature of Critical Thinking: An Outline of Critical Thinking Disposition and Abilities*. Chicago: University of Illinois.
- Fisher, A. 2009. *Berpikir Kritis Sebuah Pengantar*. Jakarta: Penerbit Erlangga.
- Fleiss, J. L., Levin, B. & Cho Paik, M. 2003. *Statistical Methods for Rates and Proportions, Third Edition*. John Wiley & Sons, Inc.
- Higgins, S., Hall, E., Baumfield, V., & Moseley, D. 2007. *Learning skills and the development of learning capabilities*. Technical Report. In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. (Online), (<https://eppi.ioe.ac.uk>), diakses 15 September 2015.
- Kemdikbud. 2015. *Permendikbud No. 53 Tahun 2015 tentang Penilaian Hasil Belajar oleh Pendidik dan Satuan Pendidikan Pada Pendidikan Dasar dan Pendidikan Menengah*. Jakarta.
- Kemendikbud. 2013. *Permendikbud No.64 Tahun 2013 tentang Standar Isi Pendidikan Dasar dan Menengah*. Jakarta.
- Kemendiknas. 2007. *Permendiknas No. 16 Tahun 2007 tentang Standar Kualifikasi Akademik dan Kompetensi Guru Mata Pelajaran*. Jakarta.
- Mullis, I.V.S, Martin, M.O. & Stanco, G.M. 2012. *TIMSS 2011 International Results in Science*. Boston: TIMSS & PIRLS International Study Center. Dari IEA Sekretariat Library, (Online), (<http://timss.bc.edu>), diakses 4 September 2015.
- Newmann, F.M, Bryk, A.S. & Nagaoka, J.K. 2001. *Authentic Intellectual Work and Standardized Tests: Conflict or Coexistence?*, (online) (<http://ccsr.uchicago.edu>), diakses 6 September 2015.
- Riduwan. 2013. *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Penerbit Alfabeta.
- Stobaugh, R. 2013. *Assessing Critical Thinking in Middle and High Schools*. New York London: Routledge Taylor & Francis Group.
- Van de Walle, J. A. 2007. *Elementary and Middle Schoolmathematics: Teaching Developmentally, Sixth Edition*. United States of America: Pearson Education, Inc.
- Wardhani, S. 2011. *Instrumen Penilaian Hasil Belajar Matematika SMP: Belajar dari PISA dan TIMSS*. Badan Pengembangan Sumber Daya Manusia Pendidikan dan Penjaminan Mutu Pendidikan.