

## PENGEMBANGAN INSTRUMEN PENILAIAN SEBAGAI CONTOH PAKET SOAL HIGHER ORDER THINKING SKILLS (HOTS) MATERI PENCEMARAN LINGKUNGAN KELAS X SMA

### *Development of Assessment Instruments as Example For Higher Order Thinking Skills (Hots) Package Questions in Environmental Pollution Topic for 10<sup>th</sup> Grade Senior High School*

**Yohanif Putri Al Avina**

Pendidikan Biologi, FMIPA, Universitas Negeri Surabaya

E-mail : [yohanifavina16030204089@mhs.unesa.ac.id](mailto:yohanifavina16030204089@mhs.unesa.ac.id)

**Winarsih**

Pendidikan Biologi, FMIPA, Universitas Negeri Surabaya

E-mail : [winarsih@unesa.ac.id](mailto:winarsih@unesa.ac.id)

#### **Abstrak**

Instrumen penilaian HOTS (*Higher Order Thinking Skills*) merupakan penilaian yang mengukur kemampuan pada ranah C-4 (melakukan proses analisis), C-5 (melakukan proses evaluasi), dan C-6 (mengkreasi). Instrumen penilaian yang dikategorikan baik yaitu instrumen yang dibuat dalam bentuk soal berbasis pengembangan HOTS. Penelitian ini memiliki tujuan guna melakukan proses pendeskripsian instrumen pengembangan paket soal HOTS pada materi pencemaran lingkungan yang valid dan reliabel ditinjau dari tingkat kesukaran dan sensitivitas. Penelitian dilakukan selama bulan Oktober 2019-Maret 2020. Instrumen diujicobkan pada 20 peserta didik pada kelas X Sekolah Menengah Atas (SMA) Negeri 1 Bangil. Metode yang digunakan yaitu metode validasi, dan tes. Teknis analisis data dilakukan secara deskriptif yang meliputi analisis hasil tes HOTS untuk mengetahui nilai reliabilitas tes, tingkat kesukaran tes, dan sensitivitas tes. Hasil penelitian menunjukkan bahwa instrumen penilaian yang dikembangkan dinyatakan valid ditinjau dari aspek materi, konstruksi, dan bahasa mendapatkan nilai 4 dengan kategori sangat valid, dengan reliabilitas yang mempunyai nilai sebesar 0,80 dan dikategorikan sangat tinggi, indeks kesukaran soal dengan besar 25% mudah, 50% sedang, dan 25% sukar, serta sensitivitas menunjukkan 100% soal sensitif. Instrumen penilaian materi pencemaran lingkungan dinyatakan valid dan reliabel sehingga dapat dijadikan sebagai contoh paket soal HOTS.

**Kata Kunci:** Instrumen, Paket Soal HOTS, Validitas, Reliabilitas, Tingkat Kesukaran, Sensitivitas.

#### **Abstract**

HOTS (*Higher Order Thinking Skills*) assessment instruments are assessments that measure abilities in the realm HOTS assessment instruments are abilities measurement in the realm of C-4 (analyzing), C-5 (evaluating), and C-6 (creating). A good assessment instrument is an instrument made in the form of questions based on the development of HOTS This study aims to describe the instrument development package HOTS on valid and reliable environmental pollution topic based on the level of difficulty and sensitivity. The study was tested during October 2019-March 2020. The instrument was tested on 20 10<sup>th</sup> students of SMAN 1 Bangil. The method used is validation and tests. Analysis Method of the data was done descriptively which included analysis of HOTS test results to determine the value of test reliability, level of difficulty, and sensitivity. The results showed that the assessment instruments developed were declared valid in terms of aspects of the topic, construction, and language getting a value of 4 with a very valid category, reliability had a value of 0.80 with a very high category, an index of difficulty level of 25% was easy, 50% moderate, and 25% difficult, and sensitivity shows 100% sensitive questions. The instrument for evaluating environmental pollution topic is declared valid and reliable so that it can be used as an example package for HOTS.

**Keywords:** Instrument, Package Questions HOTS, Validity, Reliability, Difficulty, Sensitivity.

## PENDAHULUAN

Berpikir tingkat tinggi atau juga memiliki istilah *Higher Order Thinking Skills* (HOTS) merupakan tuntutan peserta didik masa kini yang akhirnya mereka mampu menghadapi tantangan pada abad ke-21. HOTS dapat didefinisikan sebagai sebuah keterampilan berfikir tingkat tinggi yang memberikan tuntutan untuk berpikir secara analitis, kritis dan kreatif pada segala informasi yang didapatkan (Barratt, 2014). Keterampilan HOTS dalam proses berpikirnya membutuhkan kemampuan kognitif lebih tinggi dibandingkan hanya sekedar *recall* (kemampuan mengingat), *restate* (menyatakan kembali), atau *recite* (melakukan proses rujukan dengan tanpa proses pengolahan) (Widana, 2017). Menurut Brookhart (2010) HOTS merupakan proses transfer dari suatu masalah yang solusinya ditentukan dengan cara berpikir kritis.

Kemampuan peserta didik Indonesia kebanyakan berada pada tingkatan LOTS (*Lower Order Thinking Skills*), banyak materi hafalan, dan latihan soal untuk melatih berpikir tingkat tinggi masih kurang (Nugroho, 2018). Untuk melatih berpikir tingkat tinggi diperlukan alat ukur berupa instrumen penilaian (Sharif, 2015). Instrumen adalah suatu perangkat yang berfungsi untuk mengukur hasil suatu penelitian dengan cara mengumpulkan data (Setyawan, 2013). Penilaian ialah tindakan pengambilan keputusan pada hasil pengukuran sebelum melakukan suatu kegiatan yang telah direncanakan (Basuki dan Hariyanto, 2015). Instrumen penilaian yang dikategorikan baik ialah sebuah instrumen yang diciptakan dalam bentuk soal berbasis pengembangan HOTS (Devi, 2013).

Instrumen penilaian HOTS dapat diartikan sebagai sebuah penilaian yang melakukan proses pengukuran kemampuan pada ranah C-4 (proses melakukan proses analisa-*analyzing*), C-5 (melakukan proses evaluasi-*evaluating*), dan C-6 (proses mengkreasi-*creating*). Penilaian HOTS memiliki karakteristik yaitu: melakukan proses pengukuran terhadap kemampuan berfikir tingkat tinggi, dengan menggunakan soal soal permasalahan yang kontekstual, dan memiliki soal soal yang beragam. Berpikir tingkat tinggi bisa didefinisikan sebagai proses berpikir pada level yang lebih tinggi, tidak hanya mengenai proses hafalan saja tetapi adanya keterlibatan aktivitas psikis peserta didik dalam mengkoneksikan materi pembelajaran dengan pengalaman peserta didik dan bersifat kompleks, reflektif, dan kreatif untuk memperoleh sebuah pengetahuan yang meliputi tingkat berpikir analitis, sintesis, dan evaluatif (Fanani, 2018).

Penilaian HOTS memiliki manfaat dalam peningkatan motivasi belajar peserta didik karena penilaian HOTS tidak hanya terpaku terhadap teks bacaan saja, tapi mampu dihubungkan pada hal hal yang bersinggungan langsung dengan kehidupan nyata sehingga lebih mudah dipahami (Nugroho, 2018).

Cara menilai HOTS / kemampuan berpikir tingkat tinggi yakni dengan melakukan sebuah tes. Tes merupakan sebuah instrumen yang dapat dipakai oleh guru untuk bisa tahu sejauh apa peserta didik bisa menerima pemahaman terhadap materi yang telah diberikan (Yuniar, 2015). Tes memiliki dua macam jenis tes yakni tes subjektif dan tes objektif. Tes subjektif diartikan sebagai tes yang memiliki bentuk uraian. Sedangkan, tes objektif diartikan sebagai tes yang tersusun atas pilihan ganda atau biasa disebut tes *multiple choice*, jawaban benar-salah (*true false*), jawaban singkat (*short answer*) dan penjodohan (*matching*) (Suwandi, 2009).

Soal tipe *multiple choice* / pilihan ganda berguna dalam melakukan proses pengukuran hasil belajar peserta didik yang bersifat kompleks dan berhubungan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis, serta evaluasi (Arifin, 2010). Soal dengan bentuk pilihan ganda memiliki beberapa keunggulan, diantaranya yaitu dalam soal yang akan di ujikan dapat mencakup bahasan materi yang lebih luas, dan juga memiliki tingkat validitas dan reabilitas yang relatif lebih tinggi jika dibandingkan dengan model soal esay, guru bisa menilai dengan lebih cepat dan juga lebih mudah serta dapat melakukan proses pengukuran terhadap ranah kognitif (Yuniar, 2015).

Dalam pelaksanaan pembelajaran, guru memiliki peran yang penting untuk menentukan ketepatan jenis penilaian, maka dari itu diperlukan suatu pedoman mengenai karakteristik instrumen penilaian yang dikategorikan baik sebelum digunakan. Tujuannya untuk melakukan proses evaluasi atau mengadakan suatu penilaian agar dapat menghindari kesalahan dan hasil yang tidak valid (Suwono, 2012). Tes yang dikategorikan baik memiliki beberapa karakteristik yaitu valid, reliabel, tingkat kesusahan / kerumitan, dan indeks daya beda yang dikategorikan baik (Azwar, 2016).

Berdasarkan hasil wawancara terhadap dua guru Biologi kelas X Sekolah Menengah Atas Negeri 1 Bangil didapatkan hasil bahwa 1 guru Biologi kelas X belum mengembangkan instrumen penilaian HOTS pada KD. 3.11, sedangkan 1 guru yang lain sudah mengembangkan instrumen penilaian HOTS pada KD. 3.11 ini hanya memberikan soal HOTS dengan besar 30% dari jumlah

soal yang digunakan. Sebanyak 80% dari 15 peserta didik juga mengaku masih jarang dan belum terbiasa diberikan soal-soal HOTS oleh guru Biologi, sehingga dapat diketahui bahwa butir soal yang dipakai untuk melakukan proses pengukuran pada kemampuan kognitif peserta didik masih berada pada level kognitif yang rendah yaitu C-1 (mengingat) dan C-2 (memahami). Materi yang dikembangkan dalam penelitian ini yaitu pencemaran lingkungan yang mana masalah lingkungan merupakan masalah yang kompleks sehingga peserta didik membutuhkan kemampuan untuk melakukan proses analisis (C-4), melakukan proses evaluasi (C-5), dan mengkreasi (C-6) (Lee, 2017).

Penelitian yang dilakukan Indri Putri Utami & Aryeni (2018) menguatkan hasil tersebut, hasil analisis pada soal UAS ganjil mata pelajaran Biologi Sekolah Menengah Atas Negeri 1 Kisaran memberikan sebuah hasil bahwasannya sebaran dimensi proses kognitif Taksonomi Anderson dalam soal tersebut belumlah rata, soal juga didominasi oleh dimensi proses kognitif level rendah, yaitu mengingat (C1) dan memahami (C2). Pada penelitian relevan sebelumnya yang dilakukan oleh Oktaviani, *et.al* (2017), hasil analisis soal UAS genap mata pelajaran Biologi kelas X di Sekolah Menengah Atas Negeri 3 Tanjungpinang menunjukkan bahwa masih perlu peningkatan karena masih didominasi oleh C2 saja.

Hal ini menunjukkan bahwa masih banyak kesulitan di lapangan dalam mengembangkan soal HOTS dikarenakan penilaian HOTSs ini baru mulai diberlakukan setelah kurikulum 2013 diperbaharui atau direvisi sehingga peserta didik dalam mengerjakan soal-soal HOTS seperti menalar, menganalisis, dan mengevaluasi masih lemah.

Berdasarkan uraian di atas penelitian pengembangan instrumen penilaian HOTS memiliki tujuan guna menghasilkan instrumen penilaian yang reliabel dan juga valid ditinjau dari tingkat kesusahan / kerumitan dan sensitivitas sebagai contoh paket soal HOTS pada materi pencemaran lingkungan kelas X Sekolah Menengah Atas.

## METODE

Penelitian ini masuk dalam kategori penelitian pengembangan. Dalam proses penelitian ini yang dikembangkan yaitu soal HOTS dengan mengacu pada model penelitian pengembangan atau *Research and Development* yang harusnya didalamnya terdiri dari sepluh tahapan, karena beberapa penyesuaian maka pada penelitian ini hanya sampai pada tahap keenam karena keterbatasan waktu. Penelitian dilakukan selama Bulan Oktober 2019 s.d Maret 2020. Sasaran penelitian yaitu 20 butir soal HOTS yang telah melalui proses validasi oleh

seorang validator dan dinyatakan valid. Uji coba penelitian ini berobjek pada 20 peserta didik pada kelas X Sekolah Menengah Atas Negeri 1 Bangil untuk mengetahui reliabilitas instrumen HOTS tersebut.

Metode dalam proses pengumpulan data yang dipakai pada penelitian ini yakni dengan memakai metode validasi, dan tes. Penyusunan butir soal HOTS dilakukan dengan cara menguraikan kompetensi dasar (KD) menjadi beberapa indikator yang sesuai kemudian mencari sumber yang akurat dan sesuai dengan indikator untuk dijadikan stimulus soal. Butir soal HOTS yang dikembangkan yaitu tipe pilihan ganda dengan level kognitif C-4 (melakukan proses analisis), C-5 (melakukan proses evaluasi), dan C-6 (proses mengkreasi).

Teknis analisis data dilakukan secara deskriptif yang meliputi analisis hasil tes HOTS untuk mengetahui nilai reliabilitas tes, tingkat kesusahan / kerumitan tes, dan sensitivitas tes. Nilai reliabilitas didapatkan dengan penggunaan metode belah dua (*split-half*) dengan memakai rumus *Spearman Brown* pada proses perhitungannya. Nilai tingkat kesusahan / kerumitan diperoleh dengan cara menghitung jumlah peserta tes yang memiliki jawaban benar dibagi seluruh peserta yang mengikuti tes dikali 100%. Nilai sensitivitas tes diperoleh dengan cara jumlah skor ketika telah mengikuti proses pembelajaran dikurangi jumlah skor ketika belum mengikuti proses pembelajaran dibagi banyaknya peserta yang mengikuti tes kemudian dikali dengan skor maksimal yang sudah dikurangi skor minimal.

## HASIL DAN PEMBAHASAN

Penelitian pengembangan ini menghasilkan instrumen penilaian sebagai contoh paket soal HOTS yang dinyatakan valid dan juga dinyatakan reliabel. Validitas instrumen HOTS diperoleh berdasarkan validasi dari para ahli yaitu ahli pendidikan dan ahli materi yang dilihat dari beberapa aspek didalamnya, yakni : (1) materi, (2) konstruksi, dan (3) bahasa pada masing-masing butir soal. Setiap aspek tersebut mendapatkan skor modulus keseluruhan dengan besar 4 dan dikategorikan sangat valid. Berikut merupakan hasil rekapitulasi validasi instrumen penilaian HOTS.



**Gambar 1.** Hasil Validasi Butir Soal HOTS pada Instrumen Penilaian Materi Pencemaran Lingkungan

Validasi butir soal HOTS pada aspek materi mendapatkan skor modus dengan besar 4 dan dikategorikan sangat valid. Pada aspek materi terdapat 3 kriteria yang divalidasi oleh para ahli yaitu, 1) kesesuaian soal yang diujikan dengan materi yang telah diberikan, 2) kesesuaian butir soal yang diujikan dengan kebenaran konsep, 3) Adanya jawaban yang benar yakni 1 jawaban. Akan tetapi, ditemukan butir-butir soal dengan dikategorikan kurang valid yakni dalam soal ujian dengan nomor 9, 10, dan 11, hal ini disebabkan materi soal kurang sesuai dengan indikator.

Pada aspek konstruksi memperoleh skor modus dengan besar 4 yang dikategorikan sangat valid. Aspek konstruksi memiliki 7 kriteria yang divalidasi yaitu 1) petunjuk soal mudah dipahami, 2) butir soal memakai hal hal yang dianggap menarik sehingga mampu mengundang peserta didik tertarik untuk melakukan proses pembacaan 3) butir soal memakai hal hal yang kontekstual (bersentuhan langsung dengan dunia nyata yang dialami peserta didik), 4) butir soal melakukan proses pengukuran level kognitif penalaran / HOTS (melakukan proses analisis, melakukan proses evaluasi, dan mengkreasi), 5) tidak mengandung jawaban ganda, 6) butir soal tidak bergantung dengan jawaban soal sebelumnya, 7) pilihan jawaban tidak menggunakan "semua jawaban benar/salah" atau sejenisnya.

Pada aspek bahasa memperoleh skor modus dengan besar 4 yang dikategorikan sangat valid, terdiri dari 4 kriteria yang divalidasi yakni 1) adanya penggunaan bahasa Indonesia sesuai dengan ejaan yang disempurnakan, 2) adanya penggunaan bahasa yang komunikatif dan mudah dimengerti, 3) tidak mengandung bahasa yang ambigu, 4) pilihan jawaban tidak mengulang kata/kelompok kata yang sama, dikecualikan untuk awaban yang mengandung satu kesatuan pengertian. Akan tetapi, terdapat beberapa butir soal yang dikategorikan kurang valid yaitu pada soal nomor 4, 5, dan 20, hal ini diakibatkan karena bahasa yang digunakan tidak sesuai dengan kaidah, tidak komunikatif dan susah dipahami.

Suatu tes yang dikategorikan baik memiliki kategori valid dan reliabel. Validitas tes memiliki kaitannya dengan ketepatan dan kesesuaian antara tes sebagai alat ukur dengan objek yang diukur Valid dapat dilihat dari akurasi dan kecermatan pada hasil proses validasi dari

seorang ahli materi dan ahli pendidikan dan menggunakan instrumen validasi yang telah disediakan, dengan kata lain (Azwar , 2016).

Reliabilitas yaitu dapat dipercaya. Hal ini berhubungan dengan keajegan dari suatu tes yang dikembangkan dengan tipe soal pilihan ganda. Tes yang tidak dapat menghasilkan nilai yang reliabel maka tes tersebut tidak valid (Asrul, 2015).

Nilai reliabilitas yang diperoleh pada pengembangan instrumen penilaian HOTS dengan metode belah dua (*split-half*) yang dilakukan dengan memberikan angka terpisah pada soal nomor ganjil dan genap dan dengan memakai rumus persamaan *Spearman-Brown* didapatkan sebuah nilai reliabilitas tes dengan besar 0,80 dan dikategorikan sebagai sangat tinggi, akhirnya bisa disimpulkan bahwa instrumen penilaian HOTS yang dihasilkan memenuhi kriteria tes yang reliabel.

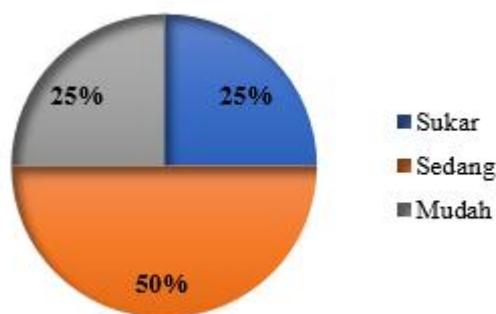
Didalam metode bentuk tes ulang dan metode bentuk paralel masih ditemukan beberapa kelemahan sehingga dalam metode belah dua dapat menjadi solusi atas kelemahan tersebut karena metode ini dapat mengestimasi reliabilitas dengan tidak diharuskan melakukan proses penyelenggaraan tes dua kali (Asrul, 2015). Suatu tes dikatakan reliabel jika peserta didik yang diukur kemampuannya menghasilkan nilai yang stabil atau relatif sama, dengan kata lain tidak mengalami perubahan meskipun dilakukan tes beberapa kali dengan tes yang sama dan waktu yang berbeda (Arikunto, 2013).

### Tingkat Kesukaran

Tingkat kesukaran ditentukan melalui perbandingan antara banyaknya peserta didik yang mengikuti tes dengan butir soal yang dijawab benar. Menurut Sukardi (2011) tingkat kesukaran yaitu angka yang menunjukkan proporsi peserta didik yang memiliki jawaban benar dalam satu soal yang dilaksanakan dengan menggunakan tes objektif. Berikut merupakan tabel hasil rekapitulasi tingkat kesukaran butir soal HOTS.

**Tabel 2.** Distribusi Tingkat Kesukaran Butir Soal HOTS Materi Pencemaran Lingkungan

No. Soal	Tingkat Kesukaran	Kategori
1	0,85	Mudah
2	0,50	Sedang
3	0,70	Mudah
4	0,80	Mudah
5	0,45	Sedang
6	0,25	Sukar
7	0,95	Mudah
8	0,40	Sedang
9	0,50	Sedang
10	0,30	Sukar
11	0,25	Sukar
12	0,60	Sedang
13	0,45	Sedang
14	0,30	Sukar
15	0,55	Sedang
16	0,75	Mudah
17	0,50	Sedang
18	0,65	Mudah
19	0,45	Sedang
20	0,20	Sukar



**Gambar 4.** Prosentase Tingkat Kesukaran dalam Butir Soal Instrumen Penilaian HOTS pada Materi Pencemaran Lingkungan

Perhitungan nilai tingkat kesukaran dalam butir soal instrumen penilaian *HOTS* diperoleh hasil masing-masing 25% mudah, 50% sedang, dan 25% sukar. Butir soal yang dikembangkan berjumlah 20 soal yang dikategorikan soal sukar berjumlah 5 butir soal, kategori soal sedang berjumlah 10 butir soal, dan kategori soal mudah berjumlah 5 butir soal.

Berdasarkan indeks tingkat kesukaran apabila indeks kesukaran yang didapatkan besar maka butir soal tersebut termasuk dalam kategori mudah, yaitu banyak peserta didik yang memiliki jawaban yang benar. Sebaliknya, apabila indeks kesukaran yang diperoleh kecil dan hampir mendekati nol maka butir soal tersebut sukar, yaitu peserta didik yang menjawab benar hanya sedikit bahkan tidak ada.

Asumsi yang dipakai untuk mendapatkan butir soal yang dikategorikan baik adalah keseimbangan dari tingkat kesukaran soal tersebut. Keseimbangan yang dimaksudkan ialah sebuah perbandingan antara butir-butir soal yang tergolong dalam kategori mudah, sedang dan sukar. Sebuah tes sebaiknya memiliki keseimbangan dengan nilai perbandingan antara mudah (3) : sukar (4) : sedang (3) (Sudjana, 2004). Tingkat kesukaran pada penelitian ini diperoleh perbandingan antara soal kategori mudah : sedang : sukar yaitu 1 : 2 : 1.

Butir soal yang dikategorikan baik dan dapat diterima adalah butir soal yang dikriteriakan sedang, sedangkan butir yang dikriteriakan mudah atau sukar dianggap sebagai soal yang tidak baik dan perlu direvisi (Asrul, 2015). Soal yang dikategorikan baik merupakan soal yang tidak terlalu mudah dan tidak terlalu sukar. Soal yang relatif mudahakan menyebabkan peserta didik tidak terangsang untuk meningkatkan usaha dalam memecahkannya. Namun berbanding terbalik dengan hal tersebut, soal yang terlalu sukar akan membuat peserta didik menjadi putus asa karena di luar jangkauannya dan

Data pada tabel 2 disimpulkan oleh peneliti dalam histogram sebagai berikut:



**Gambar 3.** Persebaran Tingkat Kesukaran dalam Butir Soal Instrumen Penilaian HOTS pada Materi Pencemaran Lingkungan

Adapun prosentase tingkat kesukaran dalam butir soal instrumen penilaian HOTS pada materi pencemaran lingkungan disajikan dalam **Gambar 4**.

tidak mempunyai semangat untuk mencoba kembali (Daryanto, 2010).

### Sensitivitas

Sensitivitas sebuah tes merupakan kemampuan tes untuk melakukan proses pengukuran efek atau dampak dari proses pembelajaran. Perhitungan sensitivitas berfungsi untuk menentukan sebaik apa soal soal tersebut dalam membedakan pengetahuan peserta didik pra dan pasca proses pembelajaran (Gronlund (dalam Hastuti, 2011)).

**Tabel 5.** Distribusi Sensitivitas Butir Soal HOTS Materi Pencemaran Lingkungan

No. Soal	Sensitivitas	Kategori
1	0,45	Sensitif
2	0,3	Sensitif
3	0,4	Sensitif
4	0,45	Sensitif
5	0,2	Sensitif
6	0,25	Sensitif
7	0,5	Sensitif
8	0,2	Sensitif
9	0,3	Sensitif
10	0,25	Sensitif
11	0,15	Sensitif
12	0,35	Sensitif
13	0,3	Sensitif
14	0,25	Sensitif
15	0,4	Sensitif
16	0,35	Sensitif
17	0,25	Sensitif
18	0,35	Sensitif
19	0,25	Sensitif
20	0,2	Sensitif

Berdasarkan data yang diperoleh didapatkan hasil perhitungan sensitivitas pada instrumen penilaian *HOTS* yang dikembangkan yaitu 100% sensitif. Nilai sensitivitas butir soal sangat beragam nilai yang terkecil yaitu pada soal nomor 11 yang memiliki nilai indeks dengan besar 0,15, namun untuk nilai sensitivitas tertinggi yaitu pada soal nomor 1 dan 4 dengan nilai indeks dengan besar 0,45.

Nomor soal 11 memiliki nilai indeks sensitivitas terkecil karena kuantitas peserta didik yang dapat memiliki jawaban yang benar dalam tes awal berjumlah 2 peserta dan tes akhir berjumlah 5 peserta, sehingga kenaikannya sangat sedikit yaitu 3 peserta. Nomor soal 1 memiliki indeks sensitivitas terbesar karena jumlah peserta didik yang dapat memiliki jawaban yang benar dalam proses tes awal yaitu 8, sedangkan dalam proses tes akhir jumlah peserta didik yang dapat memiliki jawaban yang benar yaitu 17 peserta sehingga

kenaikannya lebih banyak yaitu 9 peserta didik. Nomor soal 4 juga memiliki indeks sensitivitas terbesar karena jumlah peserta didik yang dapat memiliki jawaban yang benar dalam proses tes awal yaitu 7, sedangkan dalam proses tes akhir jumlah peserta didik yang dapat memiliki jawaban yang benar yaitu 16 peserta sehingga kenaikannya juga banyak yaitu 9 peserta didik.

Soal yang efektif yaitu soal yang dijawab benar oleh peserta didik lebih banyak setelah proses pembelajaran daripada sebelum proses pembelajaran Apabila suatu soal dapat dijawab oleh semua peserta didik sebelum dan sesudah pembelajaran maka hal ini menunjukkan bahwa proses belajar mengajar tidak berpengaruh.

### SIMPULAN

Berdasarkan hasil penelitian pengembangan instrumen penilaian *HOTS* (*High Order Thinking Skills*) pada materi pencemaran lingkungan berjumlah 20 butir soal dapat disimpulkan bahwa instrumen penilaian *HOTS* berdasarkan validitas dari para ahli ditinjau dari beberapa aspek yakni (1) materi, (2) konstruksi, dan (3) bahasa memperoleh skor modus keseluruhan dengan besar 4 yang dikategorikan sangat valid. Nilai reliabilitas diperoleh melalui hasil tes peserta didik yaitu dengan besar 0,80 yang dikategorikan sangat tinggi. Tingkat kesukaran butir soal pada instrumen penilaian *HOTS* diperoleh hasil masing-masing 25% mudah, 50% sedang, dan 25% sukar dengan perbandingan 1 : 2 : 1. Hasil sensitivitas tes pada instrumen penilaian *HOTS* yaitu 100% sensitif. Berdasarkan hasil dari proses penelitian yang telah dilaksanakan maka dapat diambil sebuah kesimpulan bahwasannya bahwa butir soal *HOTS* materi pencemaran lingkungan dinyatakan valid dan reliabel sehingga dapat dijadikan sebagai contoh paket soal *HOTS* kelas X Sekolah Menengah Atas.

### SARAN

Instrumen penilaian *HOTS* (*High Order Thinking Skills*) yang telah dikembangkan oleh penulis hanya berupa butir soal pilihan ganda dan menggunakan tes tertulis, dibutuhkan inovasi baru agar peserta didik lebih semangat dan bisa menyertakan alasan dalam pemilihan jawaban yang dipilih.

### UCAPAN TERIMA KASIH

Peneliti ingin memberikan ucapan terimakasih pada Dra. Winarsih, M.Kes. yang berperan sebagai dosen pembimbing pada penelitian ini. Terima kasih juga kepada Dra. Isnawati, M.Si dan juga Dr. Sunu Kuntjoro, M.Si sebagai dosen penguji dan validator pada penelitian ini, serta peserta didik kelas X MIA 7 Sekolah Menengah

Atas Negeri 1 Bangil atas kerjasama selama pengambilan data.

#### DAFTAR PUSTAKA

- Arifin, Z. 2010. *Evaluasi Pembelajaran Prinsip Teknik Prosedur*. Bandung : Remaja Rosdakarya.
- Arikunto, S. 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara.
- Asrul, Rusydi Ananda, Rosnita. 2015. *Evaluasi Pembelajaran*. Bandung: Citapustaka Media
- Azwar, S. 2016. *Konstruksi Tes Kemampuan Kognitif*. Yogyakarta : Pustaka Belajar.
- Barrat, Carroline, (2014). *HOTs And Assessment*. International Seminar on current issues in Primary Education: Prodi Pendidikan Guru Sekolah Dasar UM Makasar.
- Basuki, I & Hariyanto. 2015. *Asestmen Pemblajaran*. Bandung: PT. Remaja Rosdakarya.
- Brookhart, S.M. 2010. *How to Assess HOTs In Your Classroom*, (USA : ASCD).
- Daryanto. 2010. *Evaluasi Pendidikan*. Jakarta: Rineka Cipta.
- Fanani, Moh. Zainal. 2018. Strategi Pengembangan Soal HOTs dalam Kurikulum 2013. *Journal of Islamic Religious Education*. Vol 2 (1) : hlm. 57-56.
- Hastuti, Rini. 2011. Pengembangan Perangkat Pembelajaran Tabung dan Kerucut Berdasarkan Masalah di Kelas IX Sekolah Menengah Pertama Negeri 2 Madiun. Tesis Pascasarjana. Universitas Negeri Surabaya.
- Lee, K., & Lai, Y. 2017. Facilitating HOTs with the flipped classroom model: a student teacher's experience in a Hong Kong secondary school. *Research and Practice in Technology Enhanced Learning*, 12(1).
- Nugroho, R. A. 2018. *HOTS (Kemampuan Berpikir Tingkat Tinggi : Konsep, Pembelajaran, Penilaian, dan Soal-soal)*. Jakarta : PT. Gramedia.
- Setyawan, D. 2013. *Data dan Metode Pengumpulan Data Penelitian*. Surakarta : Poltekes.
- Sharif, A., & Cho, S. 2015. 21st-Century Instructional Designers: Bridging the Perceptual Gaps between Identity, Practice, Impact and Professional Development. *RUSC. Universities and Knowledge Society Journal*, 12(3), 72–85.
- Sudjana, N. 2004. *Penilaian Hasil Proses Belajar Mengajar*. Bandung : Remaja Rosdakarya Offset.
- Sukardi. 2011. *Evaluasi Pendidikan: Prinsip dan Operasionalnya*. Jakarta Timur: PT Bumi Aksara.
- Suwandi, S. 2009. *Model Asesmen Dalam Pembelajaran*. Surakarta : Mata Padi Presindo.
- Suwono, H. 2012. *Penilaian Hasil Belajar IPA*. Malang: Bayumedia Publishing.
- Utami, Indri Putri & Aryeni. 2018. Analisis Soal UAS Mata Pelajaran Biologi Berdasarkan Dimensi Proses Kognitif Taksonomi Anderson. *Jurnal Pelita Pendidikan*. Vol 6 (3) : hlm. 185
- Widana, I. W. 2017. Modul : *Penyusunan Soal HOTs*. Jakarta: Direktorat Pembinaan Sekolah Menengah Atas, DIRJEN DIKDASMEN dan DEPDIKBUD.
- Yuniar, M., Rakhmat, C., dan Saepulrohan, A. (2015). "Analisis HOTS Pada Soal Objektif Tes dalam Pelajaran IPS Kelas V SD Negeri 7 Ciamis" dalam *Prosiding Jurusan PGSD, Fakultas Ilmu Pendidikan, UPI*: hlm. 187-195.