

ANALISIS KUALITAS BUTIR SOAL UJIAN AKHIR SEMESTER GANJIL PADA MATA PELAJARAN KIMIA KELAS X DI SMA NEGERI 8 PALEMBANG TAHUN AJARAN 2020/2021

Ririn Setia Ningsi Pamungkasih¹, Effendi Nawawi^{2*}

FKIP Universitas Sriwijaya

Corresponding Author* : effendi@fkip.unsri.ac.id

ABSTRAK

Penelitian ini bertujuan untuk mengetahui kualitas butir soal yang digunakan pada pelaksanaan Ujian Akhir Semester Ganjil pada Mata Pelajaran Kimia Kelas X di SMA Negeri 8 Palembang Tahun Ajaran 2020/2021 ditinjau dari aspek Materi, Konstruksi, Bahasa, Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda dan Efektivitas Pengecoh. Metode penelitian menggunakan *mix methods* dengan teknik pengambilan data dilakukan secara dokumentasi. Teknik analisis data menggunakan teknik panel serta menggunakan bantuan program *software SPSS Versi 26* dan *Software Anates Versi 4.0.9*. Subjek dalam penelitian ini adalah seluruh siswa kelas X di SMA Negeri 8 Palembang yang berjumlah 257 siswa. Hasil penelitian menunjukkan : (1) Pada Analisis kualitatif memenuhi 91% aspek materi, 96% aspek konstruksi dan 95% aspek bahasa. (2) Pada aspek validitas butir soal, 25 kategori cukup, 65% kategori rendah, 5% kategori sangat rendah, dan 5% tidak valid, (3) Reliabilitas soal sebesar 0,494 sehingga termasuk kategori rendah atau tidak reliabel, (4) Tingkat Kesukaran menunjukkan 10% kategori sangat sukar, 35% kategori sedang, 15% kategori mudah, dan 40% kategori sangat mudah, (5) Daya Pembeda menunjukkan 5% kategori sangat jelek, 40% kategori cukup, 20% kategori baik dan 5% kategori sangat baik, (6) Efektivitas pengecoh menunjukkan 5% kategori sangat baik, 15% kategori baik, 40% kategori cukup, 20% kategori kurang baik, 20% kategori tidak baik.

Kata-Kata Kunci : Analisis Butir Soal, Aspek Materi, Aspek Konstruksi, Aspek Bahasa, Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda, Efektivitas Pengecoh

PENDAHULUAN

Kurikulum 2013 dinilai sejalan dengan pendidikan pada abad 21 yaitu membangun kemampuan belajar individu dan mendukung perkembangan peserta didik menjadi pembelajar sepanjang hayat, aktif dan pembelajar yang mandiri. Tema pengembangan kurikulum 2013 adalah kurikulum yang dapat menghasilkan insan Indonesia yang produktif, kreatif, inovatif, dan afektif melalui penguatan sikap, keterampilan, dan pengetahuan yang terintegritas (Kemendikbud, 2016). Salah satu aspek keterampilan yang harus dimiliki oleh peserta didik adalah kemampuan berpikir kritis. Menurut Purwati dkk (2016) kemampuan berpikir kritis (*critical thinking*) merupakan kemampuan individu dalam menganalisis dan mengevaluasi informasi yang didapat dari hasil pengamatan, pengalaman, penalaran, maupun

komunikasi untuk mengetahui apakah informasi yang ada merupakan informasi yang terpercaya sehingga dapat memberikan kesimpulan yang rasional dan benar.

Kemampuan berpikir kritis merupakan suatu hal yang sangat penting untuk diterapkan bersamaan dengan implementasi kurikulum 2013. Dalam kurikulum 2013 siswa dituntut untuk aktif dalam proses pembelajaran. Hal tersebut menjadikan peranan guru hanya sebagai fasilitator (Prameswari dkk, 2018). Salah satu parameter utama untuk menilai keberhasilan implementasi kurikulum 2013 adalah efektivitas pembelajaran dan penilaian (Kemendikbud, 2016). Menurut Purwanto (2009) untuk mengaktualisasikan hasil belajar tersebut maka perlu adanya serangkaian pengukuran dengan menggunakan alat evaluasi yang baik dan memenuhi syarat. Evaluasi tersebut dimaksudkan sebagai gambaran untuk meninjau apakah proses belajar mengajar yang telah berlangsung sudah efektif untuk memperoleh hasil belajar sehingga dapat diketahui apakah tujuan pembelajaran yang ditetapkan telah tercapai melalui usaha proses belajar mengajar tersebut.

Evaluasi dalam pendidikan dapat menggunakan banyak alat atau instrument, salah satunya adalah dengan tes. Menurut Widoyoko (2009) tes merupakan kumpulan pertanyaan yang harus diberikan jawaban serta tanggapan yang bertujuan untuk mengukur tingkat kemampuan seseorang. Tes lebih sering digunakan pada kegiatan pengukuran dan penilaian hasil belajar siswa terutama pada aspek kognitif. Hal tersebut dikarenakan aspek kognitif merupakan aspek yang berkaitan secara langsung terhadap penguasaan materi pengajaran atau ketuntasan siswa terhadap pelajaran yang telah diajarkan. Keberhasilan suatu pendidikan dapat dilihat dari pola penilaian hasil belajar yang telah ditentukan sesuai dengan standar kurikulum yang berlaku. Sehingga dapat dikatakan bahwa baik tidaknya kegiatan pendidikan, salah satunya ditentukan oleh penilaian hasil belajar. Ketepatan penilaian hasil belajar memberikan dampak yang sangat signifikan terhadap upaya peningkatan mutu pendidikan di sekolah (Azwar, 2002). Tes dalam dunia pendidikan dibagi menjadi dua jenis, yaitu tes standar dan tes buatan guru. Tes standar merupakan jenis tes yang telah diuji validitas dan realibilitasnya secara ketat serta telah diuji kelayakan penggunaannya secara praktis. Sedangkan tes buatan guru merupakan tes yang disusun oleh guru untuk mengukur kemampuan peserta didiknya dengan menggunakan butir-butir soal buatannya (Mulyasa, 2007).

METODE

1. Tempat dan Subjek Penelitian

Penelitian ini dilakukan di SMAN 8 Palembang yang beralamat di Jalan Pertahanan III, 16 Ulu, Kecamatan Seberang Ulu II, Kota Palembang, Provinsi Sumatera Selatan, Indonesia, Kode Pos 30265. Sampel penelitian ditentukan menggunakan Teknik Total Sampling, sehingga subjek dalam penelitian ini adalah seluruh siswa kelas X di SMA Negeri 8 Palembang dengan jumlah 257 siswa.

2. Jenis dan Prosedur Penelitian

Penelitian ini menggunakan metode kuantitatif dan kualitatif (*mix methods*) untuk mengetahui kualitas butir soal yang digunakan pada pelaksanaan ujian akhir semester ganjil di SMA Negeri 8 Palembang. Analisis kualitatif menggunakan instrumen penelitian yang disesuaikan dengan indikator kaidah penulisan butir soal. Analisis kuantitatif menggunakan bantuan program *software SPSS Versi 26* dan *ANATES Versi 4.0.9*.

Penelitian dilakukan dalam beberapa tahap : melakukan studi pendahuluan, merumuskan masalah, menentukan tujuan, sampel penelitian serta metode penelitian, menentukan menyusun proposal penelitian, mendokumentasikan data-data yang diperlukan, menganalisis data, serta menyusun laporan akhir mengenai hasil analisis.

3. Analisis Data

1) Analisis Aspek Materi, Konstruksi dan Bahasa

Analisis data pada aspek materi, konstruksi dan bahasa dilakukan dengan menggunakan instrument penilaian yang sesuai dengan kaidah penulisan butir soal. Jika butir soal memenuhi indikator maka diberi tanda (√) dan jika butir soal tidak memenuhi indikator diberi tanda (X)

a. Rumus Perhitungan

$$\text{Persentase kelayakan setiap aspek} = \frac{\text{Jumlah indikator yang sesuai}}{\text{Jumlah seluruh indikator telaah}} \times 100\%$$

Sumber: Rokhyati (2011)

b. Kriteria Kualitas Validitas Isi Soal

1.1 Tabel 1 Kriteria Validitas Isi

No	Persentase Butir Soal	Kriteria
1	90% – 100	Sangat Baik
2	80% – 89	Baik
3	70% – 79%	Cukup Baik
4	60% – 69%	Sedang
5	≤59%	Kurang Baik

Sumber: Rokhyati (2011)

2) Validitas Butir

Analisis validitas butir akan dilakukan dengan menggunakan program *software SPSS* versi 26. Hasil perhitungan atau besarnya koefisien yang didapatkan akan dikonsultasikan dengan r_{tabel} pada taraf signifikansi 5%. Apabila nilai $r_{\text{hitung}} > r_{\text{tabel}}$ maka butir soal tersebut dinyatakan valid. Sebaliknya jika $r_{\text{hitung}} < r_{\text{tabel}}$ maka butir soal tersebut dinyatakan tidak valid.

Tabel 2 Klasifikasi Validitas Butir Soal

Tes Koefisien Korelasi	Keterangan
$0,80 < r_{xy} \leq 1,00$	Validitas sangat tinggi
$0,60 < r_{xy} \leq 0,80$	Validitas tinggi
$0,40 < r_{xy} \leq 0,60$	Validitas cukup
$0,20 < r_{xy} \leq 0,40$	Validitas rendah
$r_{xy} \leq 0,20$	Validitas sangat rendah

Sumber : Arikunto (2013: 215)

3) Reliabilitas

Pada penelitian ini reliabilitas tes pilihan ganda dianalisis dengan menggunakan program *software SPSS* versi 26. Menurut Sudijono (2013) untuk menginterpretasikan reliabilitas tes diberikan kriteria penafsiran nilai r sebagai berikut :

- Apabila r_{11} sama dengan atau lebih besar dari pada 0,70 maka tes/instrument yang diuji dinyatakan memiliki reliabilitas yang tinggi.

b. Apabila r_{11} lebih kecil dari 0,70 maka tes/instrument yang diuji dinyatakan belum memiliki reliabilitas yang tinggi.

4) Tingkat Kesukaran

Daya pembeda butir soal akan dianalisis menggunakan program *software Anates* versi 4.0.9. Kemudian hasil yang didapatkan diinterpretasikan pada tabel 3.

Tabel 3 Klasifikasi Tingkat Kesukaran

Tingkat Kesukaran (P)	Keterangan
0,00 – 0,15	Sangat Sukar
0,16 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 0,85	Mudah
0,86 – 1,00	Sangat Mudah

Sumber : (Sudijono, 2013)

5) Daya Pembeda

Daya pembeda butir soal akan dianalisis menggunakan program *software Anates* versi 4.0.9. Kemudian hasil yang didapatkan diinterpretasikan pada tabel 4.

Tabel 4 Klasifikasi Daya Pembeda

Indeks Diskriminasi (D)	Keterangan
0,00 – 0,20	Jelek (<i>poor</i>)
0,21 – 0,40	Cukup (<i>satisfactory</i>)
0,41 – 0,70	Baik (<i>good</i>)
0,71 – 1,00	Sangat baik (<i>excellent</i>)
Negatif	Sangat jelek, jadi semua butir soal yang mempunyai nilai D <i>negative</i> sebaiknya Dibuang

Sumber : Arikunto (2013: 218)

6) Efektivitas Pengecoh

Efektivitas pengecoh dianalisis dengan menggunakan program *software Anates* versi 4.0.9. Kemudian hasil yang didapatkan diinterpretasikan pada tabel berikut.

Tabel 5 Klasifikasi Indeks Pengecoh

Indeks Pengecoh (IP)	Keterangan
76% -125%	Sangat baik
51% - 75% atau 126% - 150%	Baik
26% - 50% atau 151% - 175%	Kurang baik
0% - 25% atau 176% - 200%	Jelek
Lebih dari 200%	Sangat jelek

Sumber : Arifin (2009: 279)

Menurut Marzuki (2015) untuk menyimpulkan efektivitas pengecoh/*distractor* pada setiap butir soal maka digunakan kriteria yang berpatokan pada Skala Likert sebagai berikut:

Tabel 6 Kriteria Penilaian Efektivitas Pengecoh

Pengecoh yang berfungsi	Keterangan
4	Sangat baik
3	Baik
2	Cukup baik
1	Kurang baik
0	Tidak baik

Sumber: Marzuki (2015)

7) Kualitas Butir Soal

Butir soal telah dianalisis aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh maka langkah selanjutnya adalah menyimpulkan kualitas butir soal dengan menggunakan empat aspek yaitu berdasarkan validitas, daya pembeda, tingkat kesukaran dan efektivitas pengecoh/*distractor* disamping aspek reliabilitas yang merupakan syarat berbasis tes secara keseluruhan (bukan dihitung per butir soal).

Untuk menginterpretasikan kualitas butir soal digunakan kriteria kualitas butir soal yang diadaptasi dari Skala *Likert* berdasarkan penelitian yang dilakukan oleh Oktanin dkk (2015) sebagai berikut:

Tabel 8 Kriteria Kualitas Butir Soal Pilihan Ganda

Jumlah kriteria yang terpenuhi	Kualitas butir soal	Revisi	Masuk bank soal
4	Sangat baik	Tidak	Ya
3	Baik	Revisi	Belum
2	Cukup baik	Revisi	Belum
1	Kurang baik	Dibuang	Tidak
0	Tidak baik	Dibuang	Tidak

Sumber :Oktanin (2015)

Penjelasan mengenai tabel kriteria kualitas butir soal yang dikemukakan oleh Oktanin (2015), yaitu:

1. Butir soal dikatakan sangat baik apabila memenuhi 4 kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh/*distractor*. Pada kondisi ini butir soal bisa masuk ke bank soal.
 - Ditinjau dari segi validitas butir soal harus memiliki kategori valid dimana $r_{hitung} \geq r_{tabel}$, dengan koefisiensi korelasi diatas 0,3.
 - Ditinjau dari tingkat kesukaran butir soal harus memiliki tinglat kesukaran sedang yaitu anatar rentang indeks kesukaran 0,31 hingga 0,70.
 - Ditinjau dari daya pembeda butir soal harus berada pada salah satu dari ketiga kategori, yaitu cukup, baik, atau sangat baik.
 - Ditinjau dari tingkat efektivitas pengecoh, butir soal harus berada pada kategori baik atau sangat baik.
2. Butir soal dikatakan baik jika memenuhi 3 kriteria dari 4 kriteria soal yang baik (validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh/*distractor*). Pada kondisi ini butir soal belum bisa masuk ke bank soal. Soal harus direvisi hingga memenuhi empat kriteria.
3. Butir soal dikatakan sedang jika memenuhi 2 kriteria dari 4 kriteria soal yang baik (validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh/*distractor*). Pada kondisi ini soal belum bisa disimpan dalam bank soal. Soal harus direvisi hingga memnuhi empat kriteria.
4. Butir soal dikatakan tidak baik jika hanya memenuhi 1 dari 4 kriteria soal yang baik (

validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh/*distractor*). Pada kondisi ini butir soal tidak bisa masuk ke bank soal. Soal membutuhkan revisi signifikan sehingga lebih baik dibuang.

5. Butir soal dikatakan sangat tidak baik jika tidak ada kriteria yang terpenuhi (validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh/*distractor*). Pada kondisi ini butir soal tidak bisa masuk ke dalam bank soal. Soal membutuhkan revisi signifikan sehingga lebih baik dibuang.
6. Selain syarat berbasis butir, maka tes secara keseluruhan harus realibel dengan ketentuan yang sudah dijelaskan.

HASIL DAN PEMBAHASAN

1) Analisis Aspek Materi, Konstruksi dan Bahasa

Terdapat 9 butir soal (3,6,7,9,13,14,17,18,20) yang memenuhi seluruh indikator penilaian atau memiliki persentase 100% pada aspek materi, konstruksi, dan bahasa. Terdapat 1 (butir soal nomor 1) yang memenuhi indikator penilaian sebesar 76,66%. Terdapat 1 (butir soal nomor 2) yang memenuhi indikator penilaian sebesar 68,33%. Terdapat 1 (butir soal nomor 5) yang memenuhi indikator penilaian sebesar 76,66%. Terdapat 5 butir soal (4,8,12,15,16) yang memenuhi indikator penilaian sebesar 91,66%. Terdapat 3 butir soal (10,11,19) yang memenuhi indikator penilaian sebesar 96,66%.

Berdasarkan hasil analisis terhadap aspek materi, konstruksi dan bahasa Butir Soal Ujian Akhir Semester Ganjil pada Mata Pelajaran Kimia Kelas X di SMA Negeri 8 Palembang Tahun Ajaran 2020/2021 memenuhi 91% aspek materi sehingga termasuk kategori sangat baik, 96% aspek konstruksi sehingga termasuk kategori sangat baik, dan 95% aspek bahasa sehingga termasuk kategori sangat baik.

2) Validitas Butir

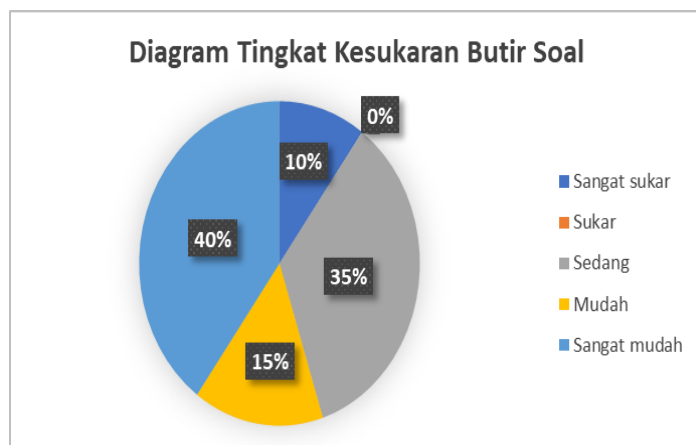
Hasil analisis menunjukkan bahwa tidak terdapat butir soal dengan kategori validitas sangat tinggi maupun validitas tinggi, 5 butir soal (2,3,4,5,8) atau 25% dengan kategori validitas cukup, 13 butir soal (1,7,9,10,11) validitas rendah, 1 butir soal atau 5% dengan kategori validitas sangat rendah dan 1 butir soal (6) atau 5% tidak valid.

3) Reliabilitas Tes

Reliabilitas tes dianalisis dengan menggunakan bantuan software SPSS Versi 26. Nilai reliabilitas tes soal pilihan ganda yang digunakan pada kegiatan penilaian akhir semester di SMA Negeri 8 Palembang adalah sebesar 0,494. Untuk menginterpretasikan reliabilitas tes diberikan kriteria penafsiran nilai r yaitu, apabila $r_{11} \geq 0,70$ maka tes/instrument yang diuji dinyatakan memiliki reliabilitas yang tinggi. Apabila $r_{11} < 0,70$ maka tes/instrument yang diuji dinyatakan belum memiliki reliabilitas yang tinggi. Berdasarkan hal tersebut maka dapat disimpulkan bahwa soal tes yang digunakan pada kegiatan penilaian akhir semester pada mata pelajaran kimia kelas X di SMA Negeri 8 Palembang tahun ajaran 2020/2021 memiliki reliabilitas pada kategori rendah.

4) Tingkat Kesukaran

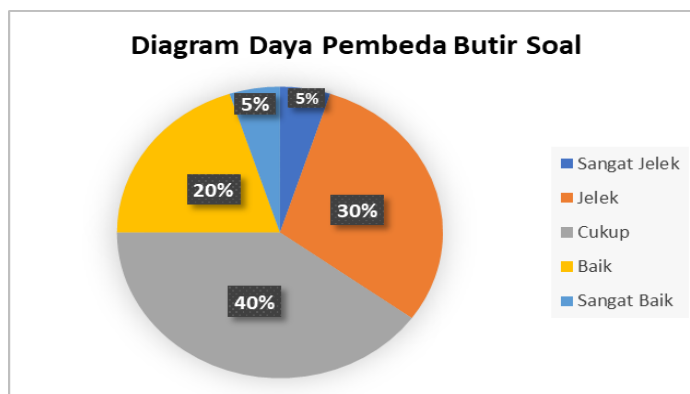
Analisis mengenai tingkat kesukaran butir-butir soal menggunakan bantuan program Anates 4.0.9. Berdasarkan analisis tersebut didapatkan bahwa pada soal tes yang digunakan dalam penilaian akhir semester ganjil pada mata pelajaran kimia di SMA Negeri 8 Palembang tahun ajaran 2020/2021 tidak ada butir soal dengan kategori sukar, terdapat 2 butir soal (14, 20) atau 10% dengan kategori sangat sukar, 7 butir soal (2,3,8,9,11,16,18) atau 35% dengan kategori sedang, 3 butir soal (12,15,19) atau 15% dengan kategori mudah, dan 8 butir soal (1,4,5,6,7,10,13,17) atau 40% dengan kategori sangat mudah.



Gambar 3. Diagram Distribusi Butir Soal Berdasarkan Tingkat Kesukaran

5) Daya Pembeda

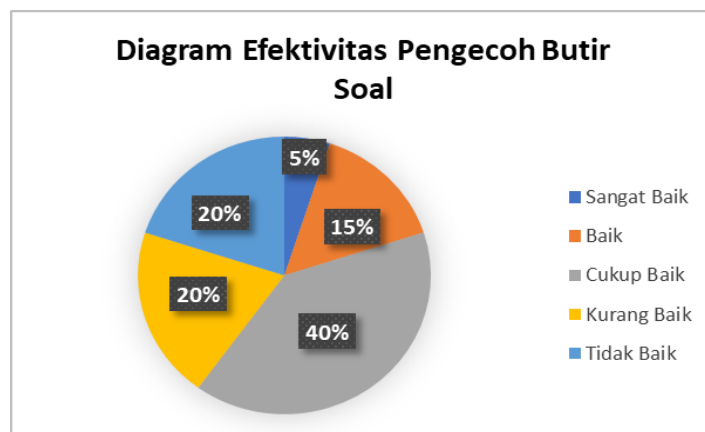
Analisis daya pembeda butir-butir soal dilakukan dengan menggunakan bantuan program Anates 4.0.9. Berdasarkan hasil analisis yang telah dilakukan terdapat 1 butir soal (20) atau 5% yang memberikan indeks daya pembeda negative dengan kata lain butir soal tersebut memiliki kategori daya pembeda sangat jelek. Terdapat 6 butir soal (1,6,7,10,11,13) atau 30% dengan kategori daya pembeda yang jelek, 8 butir soal (9,12,14,15,16,17,18,19) atau 40% dengan kategori cukup, 4 butir soal (2,3,5,8) atau 20% dengan kategori daya pembeda baik, dan 1 butir soal (4) atau 5% dengan kategori daya pembeda sangat baik. Hasil analisis daya pembeda butir tersebut dibuat ke dalam diagram persentase sebagai berikut.



Gambar 4. Diagram Distribusi Butir Soal Berdasarkan Daya Pembeda

6) Efektivitas Pengecoh

Analisis efektivitas pengecoh/*distractor* soal dilakukan dengan menggunakan bantuan program Anates 4.0.9. Berdasarkan hasil analisis yang telah dilakukan didapatkan hasil bahwa terdapat 1 butir soal (12) atau 5% dengan kategori efektivitas pengecoh sangat baik, 3 butir soal (2,4,14) atau 15% dengan kategori baik, 8 butir soal (1,5,9,10,16,17,18,19) atau 40% dengan kategori cukup baik, 4 butir soal (7,8,11,13) atau 20% dengan kategori kurang baik, serta 4 butir soal (3,6,15,20) atau 20% dengan kategori tidak baik.



Gambar 5. Diagram Distribusi Butir Soal Berdasarkan Efektivitas Pengecoh

7) Analisis Butir Soal Secara Keseluruhan Ditinjau dari Validitas Butir, Tingkat Kesukaran, Daya Pembeda dan Efektivitas Pengecoh

Setelah semua butir soal dianalisis dari segi validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh kemudian butir soal dianalisis secara keseluruhan untuk mengetahui kualitas butir soal yang digunakan pada Ujian Akhir Semester Ganjil pada Mata Pelajaran Kimia Kelas X di SMA Negeri 8 Palembang Tahun Ajaran 2020/2021. Kualitas butir soal terbagi ke dalam lima kategori yaitu sangat baik, baik, cukup, kurang baik dan tidak baik. Hasil analisis butir soal Ujian Akhir Semester pada Mata Pelajaran Kimia Kelas X di SMA Negeri 8 Palembang Tahun Ajaran 2020/2021 yaitu :

Tabel 9 Distribusi Soal Berdasarkan Aspek yang Terpenuhi

Jumlah Aspek yang Terpenuhi	Kualitas Butir Soal	Nomor Butir Soal	Jumlah	Persentase (%)
4	Sangat Baik	2	1	5%
3	Baik	3,4,8	3	15%
2	Cukup	5,9,12,14,16,18	6	30%
1	Kurang Baik	11,15,17,19	4	20%
0	Tidak Baik	1,6,7,10,13,20	6	30%

KESIMPULAN

Analisis butir soal secara kualitatif menunjukkan bahwa 91% aspek materi terpenuhi, 96% aspek konstruksi terpenuhi dan 95% aspek bahasa terpenuhi. Validitas soal menunjukkan bahwa tidak terdapat butir soal dengan kategori validitas sangat tinggi maupun validitas tinggi, 5 soal (25%) cukup, 13 soal (65%) rendah, 1 soal (5%) sangat rendah dan 1 soal atau (5%) tidak valid. Nilai reliabilitas tes soal pilihan ganda adalah sebesar 0,494 atau tidak reliabel. Tingkat kesukaran soal pilihan ganda menunjukkan bahwa tidak ada butir soal dengan kategori sukar, terdapat 2 soal (10%) sangat sukar, 7 soal (35%) sedang, 3 soal (15%) dengan mudah, dan 8 soal (40%) sangat mudah. Daya pembeda menunjukkan terdapat 1 soal (5%) memiliki kategori daya pembeda sangat jelek, 6 soal (30%) jelek, 8 soal (40%) cukup, 4 soal (20%) baik, dan 1 soal (5%) sangat baik. Efektivitas pengecoh/*distractor* menunjukkan terdapat 1 soal (5%) sangat baik, 3 soal (15%) baik, 8 soal (40%) cukup baik, 4 soal (20%) kurang baik, serta 4 soal (20%) tidak baik. Kualitas butir soal secara keseluruhan yaitu terdapat soal (30%) memiliki kategori tidak baik sehingga harus dibuang, 4 soal (20%) kategori kurang baik sehingga harus dibuang, 6 soal (30%) kategori cukup baik sehingga harus melewati tahap revisi sedang, 3 soal (15%) dengan kategori baik memerlukan revisi kecil, dan 1 soal (5%) kategori sangat baik sehingga dapat langsung masuk ke bank soal dan bisa digunakan kembali pada tes berikutnya.

DAFTAR PUSTAKA

- Agustina, M., Hastari, M. & Andina, M. (2018). Analisis Butir Soal Ulangan Akhir Semester Mata Pelajaran Bahasa Indonesia kelas XI. *Jurnal Ilmiah Bina Edukasi*. 11(1) : 26.35.
- Amirono dan Daryanto. (2016). *Evaluasi dan Penilaian Pembelajaran Kurikulum 2013*. Jakarta: Gava Media.
- Arifin, Z. (2014). *Evaluasi Pembelajaran*. Bandung : PT Remaja Rosdakarya.
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara.
- Asrul., Rusydi, A., & Rosnita. (2015). *Evaluasi Pembelajaran*. Bandung: Cita Pustaka Media.
- Azwar, S. (1987). *Reliabilitas dan Validitas Edisi Ketiga*. Yogyakarta : Pustaka Belajar.
- Daryanto. (2010). *Evaluasi Pendidikan*. Jakarta : PT Rineka Cipta.

- E, Mulyasa. (2007). *Standar Kompetensi dan Sertifikasi Guru*. Bandung : PT Remaja Rosdakarya.
- Fatimah, L. U. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Jurnal Komunikasi dan Pendidikan Islam*. 8(2): 37-64.
- Firmansyah, H. (2018). Analisis Butir Soal Tes Pilihan Ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Wonosari Tahun 2017/2018. Yogyakarta : *Skripsi*.
- Irawati, R., Elvin, Y.E & Sri, B . (2020). Analisis Butir Soal Ujian Akhir Semester Gasal Menggunakan Program Anbuso di SMA Negeri 1 Boyolali Tahun Ajaran 2019/2020. *Jurnal Materi dan Pembelajaran Fisika*. 10(1) : 11 – 19.
- Iskandar, D. & Senam. (2015). Studi Kemampuan Guru Kimia SMA Lulusan UNY dalam Mengembangkan Soal UAS Berbasis HOTS. *Jurnal Inovasi Pendidikan IPA* 1(1): 65-72.
- Kemendikbud. (2016). *Panduan Penilaian Oleh Pendidik dan satuan Pendidikan untuk Sekolah Menengah Atas*. Jakarta : Kemendikbud.
- Keumala, M., Ibnu, K., dan Latifah, H. (2019). Analisis Butir Soal Ujian Semester Genap Pelajaran Kimia Kelas X SMK SMTI Banda Aceh Tahun Ajaran 2018/2019 Menggunakan Program Proanaltes. *Jurnal Ilmiah Mahasiswa Jurusan Pendidikan Kimia (JIMKPK)*. 4(2): 27-32.
- Kusaeri & Suprananto.(2012). *Pengukuran dan Penilaian Pendidikan*. Yogyakarta : Graha Ilmu.
- M, Marthunis., Ibnu, K., & Zulfadli. (2015). Analisis Kualitas Butir Soal Ujian Semester Genap Mata Pelajaran Kimia Kelas X MAN Model Banda Aceh Tahun Pelajaran 2014/2015 Menggunakan Program Analtes. *Jurnal Ilmiah Mahasiswa Pendidikan Kimia (JMIK)*. 1(4): 70-78.
- Marzuki. (2015). Analisis Butir Soal Ulangan Tengah Semester Genap Mata Pelajaran Produktif Akuntansi Kelas X Program Keahlian Akuntansi SMKN 1 Yogyakarta Tahun Ajaran 2014/2015. *Skripsi*. Yogyakarta: UNY.
- Oktanin, W.S. & Sukirno. (2015). Analisis Butir Soal Ujian Akhir Mata Pelajaran Ekonomi Akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*. 9(1): 35-34.
- Prabayanti, N. M. D., I. K. Sudiana, & N. M. Wiranti. (2018). Analisis Tes Ulangan Kenaikan Kelas Buatan Guru Mata Pelajaran Kimia. *Jurnal Pendidikan Kimia Indonesia*. 2(1): 1-10.
- Prameswari, S. W., Suharno., & Sarwanto. (2018). Inculcate Critical Thinking Skills in Primary School. *Jurnal Universitas Sebelas Maret*. 1(1): 742- 750.
- Purwanto. (2009). *Evaluasi Hasil Belajar*. Yogyakarta : Pustaka Belajar.

- Purwati, R., Hobri., & Arif,F. (2016). Analisis Kemampuan Berpikir Kritis Siswa dalam Menyelesaikan Masalah Persamaan Kuadrat Pada Pembelajaran Model *Creative Problem Solving*. *Jurnal Kadikma*. 7(1): 84-93.
- Putra, S.R. (2013). *Desain Evaluasi Belajar Berbasis Kinerja*. Yogyakarta : DIVA Press.
- Rahmasari, D & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. *Economic Education Analysis Journal*. 5(1) : 317 –330.
- Ratnawulan, E. & A. Rusdiana. (2014). *Evaluasi Pembelajaran*. Bandung : Pustaka Setia.
- Rokhyati. (2011). Karakteristik Secara Kualitatif dan Kuantitatif Soal Ulangan Akhir Semester Genap Bahasa Indonesia Kelas XII SMA Negeri di Kabupaten Purbalingga Tahun Pelajaran 2010/2011. *Skripsi*. Yogyakarta: UNY.
- Septiana, N . (2016). Analisis Butir Soal Ulangan Akhir Semester (UAS) Biologi Tahun Ajaran 2015/2016 Kelas X dan XI pada MAN Sampit. *EduSains* . 4(2) :115 – 121.
- Silverius, S. (1991). *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT Grasindo.
- Sofiana, S. (2010). Analisis Butir Soal Ulangan Kenaikan Kelas Mata Pelajaran Kimia Kelas X SMA Negeri 8 Surakarta Tahun Ajaran 2009/2010. *Skripsi*. Surakarta : FKIP UNS.
- Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta : PT Raja Grafindo Persada.
- Sudjana, N. (2012). *Penilaian Hasil Proses Belajar Mengajar*. Bandung : PT Remaja Rosdakarya.
- Sugiyono. (2012). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif dan R&D)*. Jakarta : Alfabeta.
- Sukardi. (2008). *Evaluasi Pendidikan Prinsip & Operasionalnya*. Jakarta : Bumi Aksara.
- Widoyoko, E.P. (2009). *Evaluasi Program Pembelajaran*. Yogyakarta : Pustaka Pelajar.