



PENGEMBANGAN BUTIR SOAL HIGHER ORDER THINKING SKILL (HOTS) MATERI SISTEM KOORDINASI PADA SISWA SMA KELAS XI

Ali Mustofa¹, Murni Saptasari²

S2 Pendidikan Biologi/ Jurusan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri
Malang, Jalan Semarang No 05 Malang 1,2

Email Penulis Korespondensi: alimustofa0507@gmail.com

Info Artikel

Sejarah Artikel:

Diterima 18 Des 2020

Direvisi 13 Feb 2021

Disetujui 13 Feb 2021

Keywords: (max 5 words)

Hots,

Question Item Analysis,

Question development.

Abstract

Conducting an assessment is one of the teacher's duties apart from compiling a learning program and implementing it in the classroom. The teacher must also be able to determine what can be obtained or achieved from the learning process that has been organized. The purpose of this study was to develop HOTS questions on the coordination system material and to test the quality of HOTS questions in the coordination system material for high school students in grade X. The subjects of this study consisted of 18 students who were selected purposively. Analysis of the items consisted of validity, reliability, difficulty level, differentiation, and trick effectiveness. The results of the analysis of the validity of the questions with a significant level of 5% for 18 students obtained 50 questions (92%) valid. The results of the reliability test of the items obtained the reliability index in the "high" criteria. So the items have high consistency in measuring the ability of students regarding the material of the coordination system to distinguish high-skilled students from low-ability students. The results of the analysis of the difficulty level of the questions were: 26 questions (5%) had an easy difficulty level, 9 questions (75%) had a moderate difficulty level and 15 questions (20%) had a difficult difficulty level. Overall, there were 4 distractor test results that did not function. A bully is said not to function if less than 5% of all test takers are selected. The crooks that didn't work need to be revised again. The results of the 84% difference power test were rejected so it needed to be revised

Abstrak

Penilaian adalah tugas guru dan merancang pembelajaran dan menerapkannya di kelas., menerapkan apa yang dapat dicapai selama proses pembelajaran yang telah dilakukan. Tujuan dari penelitian ini adalah mengembangkan soal HOTS pada materi sistem koordinasi dan menguji kualitas soal HOTS materi sistem koordinasi pada siswa SMA kelas X. Subyek penelitian ini terdiri atas 18 siswa yang dipilih secara purposive. Menganalisis soal terdiri atas uji validitas, reabilitas, tingkat kesukaran, daya beda, efektivitas pengecoh. Hasil analisis validitas soal dengan taraf signifikan 5% untuk siswa berjumlah 18 diperoleh 50 soal (92%) valid. Hasil uji reliabilitas soal diperoleh indeks reliabilitas dalam kriteria "tinggi", sehingga butir soal memiliki konsistensi yang tinggi dalam mengukur kemampuan peserta didik mengenai materi sistem koordinasi dapat membedakan peserta didik berkemampuan tinggi dengan peserta didik berkemampuan rendah. Hasil uji analisis tingkat kesukaran soal yaitu: 26 soal (5%) memiliki tingkat kesukaran kategori mudah, 9 soal (75%) memiliki tingkat kesukaran kategori sedang dan 15 soal (20%) yang memiliki tingkat kesukaran kategori sukar. Hasil uji pengecoh pada soal secara keseluruhan ada 4 pengecoh tidak berfungsi. Hasil uji daya beda 84% soal ditolak sehingga perlu direvisi.

PENDAHULUAN

Pendidikan merupakan ujung tombak kemajuan suatu bangsa, jika pendidikan baik maka kemajuan bangsa tersebut akan baik pula. Kemajuan suatu bangsa dipengaruhi oleh faktor kualitas generasi yang tangguh, kreatif, inovatif dan berakhlak mulia. Ketercapaian tujuan pendidikan dipengaruhi pada proses kegiatan belajar mengajar, sarana prasarana dan guru profesional. Untuk mengetahui ketercapaian kemampuan hasil belajar perlu dilakukan evaluasi dari hasil belajar peserta didik. Penilaian digunakan sebagai indikator keberhasilan suatu proses kegiatan belajar mengajar baik ranah afektif, kognitif maupun psikomotor (Rusdiana, Sumardi and Arifiyanto, 2014).

Penilaian pembelajaran dikatakan berhasil jika diukur dengan alat ukur yang sesuai dengan tujuan pembelajarannya dan kompetensi yang dicapainya. Informasi yang diperoleh dari penilaian harus komprehensif dan telah dilakukan pada saat-saat yang tepat selama dan setelah peserta didik belajar. Pengukuran harus dilakukan di sepanjang proses belajar yang dijalani peserta didik (Jurs and Wiersma, 1990). Pengukuran adalah proses untuk memperoleh deskripsi angka tentang derajat karakteristik tertentu yang dimiliki oleh individu (Gronlund, 1998). Pengumpulan informasi ini dilakukan selain menggunakan tes, juga dilakukan dengan mengobservasi peserta didik ketika sedang belajar, mewawancarai, atau memeriksa produk siswa. Hasil penilaian sangat diperlukan dalam melakukan evaluasi, hal ini terkait dengan kebutuhan untuk membuat keputusan, seperti dinyatakan oleh (Brown and Shavelson, 1996) bahwa evaluasi adalah kegiatan melakukan keputusan berdasarkan informasi yang telah diperoleh dalam penilaian tersebut. Gronlund (1998) menyatakan evaluasi menentukan tingkat ketercapaian tujuan melalui proses yang sistematis mulai dari mengumpulkan data, menganalisis dan memberikan penilaian. Evaluasi juga memberikan keputusan melalui proses sistematis dari mengumpulkan data, menganalisis, dan mengambil keputusan (Gronlund, 1998).

Ketercapaian pembelajaran akan mendorong terlaksananya *life long learning*. Namun, *life long learning* tidak didukung dengan pengembangan soal-soal HOTS yang dapat meningkatkan kemampuan siswa berfikir analitis. HOTS merupakan kemampuan berpikir yang terdiri atas berpikir kritis, berpikir kreatif, dan pemecahan masalah (Brookhart, 2010). Tujuan pembelajaran yang mengembangkan HOTS adalah untuk membekali siswa terampil memberi alasan dan membuat keputusan (Brookhart, 2010). Pentingnya HOTS dalam pembelajaran juga ditunjukkan oleh hasil penelitian (Murray, 2011) yang menyebutkan bahwa ketika siswa menggunakan HOTS maka siswa memutuskan apa yang harus dipercayai dan apa yang harus dilakukan, menciptakan ide-ide baru, membuat prediksi dan memecahkan masalah nonrutin. Tujuan dari penelitian ini adalah mengembangkan soal HOTS pada materi sistem koordinasi dan mengetahui kualitas soal HOTS materi sistem koordinasi pada siswa SMA kelas X melalui uji validitas, reabilitas, uji kesukaran soal, efektivitas pengecoh, dan uji daya beda soal.

METODE PENELITIAN

1. Jenis Penelitian

Penelitian ini termasuk dalam penelitian deskriptif kuantitatif, yaitu penelitian yang deskripsi atau gambarnya menggunakan ukuran, jumlah atau frekuensi. Penelitian yang dilakukan tidak memberikan perlakuan, perubahan pada variabel bebas, tetapi menggambarkan suatu kondisi apa adanya.

2. *Populasi, Sampling Dan Teknik Sampling*

Populasi penelitian ini adalah siswa kelas XI MAM 1 Malang pada tahun ajaran 2018/2019. Penentuan sampel dilakukan dengan menggunakan *teknik purposive sampling* dengan jumlah responden adalah 18 siswa.

3. *Waktu dan tempat penelitian*

Penelitian dilakukan pada bulan April 2019 di MAM 1 Malang

4. *Instrumen penelitian*

a. *Tes Soal Hots Materi Sistem Koordinasi*

Tes Soal Hots Materi Sistem Koordinasi terdiri terdiri atas 50 soal pilihan ganda dengan kriteria Hots yang disesuaikan dengan Kompetensi Dasar, Indikator dan Tujuan pembelajaran yang akan dicapai yang dibagikan kepada setiap responden. Materi sistem koordinasi pada tes hasil belajar terdiri dari beberapa indikator yaitu: 3.10.1 Mendeskripsikan struktur jaringan dan menjelaskan fungsi penyusun organ pada sistem koordinasi. 3.10.2 Mengaitkan hubungan antara struktur jaringan penyusun organ dan proses koordinasi pada manusia. 3.10.3 Menganalisis peran saraf dalam mekanisme koordinasi dan regulasi pada manusia. 3.10.4 Menganalisis peran hormon dalam mekanisme koordinasi dan regulasi pada manusia. 3.10.4 Menganalisis gangguan fungsi yang terjadi pada sistem koordinasi manusia.

b. Uji Validitas Soal

Uji validitas merupakan cara memperoleh data untuk tingkat kelayakan *butir soal* sebelum diujicobakan. Lembar uji validitas digunakan untuk merevisi *butir soal* hingga layak untuk diujicobakan. Lembar uji validitas soal dilengkapi dengan kriteria-kriteria untuk menentukan diterima dan tidaknya soal.

5. Pengumpulan Data

Pengukuran hasil tes *hots* soal materi sistem koordinasi terdiri atas 50 butir soal dengan kriteria C4, C3 dan C2 setiap item bernilai 1 dan salah 0. Pertanyaan pada tes disesuaikan dengan tujuan ketercapaian pembelajaran.

6. Teknik Analisis data

a. Analisis Tingkat Kesukaran

Bilangan yang menunjukkan sukar dan mudahnya soal disebut indeks kesukaran (*difficulty index*). Besarnya indeks kesukaran antara 0,0 sampai dengan indeks 1,0. Soal dengan indeks kesukaran 0,0 meunjukkan bahwa soal terlalu sukar, sedangkan soal dengan indeks 1,0 menunjukkan soal terlalu mudah (Arikunto, 2012).

Tingkat kesukaran dapat dicari dengan rumus (Sunarmi, 2016):

$$TK = \frac{WB + WA}{nB + nA} \times 100\%$$

Keterangan:

WB = jumlah peserta tes kelompok bawah yang menjawab benar pada butir soal

WA = jumlah peserta tes kelompok atas yang menjawab benar pada butir soal

TK = tingkat kesukaran

nB = jumlah peserta tes kelompok bawah

nA = jumlah peserta tes kelompok atas

Kategori tingkat kesukaran dapat dilihat pada Tabel 1.

Tabel 1 Klasifikasi Kategori Soal Berdasarkan Tingkat Kesukaran

| <u>Kategori soal</u> | <u>Indeks kesukaran</u> |
|----------------------|-------------------------|
| <u>Sukar</u> | 0,00-0,25 |
| <u>Sedang</u> | 0,26-0,75 |
| <u>Mudah</u> | 0,76-1,00 |

Sumber: Yusuf (2015)

b. Analisis Daya Beda Soal

Analisis daya beda mengkaji butir soal dengan tujuan untuk mengetahui kesanggupan soal dalam membedakan siswa yang tergolong mampu dengan siswa yang tergolong kurang atau lemah prestasinya. Daya beda soal adalah kemampuan sesuatu soal untuk membedakan antara siswa yang pandai dan siswa yang berkemampuan rendah (Arikunto, 2012).

Rumus pengukuran daya beda yaitu :

$$DB = \frac{WA - WB}{n}$$

Keterangan:

WB = jumlah peserta tes kelompok bawah yang menjawab benar pada butir soal

WA = jumlah peserta tes kelompok atas yang menjawab benar pada butir soal

DB = daya beda

N = jumlah peserta kelompok atas atau bawah

Kategori daya beda dapat dilihat pada Tabel 2.

Tabel 2 Klasifikasi Kategori Soal Berdasarkan Daya Beda

| <u>Kategori soal</u> | <u>Indeks kesukaran</u> |
|----------------------|-------------------------|
| <u>Baik sekali</u> | 0,40 – 1,00 |
| <u>Baik</u> | 0,30 – 0,39 |
| <u>Sedang</u> | 0,20 – 0,29 |
| <u>Kurang</u> | 0,01 – 0,19 |

Sumber: Yusuf (2015)

Penentuan kriteria soal dapat diterima atau tidak dapat mengacu pada hasil pengukuran daya beda soal. Kriteria soal dapat dilihat pada Tabel 3.

Tabel 3 Kriteria untuk Menentukan Soal Diterima atau Tidak

| Kategori soal | Indeks kesukaran |
|-------------------------|------------------|
| Sangat baik | 0,40 – 1,00 |
| Diterima dan diperbaiki | 0,30 – 0,39 |
| Diperbaiki | 0,20 – 0,29 |
| Ditolak | 0,01 – 0,19 |

Sumber: Sunarmi, dkk (2013)

c. Analisis Efektivitas Pengecoh Soal

Penyebaran pilihan jawaban dijadikan dasar dalam penelaah soal. Hal ini dimaksudkan untuk mengetahui berfungsi tidaknya jawaban yang tersedia. Suatu pilihan jawaban pengecoh berfungsi apabila pengecoh paling tidak dipilih 5% peserta tes atau siswa dan lebih banyak dipilih oleh kelompok siswa yang belum paham materi. Patokan yang digunakan untuk interpretasi keefektifan penggunaan distraktor diadaptasi dari *Skala Likert* yang dapat dilihat pada Tabel 2.4.

Tabel 2.4 Kriteria Interpretasi Keefektifan Penggunaan Distraktor

| Pengecoh yang berfungsi | Kriteria |
|-------------------------|--------------------------------|
| 0 | Semua pengecoh tidak berfungsi |
| 1 | Kurang baik |
| 2 | Cukup |
| 3 | Baik |
| 4 | Sangat baik |

Sumber: Wahyuni (2017)

d. Analisis Reliabilitas

Setelah dilakukan pembelahan ganjil-genap maka mencari realibilitas tes dengan jalan mengorelasikan kedua skor menggunakan rumus *Product Moment Person* yaitu:

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\{(N\sum X^2 - (\sum X)^2)\{(N\sum Y^2 - (\sum Y)^2)\}}}$$

Keterangan:

r_{xy} = korelasi *product moment*

N = jumlah responden

$\sum X$ = jumlah skor variabel belahan pertama (butir soal ganjil)

$\sum Y$ = jumlah skor variabel belahan kedua (butir soal genap)

Koefisien korelasi dari perhitungan dengan rumus *Product Moment Person* menunjukkan realibilitas separuh soal, sehingga untuk mencari reliabilitas seluruh tes menggunakan rumus *Spearman Brown* sebagai berikut.

$$r_{11} = \frac{2r^{1/2}}{1 + 1/2}$$

Keterangan:

r_{11} = reliabilitas seluruh tes

$r^{1/2}$ = reliabilitas separuh tes

Berikut ini merupakan tabel klasifikasi reliabilitas soal. Nilai reliabilitas yang didapatkan dari perhitungan dengan rumus *Spearman Brown* kemudian diklasifikasikan menggunakan klasifikasi reliabilitas soal yang dapat dilihat pada Tabel 5.

Tabel 5 Klasifikasi Reliabilitas Soal

| No. | Koefisien realibilitas | Tingkat reliabilitas |
|-----|------------------------|----------------------|
| 1 | 0,800-1,000 | Sangat tinggi |
| 2 | 0,600-0,799 | Tinggi |
| 3 | 0,400-0,599 | Cukup |
| 4 | 0,200-0,399 | Rendah |
| 5 | 0,00-0,199 | Sangat rendah |

Sumber: Sunarti dan Rahmawati (2014)

HASIL DAN PEMBAHASAN

Berdasarkan hasil implementasi yang dikembangkan dapat di hasilkan hasil analisis berikut ini :

1) Uji Validitas Soal

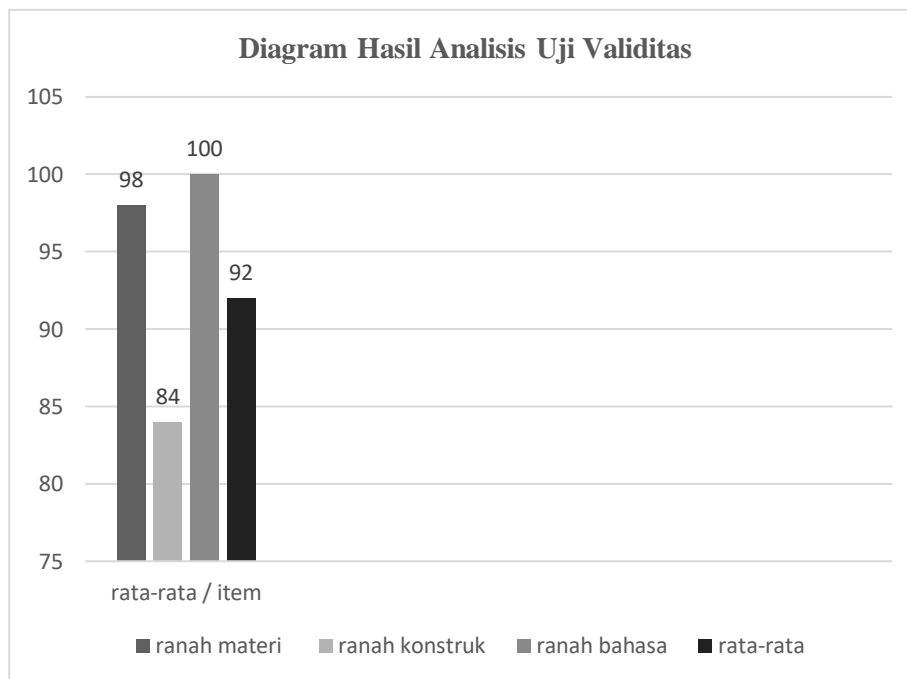
Berdasarkan hasil analisis data, uji validitas soal mencakup 3 aspek pertanyaan utama yaitu ranah materi, ranah konstruksi dan ranah bahasa. Hasil uji validitas disajikan pada tabel 3.1.

Tabel 3.1Aspek penilaian uji validitas

| No. | Jenis Pertanyaan | Prosentase |
|----------------------------|--|------------|
| A. RANAH MATERI | | |
| 1. | Butir soal sesuai dengan indikator | 98% |
| 2. | Batasan pertanyaan dan jawaban yang diharapkan jelas | 100% |
| 3. | Isi materi sesuai dengan tujuan pengukuran | 98% |
| 4. | Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah, dan tingkat kelas | 100% |
| 5. | Pilihan/jawaban benar-benar berfungsi, pengecoh bukanlah hal yang terlibat benar-benar salah | 96% |
| Rata-rata | | 98% |
| B. RANAH KONSTRUKSI | | |
| 6. | Pokok soal (stem) dirumuskan dengan jelas | 98% |
| 7. | Rumusan soal dan pilihan dirumuskan dengan tegas | 100% |
| 8. | Pokok soal tidak memberi petunjuk/mengarah kepada pilihan jawaban yang benar | 100% |
| 9. | Pokok soal tidak mengandung pernyataan negatif ganda | 96% |
| 10. | Bila terpaksa menggunakan kata negatif, maka harus digarisbawahi atau dicetak miring | 98 |
| 11. | Pilihan jawaban homogen | 100% |
| 12. | Hindari adanya alternatif jawaban "seluruh jawaban di atas benar" atau "tak satu jawaban di atas yang benar", dan sejenisnya | 100% |
| 13. | Panjang alternatif/pilihan jawaban relatif sama, tidak ada yang sangat panjang dan tidak ada yang sangat pendek | 96% |
| 14. | Pilihan jawaban dalam bentuk angka/waktu yang diurutkan | 96% |
| 15. | Wacana, gambar, atau grafik benar-benar berfungsi | 8% |
| 16. | Antarbutir soal tidak bergantung satu sama lain | 36% |
| Rata-rata | | 84,3% |

| C. RANAH BAHASA | | |
|----------------------|---|------|
| 17. | Rumusan kalimat komunikatif | 100% |
| 18. | Kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya | 100 |
| 19. | Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian | 100% |
| 20. | Menggunakan bahasa/kata yang umum (bukan bahasa lokal) | 100% |
| 21. | Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan peserta didik | 100% |
| Rata-rata | | 100% |
| Rata-rata semua item | | 92% |

Berdasarkan hasil analisis uji validitas Tabel 1, menunjukkan bahwa soal Tes Semester Genap **Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang** memiliki kriteria yang sangat baik. Hal ini dapat dilihat dari persentase hasil uji validitas dengan angka 92%. Ringkasan hasil analisis daya beda soal pada masing-masing sekolah dapat dilihat pada Gambar 3.1



Gambar 3.1 Hasil Analisis Uji Validitas

2) Analisis Tingkat Kesukaran

Berdasarkan hasil analisis data, tingkat kesukaran soal dikategorikan menjadi tiga kategori yaitu soal kategori mudah, sedang, dan sukar. Hasil analisis data untuk tingkat kesukaran soal dapat dilihat pada Tabel 3.2.

Tabel 3.2 Tingkat Kesukaran Butir Soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang

| Kategori | Jumlah | Persentase | No. Soal |
|----------|--------|------------|---|
| Mudah | 26 | 52% | 2,4,10,11,18,19,20, 23,24,25,26,28,30,31,32,35,36,37,38,42,43,44,46,47,48,50 |
| Sedang | 9 | 18% | 3,8,9,16,17,27,29,45,49 |
| Sukar | 15 | 30% | 1,5,6,7,12,13,14,15,21,22,33,34,39,40,41, |

Berdasarkan Tabel 3.2 tingkat kesukaran soal pilihan ganda jawaban tunggal materi system koordinasi KD 3.10 Semester Genap Biologi Kelas XI MAM 1 Malang dapat dikategorikan memiliki tingkat kesukaran mudah. Hal ini dapat dilihat dari hasil persentase soal kategori mudah lebih tinggi daripada soal kategori sedang dan sukar.

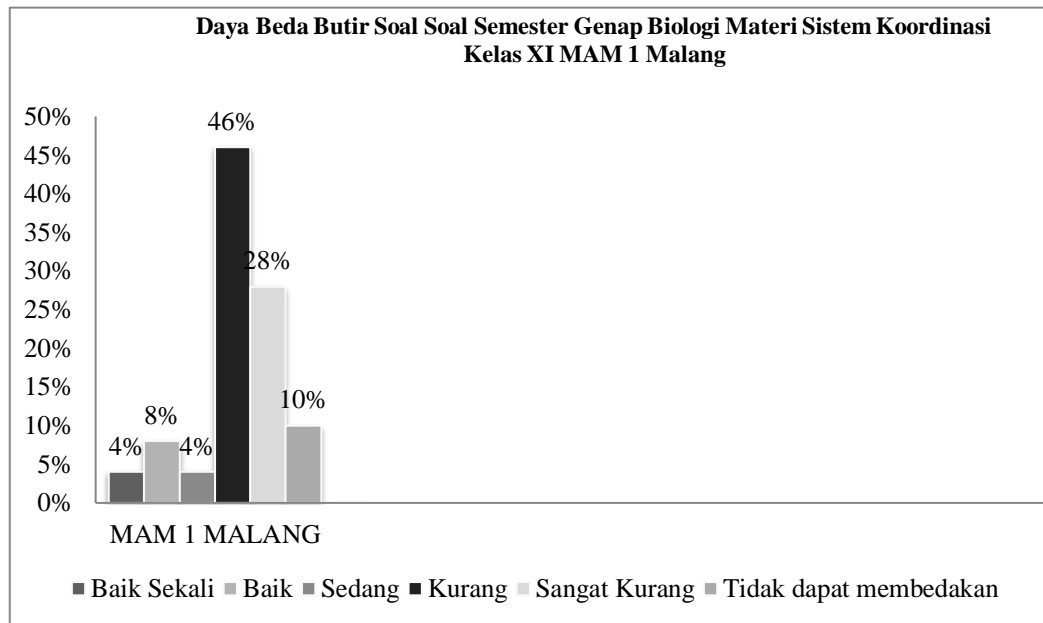
3) Analisis Daya Beda Soal

Berdasarkan hasil analisis data, daya beda soal dikategorikan menjadi lima kategori yaitu soal dengan daya beda baik sekali, soal dengan daya beda baik, soal dengan daya beda sedang, soal dengan daya beda kurang dan soal dengan daya beda sangat kurang. Hasil analisis data untuk tingkat kesukaran soal dapat dilihat pada Tabel 3.3.

Tabel 3.3 Daya Beda Butir Soal Soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang

| Kategori | Jumlah | Persentase | No. Soal |
|-----------------------|--------|------------|--|
| Baik Sekali | 2 | 4% | 16, dan 29. |
| Baik | 4 | 8% | 3,9,17, dan 27. |
| Sedang | 2 | 4% | 10 dan 45 |
| Kurang | 23 | 46% | 2,4,11,14,18,19,20,22,24,25,28,30,31,32,36,37,39,43,44,46,47,49, dan 50. |
| Sangat Kurang | 14 | 28% | 1,5,6,7,15,21,23,26,33,34,35,41,42,dan 48. |
| Tidak bisa membedakan | 5 | 10% | 8,12,13,38, dan 40. |

Berdasarkan hasil analisis daya beda pada Tabel 3.3. menunjukkan bahwa soal Tes Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang memiliki daya beda kurang. Hal ini dapat dilihat dari persentase soal dengan daya beda kurang paling tinggi daripada soal dengan kategori daya beda baik sekali, baik, sedang, sangat kurang dan tidak bisa membedakan. Ringkasan hasil analisis daya beda soal pada masing-masing sekolah dapat dilihat pada Gambar 3.2.



Gambar 3.2 Daya Beda Butir Soal Soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang

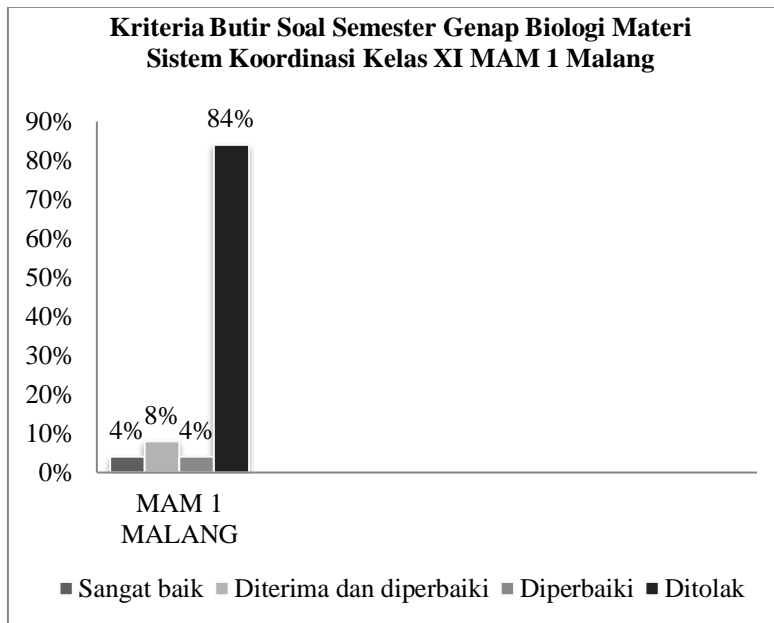
Berdasarkan hasil analisis data, kriteria soal dikategorikan menjadi empat kategori yaitu soal dengan kriteria soal sangat baik, soal diterima dan diperbaiki, soal diperbaiki, dan soal ditolak. Hasil analisis data untuk kriteria soal dapat dilihat pada Tabel 3.4.

Tabel 3.4 Kriteria Butir Soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang

| Kategori | Jumlah | Persentase | No. Soal |
|-------------------------|--------|------------|--|
| Sangat Baik | 2 | 4% | 16, dan 29. |
| Diterima dan Diperbaiki | 4 | 8% | 3,9,17, dan 27. |
| Diperbaiki | 2 | 4% | 10 dan 45 |
| Ditolak | 42 | 84% | 1,2,4,5,7,8,11,12,13,14,15,18,19,20,21,22,23,24,25,26,28,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,46,47,48,49, dan 50. |

**jumlah ditolak ditambahkan dengan jumlah soal sangat kurang dan yang tidak dapat membedakan*

Berdasarkan hasil analisis daya beda pada Tabel 3.4 menunjukkan bahwa soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang memiliki kriteria butir soal ditolak. Hal ini dapat dilihat dari persentase soal dengan kriteria butir soal ditolak paling tinggi daripada soal dengan kriteria soal sangat baik, kriteria soal diterima dan diperbaiki, serta kriteria soal diperbaiki. Ringkasan hasil analisis kriteria soal pada masing-masing sekolah dapat dilihat pada Gambar 3.3.



Gambar 3.3 Kriteria Butir Soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang

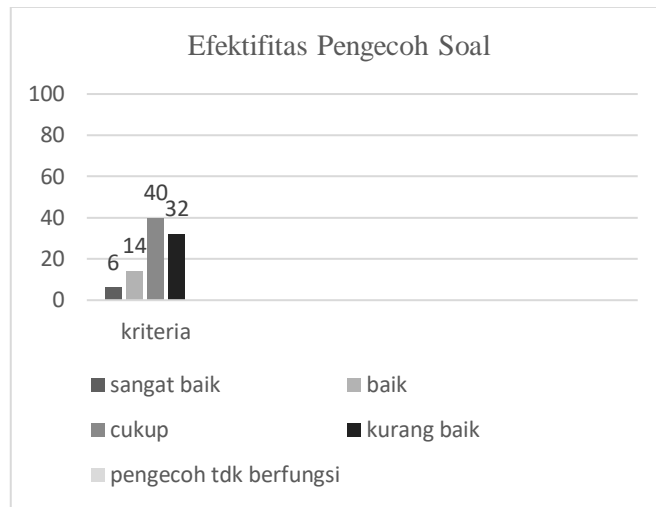
4) Analisis Efektivitas Pengecoh Soal

Berdasarkan hasil analisis data, daya beda soal dikategorikan menjadi lima kategori yaitu Kriteria sangat baik semua pengecoh berfungsi, baik, cukup, kurang baik, semua pengecoh tidak berfungsi. Hasil analisis data ditunjukkan pada tabel 3.5.

Tabel 3.5 Analisis efektivitas pengecoh soal

| Kategori | Jumlah | Persentase | No. Soal |
|--|--------|------------|---|
| Sangat baik, semua pengecoh berfungsi | 3 | 6% | 6, 18, 39 |
| Baik | 7 | 14% | 1, 8, 10, 16, 33, 34, 49 |
| Cukup | 20 | 40% | 3, 5, 7, 9, 11, 12,13, 14, 15, 17, 19, 21, 22, 23, 27, 28, 29, 31, 41, 48 |
| Kurang baik | 16 | 32% | 2, 4, 20, 24, 24, 30, 32, 36, 37, 38, 42, 43, 44, 45, 47, 50 |
| Semua pengecoh tidak berfungsi | 4 | 8% | 26, 35, 40, 46 |

Berdasarkan hasil analisis efektivitas pengecoh pada Tabel 2, menunjukkan bahwa soal **Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang** memiliki kriteria efektifitas pengecoh soal kategori cukup. Hal ini dapat dilihat dari persentase efektivitas pengecoh dengan kriteria cukup paling tinggi daripada Kriteria sangat baik semua pengecoh berfungsi, baik, kurang baik, dan semua pengecoh tidak berfungsi. Ringkasan hasil analisis kriteria soal pada masing-masing sekolah dapat dilihat pada Gambar 3.4.



Gambar 3.4. Diagram Efektifitas Pengecoh Soal

5) Reliabilitas

Berdasarkan hasil analisis data, Reliabilitas didapatkan pada skor 0,74 yang mengindikasikan bahwa Reliabilitasnya tinggi, Hasil analisis data untuk tingkat kesukaran soal dapat dilihat pada Tabel 3.6.

Tabel 3.6 Analisis reliabilitas soal

Rata2= 27,83
 Simpang Baku= 7,01
 KorelasiXY= 0,58
 Reliabilitas Tes= 0,74
 Nama berkas: D:\PASCAS T\DATA ANALISIS (1).ANA

| No.Urut | No. Subyek | Skor Ganjil | Skor Genap | Skor Total |
|---------|------------|-------------|------------|------------|
| 1 | 1 | 15 | 19 | 34 |
| 2 | 2 | 8 | 15 | 23 |
| 3 | 3 | 16 | 19 | 35 |
| 4 | 4 | 14 | 17 | 31 |
| 5 | 5 | 11 | 18 | 29 |
| 6 | 6 | 4 | 7 | 11 |
| 7 | 7 | 5 | 6 | 11 |
| 8 | 8 | 10 | 18 | 28 |
| 9 | 9 | 10 | 18 | 28 |
| 10 | 10 | 13 | 17 | 30 |
| 11 | 11 | 10 | 18 | 28 |
| 12 | 12 | 17 | 17 | 34 |
| 13 | 13 | 16 | 19 | 35 |
| 14 | 14 | 14 | 16 | 30 |
| 15 | 15 | 8 | 18 | 26 |
| 16 | 16 | 15 | 13 | 28 |
| 17 | 17 | 17 | 17 | 34 |
| 18 | 18 | 8 | 18 | 26 |

Berdasarkan hasil analisis reliabilitas pada Tabel 3.6, menunjukkan bahwa soal Semester Genap Biologi Materi Sistem Koordinasi Kelas XI MAM 1 Malang memiliki kriteria reliabilitas tinggi.

Kompetensi yang perlu dimiliki oleh guru adalah kemampuan menyusun soal sebagai instrumen tes peserta didik dengan baik dan analisis hasil tes. Tes yang dikonstruksi tanpa memperhatikan kualitasnya dimungkinkan tidak dapat mengungkapkan kemampuan peserta didik dengan tepat. Kualitas tes hasil belajar yang baik mencakup:

1) Validitas.

Suatu tes atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya sesuai dengan maksud dilakukannya pengukuran tersebut (Masidjo, 1995). Berdasarkan hasil analisis butir soal HOTS yang telah dikembangkan telah memiliki validitas yang tinggi yaitu 92%.

2) Reliabilitas.

Reliabilitas adalah taraf kemampuan tes dalam menunjukkan konsistensi hasil pengukurannya yang diperlihatkan dalam taraf ketepatan dan ketelitian hasil. Dari hasil analisis menunjukkan hasil reliabilitas yang tinggi yaitu 0,74. Sehingga butir soal HOTS yang dikembangkan dapat digunakan dalam tes pembelajaran di SMA.

3) Daya Pembeda.

Daya pembeda adalah taraf jumlah jawaban benar siswa yang tergolong kelompok (pandai = *upper group*) berbeda dari siswa yang tergolong kelompok bawah (kurang pandai = *lower group*) untuk suatu item (Pantiwati, 2016). Berdasarkan hasil analisis butir daya beda soal HOTS memiliki kriteria ditolak sebesar 84% sehingga soal-soal ini tidak dapat diimplementasikan.

4) Tingkat kesukaran.

Tingkat kesulitan merupakan kemampuan siswa untuk menjawab soal dengan kriteria soal mudah, sedang dan sukar (Sulistiyorini, 2009). Widoyoko (2014) mengungkapkan bahwa tingkat kesukaran yang baik pada suatu tes adalah 25% mudah, 50% sedang, dan 25% sukar. Namun, berdasarkan fakta menunjukkan bahwa soal HOTS yang dikembangkan proporsi soal mudah lebih banyak yaitu sekitar 52% sehingga perlu dikaji ulang terkait proporsi tingkat kesukaran soal (Widoyoko, 2014).

5) Analisis Pengecoh.

Pengecoh (*distractor*) adalah pilihan yang bukan merupakan kunci jawaban (Pantiwati, 2016). Arikunto (2012) memaparkan bahwa pengecoh dapat berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik bagi peserta tes yang kurang memahami materi. Berdasarkan hasil analisis menunjukkan bahwa efektifitas pengecoh dalam kategori cukup 40% maka perlu dikaji ulang dalam mengembangkan butir soal berbasis HOT (Arikunto, 2012). Mengembangkan HOTS diperlukan kemampuan guru untuk merencanakan dan mengelola pembelajaran yang efektif dalam membelajarkan siswa baik dalam berpikir secara logis, sikap, maupun keterampilan. Guru yang efektif adalah guru yang mempunyai persiapan dan pelaksanaan pembelajaran yang sistematis (Moon, Mayes and Hutchinson, 2002).

SIMPULAN

Hasil analisis validitas soal dengan taraf signifikan 5% untuk siswa berjumlah 18 diperoleh 50 soal (92%) valid. 2) Hasil uji reliabilitas soal diperoleh indeks reliabilitas dalam kriteria “tinggi”. Jadi butir soal memiliki konsistensi yang tinggi dalam mengukur kemampuan peserta didik mengenai materi sistem koordinasi dapat membedakan peserta didik berkemampuan tinggi dengan peserta didik berkemampuan rendah. 4) Hasil uji analisis tingkat kesukaran soal yaitu: 26 soal (5%) memiliki tingkat kesukaran kategori mudah, 9 soal (75%) memiliki tingkat kesukaran kategori sedang dan 15 soal (20%) yang memiliki tingkat kesukaran kategori sukar. 5) Hasil uji pengecoh pada soal secara keseluruhan ada 4 pengecoh tidak berfungsi. Pengecoh disebut tidak berfungsi jika dipilih kurang dari 5% keseluruhan peserta tes. Pengecoh yang tidak berfungsi perlu direvisi kembali. Hasil Uji daya beda 84% soal ditolak sehingga perlu direvisi.

DAFTAR PUSTAKA

- Arikunto, S. (2012) ‘Dasar-dasar evaluasi pendidikan edisi 2’, Jakarta: Bumi Aksara.
- Brookhart, S. M. (2010) *How to assess higher-order thinking skills in your classroom*. ASCD.
- Brown, J. H. and Shavelson, R. J. (1996) *Assessing Hands-On Science: A Teacher’s Guide to Performance Assessment*. ERIC.
- Gronlund, G. (1998) ‘Portfolios as an Assessment Tool: Is Collection of Work Enough?’, *Young Children*, 53(3), pp. 4–10.
- Jurs, S. G. and Wiersma, W. (1990) ‘Educational measurement and testing’, *United States: A Division of Simon & Schuster*.
- Masidjo, I. (1995) *Penilaian pencapaian hasil belajar siswa di sekolah*. Walter de Gruyter.
- Moon, B., Mayes, A. S. and Hutchinson, S. (2002) *Teaching, learning and the curriculum in secondary schools: a reader*. Psychology Press.

Murray, E. C. (2011) 'Implementing higher-order thinking in middle school mathematics classrooms'. University of Georgia.

Pantiwati, Y. (2016) 'Hakekat asesmen autentik dan penerapannya dalam pembelajaran biologi', *Jurnal Edukasi Matematika dan Sains*, 1(1), pp. 18–27.

Rusdiana, H., Sumardi, K. and Arifiyanto, E. S. (2014) 'Evaluasi Hasil Belajar Menggunakan Penilaian Autentik Pada Mata Pelajaran Kelistrikan Sistem Refrigerasi', *Journal of Mechanical Engineering Education*, 1(2), pp. 274–283.

Sulistyorini, E. P. (2009) 'dalam Meningkatkan Mutu Pendidikan', *Yogyakarta: Teras*.

Widoyoko, E. P. (2014) 'Penilaian hasil pembelajaran di sekolah', *Yogyakarta: Pustaka Pelajar*, 1(2), p. 8.