

**VALIDITAS EMPIRIS BUTIR SOAL *HIGH ORDER THINKING* (HOT) BERBASIS *COMPUTER BASED TEST* (CBT) PADA SUB MATERI SISTEM INDERA SISWA KELAS XI SMA**

**EMPIRICAL VALIDITY QUESTIONS OF *HIGH ORDER THINKING* (HOT) EVALUATION INSTRUMENT BASED ON *COMPUTER BASED TEST* (CBT) AT SENSORY SYSTEM SUB TOPIC OF STUDENT CLASS XI SENIOR HIGH SCHOOL**

**Betty Asrotaja Yonika Putri**

Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya  
Jalan Ketintang, Gedung C3 Lt.2 Surabaya 60231  
Email: yonika.betty@gmail.com

**Raharjo**

Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya  
Jalan Ketintang, Gedung C3 Lt.2 Surabaya 60231  
e-mail: raharjo@unesa.ac.id

**Abstrak**

Berpikir tingkat tinggi merupakan keterampilan yang membutuhkan kemampuan seperti menganalisis, mengevaluasi, dan menciptakan. Salah satu cara untuk mengembangkan kemampuan berpikir tingkat tinggi pada siswa adalah dengan membiasakan memberikan soal dalam bentuk pemecahan masalah, berpikir kreatif dan berpikir kritis. Penelitian ini bertujuan untuk mengembangkan serta menghasilkan instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT). Penelitian ini dilakukan pada Bulan Januari - April 2017. Butir soal yang dikembangkan sudah divalidasi secara teoritis dan dinyatakan valid. Instrumen yang dikembangkan diujicobakan pada 20 siswa kelas XI SMA Negeri 4 Surabaya. Metode pengumpulan data yang digunakan adalah metode tes dan metode angket. Teknik analisis data menggunakan teknik analisis secara deskriptif yang meliputi analisis hasil Tes *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) untuk mengetahui nilai reliabilitas, tingkat kesulitan, serta indeks daya beda dan analisis respon siswa untuk mengetahui kepraktisan butir soal *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT). Hasil analisis butir soal memiliki nilai reliabilitas sebesar 0,56%, tingkat kesukaran soal mudah sebesar 20%, sedang 55%, dan sulit sebesar 25%. Kepraktisan rata-rata instrumen evaluasi berdasarkan respon siswa sebesar 87,5 dengan interpretasi sangat layak. Berdasarkan hasil penelitian yang telah dilakukan maka dapat disimpulkan bahwa instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) pada sub materi Sistem Indera valid dan efektif.

**Kata kunci:** Validitas Empiris, *High Order Thinking* (HOT), *Computer Based Test* (CBT), sub materi Sistem Indera

**Abstract**

Higher Order Thinking is a skill that requires skills such as analyzing, evaluating, and creating. One way to develop high-order thinking skills in students is to get the problem solved in the form of problem solving, creative thinking and critical thinking. The good from this reasearch to develop and create evaluation instrument of *High Order Thinking* (HOT) with *Computer Based Test* . This research was conducted on Januray until April 2017. The items developed have been validated theoretically and valid. Developed Instrument was tested on twenty Students from class XI SMA N 4 Sidoarjo. Data Collection method used test method and questionnaire method. The data analysis technique used descriptive analysis technique which included analysis of High Order Thinking (HOT) test based on Computer Based Test (CBT) To know the value of reliability, difficulty level, and different power index and student response analysis to know the practicality of High Order Thinking (HOT) items based on Computer Based Test (CBT). The result of question analysis have reliability value of 0,56%, easy question level are 20%, medium level are 55%, hard level are 25%. The average of evaluation instrument practically based on student response are 87,5 withvery feasible interpretaion. Based on the result from research that done, so it can be concluded that evaluation instrument of *High Order Thinking* (HOT) with *Computer Based Test* (CBT) at sensory system sub topic are valid and effective.

**Keywords:** Empirical Validity, *High Order Thinking* (HOT), *Computer BasedTest* (CBT), Subtopic of the Sensory System

## PENDAHULUAN

Tujuan penilaian yang terdapat dalam Peraturan Menteri Pendidikan dan Kebudayaan Nomor 23 Tahun 2016 adalah untuk mengukur kemampuan hasil belajar peserta didik dan menyusun laporan perkembangan hasil belajar peserta didik. Penilaian kognitif dapat dilakukan dengan cara tertulis, lisan dan penugasan. Secara garis besar terdapat dua bentuk soal tes tertulis yaitu pilihan ganda dan esay. Menurut Djaali & Muljono (2007) tes tertulis atau *paper and pencil test* merupakan tes dalam bentuk pertanyaan yang jawabannya ditulis dalam bentuk tulisan oleh peserta didik.

Berpikir tingkat tinggi merupakan gabungan dari beberapa aspek yaitu kemampuan berpikir kritis, kreatif, serta pemecahan masalah. Menurut Sudargo, *et.al* 2010 pembelajaran Biologi di SMA selama ini jenjang kognitif yang dicapai adalah hanya pada taraf C1, C2 serta C3. Pengembangan soal untuk jenjang C4, C5 serta C6 sangat jarang. Keterbatasan kemampuan guru dalam membuat variasi soal dikarenakan guru tidak dapat memaksimalkan tuntutan yang terdapat pada Kurikulum 2013. Guru sering kali membuat soal menggunakan konteks yang teoritis dan tidak ada keterkaitan antara pengetahuan yang diperoleh dalam pembelajaran dengan kondisi yang sebenarnya dalam kehidupan sehari-hari. Kondisi saat ini kebanyakan siswa memiliki kemampuan yang rendah dalam memahami informasi yang bersifat kompleks, pemahaman teori, analisis dan pemecahan masalah. Untuk mengatasi hal ini siswa perlu dilatihkan dalam mengasah kemampuan berpikir tingkat tinggi dengan membiasakan memberikan soal-soal yang dapat memecahkan permasalahan.

Hasil pra penelitian yang dilakukan di SMA Negeri 4 Sidoarjo pada siswa kelas XII MIPA 5 rata-rata menganggap sub materi Sistem Indera banyak dan sulit dipahami. Hal tersebut berkaitan juga pada saat ulangan harian. Siswa sulit mengidentifikasi soal ulangan yang berkaitan dengan sistem indera. Setiap gambar pada soal ulangan harian dianggap oleh siswa menjebak, karena menurut siswa gambar yang disediakan tidak jelas. Ketidakjelasan gambar yang terdapat pada soal ulangan adalah karena proses perbanyak soal yang menggunakan mesin *fotocopy*. Gambar yang dihasilkan menyebabkan siswa tidak dapat membedakan antara gambar satu dengan gambar yang lain.

Pembuatan butir soal *High Order Thinking* (HOT) dapat disajikan dengan memanfaatkan teknologi informasi dan komunikasi. Butir soal yang telah dibuat

sebelumnya telah divalidasi secara teoritis dan dinyatakan valid. Penggunaan teknologi informasi dan komunikasi meningkat sejalan dengan pemanfaatan internet sebagai aspek pendukung. Hal ini berdampak pada adanya banyak variasi dalam proses pembelajaran. Salah satu variasi yang dapat diberikan adalah dengan pembuatan butir soal. Butir soal dibuat berbasis internet dengan media komputer untuk mendukung. Pemanfaatan teknologi informasi dan komunikasi dimanfaatkan di dunia pendidikan, termasuk penilaian berbasis komputer dengan menggunakan tes dengan komputer (De-Siqueira, 2009). Menurut Hosseini, *et.al* (2014) munculnya teknologi baru, ujian menggunakan komputer menjadi luas dan dalam pengujiannya dilaksanakan dalam skala besar.

Pemanfaatan media informasi dan telekomunikasi dalam proses perakitan butir soal dapat memberikan banyak manfaat. Menurut Martin (2008) kelebihan ujian berbasis komputer adalah dalam hal pengumpulan tes tidak perlu mengumpulkan dengan banyak kertas dan saat proses mengoreksi guru tidak perlu membutuhkan waktu yang lama. Bagi para siswa hal ini juga dapat membiasakan dalam berperilaku jujur dengan tidak mencontek. Butir soal pada *Computer Based Test* (CBT) dapat digunakan dengan dua sistem, yaitu sistem *online* dan sistem *offline*.

*Computer Based Test* (CBT) dengan sistem online ini memiliki kelebihan yaitu soal dan opsi jawaban yang terdapat di dalamnya otomatis akan diacak oleh sistem. Bagi guru dengan adanya sistem online dapat membantu dalam melihat grafik berupa urutan siswa yang banyak menjawab benar. Namun selain memiliki kelebihan, sistem online ini juga memiliki kekurangan yaitu server internet yang digunakan jika lemah dapat menghambat proses pengerjaan butir soal. Pada *Computer Based Test* (CBT) dengan sistem *offline* ini dapat membantu untuk sekolah-sekolah yang berada di daerah yang sulit untuk mengakses internet. Pengembangan *Computer Based Test* (CBT) ini dapat mengurangi dalam pemakaian kertas saat ujian dan dapat meminimalisir waktu dalam proses pengoreksian.

Berdasarkan paparan diatas tujuan dari penelitian ini adalah untuk mengembangkan Instrumen evaluasi *High Order Thinking* (HOT) dalam bentuk pemecahan masalah, berpikir kreatif dan berpikir kritis. berbasis *Computer Based Test* (CBT) pada sub materi Sistem Indera Kelas XI SMA.

## METODE PENELITIAN

Jenis penelitian yang digunakan merupakan penelitian pengembangan, karena dalam penelitian ini mengembangkan butir soal *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT). Penelitian ini dilakukan pada Bulan Januari - April 2017. Sasaran dalam penelitian ini adalah butir soal *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) yang sebelumnya telah divalidasi secara teoritis dan dinyatakan valid serta diujicobakan kepada 20 siswa Kelas XI SMA Negeri 4 Sidoarjo yang heterogen.

Metode Pengumpulan data dalam penelitian ini menggunakan metode metode tes, serta metode angket. Teknik analisis data menggunakan teknik analisis secara deskriptif yang meliputi analisis hasil Tes *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) untuk mengetahui nilai reliabilitas, tingkat kesulitan, serta indeks daya beda dan analisis respon siswa untuk mengetahui kepraktisan butir soal *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT).

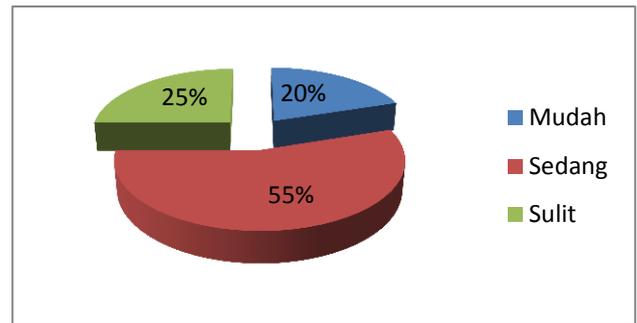
**HASIL DAN PEMBAHASAN**

Hasil uji coba terbatas pada instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) didapatkan data yang digunakan untuk menentukan reliabilitas, tingkat kesukaran soal serta indeks daya beda soal. Hasil reliabilitas soal dihitung dengan menggunakan rumus persamaan Kuder Rinchadson (KR-20) dan diperoleh hasil 0,56 dengan kategori sedang.

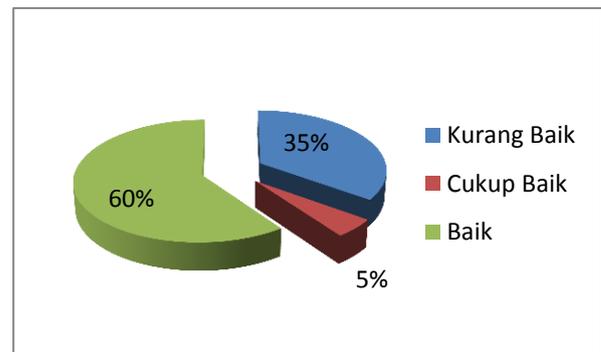
**Tabel 1.** Tingkat Kesulitan dan Indeks Daya Beda Instrumen Evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT).

No. soal	Tingkat kesulitan	Ket	Indeks Daya Beda	Ket
1	0,250	Sulit	0,573	Baik
2	0,750	Mudah	0,497	Baik
3	0,600	Sedang	0,567	Baik
4	0,250	Sulit	0,076	Kurang baik
5	0,650	Sedang	-0,062	Kurang baik
6	0,900	Mudah	-0,430	Kurang Baik
7	0,300	Sedang	0,628	Baik
8	0,250	Sulit	0,458	Baik
9	0,750	Mudah	0,497	Baik
10	0,400	Sedang	0,648	Baik
11	0,600	Sedang	-0,108	Kurang Baik
12	0,400	Sedang	0,716	Baik
13	0,700	Sedang	0,238	Cukup Baik
14	0,750	Mudah	0,191	Kurang baik
15	0,300	Sedang	0,448	Baik
16	0,600	Sedang	0,095	Kurang Baik
17	0,650	Sedang	0,354	Baik

No. soal	Tingkat kesulitan	Ket	Indeks Daya Beda	Ket
18	0,900	Mudah	0,176	Kurang baik
19	0,350	Sedang	0,583	Baik
20	0,250	Sulit	0,344	Baik



**Gambar 1.** Persentase Tingkat Kesulitan Butir Soal pada Instrumen Evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT).



**Gambar 2.** Persentase Indeks Daya Beda Butir Soal pada Instrumen Evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT).

Kepraktisan Pengembangan Instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) pada sub materi Sistem Indera berupa angket respon yang dibagikan kepada siswa setelah melakukan uji coba. Hasil respon siswa disajikan dalam Tabel 3:

**Tabel 2.** Respon Siswa Terhadap Pengembangan Instrumen Evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT).

No	Pernyataan	Presentase (%)	
		YA	TIDAK

No	Pernyataan	Presentase (%)	
		YA	TIDAK
1	Tampilan yang terdapat pada <i>Computer Based Test</i> menarik.	100	0
2	Siswa dapat dengan mudah mengoperasikan <i>Computer Based Test</i> .	90	10
3	Alokasi waktu yang diberikan cukup untuk mengerjakan soal.	55	45
4	Sistem pengacakan soal dan jawaban, dapat membuat siswa mengerjakan soal tes dengan mandiri dan jujur.	90	10
5	Soal-soal yang terdapat pada tes dapat dengan mudah dipahami.	90	10
6	Bobot soal yang diberikan dapat melatih siswa dalam berpikir tingkat tinggi.	95	5
7.	Setelah mengerjakan butir soal <i>High Order Thinking</i> , siswa lebih termotivasi untuk belajar.	100	0
8	Ujian biologi berbasis <i>Computer Based Test</i> (CBT) merupakan hal yang baru bagi siswa.	80	20
Jumlah rata-rata		87,5	12,5

Penelitian ini bertujuan untuk menghasilkan produk berupa instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) pada sub materi Sistem Indera berdasarkan validitas, kualitas soal yang berupa nilai reliabilitas, tingkat kesukaran, indeks daya beda, serta kepraktisan. Butir soal yang disajikan sebelumnya telah divalidasi secara teoritis dan dinyatakan valid. Terdapat 3 aspek yang sebelumnya telah divalidasi yaitu aspek materi, konstruksi dan bahasa. Rata-rata persentase aspek yang diperoleh adalah sebesar 97,5% yang secara teoritis dinyatakan sangat layak dan valid. Menurut Widoyoko (2016) jika data yang dihasilkan dari sebuah instrumen valid, maka instrumen tersebut juga valid. Menurut Kunandar (2013) validitas yang tinggi mempunyai arti bahwa mampu mengungkapkan aspek hasil belajar tertentu secara tepat.

Instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) yang dinyatakan valid kemudian dilakukan uji coba secara terbatas kepada 20 siswa untuk mengetahui nilai reliabilitas, tingkat kesukaran serta indeks daya beda. Nilai reliabilitas yang didapatkan pada pengembangan instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) adalah sebesar 0,56 dengan interpretasi reliabilitas yaitu sedang,

sehingga dapat dinyatakan bahwa instrumen yang dihasilkan memenuhi kriteria tes yang baik yaitu sudah reliabel. Menurut Azwar (2016), Suatu tes dinyatakan baik apabila tes tersebut valid dan reliabel. Menurut Matondang (2009) reliabilitas mengacu pada konsistensi hasil ukur tes jika dilakukan pengukuran ulang pada kelompok yang sama. Perhitungan tingkat kesulitan butir soal pada instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) menghasilkan perbandingan tingkat kesulitan yaitu 4:11:5. Menurut Rachma (2015), suatu instrumen tes dapat dikatakan proporsional yang baik apabila memiliki perbandingan sebesar 30%,50%, dan 20% pada tingkatan soal mulai dari mudah, sedang dan sulit. Butir soal dalam kategori mudah terdapat pada butir soal pada nomor 2, 6, 9, 14, dan 18. Sedangkan untuk butir soal dalam kategori sedang terdapat pada nomor 3, 5, 7, 10, 11, 12, 13, 15,16, 17, dan 19. Untuk butir soal dalam kategori sulit, terdapat pada nomor 1, 4, 8, dan 20.

Butir soal yang dianggap sulit bagi siswa adalah butir soal mengenai mata, lidah. Terdapat 2 butir soal yang dianggap sulit oleh siswa mengenai mata yaitu pada indikator memprediksi urutan mekanisme penghantar implus pada mata dan mengaitkan struktur dan fungsi dengan kasus mengenai masalah penglihatan melalui studi literatur. Dua butir soal selanjutnya yaitu mengenai lidah. Pada butir soal mengenai lidah ini indikator soalnya yaitu tentang menghubungkan struktur lidah dengan proses kerja lidah dan menganalisis kasus mengenai kelainan yang terjadi pada indera pengecap.

Fungsi dari tingkat kesulitan pada suatu instrumen menurut Maulida dan Yusrizal Muhibbuddin (2015) adalah dapat memperoleh informasi mengenai penekanan pengajaran, melakukan tindak lanjut terhadap butir soal yang dianggap biasa, dan dapat digunakan sebagai pengenalan terhadap suatu konsep yang diajarkan ulang pada siswa serta memberikan masukan terhadap siswa terhadap hasil belajar yang diperoleh.

Perhitungan nilai indeks daya beda butir soal pada instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) diperoleh hasil yaitu 35% kurang baik, 5% cukup baik, dan 60% baik. Menurut Syamsudin (2012), tujuan dari perhitungan daya beda adalah untuk mengetahui apakah suatu butir soal dapat membedakan antara kelompok atas dan kelompok bawah yang dapat diketahui melalui selisih proporsi dari yang menjawab pada masing-masing kelompok. Terdapat beberapa untuk mengetahui

apakah suatu butir soal dapat membedakan antara kelompok atas dan kelompok bawah yang dapat diketahui melalui selisih proporsi dari yang menjawab pada masing-masing kelompok. Terdapat beberapa faktor yang dapat mempengaruhi rendahnya nilai indeks daya beda pada suatu instrumen evaluasi yaitu pada gambar yang terdapat pada beberapa butir soal tidak berfungsi sehingga dapat menimbulkan siswa untuk menebak jawaban, selain itu faktor lain yang dapat mempengaruhi adalah alokasi waktu yang diberikan terlalu singkat sehingga membuat siswa merasa terburu-buru dan tidak memahami soal dengan teliti.

Tingkat kesulitan soal dapat berpeluang pada nilai beberapa indeks daya beda soal. Hal ini dikarenakan jika instrumen evaluasi yang memiliki tingkat kesulitan tinggi (sulit) tidak dapat dipastikan jika nilai indeks daya bedanya juga tinggi. Sebaliknya jika instrumen evaluasi memiliki tingkat kesulitan sedang dapat memiliki nilai indeks daya beda tinggi. Menurut Zein (2013), tidak ada pengaruh mengenai tinggi rendahnya nilai tingkat kesulitan dengan indeks daya beda pada taraf yang signifikan 1%.

Berdasarkan hasil analisis pada reliabilitas, tingkat kesulitan dan indeks daya beda pada instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) dengan menggunakan program ANBUSO versi 6.0 diperoleh data setelah dilakukan ujicoba mengenai instrumen tersebut bahwa terdapat 17 butir soal yang diterima dan 3 butir soal yang perlu dilakukan revisi. Instrumen evaluasi *High Order Thinking* (HOT) ini dikembangkan menggunakan indikator-indikator yang berasal dari tiga aspek dalam berpikir tingkat tinggi yaitu, kemampuan berpikir kritis, kreatif, serta pemecahan masalah. Menurut Rofiah, dkk (2013) instrumen evaluasi berpikir tingkat tinggi dapat menjadi instrumen yang baik apabila dalam penyusunan instrumen tersebut terdapat tiga aspek dalam berpikir tingkat tinggi yang terdiri dari berpikir kritis, kreatif dan pemecahan masalah. Terdapat 6 aspek indikator dalam kemampuan berpikir kritis yang dikembangkan dari teori yang diajukan oleh Ennis dalam Sunaryo (2014) diantaranya adalah mampu dalam mengajukan masalah, merevisi konsep yang salah, merencanakan strategi, mengevaluasi keputusan mengkritik, serta mampu mengevaluasi keputusan. Instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) yang disusun memiliki pokok soal berupa pernyataan yang berisi masalah untuk merangsang stimulus pada siswa agar dapat berpikir kritis dalam

menyelesaikan sebuah masalah tersebut dalam bentuk soal. Indikator soal yang dapat membiasakan siswa untuk berpikir kritis adalah memprediksi urutan mengenai mekanisme penghantar impuls pada mata, mengaitkan hubungan struktur telinga manusia dengan fungsinya, mengurutkan proses penerimaan rangsang bunyi pada telinga manusia. Instrumen evaluasi yang digunakan untuk menguji kemampuan berpikir kreatif terdapat soal dalam bentuk gambar yang disediakan serta menyajikan masalah yang dapat memunculkan kreativitas siswa dalam memecahkannya. Indikator soal yang dapat menguji kemampuan berpikir kreatif siswa adalah mengaitkan kelainan kulit dengan struktur dan fungsinya, menghubungkan hubungan antara indera pembau dengan indera pengecap, menghubungkan struktur lidah dengan proses kerja lidah, mengaitkan struktur dan fungsi mata dengan proses penglihatan, mengaitkan struktur dan fungsi mata dengan kasus mengenai masalah penglihatan, mengurutkan proses penerimaan rangsang bunyi pada telinga manusia. Menurut Woods dalam Mourtos, *et.al* (2004) dalam pemecahan masalah terdapat 11 aspek indikator diantaranya siswa mampu mengidentifikasi masalah, menghubungkan antara sebab-akibat, menerapkan konsep yang sesuai dengan masalah, memiliki rasa ingin tahu yang tinggi, dapat membuat suatu gambar dalam menyelesaikan masalah, menjelaskan berbagai solusi pada suatu masalah, memiliki pemikiran yang terbuka, dapat membuat keputusan, teliti dalam bekerja, berani berargumentasi, serta mampu merefleksikan keefektifan dalam proses pemecahan masalah. Indikator soal yang dapat membiasakan siswa dalam pemecahan masalah adalah menganalisis kasus yang terjadi akibat gangguan pada pendengaran, menganalisis kasus mengenai struktur dan fungsi lidah, menganalisis kasus mengenai struktur dan fungsi pada indera pembau, menganalisis kasus mengenai kelainan pada indera pembau, menganalisis kasus berdasarkan pada struktur dan fungsi indera pembau.

Pengembangan instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) dinilai kepraktisannya berdasarkan tanggapan siswa setelah mengerjakan butir soal tersebut. Para siswa merespon dengan baik pada instrumen evaluasi berbasis komputer tersebut, hal ini dapat diketahui bahwa dengan presentase yang didapatkan memiliki rata-rata 87,5% yang diinterpretasikan dalam kategori sangat layak.

Instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) dimuat

dalam bentuk ujian online dengan menggunakan sistem pengacakan butir soal serta opsi jawaban secara otomatis disertai dengan proses *scoring* otomatis dan terdapat *feedback* berupa jawaban yang benar. Para siswa dapat dengan mudah mengoperasikannya dengan mudah. Hal ini dibuktikan dengan nilai presentase sebesar 95%. Menurut Piaw (2012) *Computer Based Test* merupakan sistem penilaian secara otomatis dan dinilai dapat menjadi efektif karena dapat dengan mudah dioperasikan oleh siapa saja. Menurut Jamil, *et.al* (2012) penilaian secara online dianggap telah mengurangi beban guru dalam hal mengoreksi jawaban siswa, karena siswa dengan mudah dapat mengetahui nilainya secara langsung tanpa menunggu diberitahu oleh guru.

#### SIMPULAN

Berdasarkan hasil analisis dan pembahasan, dapat disimpulkan bahwa instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) memiliki nilai reliabilitas sebesar 0,56 dalam kategori sedang, tingkat kesulitan yang beragam yaitu 25% soal mudah, 55% soal sedang, 20% soal sulit dan nilai indeks daya beda yaitu 5% cukup baik, 35% kurang baik, dan 60% baik. Efektivitas instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) yang ditinjau dari respon siswa setelah mengerjakan tes adalah sebesar 87,5 dengan interpretasi sangat baik.

#### SARAN

Instrumen evaluasi *High Order Thinking* (HOT) berbasis *Computer Based Test* (CBT) yang dikembangkan oleh penulis hanya berupa pilihan ganda, perlu adanya inovasi baru agar siswa dapat menyertakan alasan dalam pemilihan jawaban yang dipilih. Perlu adanya penelitian lebih lanjut mengenai penambahan fitur berupa multimedia agar siswa dapat memecahkan masalah melalui video dan mengenai server *Computer Based Test* (CBT) agar dapat digunakan dalam keadaan offline yang memanfaatkan banyak komputer.

#### UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada Dr. Widowati Budijastuti M.Si. dan Dr. Sifak Indana, M.Pd. selaku dosen penguji dan Erlin R Purnama M.Si. selaku dosen validator serta penelaah pada penelitian. Siswa-siswi Kelas XI MIPA 6 SMA Negeri 4 Sidoarjo atas kerjasamanya selama pengambilan data

#### DAFTAR PUSTAKA

- Azwar, Saifuddin. 2016. *Konstruksi Tes Kemampuan Kognitif*. Yogyakarta : Pustaka Belajar.
- De-Siqueira. Jose Macario. Peris Fajames. Guillermo. Gimenez. Fernando. Magal Royo dan Teresa. 2009. Spanish students and teachers' preferences towards computer-based and paper-and-pencil tests at universities. *Jurnal ScienceDirect*. Vol 01: hal 814-817.
- Djaali dan Muljono. 2007. *Pengukuran Dalam Bidang Pendidikan*. Jakarta.Grasindo
- Hosseini, Monirosadat., Mohamad J Zairol Abidin, dan Mostafa Baghdania. 2014. Comparability of Test Results of Computer Based Tests (CBT) and Paper and Pencil Tests (PPT) among English Language Learners in Iran. *Jurnal ScienceDirect*. 98:659-667.
- Jamil, Mubashrah., Fazzlijan Mohamed Adnan Khan, dan Mona Masood. 2012. Computer Based VS Paper Based Examinations: Perceptions of University Teacher. *The turkish online journal of educational technology*. Volume 11 issue 4.
- Mourtos, N.J., N.D Okamoto, dan J Rhee. 2004. Defining, Teaching, and Assessing Problem Solving Skills. *7th UICEE Annual Conference on Engineering Education* Mumbai. Di akses pada tanggal 23 April 2017
- Piaw, Chua Yan. 2012. Replacing Paper-Based Testing with Computer-Based Testing Inassessment: Are We Doing Wrong?. *Jurnal ScienceDirect*. 64: 655-664.
- Rachma,Nurmaidia. 2015. Pengembangan Tes Elektronik (*E.Test*) Berbasis Komputer Pada Materi Bioteknologi di SMA Negeri 1 Surabaya. *Jurnal BioEdu*. Volume 4 No.3.
- Rofiah, Emi.,Nonoh Siti Aminah, dan Elvin Yusliana Ekawati. 2013. Penyusunan Instrumen Tes Kemampuan Berpikir Tingkat Tinggi Fisika Pada Siswa SMP. *Jurnal Pendidikan Fisika*, 1(2):21.
- Sudargo, Fransisca., dan Soesy Asiah S,. 2010. Kemampuan Pedagogik Calon Guru dalam Meningkatkan Kemampuan Berpikir Kritis dan Ketrampilan Proses Siswa Melalui Pembelajaran Berbasis Praktikum. *Jurnal Pengajaran MIPA*. 16 (1), 4-12.

- Sunaryo, Yono. Model Pembelajaran Berbasis Masalah Untuk Meningkatkan Kemampuan Berpikir Kritis dan Kreatif Matematika Siswa di Kota Tasikmalaya. *Jurnal Pendidikan dan Keguruan*. Vol 1. No.2.
- Syamsudin. 2012. Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pola Jawaban Tes. *Jurnal Ilmu Tarbiyah "At-Tajdid"*. Volume 1 No.2.
- Widoyoko, Eko Putro. 2016. *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Belajar
- Zein, Anizam., Muhyiatul Fadillah, dan Rahma Novianti. Hubungan Antara Validitas Butir, Reliabilitas, Tingkat Kesukaran dan Daya Pembeda Soal Ujian Semester Genap Bidang Studi Biologi Kelas XI SMA/MA Negeri di Kota Padang Tahun Ajaran 2010/2011. *Prosiding Semirata FMIPA UNIVERSITAS LAMPUNG*.