

## IDENTIFIKASI BUTIR SOAL BIOLOGI PADA SMK KESEHATAN NURUL ISLAM HUSADA LAMPUNG SELATAN

**Yulius Restadi Kurniawan\***, Rini Rita T. Marpaung, Berti Yolida  
Pendidikan Biologi, FKIP Universitas Lampung

\* *Corresponding author*, Hp: 085279298554, E-mail: Yulius\_Resta@gmail.com

**Abstract: Identification Of Biology Exam In Health Vocational High School Nurul Islam Husada South Lampung.** The aim of this research was to find out the validity, difficulty level, discrimination power, humbug effectivity, and the reliability of XI grade Biology exam at health vocational high school of Nurul Islam Husada. The sampling method was purposive sampling techniques, data collecting method was using documentation. The result of validity was 20% question valid and 80% question was not valid. The criteria of the question was 0% very high validity, 2,5% was high validity, 10% was average validity, 7% was low validity, and 0% was very low validity. The difficulty level was 5% very hard, 30% was hard, 45% was average, 20% was easy, and 0% was very easy. The difference effort from 40 question was 20% low, 42,5% was average, and 37,5% was high. 80% of the humbug was work, but 20% of the humbug was failed. The reliability was 0,584, it means the exam was reliable.

**Keywords:** ITEMAN program version 3.50, question analysis, validity

**Abstrak: Identifikasi Butir Soal Biologi Pada Smk Kesehatan Nurul Islam Husada Lampung Selatan.** Penelitian ini bertujuan untuk mengukur validitas, tingkat kesukaran, daya beda, efektifitas pengecoh, dan reliabilitas soal ujian Biologi kelas XI pada SMK Kesehatan Nurul Islam Husada. Metode pengumpulan data dengan dokumentasi. Hasilnya validitas diketahui sebanyak 20% soal valid dan 80% soal tidak valid. Dengan kriteria sebanyak 0% soal memiliki validitas sangat tinggi, 2,5% soal memiliki validitas tinggi, 10% soal memiliki validitas sedang, 7,5% memiliki validitas rendah dan 0% soal memiliki validitas sangat rendah. Tingkat kesukaran soal sebanyak 5% sangat sukar, 30% sukar, 45% sedang, 20% mudah dan 0% sangat mudah. Daya beda dari 40 butir soal diketahui sebanyak 20% rendah, 42,5% sedang dan 37,5% tinggi. Sebanyak 80% pengecoh berfungsi, sedangkan sebanyak 20% pengecoh tidak berfungsi. Reliabilitas adalah 0,584 artinya soal cukup reliabel.

**Kata kunci:** analisis soal, program ITEMAN versi 3.50, validitas

## PENDAHULUAN

Pendidikan adalah salah satu bentuk perwujudan kebudayaan manusia yang dinamis dan sarat perkembangan. Pendidikan di Indonesia bertujuan untuk menyampaikan pengetahuan kepada peserta didik dan meningkatkan kualitas peserta didik sehingga menjadi manusia yang kreatif, terampil serta professional (Trianto, 2010: 51). Pendidikan merupakan hal penting dalam suatu negara, sebab jika suatu negara memiliki kualitas pendidikan yang baik, maka negara tersebut akan menghasilkan sumber daya manusia yang berkualitas dan memiliki daya saing.

Saat ini, pemerintah berusaha meningkatkan kualitas pendidikan di Indonesia secara terus menerus. Hal tersebut diterapkan melalui Kurikulum Tingkat Satuan Pendidikan (KTSP), yang merupakan pengembangan dan penyempurnaan dari Kurikulum Berbasis Kompetensi (KBK). Kurikulum Tingkat Satuan Pendidikan (KTSP) merupakan hasil pembaharuan Kurikulum Berbasis Kompetensi (KBK) yang menghendaki suatu pembelajaran yang tidak hanya mempelajari tentang konsep, teori dan fakta, tetapi juga penerapannya dalam kehidupan sehari-hari (Trianto, 2010: 54).

Perubahan kurikulum yang terjadi membuat proses evaluasi dalam dunia pendidikan mengalami perubahan waktu pelaksanaannya, dari kurikulum yang menggunakan sistem evaluasi catur wulan sampai dengan sekarang yang menggunakan sistem evaluasi per semester. Dengan di-berlakukannya Kurikulum Tingkat Satuan Pendidikan (KTSP) yang merupakan pengembangan dan penyempurnaan Kurikulum Berbasis

Kompetensi (KBK) maka program pengajaran dalam satu tahun dibagi menjadi dua semester sehingga untuk mengukur tingkat keberhasilan pengajaran dilakukan evaluasi tiap semester.

Kegiatan evaluasi mempunyai peranan penting dalam pendidikan, begitu pula dalam proses pembelajaran karena dengan evaluasi dapat diketahui hasil dari kegiatan pembelajaran yang dilaksanakan dan dari hasil tersebut dapat ditentukan tindak lanjut yang akan dilakukan. Evaluasi sendiri merupakan proses penentuan sampai sejauh mana kemampuan yang dapat dicapai oleh siswa dalam proses pembelajaran yang sebelumnya telah ditentukan besaran patokan nilainya, sehingga dapat dilakukan penilaian (*value judgment*).

Proses evaluasi seperti yang telah kita ketahui dapat dilakukan dengan kegiatan tes maupun non-tes. Alat evaluasi yang digunakan pada evaluasi per semester adalah tes obyektif pilihan ganda, karena bentuk ini dapat mencakup banyak materi pelajaran, penskoran obyektif dapat dikoreksi siapa saja dengan mudah, tingkat berpikir yang terlibat bisa dari tingkat pengetahuan sampai tingkat sintesis dan analisis.

Hasil penelitian Arrizqi (2010: 71) yaitu analisis UAS Biologi Kelas XI di Kabupaten Tegal pada soal pilihan ganda menunjukkan rata-rata persentase tingkat kesukaran mudah 29%, sedang 44% dan sulit 26% ; reliabilitas sebesar 0,467; daya beda soal 80%; efektifitas pengecoh sebanyak 49% kurang berfungsi. Sedangkan pada analisis secara kualitatif menunjukkan bahwa 33 butir sesuai (aspek materi dan bahasa), tapi aspek konstruksi perlu direvisi. Penelitian

yang dilakukan oleh Afyana (2010: 58) mengenai analisis soal latihan UN IPA di Kabupaten Batang, menunjukkan kualitas soal dalam kategori kurang baik. Karena soal didominasi oleh tingkat kesukaran soal yang tinggi.

Hasil belajar adalah kemampuan yang dimiliki oleh siswa setelah menerima pengalaman belajarnya. Evaluasi terhadap hasil belajar siswa selama satu semester diujikan melalui UAS. Nilai dari UAS ini merupakan gambaran penguasaan kompetensi yang dipelajari siswa selama satu semester, sehingga diperlukan soal yang berkualitas baik. Untuk mendapatkan soal yang berkualitas maka harus dilakukan langkah pengembangan soal yang sesuai standar agar dapat mengevaluasi secara tepat apa yang akan diukur (Sudjana, 2001: 47).

Tujuan penelitian ini adalah untuk mengetahui validitas, tingkat kesukaran, daya pembeda, efektifitas pengecoh, dan reliabilitas dari butir soal ujian Biologi semester ganjil pada SMK Kesehatan Nurul Islam Husada Lampung Selatan tahun pelajaran 2015/ 2016.

## **METODE**

Penelitian ini dilaksanakan pada bulan Desember 2015 di SMK Kesehatan Nurul Islam Husada tepatnya saat diadakan ujian semester ganjil pada tahun pelajaran 2015/ 2016. Populasi dari penelitian ini adalah lembar jawaban ujian Biologi semester ganjil pada SMK Kesehatan Nurul Islam Husada. Penentuan sampel soal dengan *purposive sampling* yaitu soal (berserta kunci jawaban) ujian Biologi semester ganjil pada SMK Kesehatan Nurul Islam Husada tahun ajaran 2015/

2016, sedangkan lembar jawaban siswa juga diperoleh dari SMK Kesehatan Nurul Islam Husada. Desain penelitian ini adalah deskriptif sederhana. Metode deskriptif adalah metode penelitian untuk membuat gambaran mengenai situasi atau kejadian, sehingga metode ini akan mengadakan akumulasi data dasar belaka (Nazir, 2005: 84).

Terdapat dua pendekatan yang dapat digunakan untuk mengetahui kualitas tes yang dibuat yaitu pendekatan kualitatif dan kuantitatif. Pendekatan kualitatif dilakukan dengan cara menelaah butir soal, sebaiknya dilakukan sebelum perangkat tes tersebut diujikan. Hal yang ditekankan adalah penilaian dari aspek materi, konstruksi, dan bahasa. Sedangkan pendekatan kuantitatif merupakan metode penelaahan butir tes yang didasarkan pada data empiris yang diperoleh melalui respons peserta tes. Dengan demikian, pendekatan kuantitatif dilakukan setelah tes diberikan kepada peserta tes atau pernah diujikan pada kelompok peserta lain yang memiliki karakteristik hampir identik.

Analisis kualitatif dinamakan validitas logis (*logical validity*) dan analisis kuantitatif dinamakan validitas empiris (*empirical validity*). Melalui validitas empiris akan diperoleh informasi butir soal yang meliputi tingkat kesukaran, daya pembeda dan efektifitas pengecoh.

Dari penjelasan di atas dapat diketahui bahwa kualitas soal yang baik dapat dilihat dari validitas logis dan validitas empiris. Validitas dapat berkenaan dengan ketepatan alat penilaian terhadap konsep yang

dinilai sehingga betul-betul menilai apa yang seharusnya dinilai.

**Validitas logis**, validitas ini terdiri atas:

- a. **Validitas isi**, sebuah tes dikatakan memiliki validitas isi apabila mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran yang diberikan (Arikunto, 2007: 67). Validitas isi menurut menunjuk pada sejauh mana instrumen tersebut mencerminkan isi yang dikehendaki.
- b. **Validitas konstruksi**, sebuah tes dikatakan memiliki validitas konstruksi jika butir soal yang membangun tes tersebut mengukur setiap aspek berpikir seperti yang dirumuskan dalam indikator.

**Validitas empiris**, validitas ini terdiri atas:

1. **Tingkat kesukaran**, adalah pengukuran seberapa besar derajat kesukaran suatu soal. Jika pada suatu soal memiliki tingkat kesukaran seimbang, maka dapat dikatakan bahwa soal tersebut baik. Sebaiknya dalam penyusunan tes tidak terlalu sukar dan tidak pula terlalu mudah. Angka yang dapat menunjukkan mengenai tingkat kesukaran dikenal dengan "*Difficulty Index*" yang diberi lambang P (*Proportion*). Besarnya tingkat kesukaran antara 0,00 sampai dengan 1,0. Klasifikasi indeks kesukaran disajikan pada Tabel 1.

Tabel 1. Klasifikasi indeks kesukaran

No.	Rentang	Kriteria
1.	P antara 0,00 - 0,10	Sangat sukar
2.	P antara 0,11 - 0,30	Sukar
3.	P antara 0,31 - 0,70	Sedang
4.	P antara 0,71 - 0,90	Mudah
5.	P > 0,90	Sangat mudah

2. **Daya pembeda**, adalah pengukuran sejauh mana suatu butir soal yang mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu. Semakin tinggi koefisien daya pembeda suatu butir soal, semakin mampu butir soal tersebut membedakan antara peserta didik yang menguasai kompetensi dengan peserta didik yang kurang menguasai kompetensi. Seluruh peserta tes dikelompokkan menjadi 2 yaitu kelompok atas dan kelompok bawah. Angka yang menunjukkan besarnya daya pembeda disebut "*Discriminating Power*" yang diberi lambang D. Besarnya daya pembeda yang berkisar antara 0,00 sampai 1,00. Klasifikasi daya pembeda dapat disajikan pada Tabel 2.

Tabel 2. Klasifikasi daya pembeda

No.	Rentang	Kriteria
1.	D < 0,10	Daya pembeda soal rendah
2.	D = 0,10 - 0,29	Daya pembeda soal sedang
3.	D > 0,3	Daya pembeda soal tinggi

3. **Analisis pengecoh**, pada soal pilihan ganda terdapat alternatif jawaban/ *option* yang merupakan pengecoh (distraktor). Butir soal yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya butir soal yang kurang baik, pengecohnya akan dipilih secara tidak merata. Pengecoh dianggap baik apabila jumlah peserta didik yang memilih pengecoh itu sama atau mendekati jumlah ideal. Menurut Surapranata (2005: 24) suatu pengecoh dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 5% peserta tes.

4. **Reliabilitas**, adalah ketetapan atau keajegan suatu alat evaluasi (Sudjana, 2001: 219). Reliabilitas merupakan indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya dan dapat diandalkan. Suatu tes dikatakan memiliki ketetapan jika dapat dipercaya, konsisten/ stabil dan produktif kapanpun tes tersebut digunakan.

Terdapat tiga cara untuk mengetahui reliabilitas suatu tes yaitu:

- a. Dengan metode dua tes, dua tes yang paralel dan setaraf di berikan kepada sekelompok siswa, kemudian kedua hasilnya dicari korelasinya.
- b. Dengan metode satu tes, sebuah tes diberikan dua kali kepada sekelompok siswa yang sama tapi dalam waktu yang berbeda. Kemudian pada kedua hasilnya dicari korelasinya.
- c. Metode *split-half*, suatu tes dibagi menjadi dua bagian yang sama tingkat kesukarannya, sama isi dan bentuknya. Kemudian dilihat skor masing-masing bagian

paruhan tes tersebut dan mencari korelasinya.

Nilai dari reliabilitas diberi lambang  $r$  yang dapat mencari besarnya dengan menggunakan rumus KR 20 atau Spearman-Brown. Harga  $r$  yang diperoleh dikonsultasikan dengan  $r$  tabel *product moment* dengan taraf signifikan 5%. Jika harga  $r$  hitung  $>$   $r$  tabel, maka soal tersebut reliabel.

Klasifikasi reliabilitas disajikan pada Tabel 3.

Tabel 3. Klasifikasi reliabilitas

No.	Rentang	Kriteria
1.	$0,800 < r \leq 1,000$	Sangat tinggi
2.	$0,600 < r \leq 0,800$	Tinggi
3.	$0,400 < r \leq 0,600$	Sedang
4.	$0,200 < r \leq 0,400$	Rendah
5.	$0,000 < r \leq 0,200$	Sangat rendah

Obyek penelitian ini adalah soal (termasuk kunci jawaban) dan lembar jawaban siswa pada Ujian Biologi Semester Ganjil tahun ajaran 2015/ 2016 pada SMK Kesehatan Nurul Islam Husada Lampung Selatan. Hasil tes butir soal Ujian Biologi Semester Ganjil Kelas XI meliputi validitas, tingkat kesukaran, daya pembeda, efektifitas pengecoh, serta reliabilitas.

Analisis dilakukan pada soal bentuk pilihan jamak dengan menggunakan program ITEMAN versi 3.50 dan SPSS versi 16.0. Alasan digunakannya ITEMAN dan SPSS 16.0 sebagai program untuk mengolah data secara empiris adalah lebih tepat, efektif, dan mudah digunakan yang sesuai dengan teori tes klasik.

Program ITEMAN (*ITEM AND TEST ANALYSIS*) versi 3.50 merupakan salah satu program klasik yang digunakan untuk menganalisis butir soal yaitu menelaah butir soal melalui informasi dari jawaban

peserta tes (*testee*) untuk meningkatkan kualitas butir soal yang bersangkutan dengan teori klasik. Alasan dipilihnya program ini adalah mudah dilakukan dengan komputer, murah, dan sederhana.

Dalam program ini dimasukkan data berupa nama siswa (bisa diganti dengan kode siswa) dan jawaban/ hasil tes siswa. Pada baris ketiga dituliskan pilihan jawaban (biasanya adalah 4 atau 5 pilihan). Pilihan jawaban ini ditulis sebanyak soal yang digunakan dalam tes.

ITEMAN merupakan perangkat/ program untuk menganalisis butir soal dan tes. Program ini didasarkan pada teori tes klasik. Analisis soal secara klasik adalah proses penelaahan butir soal melalui informasi dari jawaban peserta didik guna meningkatkan mutu butir soal yang bersangkutan dengan menggunakan teori klasik. Melalui data empiris butir soal yang ditampilkan akan dapat menjelaskan kualitas soal tersebut.

Terdapat kelemahan utama dari program ini yaitu sangat dipengaruhi oleh kemampuan responden. Artinya jika soal diujikan pada anak berkemampuan tinggi dengan anak berkemampuan rendah maka akan terjadi perbedaan hasil analisis. Untuk mengantisipasi hal tersebut maka biasanya analisis soal dengan menggunakan ITEMAN dilakukan secara sampling. Semakin besar sampling dan semakin baik teknik samplingnya maka semakin baik kualitas hasil analisis.

Pada analisis butir menggunakan teori tes klasik tipe objektif, kualitas butir dilihat dari paling tidak dua parameter yaitu tingkat kesukaran dan daya pembeda. Selain itu juga menguji efektifitas distraktor-

distraktor pada setiap butir untuk menentukan apakah distraktor tersebut berfungsi atau belum. (Rudyatmi dan Anni, 2010: 61).

Program ini termasuk satu paket dalam MicroCat yang dikembangkan oleh perusahaan *Assessment System Corporation* mulai tahun 1982 dan mengalami revisi pada tahun 1984, 1986, 1988, dan 1993. Mulai dari versi 2.00 sampai dengan versi 3.50. Fungsi dari program ITEMAN adalah:

1. Untuk menganalisis data file (format ASCII) jawaban butir soal yang dihasilkan manual melalui *manual entry* data atau dari mesin *scanner*.
2. Menskor dan menganalisis data soal pilihan ganda untuk 30.000 siswa dan 250 butir soal.
3. Menganalisis sebuah tes yang terdiri dari 10 skala (subtes) dan memberikan informasi tentang validitas setiap butir (daya pembeda, tingkat kesukaran, proporsi jawaban pada setiap *option*), reliabilitas (KR-20/Alpha), standar *error measurement*, *mean*, *variance*, standar deviasi, *skew*, kurtosis untuk jumlah skor pada jawaban benar, skor minimum dan maksimum, skor median dan frekuensi distribusi skor (Rudyatmi dan Anni, 2010: 61).

Program ITEMAN juga memberikan hasil skor untuk setiap peserta tes yang menunjukkan jumlah benar dari seluruh jawaban. Sebelum menggunakan program ITEMAN perlu diketahui bahwa terdapat 5 baris utama yang harus dientrykan. Data yang akan dianalisis diketik melalui *notepad* atau *Microsoft Office Word* dengan jenis font *Courier New*. File data yang

akan dientrykan ke program ITEMAN terdiri atas 5 baris.

- a. Baris pertama adalah baris pengontrol yang mendeskripsikan data.
- b. Baris kedua adalah daftar kunci jawaban setiap butir soal.
- c. Baris ketiga adalah daftar jumlah option untuk setiap butir soal.
- d. Baris keempat adalah daftar butir soal yang akan dianalisis (jika butir yang akan dianalisis diberi tanda Y, jika tidak diikuti dalam analisis diberi tanda N).
- e. Baris kelima dan seterusnya adalah data siswa dan jawaban pilihan siswa.

Untuk menentukan tipe hasil belajar atau tingkat kemampuan berpikir yang akan dinilai, penyusun tes dapat berpedoman pada indikator pembelajaran atau tujuan evaluasi itu sendiri. Sehingga pemilihan alat evaluasi dan penyusunan instrumen tes akan tepat sesuai dengan tingkat kemampuan peserta didik.

Bloom membagi tingkat kemampuan atau tipe hasil belajar yang termasuk aspek kognitif menjadi enam yaitu pengetahuan hafalan, pemahaman atau komprehensif, penerapan atau aplikasi, analisis dan sintesis, evaluasi serta kreativitas. Berikut adalah penjelasannya yaitu:

1. Pengetahuan hafalan (C1) atau *knowledge* ialah tingkat kemampuan yang hanya meminta responden atau *testee* untuk mengenal atau mengetahui adanya konsep, fakta, atau istilah-istilah tanpa harus mengerti atau dapat menilai atau dapat menggunakannya. Dalam hal ini biasanya *testee* hanya dituntut untuk menyebutkan kembali (*recall*) atau menghafal saja.
2. Pemahaman atau komprehensif (C2) adalah tingkat kemampuan

yang mengharapkan *testee* mampu memahami arti atau konsep, situasi serta fakta yang diketahuinya. Dalam hal ini *testee* tidak hanya hafal secara verbal akan tetapi juga memahami konsep dari masalah atau fakta yang ditanyakan.

3. Aplikasi atau penerapan (C3) adalah penggunaan abstraksi pada situasi konkret atau situasi khusus. *Testee* dituntut kemampuannya untuk menerapkan atau menggunakan apa yang telah diketahuinya dalam situasi baru baginya (diabstrakkan). Abstraksi ini dapat berupa ide, teori, atau petunjuk praktis.
4. Kemampuan analisis dan sintesis (C4) adalah kemampuan yang mengukur *testee* untuk meng-analisis atau menguraikan suatu integritas atau situasi tertentu ke dalam komponen atau unsur pembentuknya. Diharapkan siswa dapat memahami dan sekaligus mampu memilah-milahnya menjadi bagian-bagian, termasuk juga menguraikan bagaimana proses terjadinya sesuatu, cara bekerjanya sesuatu, atau mungkin juga sistematikanya. Sedangkan kemampuan sintesis adalah pe-nyatuan unsur-unsur atau bagian-bagian ke dalam suatu bentuk yang menyeluruh. Diharapkan *testee* mampu menemukan hubungan kausal atau urutan tertentu atau menemukan abstraksinya yang berupa integritas.
5. Kemampuan evaluasi (C5) adalah kemampuan dari *testee* atau peserta tes untuk membuat suatu penilaian tentang suatu hipotesis atau sebagainya berdasarkan suatu kriteria tertentu. Kegiatan penilaian dapat dilihat dari segi

tujuannya, gagasannya, cara bekerjanya, cara me-mecahannya, metodenya, dan materinya lainnya.

6. Kemampuan berkreasi atau ber-kreativitas (C6) adalah tingkat kemampuan untuk merancang, membangun, merencanakan, mem-produksi, menemukan, mem-perbaharui, menyempurnakan, memperkuat dan mengubah sesuatu menjadi baru.

## HASIL PENELITIAN

Hasil analisis secara kuantitatif soal pilihan jamak sebanyak 40 butir pada ujian Biologi semester ganjil pada SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016 dengan program ITEMAN versi 3.50 dapat diketahui validitas soal, kesukaran soal, daya pembeda, efektifitas pengecoh dan reliabilitas soal.

Berdasarkan analisis dengan program SPSS versi 16.0 secara keseluruhan dapat diketahui bahwa 20% soal valid dan 80% soal tidak valid. Dengan kategori validitas soal sebanyak 0% soal memiliki validitas sangat tinggi, 2,5% soal memiliki validitas tinggi, 10% soal memiliki validitas sedang, 7,5% memiliki validitas rendah dan 0% soal memiliki validitas sangat rendah.

Tabel 4. Validitas Soal.

Kategori	Jumlah	Persen-tase	No Soal
Sangat Tinggi	0	0%	-
Tinggi	1	2,5%	16
Sedang	4	10%	7, 12, 14, 20
Rendah	3	7,5%	2, 18, 38
Sangat Rendah	0	0%	-

Berdasarkan analisis butir soal secara keseluruhan dengan menggunakan program ITEMAN 3.50 tingkat kesukaran soal pilihan jamak dapat diketahui 5% sangat sukar, 30% sukar, 45% sedang, 20% mudah dan 0% sangat mudah.

Tabel 5. Tingkat Kesukaran Soal.

Kategori	Jumlah	Persen-tase	No Soal
Sangat Sukar	0	5%	-
Sukar	8	30%	9, 13, 24, 25, 28, 29, 33, 35
Sedang	18	45%	1, 2, 3, 4, 6, 11, 15, 19, 21, 22, 23, 26, 27, 31, 34, 36, 37, 40
Mudah	12	20%	5, 7, 8, 10, 12, 14, 17, 18, 20, 30, 32, 39
Sangat Mudah	2	0%	16, 38

Berdasarkan analisis butir soal secara keseluruhan dengan program ITEMAN versi 3.50, daya beda soal pilihan jamak dapat diketahui sebanyak 20% rendah, 42,5% sedang dan 37,5% tinggi.

Tabel 6. Daya Beda Soal.

Kategori	Jumlah	Persentase	No Soal
Rendah	8	20%	1, 3, 9, 21, 24, 32, 34, 40
Sedang	17	42,5%	4, 5, 6, 10, 15, 19, 22, 23, 26, 27, 28, 29, 30, 33, 35, 36, 37
Tinggi	15	37,5%	2, 7, 8, 11, 12, 13, 14, 16, 17, 18, 20, 25, 31, 38, 39

Berdasarkan analisis secara keseluruhan pengecoh pada soal pilihan jamak dapat diketahui bahwa sebanyak 80% pengecoh berfungsi, sedangkan sebanyak 20% pengecoh tidak berfungsi.

Tabel 7. Efektifitas Pengecoh.

Kategori	Jumlah	Persentase	No Soal
Berfungsi	32	80%	1, 2, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 31, 33, 35, 36, 37, 38, 39
Tidak Berfungsi	8	20%	3, 4, 9, 21, 30, 32, 34, 40

Pada statistik tes dapat dilihat bahwa koefisien reliabilitas *alpha* sebesar 0,584. Hal ini berarti reliabilitas ujian Biologi semester

ganjil pada SMK Kesehatan Nurul Islam Husada cukup.

## PEMBAHASAN

Pada pengukuran validitas soal yang dilakukan pada soal ujian Biologi kelas XI semester ganjil SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016 diketahui hasil penelitian Berdasarkan analisis dengan program SPSS versi 16.0 (dilihat dari *r* hitung atau *Corrected Item-Total Correlation*) untuk kemudian dibandingkan dengan *r* tabel, secara keseluruhan dapat diketahui bahwa 20% soal valid dan 80% soal tidak valid. Dengan kategori validitas soal sebanyak 0% soal memiliki validitas sangat tinggi, 2,5% soal memiliki validitas tinggi, 10% soal memiliki validitas sedang, 7,5% memiliki validitas rendah dan 0% soal memiliki validitas sangat rendah.

Pada pengukuran tingkat kesukaran soal yang dilakukan pada soal ujian Biologi semester ganjil SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016 (pada program ITEMAN ditunjukkan dengan proporsi siswa menjawab benar/ *proportion correct*), hasilnya diketahui bahwa sebanyak 0% soal memiliki tingkat kesukaran sangat mudah, 20% soal memiliki tingkat kesukaran mudah, 45% soal memiliki tingkat kesukaran sedang, 30% soal memiliki tingkat kesukaran sukar dan 5% soal memiliki tingkat kesukaran sangat sukar.

Pada pengukuran daya pembeda yang dilakukan pada soal ujian Biologi semester ganjil SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016 (pada program ITEMAN ditunjukkan dengan *biser*) didapatkan hasil sebanyak 20% daya

beda soal termasuk kedalam kategori rendah. 42,5% daya beda soal termasuk kedalam kategori sedang. Dan 37,5% daya beda soal termasuk kedalam kategori tinggi.

Pada pengukuran keefektifan pengecoh yang dilakukan pada soal ujian Biologi semester ganjil SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016, didapatkan hasil sebanyak 80% soal memiliki pengecoh yang berfungsi. Dan sebanyak 20% soal memiliki pengecoh yang tidak berfungsi.

Pada soal pilihan jamak, terdapat alternatif jawaban/ *option* yang merupakan pengecoh/ distraktor. Sehingga dapat diketahui berfungsi atau tidaknya pengecoh pada soal ujian Biologi semester ganjil SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/2016 dengan melihat seberapa banyak dan seberapa merata peserta didik memilih pengecoh.

Butir soal dengan pengecoh baik, pengecohnya akan dipilih secara merata oleh peserta didik. Sedangkan butir soal dengan pengecoh kurang baik pengecohnya akan dipilih secara tidak merata (Surapranata, 2005: 24).

Pada pengukuran reliabilitas yang dilakukan pada soal ujian Biologi semester ganjil SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016 (pada program ITEMAN dapat dilihat pada skala statistik dan ditunjukkan dengan *alpha*) dapat dilihat bahwa koefisien reliabilitas *alpha* sebesar 0,584. Sehingga dapat diketahui bahwa reliabilitas pada soal ujian Biologi semester ganjil SMK Kesehatan Nurul Islam tahun pelajaran 2015/2016 masuk dalam kategori cukup.

Setelah dilakukan penelaahan soal menggunakan format penelaahan yang mencakup materi, konstruksi

dan bahasa terdapat beberapa soal yang perlu diperbaiki. Dari segi materi 100% soal telah sesuai dengan aspek penelaahan yang meliputi kesesuaian dengan kom-petensi dasar, kesesuaian dengan kompetensi (urgensi, kontinuitas, keterpakaian sehari-hari), mengacu pada ranah kognitif dan hanya ada satu kunci jawaban. Semua soal telah sesuai dengan kompetensi yang diujikan. Proporsi pembagian soal pada masing-masing kompetensi sudah seimbang sehingga soal dapat dikatakan sesuai dengan kompetensi yang diujikan.

Soal yang mengacu pada ranah kognitif C1 sebesar 47,5% (sebanyak 19 soal), C2 sebesar 42,5% (sebanyak 17 soal), dan soal C3 sebesar 10% (sebanyak 4 soal), sedangkan untuk soal C4, C5 dan C6 tidak ada sama sekali. Dari persentase tersebut maka dapat diperoleh perbandingan C1 : C2 : C3 : C4 : C5 : C6 = 4 : 4 : 1 : 0 : 0 : 0. Dalam soal pilihan ganda ranah kognitif C5 dan C6 tidak ada, hal ini karena alat evaluasi berupa tes pilihan ganda kurang sesuai.

Kemampuan evaluasi dan kreativitas lebih sesuai jika diterapkan pada soal uraian, sehingga akan dengan mudah mengetahui daya berpikir siswa yang sebenarnya.

Berdasarkan hasil analisis kuantitatif diketahui bahwa tingkat kesukaran soal yang sedang. Ranah C1 (hafalan), C2 (pemahaman) dan C3 (penerapan) termasuk pada tingkat kemampuan berpikir menengah. Untuk ranah kognitif C4 (analisis), C5 (evaluasi) dan C6 (berkreativitas) tidak ada sama sekali karena memiliki tingkat kemampuan berpikir tinggi. Hal ini mengakibatkan tingkat kesukaran sangat sukar hanya 0%. Dari perbandingan diatas dapat di

simpulkan bahwa perlu dilakukan peninjauan ulang pada soal mengenai daya berpikir siswa.

Dari segi konstruksi soal nomor 26 dan 27 sebaiknya direvisi karena salah dalam penulisan pilihan pertanyaan dan jawaban sehingga membuat pilihan jawaban menjadi tidak homogen dan tidak logis (dalam hal ini pengecoh), sedangkan soal nomor 36 memiliki gambar yang kurang jelas karena buram dan beberapa bagian yang ditunjuk tidak jelas.

Ditinjau dari segi bahasa soal pilihan jamak ujian Biologi kelas XI semester ganjil pada SMK Kesehatan Nurul Islam Husada tahun pelajaran 2015/ 2016 secara keseluruhan sudah sesuai dengan aspek penelaahan yang meliputi kesesuaian bahasa dengan kaidah bahasa Indonesia, penggunaan bahasa yang komunikatif, tidak menggunakan bahasa yang berlaku setempat/tabu dan pilihan jawaban tidak mengulang kelompok kata yang sama.

Analisis soal secara kuantitatif menunjukkan bahwa soal memiliki tingkat kesukaran sedang, daya pembeda baik, efektifitas pengecoh berfungsi sebanyak 80% dan reliabilitas cukup. Pada analisis soal secara kualitatif diketahui bahwa soal telah memenuhi aspek materi, konstruksi dan bahasa. Hanya terdapat beberapa soal yang harus diperbaiki yaitu pada aspek konstruksi.

Syarat kemampuan khusus yang dimiliki oleh seorang guru dalam menyusun soal sesuai standar adalah menguasai materi yang akan di ujikan, mampu membahasakan gagasan, memahami karakteristik peserta tes dan menguasai teknik penulisan soal. Dalam mengembangkan instrumen penilaian, umumnya guru menyusun kisi - kisi secara mandiri

tanpa ditelaah, direvisi dan divalidasi. Oleh karena itu kemampuan guru dalam me-nyusun kisi-kisi perlu secara terus menerus dievaluasi dan ditingkatkan.

Dalam menentukan tingkat kesukaran sedang pada soal dilakukan dengan menyusun kisi-kisi soal yang memperhatikan proporsi ranah kognitif. Ranah C1 dan C2 lebih banyak dari yang lain sehingga soal masuk dalam kategori mudah. Sehingga perlu diadakannya peningkatan kemampuan guru dalam penguasaan teknik penulisan soal, kisi-kisi yang baik serta mengembangkan instrumen pe-nilaian. Selain faktor diatas, peserta tes juga mempengaruhi hasil dari analisis soal. Jumlah peserta tes atau *testee* sangat berpengaruh terhadap hasil analisis, begitu juga perbedaan tingkat kemampuan berpikir peserta tes. Semakin banyak dan bervariasi kemampuan peserta tes maka hasilnya akan lebih akurat dan representatif. Oleh karena itu dalam pengambilan sampel perlu dipilih secara berstrata sehingga diperoleh data yang akurat.

## **SIMPULAN**

Butir soal soal ujian semester ganjil Biologi kelas XI pada SMK Kesehatan Nurul Islam Husada Lampung Selatan tahun pelajaran 2015/ 2016 sebanyak 0% soal memiliki validitas sangat tinggi, 2,5% soal memiliki validitas tinggi, 10% soal memiliki validitas sedang, 7,5% memiliki validitas rendah dan 0% soal memiliki validitas sangat rendah.

Tingkat kesukaran dari 40 butir soal soal ujian semester ganjil Biologi kelas XI pada SMK Kesehatan Nurul Islam Husada

Lampung Selatan tahun pelajaran 2015/ 2016 diketahui sebanyak 5% sangat sukar, 30% sukar, 45% sedang, 20% mudah dan 0% sangat mudah

Daya beda dari 40 butir soal ujian semester ganjil Biologi kelas XI pada SMK Kesehatan Nurul Islam Husada Lampung Selatan tahun pelajaran 2015/ 2016 diketahui sebanyak 20% rendah, 42,5% sedang dan 37,5% tinggi

Pengecoh dari 40 butir soal ujian semester ganjil Biologi kelas XI pada SMK Kesehatan Nurul Islam Husada Lampung Selatan tahun pelajaran 2015/ 2016 diketahui sebanyak 80% pengecoh berfungsi, sedangkan sebanyak 20% pengecoh tidak berfungsi.

Reliabilitas dari 40 butir soal soal ujian semester ganjil Biologi kelas XI pada SMK Kesehatan Nurul Islam Husada Lampung Selatan tahun pelajaran 2015/ 2016 cukup.

#### **DAFTAR RUJUKAN**

Afiyana, I. 2010. *Analisis kualitas soal latihan UN mata pelajaran IPA menggunakan program ITEMAN di Kabupaten Batang*. Skripsi. Semarang: Universitas Negeri Semarang.

Arikunto. 2007. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Arrizqi, SN. 2010. *Analisis soal UAS Biologi Kelas XI di Kabupaten Tegal Tahun 2008/2009*. Skripsi. Semarang: Universitas Negeri Semarang.

Nazir, M. 2005. *Metode Penelitian*. Jakarta: Ghalia Indonesia.

Rudyatmi E dan Anni R. 2010. *Evaluasi Pembelajaran (Bahan Ajar)*. Semarang: FMIPA UNNES.

Sudjana. 2001. *Dasar-Dasar Proses Belajar Mengajar*. Bandung: Sinar Baru Algesindo.

Surapranata, S. 2005. *Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2004*. Bandung: Remaja Rosdakarya.

Trianto. 2010. *Model Pembelajaran Terpadu*. Jakarta: Bumi Aksara.