

Kelayakan Instrumen Penilaian Higher Order Thinking Skills pada Materi Lingkungan

Kusuma Wardany¹, Sulis Anjarwati²
 Universitas Nahdlatul Ulama Lampung¹²
kusuma.wardany@gmail.com¹, sulisanjarwati@gmail.com²

ABSTRAK

Penilaian adalah alat ukur yang digunakan dalam penguasaan Higher Order Thinking Skills (HOTS) pada materi lingkungan. Penelitian pengembangan instrumen penilaian ini bertujuan untuk mengetahui kelayakan penilaian Higher Order Thinking Skills pada materi lingkungan di SMA Surakarta. Penelitian pengembangan penilaian Higher Order Thinking Skills (HOTS) menggunakan model Research and Development (R&D) mengacu pada Borg and Gall (1983) yang telah dimodifikasi. Sampel untuk menguji kelayakan penilaian meliputi 2 validator ahli, yaitu ahli materi dan ahli instrumen evaluasi. Fungsi ahli materi dan ahli instrumen evaluasi adalah untuk memvalidasi kesesuaian soal dengan indikator (validitas konstruk) sehingga menjamin soal yang dikembangkan dapat mengukur Higher Order Thinking Skills siswa. Hasil validasi ahli materi menunjukkan skor rata-rata keseluruhan aspek materi adalah 100% dengan kategori "Sangat Baik". Hasil validasi ahli instrumen evaluasi menunjukkan skor rata-rata keseluruhan aspek materi adalah 80,95% dengan kategori "Baik", sehingga penilaian Higher Order Thinking Skills (HOTS) pada materi lingkungan dinilai layak untuk diujikan pada kelompok uji terbatas maupun uji lapangan terhadap siswa.

Kata Kunci: Kelayakan, Instrumen, Penilaian, Higher Order Thinking Skills (HOTS)

ABSTRACT

Assessment is a measuring tool used in mastery of Higher Order Thinking Skills (HOTS) in environmental materials. Research on the development of this assessment instrument aims to determine the feasibility of assessing Higher Order Thinking Skills on environmental materials in SMA Surakarta. Research on the development of the Higher Order Thinking Skills (HOTS) assessment using the Research and Development (R&D) model refers to the modified Borg and Gall (1983). The sample to test the feasibility of the assessment includes 2 expert validators, namely material experts and evaluation instrument experts. The function of material experts and evaluation instrument experts is to validate the suitability of questions with indicators (construct validity) so as to ensure that the questions developed can measure students' Higher Order Thinking Skills. The results of the material expert validation show that the average score of all aspects of the material is 100% in the "Very Good" category. The results of the evaluation instrument expert validation show that the average score of all aspects of the material is 80.95% in the "Good" category, so that the Higher Order Thinking Skills (HOTS) assessment on environmental material is considered feasible to be tested in the limited test group or field tests on students.

Keywords: Feasibility, Instruments, Assessment, Higher Order Thinking Skills (HOTS)

PENDAHULUAN

Permendikbud No. 81 Tahun 2013 tentang implementasi kurikulum menyebutkan bahwa kebutuhan kompetisi masa depan diperlukan Higher Order Thinking Skills secara kritis, keterampilan komunikasi, dan kreatif (Kemendikbud, 2013). Hal ini sejalan dengan karakteristik skills masyarakat abad ke-21 yang dipublikasikan oleh Partnership of 21st Century Skill mengidentifikasi bahwa pebelajar pada abad ke-21 harus mampu mengembangkan keterampilan kompetitif yang diperlukan pada abad ke-21 yang berfokus pada pengembangan Higher Order Thinking Skills, seperti: berpikir kritis (*critical thinking*), pemecahan masalah (*problem solving*), keterampilan berkomunikasi (*communication skills*), melek TIK, teknologi informasi dan komunikasi (ICT, *information and Communication Technology*), melek informasi (*information literacy*), dan melek media (*media literacy*) (Basuki & Haryanto, 2012).

Higher Order Thinking Skills merupakan suatu keterampilan berpikir yang tidak hanya membutuhkan keterampilan mengingat, tetapi membutuhkan keterampilan lain yang lebih tinggi. Indikator untuk mengukur Higher Order Thinking Skills meliputi keterampilan menganalisa (C4), mengevaluasi (C5), dan menciptakan (C6) (Anderson & Krathwohl, 2001). Higher Order Thinking Skills sebagai keterampilan berpikir yang terjadi ketika seseorang mengambil informasi baru dan informasi yang sudah tersimpan dalam ingatannya, selanjutnya menghubungkan informasi tersebut dan menyampaikannya untuk mencapai tujuan atau jawaban yang dibutuhkan (Lewis & Smith, 1993).

Assesmen atau penilaian merupakan istilah umum yang mencakup keseluruhan prosedur yang digunakan untuk mendapatkan informasi tentang hasil belajar siswa (pengamatan, peringkat, pengujian menggunakan kertas dan pensil) dan membuat penilaian mengenai proses pembelajaran (Gronlund & Linn, 1995). Dalam dunia pendidikan, penilaian diartikan sebagai prosedur yang digunakan untuk mendapatkan informasi untuk mengukur taraf pengetahuan dan keterampilan subjek didik yang hasilnya akan digunakan untuk keperluan evaluasi (Reynold, C.R., Livingston, R. B., & Willson, V, 2010).

Dalam Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan dinyatakan bahwa ulangan adalah proses yang dilakukan untuk mengukur pencapaian kompetensi peserta didik secara berkelanjutan dalam proses pembelajaran, untuk memantau kemajuan, melakukan perbaikan pembelajaran, dan menentukan keberhasilan belajar peserta didik. Cakupan ulangan meliputi seluruh indikator yang mempresentasikan seluruh KD pada periode tersebut. Hasil penelitian tentang pemetaan pendidikan melalui pemenuhan delapan (8) Standar Nasional Pendidikan (SNP) berbagai Sekolah Menengah Atas (SMA) Negeri di Surakarta menunjukkan bahwa capaian pemenuhan standar proses (standar 2) dan standar penilaian (standar 8) menunjukkan adanya gap yang terbesar antara skor ideal dan skor riil, yaitu sebesar 3,70 %. Hal ini terjadi karena ketercapaian kompetensi pembelajaran di SMA Surakarta melalui aspek kognitif belum dapat meningkatkan Higher Order Thinking Skills siswa (Sajidan, Sugiharto & Prasetyani, 2013).

Hasil Observasi SMA di Surakarta yang dipilih secara acak melalui Ulangan dan Ujian seperti Ujian Nasional, Ujian Akhir Semester, Ujian Tengah Semester, Ujian Sekolah, Ulangan Harian, maupun dari buku paket yang guru dan siswa gunakan pada materi lingkungan menunjukkan bahwa soal dan pertanyaan masih dalam ranah kognitif yang rendah (*Lower Order Thinking Skills*). Hasil

rerata presentase pada materi lingkungan di 5 sekolah sebanyak 81,4% Lower Order Thinking Skills, pada ranah kognitif C1 (mengetahui), C2 (memahami), C3 (mengaplikasikan) dan pada dimensi pengetahuan faktual, konseptual dan prosedural serta hanya 18,58 % Higher Order Thinking Skills, baik ranah kognitif C4 (menganalisis) pada dimensi pengetahuan faktual, konseptual dan metakognisi (Lihat tabel lampiran).

Soal-soal penilaian di atas diidentifikasi berdasarkan Taksonomi Bloom yang meliputi dimensi proses berpikir pada dimensi pengetahuan. Dimensi proses berpikir terdiri dari enam kategori, yaitu: C1 (mengingat), C2 (memahami), C3 (mengaplikasikan), C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta). Sedangkan pada dimensi pengetahuan, yaitu fakta, konsep, prosedural, dan metakognisi. Ranah C1-C3 merupakan Lower Order Thinking Skills dan C3-C6 merupakan Higher Order Thinking Skills (Anderson & Krathwohl, 2001).

Penilaian yang mengukur Higher Order Thinking Skills dapat menggunakan bentuk tes subjektif dan tes objektif. Tes subjektif merupakan tes bentuk esai. Tes esai adalah suatu bentuk uraian dengan mempergunakan bahasa sendiri. Dalam tes bentuk esai siswa dituntut untuk berpikir tentang dan mempergunakan apa yang diketahui yang berkenaan dengan pertanyaan yang harus dijawab. Tes objektif merupakan bentuk tes yang terdiri dari tes jawaban benar-salah (true false), pilihan ganda (multiple choice), isian (completion), dan penjodohan (matching). Dari berbagai alat penilaian tertulis, tes memilih jawaban benar salah, isian singkat, menjodohkan dan sebab akibat merupakan alat yang hanya menilai keterampilan berpikir rendah, yaitu keterampilan mengingat (pengetahuan). Tes pilihan ganda (multiple choice) dapat digunakan untuk menilai keterampilan mengingat dan memahami dengan cakupan materi yang luas (Suwandi, 2009).

Untuk menyatakan hasil belajar baik atau buruk, berhasil atau gagal, sukses atau tidaknya sesuatu, maka data harus benar-benar dapat dipercaya/akurat agar ketetapan yang di ambil tidak salah. Jika salah datanya salah pula hasil penilaiannya dan akibatnya salah pula keputusannya. Oleh karena itu diperlukan adanya alat tes yang baik agar data dapat akurat (Subali, 2010). Alat tes yang baik harus memenuhi sejumlah kriteria, antara lain bahwa tes haruslah tidak terlalu mudah dan sebaliknya tidak terlalu sulit. Alat tes yang baik harus dapat dipertanggungjawabkan dari segi kelayakan (appropriateness), kesahihan (validity), keajegan (reliability), ketertafsiran (interpretability), dan kebergunaan (usability) (Suwandi, 2009).

Berdasarkan pernyataan di atas, maka perlu dilakukannya uji kelayakan penilaian yang mengukur kemampuan berpikir tingkat tinggi (Higher Order Thinking Skills) pada materi lingkungan. Uji kelayakan ini bertujuan untuk mengetahui apakah alat tes ini benar-benar akurat dan dipercaya untuk menyatakan hasil belajar siswa berhasil atau tidak. Uji kelayakan pada penilaian ini digunakan untuk mengukur kognitif produk Higher Order Thinking Skills yang dilakukan.

METODE PENELITIAN

Model penelitian dan pengembangan (Research & DevelopmentI) yang dilakukan mengacu pada Borg & Gall (1983). Sedangkan prosedur pengembangan pada penelitian pengembangan penilaian ini adalah dengan mengadakan penelitian pendahuluan yang merupakan suatu analisis kebutuhan yang dilakukan di sekolah untuk mencari permasalahan yang akan diselesaikan melalui produk pengembangan. Pada tahap analisis kebutuhan antara lain yaitu standarisasi yang dilakukan berdasar-

kan pencapaian delapan Standar Nasional Pendidikan; menganalisis proses pembelajaran yang dilakukan guru melalui wawancara pada guru terkait proses pembelajaran yang biasanya dilakukan; menganalisis hasil butir soal ulangan harian, Ujian Tengah Semester (UTS), Ujian Akhir Semester (UAS), Ujian Nasional (UN), serta analisis bank soal yang digunakan guru untuk mengetahui persentase penggunaan tingkat Taksonomi Bloom dalam soal; serta studi pustaka dilakukan dengan membaca buku dan jurnal penunjang. Kemudian dengan membuat desain perencanaan produk awal (prototype) soal materi lingkungan dengan jumlah kurang lebih 25 soal. Soal dibuat menyesuaikan dengan dimensi kognitif Taksonomi Bloom dengan dimensi pengetahuan. Kemudian melakukan penyusunan instrumen validasi soal. Instrumen validasi soal yang digunakan adalah angket yang nantinya diberikan kepada ahli materi, ahli instrument evaluasi.

Validasi Produk

Validasi produk soal yang dilakukan mengacu pada langkah-langkah validasi dalam Borg & Gall (1983) sebagai berikut :

Soal kepada Ahli

Ahli yang digunakan dalam penelitian adalah orang yang memiliki bidang ilmu linear dengan tujuan pengembangan produk. Ahli materi yang dilibatkan adalah dosen ekosistem dan lingkungan. Penelitian pengembangan penilaian melibatkan satu orang ahli materi yakni Prof. Dr. Maridi, M.Pd yang merupakan dosen bidang ilmu lingkungan di Program Studi Pendidikan Sains Pascasarjana Universitas Sebelas Maret. Ahli penilaian yang dilibatkan adalah dosen pengampu mata kuliah penilaian dan evaluasi pembelajaran. Ahli instrumen evaluasi dalam penelitian pengembangan ini adalah Dr. Nonoh Siti Aminah, M.Pd yang merupakan dosen bidang ilmu Penilaian Evaluasi Pendidikan program studi fisika FKIP Universitas Sebelas Maret. Data kelayakan soal diukur menggunakan kriteria kelayakan soal. Data yang diperoleh kemudian dianalisis, jika hasil validasi dari para validator belum layak maka dilakukan revisi terhadap produk sesuai saran validator, selanjutnya dilakukan validasi kembali ke validator.

Teknik Analisis Data

Teknik analisis data dimulai dari tahap analisis kebutuhan, validasi ahli dianalisis dengan teknik sebagai berikut:

Data Validasi Ahli

Data penilaian ahli dan guru senior terhadap soal dianalisis dengan teknik deskriptif dengan menggunakan persentase (Purwanto, 2010). Analisis data dilakukan dengan cara menghitung skor yang dicapai dari seluruh aspek yang dinilai kemudian menghitungnya dengan rumus sebagai berikut:

$$N = \frac{k}{Nk} \times 100\%$$

Nk

Keterangan:

N : persentase kelayakan aspek

K : skor hasil pengumpulan data

Nk : skor maksimal (skor kriteria tertinggi x jumlah aspek x jumlah validator)

Tabel 3.3 Kriteria Interpretasi Skor Validasi Ahli

Interval kriteria	Kriteria
$86 \% \leq N < 100\%$	Sangat baik
$72 \% \leq N < 85\%$	Baik
$58 \% \leq N < 71\%$	Cukup
$44 \% \leq N < 57\%$	Kurang
$N \leq 44\%$	Sangat Kurang

Sumber: Sudjana (2009)

HASIL DAN PEMBAHASAN

Hasil

1. Hasil Uji Produk Instrumen Penilaian

a. Hasil Validasi Produk Kepada Ahli Materi

Ahli yang dimaksud dalam penelitian pengembangan penilaian adalah orang yang memiliki bidang ilmu linier dengan pengembangan produk. Hasil dan indikator instrumen penilaian oleh ahli materi disajikan pada Tabel 1.

Tabel 2. Hasil Penilaian Aspek Materi Higher Order Thinking Skills dari Ahli Materi Lingkungan

No	Aspek yang dinilai	Skor (%)	Konversi	Kriteria
1.	Kebenaran Materi	100	A	Sangat Baik
2.	Soal	100	A	Sangat Baik
3.	Kedalaman Materi	100	A	Sangat Baik
	Keterbacaan			
	Rata-Rata	100		

Sumber: Lembar Penilaian Ahli Materi

Tabel 2 menunjukkan skor rata-rata keseluruhan aspek materi adalah 100 dan kategori termasuk “A” (Sangat Baik). Masing-masing aspek yang dinilai seperti kebenaran materi, soal dan kedalaman dan keterbacaan materi seluruhnya 100% ber kriteria sangat baik.

b. Hasil Validasi Produk Kepada Ahli Instrumen Evaluasi

Fungsi ahli instrumen evaluasi adalah untuk memvalidasi kesesuaian soal dengan indikator (validitas konstruk) sehingga menjamin soal yang dikembangkan dapat mengukur Higher Order Thinking Skills siswa. Hasil penilaian aspek dan indikator instrument evaluasi oleh ahli instrumen evaluasi disajikan pada Tabel berikut.

Tabel 1. Penilaian Indikator Materi Higher Order Thinking Skills dari Ahli Materi (Materi Lingkungan)

No	Indikator yang dinilai	Skor %	Konversi	Kriteria
1	Konsep materi soal Benar	100	A	Sangat Baik
2	Cakupan Materi Sesuai Tingkatan Siswa	100	A	Sangat Baik
3	Istilah Yang digunakan jelas	100	A	Sangat Baik
4	Materi soal Mudah dipahami	100	A	Sangat Baik
5	Materi soal Ditulis sistematis, runtut, dan alur logika jelas	100	A	Sangat Baik
Rata-rata		100	A	Sangat Baik

Sumber: Lembar penilaian indikator Ahli Materi

Tabel 1 menunjukkan bahwa skor rata-rata keseluruhan indikator materi adalah 100 % dan termasuk kategori “A yaitu sangat baik. Aspek indikator yang dinilai memiliki skor 100% seperti konsep materi soal benar, cakupan materi sesuai tingkatan siswa, istilah yang digunakan jelas, materi soal mudah dipahami, dan materi soal ditulis sistematis, runtut, dan alur logika jelas.

Tabel 2. Hasil Penilaian Aspek Materi Higher Order Thinking Skills dari Ahli Instrumen (Materi Lingkungan)

No	Aspek yang dinilai	Skor (%)	Konversi	Kriteria
1.	Ranah Materi	87,98	A	Sangat Baik
2.	Ranah Kontruksi	60,84	C	Cukup
3.	Ranah Bahasa	75	B	Baik
4.	Rubrik Penilaian	100	A	Sangat Baik
Rata-Rata		80,95		

Sumber: Lembar Penilaian Ahli Insrumen Evaluasi

Tabel 2 menunjukkan skor rata-rata keseluruhan aspek materi adalah 80,95 % dan kategori termasuk “B” (Baik). Aspek yang dinilai seperti ranah materi sebesar 87,98% dengan kriteria sangat baik, begitupula dengan rubrik penilaian sebesar 100% dengan kriteria sangat baik, ranah konstruksi sebesar 60,84% dengan kriteria cukup, dan ranah bahasa sebesar 75% dengan kriteria baik.

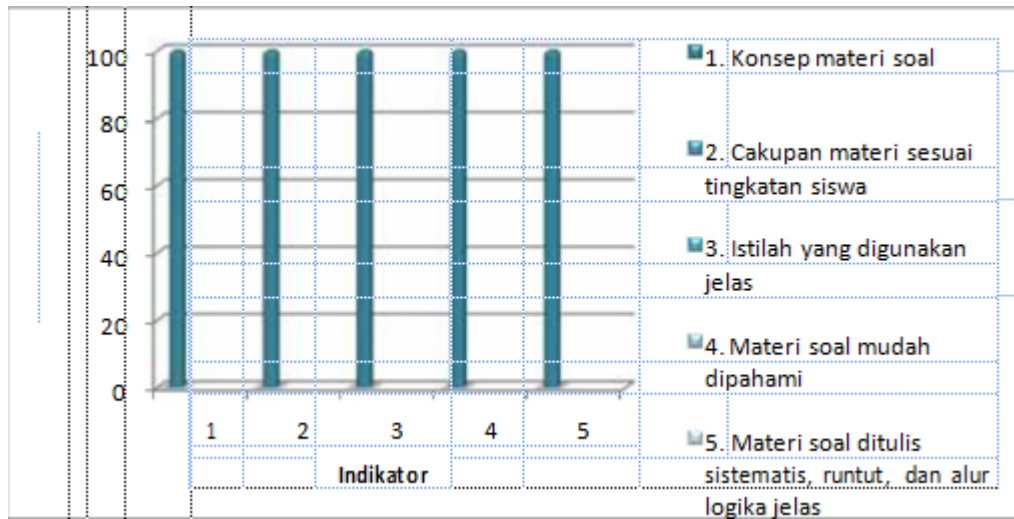
Ahli instrumen memberikan saran-saran untuk merevisi produk instrumen evaluasi hasil pengembangan sebagai berikut:

- 1) Perhatikan koherensi antara pertanyaan dengan jawaban
- 2) Perhatikan kesesuaian taksonomi Bloom dengan soal
- 3) Penulisan pilihan jawaban soal lebih disederhanakan
- 4) Perhatikan tata cara penulisan istilah, tata tulis dan sebagainya

B. Kelayakan Produk Instrumen Penilaian

1. Analisis Hasil Validasi Ahli

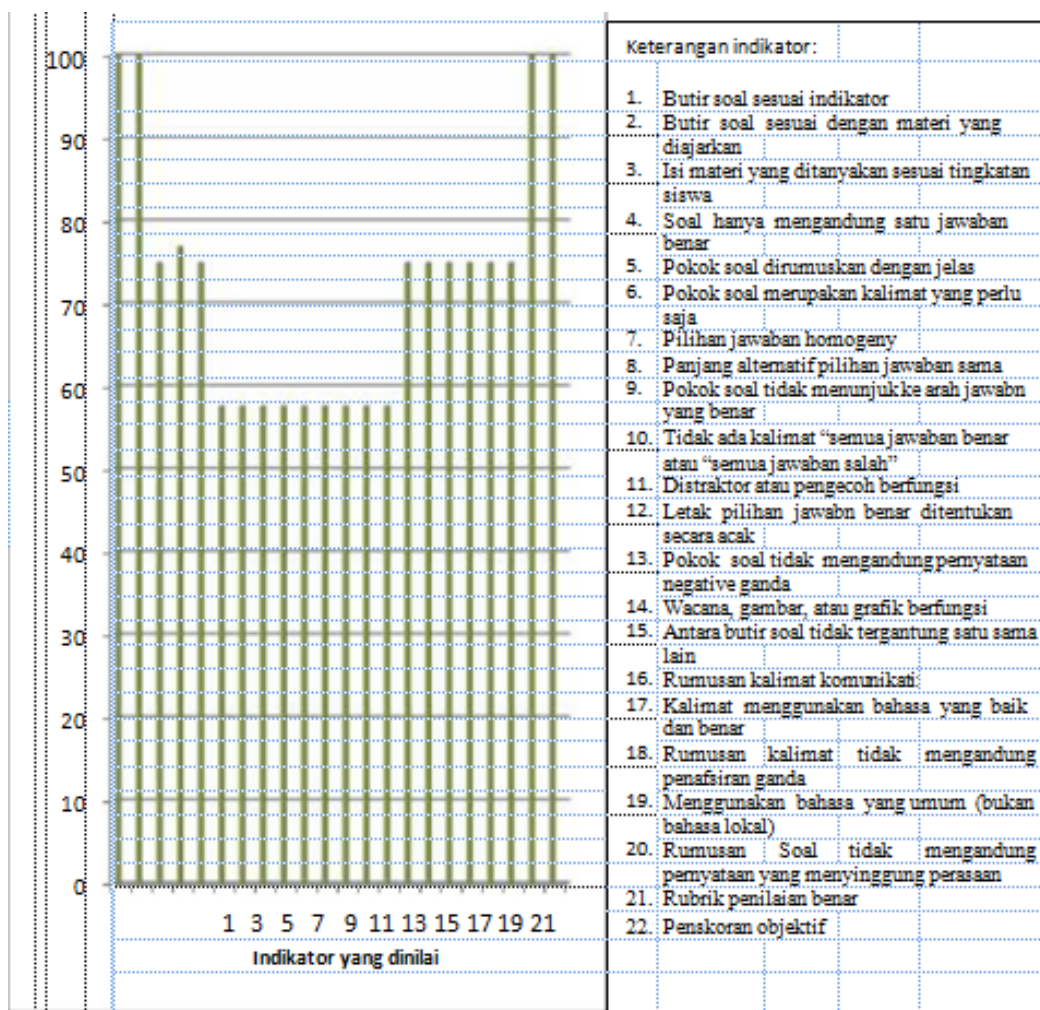
Data hasil validasi ahli dianalisis untuk menilai kelayakan instrumen penilaian hasil pengembangan. Analisis hasil penilaian ahli materi ditunjukkan pada Gambar 1.



Gambar 1. Persentase Indikator Penilaian Aspek Materi Instrumen Penilaian Higher Order Thinking Skills Materi Lingkungan

Analisis terhadap gambar 1 menunjukkan tidak ada kesalahan konsep dalam materi soal, cakupan materi sesuai dengan tingkatan siswa, istilah yang digunakan dalam soal mudah dipahami siswa, materi soal mudah dipahami siswa, materi soal ditulis sistematis, runtut, serta alur logika jelas. Kesimpulan yang diperoleh adalah instrumen penilaian memiliki validitas isi yang sangat baik. Validitas isi (content validity) adalah validitas yang mempertanyakan bagaimana kesesuaian butir soal dalam tes dengan deskripsi bahan yang diajarkan (materi) (Munthe, 2012). Penjaminan validitas isi instrumen penilaian dilakukan dengan memberikan prototype pengembangan instrumen penilaian Higher Order Thinking Skills kepada ahli bidang materi lingkungan.

Validitas instrumen penilaian meliputi validitas isi dan validitas konstruk. Validitas konstruk instrumen penilaian dinilai oleh ahli instrumen evaluasi. Hasil penilaian dari ahli instrumen evaluasi ditunjukkan pada Gambar 2.



Gambar 2. Gambar 4.5 Persentase Indikator Penilaian Aspek Konstruk Instrumen Penilaian Higher Order Thinking Skills Materi Lingkungan

Analisis terhadap gambar 2 menunjukkan instrument yang dikembangkan sesuai dengan materi yang diajarkan isi materi sesuai tingkatan siswa, soal hanya mengandung satu jawaban benar, pokok soal dirumuskan dengan jelas pokok soal merupakan kalimat yang diperlukan saja, pilihan jawaban homogeny, panjang alternatif jawaban sama, pokok soal tidak menunjuk ke arah jawaban benar, tidak ada kalimat "semua jawaban benar" atau "semua jawaban salah", distraktor atau pengecoh berfungsi, letak pilihan jawaban ditentukan secara acak, pokok soal tidak mengandung pernyataan negatif ganda, wacana, gambar berfungsi, antara butir soal tidak tergantung satu sama lain, rumusan kalimat komunikatif, kalimat menggunakan bahasa yang baik dan benar, rumusan kalimat tidak menimbulkan penafsiran ganda, menggunakan bahasa yang umum (bukan bahasa lokal), rumusan soal tidak menyinggung perasaan, rubrik penilaian soal benar, serta penskoran objektif. Kesimpulan yang diperoleh adalah instrument penilaian yang memiliki validitas konstruk yang baik.

Validitas konstruk (construct validity) adalah validitas yang mempertanyakan apakah butir-butir soal dalam tes sesuai dengan tingkatan kompetensi atau ranah yang dituntut (Munthe, 2012). Analisis validitas konstruk suatu tes dapat dilakukan dengan mencocokkan kemampuan berpikir yang ada pada butir soal dengan kemampuan berpikir yang tercantum dalam rumusan indikator yang diukur. Arikunto (2007) menyebutkan sebuah tes dikatakan memiliki validitas konstruksi apabila butir-butir soal yang membangun tes tersebut mengukur setiap aspek berpikir seperti yang disebutkan dalam tujuan instruksional khusus atau sesuai dengan aspek berpikir yang menjadi tujuan instruksi khusus. Indikator Higher Order Thinking Skills yang digunakan diadaptasi dari Anderson dan Krathwohl (2001) meliputi kemampuan menganalisis, mengevaluasi, dan menciptakan. Pencocokan validitas konstruk dilakukan dengan memasang setiap butir soal dengan aspek indikator Higher Order Thinking Skills.

Pembahasan

Higher Order Thinking Skills merupakan suatu keterampilan berpikir yang tidak hanya membutuhkan kemampuan mengingat, tetapi membutuhkan kemampuan lain yang lebih tinggi. Keterampilan Higher Order Thinking Skills sebagai keterampilan berpikir yang terjadi ketika seseorang mengambil informasi baru dan informasi yang sudah tersimpan dalam ingatannya, selanjutnya menghubungkan informasi tersebut dan menyampaikannya untuk mencapai tujuan atau jawaban yang dibutuhkan (Lewis & Smith, 1993). Keterampilan berpikir tingkat tinggi pada siswa dapat diberdayakan dengan memberikan masalah yang tidak biasa dan tidak menentu seperti pertanyaan atau dilema, sehingga penerapan yang sukses dari kemampuan ini adalah ketika siswa berhasil menjelaskan, memutuskan, menunjukkan, dan menghasilkan penyelesaian masalah dalam konteks pengetahuan dan pengalaman (King, Goodson, Ludwika, dan Rohani, 2010).

Anderson & Krathwohl (2001) menyatakan Higher Order Thinking Skills memiliki tingkatan taksonomi Bloom seperti analisis, sintesis dan evaluasi (Taksonomi Bloom lama) dan tingkat menganalisis, mengevaluasi dan mencipta (Revisi Taksonomi Bloom). Pada Taksonomi Bloom Revisi, yang termasuk ke dalam kategori Higher Order Thinking Skills adalah pada tingkat Analyze (Menganalisis), Evaluate (Mengevaluasi) dan Create (Mencipta). Taksonomi Bloom dianggap merupakan dasar bagi berpikir tingkat tinggi. Pemikiran ini didasarkan bahwa beberapa jenis pembelajaran memerlukan proses kognisi yang lebih daripada yang lain, tetapi memiliki manfaat-manfaat lebih umum (Pohl, 2000).

Higher Order Thinking Skills perlu dilatihkan untuk siswa di sekolah. Ini penting sebab siswa akan berhasil menjelaskan, memutuskan, menunjukkan dan menghasilkan penyelesaian masalah dalam konteks pengetahuan dan pengalaman (King, et.al, 2010). Higher Order Thinking Skills perlu diukur menggunakan suatu penilaian yang jelas dan valid. Pengukuran Higher Order Thinking Skills belum sering dilakukan oleh praktisi pendidikan serta sistem penilaian yang baik terdapat di sekolah juga masih sangat dibutuhkan. Dengan demikian perlu dilakukan pengembangan penilaian guna mengukur Higher Order Thinking Skills dengan bentuk soal pilihan ganda dan essay.

Penilaian sendiri diartikan sebagai prosedur yang digunakan untuk mendapatkan informasi untuk mengukur taraf pengetahuan dan keterampilan subjek didik yang hasilnya akan digunakan untuk keperluan evaluasi. Informasi adalah data yang diperoleh melalui pengukuran dan nonpengukuran termasuk di dalamnya dengan melakukan observasi kelas, menggunakan tes yang standar atau buatan tes guru, proyek dan portofolio subjek belajar. Untuk memperoleh informasi yang berupa data kuantitatif dilakukan pengukuran. Pengukuran merupakan proses untuk memperoleh deskripsi numeric atau kuantitatif tentang tingkatan karakteristik yang dimiliki seseorang dengan aturan tertentu (Subali, 2010).

Diperlukan teknik penilaian untuk mengukur keterampilan kognitif siswa yaitu menurut Grounlund & Linn (1995) mengklasifikasikan teknik penilaian tes menjadi beberapa kategori, yakni tes bentuk pilihan, tes bentuk mengkonstruksi jawaban, dan penilaian yang diperluas. Tes bentuk pilihan dapat berupa pilihan ganda, salah benar, menjodohkan/memasangkan, tes bentuk mengkonstruksi jawaban dapat berupa tes isian, uraian terstruktur, dan uraian terbuka, penilaian yang diperluas dapat berupa proyek atau portofolio.

Pengembangan instrumen penilaian dimulai dengan membuat prototype pengembangan. Prototype pengembangan merupakan bentuk awal dari produk yang dibuat berisi pendahuluan, langkah-langkah pengembangan instrumen penilaian, perangkat instrumen penilaian yang terdiri dari kisi-kisi, butir-butir soal, lembar jawaban, kunci jawaban, serta rubrik penilaian soal. Semua perangkat kemudian diajukan dan divalidasi oleh ahli validator meliputi ahli materi, dan ahli instrument.

Tujuan validasi ke ahli materi adalah menjamin validitas isi dan butir instrument penilaian yang dikembangkan sehingga tidak terjadi kesalahan konsep. Tujuan validasi ke ahli instrumen evaluasi adalah untuk menjamin validitas konstruk instrumen yang dikembangkan sehingga instrumen yang dikembangkan dapat mengukur Higher Order Thinking Skills siswa. Tujuan validasi ke ahli guru senior yaitu untuk menjamin kelayakan instrumen penilaian sebelum penerapan. Hasil validasi perangkat instrumen penilaian dimasing-masing ahli validator ialah memiliki validasi isi dan konstruk yang baik dan layak untuk diterapkan di sekolah tentunya setelah dilakukan revisi, dimana revisi yang meliputi memperbaiki konsep yang masih salah, memperbaiki kesesuaian dengan konsep taksonomi Bloom, memperbaiki penggunaan istilah kata yang masih banyak kesalahan, serta alokasi waktu. Revisi tersebut guna untuk memperbaiki instrumen penilaian yang akan digunakan pada uji kelompok kecil.

Berdasarkan Hasil Pengamatan di atas, maka instrument penilaian Higher Order Thingking Skills (HOTS) yang telah di uji validasi oleh ahli materi dan ahli instrument evaluasi, maka instrument tersebut layak untuk di ujikan pada kelompok uji terbatas maupun uji lapangan terhadap siswa SMA.

KESIMPULAN

Adapun kesimpulan yang dapat disimpulkan dari hasil penelitian adalah sebagai berikut:

1. Hasil validasi ahli materi menunjukkan skor rata-rata keseluruhan aspek materi adalah 100% dengan kategori “Sangat Baik”. Hasil validasi ahli instrument evaluasi menunjukkan skor rata-rata keseluruhan aspek materi adalah 80,95% dengan kategori “Baik”.

2. Instrumen Penilaian ini dapat mengukur Higher Order Thinking Skills siswa SMA yang memiliki tingkatan seperti menganalisis, mengevaluasi dan mencipta (Taksonomi Bloom yang direvisi) dengan soal berbentuk pilihan ganda dan essay berdasarkan kasus.
3. Instrumen penilaian Higher Order Thinking Skills pada materi lingkungan dinilai layak untuk diujikan pada kelompok uji terbatas maupun uji lapangan terhadap siswa SMA.

REFERENSI

- Aiken, Lewis R. (1994). *Psychological Testing and Assessment*, (Eight Edition)
- Anderson, W. L & Krathwohl, R, D. (2001). *A Taxonomy for Learning Teaching and Assessing A Revision of Bloom's Taxonomy of Educational Objectives*. USA: Addison Wesley Longman.
- Aprianto. (2008). *Implementasi Sistem Sistem Manajemen Mutu ISO 9001: 2008 Berdasarkan Pada 8 (delapan) prinsip manajemen mutu pada SMK Negeri 13 Bandung*. Universitas Teknologi Bandung: Bandung
- Arikunto, Suharsimi. (2007). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara
- Basuki, Ismet, & Hariyanto. (2012). *Assesmen Penelitian*. Bandung: PT Remaja Rosdakarya Offset
- Bloom. (1964). *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain* Butkowski, Jean. 1994. *Improving Student Higher Order Thinking Skills in Mathematics*. Tesis, Educational Resources Information Center
- Borg and Gall. (1983). *Educational Research, An Introduction* (4th Ed). Newyork and London. Longman Inc. Boston: Allyn and Bacon.
- Gronlund, M. S and Linn, Joyce. (1995). *Measurement And Assesment In Teaching*. Prentice-Hall, Pearson Education Upper Saddle River, New Jersey
- Kemendikbud. (2013) . *Permendikbud No. 81A tentang Implementasi Kurikulum*. Jakarta: Kementerian Pendidikan dan Kebudayaan
- King, JF; Goodson, Ludwika, dan Rohani, Faranak, (2010). *Higher Order Thinkin Skills, Definition, Teaching Strategis, Assesment*. A Publication of The Educational Services Program. www.Cala.fsu.edu
- Krathwohl, Bloom & Masia. (1964). *The Taxonomy of Educational Objectives: Handbook II* Krathwohl, D. R. 2002. *A revision of Bloom's Taxonomy: an overview - Theory Into Practice*, College of Education, The Ohio State University Learning Domains or Bloom's Taxonomy: *The Three Types of Learning*, tersedia di www.nwlink.com/~donclark/hrd/bloom.html
- Lewis, A., & Smith, D. (1993). *Defining Higher Order Thingking. Theory Into Practice*, XXXII (3), 131-137
- Munthe, Bermawi. (2012). *Desain Pembelajaran*. Yogyakarta: Pustaka Insan Madani
- Purwanto. (2010). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar
- Reynold, C.R., Livingston, R. B., & Willson, V. (2010). *Measurement and Assesment in Education*. Pearson Education, Inc., Upper Saddle River, New Jersey 07458. Pearson.
- Sajidan, Sugiharto B. & Prasetyani, N.M. (2013). *Penerapan Dan Pengembangan Mutu Pendidikan Melalui Lesson Studi Pada Mata Pelajaran Biologi Berbasis Pemetaan Ketuntasan UN Biologi, Laporan Hibah Guru Besar*. LPPM : UNS
- Subali, Bambang. (2010). *Penilaian, Evaluasi dan Remediasi Pembelajaran Biologi*. UNY: Yogyakarta

Kusuma Wardani, dkk

Sudjana, Nana. (2010). Penilaian Hasil Proses Belajar Mengajar. Bandung: Remaja Rosdakarya
Suwandi, Sarwiji. (2009). Model Asesmen Dalam Pembelajaran. Mata Padi Presindo: Surakarta.