

UJI KRITERIA INSTRUMEN PENILAIAN HASIL BELAJAR KIMIA

Ana Yustika*, Eko Budi Susatyo dan Murbangun Nuswowati

Jurusan Kimia FMIPA Universitas Negeri Semarang
Gedung D6 Lantai 2 Kampus Sekaran Gunungpati Semarang, 50229, Telp. (024)8508035
E-mail: anayustika@yahoo.com

ABSTRAK

Penelitian ini dilaksanakan dengan metode deskriptif untuk mengetahui hasil uji kriteria instrumen penilaian hasil belajar kimia kelas XI. Uji kriteria tersebut dilaksanakan melalui analisis validitas butir, indeks kesukaran, jenjang soal, efektifitas *distractor* dan *reliabilitas* soal Ujian Akhir Semester (UAS) Gasal tahun ajaran 2013/2014 mata pelajaran kimia. Sebagai sekolah sampel digunakan Menengah Atas (SMA) Negeri 1 Ambarawa dan Madrasah Aliyah (MA) Negeri Suruh. Bentuk soal yang dianalisis adalah pilihan ganda, dengan total soal sebanyak 40 butir. Berdasarkan hasil uji kriteria instrumen penilaian hasil belajar di salah satu SMA Negeri di Ambarawa, diketahui bahwa sebanyak 27 butir soal tergolong valid dengan jenjang soal C1/pengetahuan sampai dengan C3/aplikasi. Terdapat 3 butir soal sukar, 12 butir soal sedang, dan 25 butir soal mudah. Dari total 160 butir *distractor* yang digunakan, 89 butir diantaranya tergolong efektif. Instrumen tergolong reliabel karena memiliki koefisien reliabilitas 0,70. Sedangkan untuk hasil uji kriteria instrumen penilaian hasil belajar di salah satu MA Negeri Suruh, sebanyak 28 butir soal tergolong valid dengan jenjang soal C1/pengetahuan sampai dengan C4/analisis. Terdapat 10 butir soal sedang dan 30 butir soal mudah. *Distractor* (pengecoh) yang tergolong efektif berjumlah 91 butir. Instrumen tergolong reliabel karena memiliki koefisien reliabilitas 0,81.

Kata kunci: instrumen, kriteria, penilaian

ABSTRACT

This research was conducted with descriptive method to determine the results of assessment criteria for learning outcomes instrument in chemistry subject of class XI. The test include the analysis of the validity, difficulty index, level of matter, distractor effectiveness and reliability of the odd semester final examination in the academic year 2013/2014 chemistry subjects at a Public Senior High School of Ambarawa and a Public Islamic Senior High School of Suruh. Form of matter that is analyzed is multiple choice, with total 40 items. Based on the test results in Public Senior High School of Ambarawa, it is known that 27 items valid, by about C1/know up to C3/apply level. There are 3 items was difficult, 12 items medium, and 25 items easy. Of the total 160 existing distractor items, 89 items classified distractor effective. Instruments classified as reliable because it has a reliability coefficient of 0.70. While the test results in Public Islamic Senior High School of Suruh, 28 items were classified as valid items by about C1/know up to C4/analyze level. There are 10 items was medium and 30 items easy. Effective distractor was 91 items. Instruments classified as reliable because it has a reliability coefficient of 0.81.

Keywords: assessment, criteria, instrument

PENDAHULUAN

Ujian Akhir Semester (UAS) merupakan alat evaluasi pendidikan yang digunakan guru untuk mengetahui tingkat penca-

paian kompetensi siswa di akhir pembelajaran. Sebagai alat evaluasi pendidikan, paling tidak UAS memiliki empat fungsi yaitu (1)

Ana Yustika, dkk, Uji Kriteria Instrumen

untuk membantu guru dalam mengevaluasi siswa; (2) untuk menaksir apakah siswa benar-benar memahami pembelajaran seperti yang diharapkan; (3) untuk memotivasi siswa; dan (4) untuk membantu siswa dalam usaha atau karya bidang akademik (Jandaghi dan Fatemeh, 2008). Adapun berbagai macam bentuk soal yang dapat digunakan untuk menaksir penguasaan materi siswa, diantaranya dengan ujian jawaban bebas (pertanyaan *esai* panjang, pertanyaan dengan jawaban singkat, pertanyaan *esai* modifikasi), pertanyaan pilihan ganda, tes individu, dan tes kelompok. Namun, pada dasarnya tidak ada bentuk soal yang lebih unggul dibandingkan yang lain (Khan dan Badr, 2011).

Bentuk soal pilihan ganda merupakan soal yang umum digunakan pada UAS. Sebagaimana yang dinyatakan oleh Zaman, *et al.*, (2010) bahwa penggunaan soal pilihan ganda sebagai alat evaluasi pendidikan merupakan tren yang umum digunakan di seluruh dunia. Penggunaan tes pilihan ganda ternyata memiliki banyak keuntungan, diantaranya banyak sekali materi yang dapat dicakup (Suharsimi, 2009) dan sistem skoringnya sangat mudah serta reliabel. Selain itu, untuk menilai hasil tes pilihan ganda, guru juga bisa menggunakan program komputer (Khan dan Badr, 2011).

Uji kriteria instrumen penilaian hasil belajar melalui analisis butir soal penting dilaksanakan untuk mengetahui baik tidaknya butir-butir soal yang diujikan untuk mengukur kemampuan siswa. Hal ini didukung oleh Purwati dan Irni (2009) yang

menyatakan bahwa analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai. Selain itu, analisis butir soal ini digunakan untuk mengamati karakteristik khusus dari butir-butir soal dan digunakan untuk menjamin bahwa pertanyaan yang diberikan sesuai dengan materi ujian (Zaman, *et al.*, 2010).

Analisis butir soal yang dilaksanakan dalam penelitian ini mencakup beberapa hal, diantaranya adalah analisis validitas butir, indeks kesukaran, jenjang soal, efektifitas *distractor* dan reliabilitas instrumen penilaian hasil belajar. Validitas adalah ketepatan interpretasi hasil prosedur pengukuran (Ratnaningsih, 2011). Validitas butir menandai bahwa butir tes dapat menjalankan fungsi dan pengukurannya dengan baik. Hal ini diketahui dari seberapa besar peran yang diberikan oleh butir soal tes tersebut dalam mencapai skor seluruh tes (Nuswowati, *et al.*, 2010).

Selain valid, tes juga harus tetap apabila digunakan beberapa kali. Karakteristik ini biasanya disebut sebagai reliabilitas (Jandaghi dan Fatemeh, 2008). Ajeg yang dimaksudkan disini bukan berarti harus sama, tetapi sama dalam kedudukan siswa di antara anggota kelompok yang lain. Reliabilitas sebuah instrumen harus memenuhi minimal 0,70 (Suparji, 2010).

Berdasarkan penelitian, disebutkan bahwa indeks kesukaran merupakan alat yang sangat baik digunakan untuk menilai kualitas soal tipe pilihan ganda (Patel dan Neeraj, 2013). Indeks kesukaran merupakan

bilangan yang menunjukkan sukar dan mudahnya suatu soal. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal dengan indeks kesukaran 0,0 menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,0 menunjukkan bahwa soalnya terlalu mudah (Suharsimi, 2009).

Berkenaan dengan analisis mengenai proporsi setiap jenjang pada kedua paket soal UAS, peneliti menggunakan taksonomi *Bloom* versi terbaru menurut Peggy Dettmer. Di dalam taksonomi *Bloom* versi terbaru ini dikenal 8 jenjang (level) dalam ranah kognitif. Jenjang tersebut diantaranya pengetahuan (C1), pemahaman (C2), aplikasi (C3), analisis (C4), evaluasi (C5), dan sintesis (C6), imajinasi (C7) dan kreasi (C8). Berdasarkan taksonomi tersebut, dapat dikatakan bahwa soal dengan jenjang C4 sampai C8 merupakan soal-soal berpikir tingkat tinggi. Semakin banyak jenjang soal tingkat tinggi tersebut, semakin baik pula kualitas soal.

Dengan menganalisis butir soal, dapat pula ditentukan baik tidaknya pengecoh (*distractor*) yang dibuat oleh guru. Pengecoh yang tidak dipilih sama sekali oleh *testee* berarti bahwa pengecoh itu jelek. Suatu *distractor* dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 5% pengikut tes (Suharsimi, 2009:220).

Rumusan masalah dalam penelitian ini adalah bagaimana hasil uji kriteria instrumen penilaian hasil belajar di SMA Negeri 1 Ambarawa dan MA Negeri Suruh, yang dilaksanakan melalui analisis butir soal Ujian Akhir Sekolah (UAS) kimia kelas XI semester gasal tahun ajaran 2013/2014.

Adapun tujuan dari penelitian ini adalah untuk mengetahui hasil uji kriteria instrumen penilaian hasil belajar di SMA Negeri 1 Ambarawa dan MA Negeri Suruh, yang dilaksanakan melalui analisis butir soal Ujian Akhir Sekolah (UAS) kimia kelas XI semester gasal tahun ajaran 2013/2014.

METODE PENELITIAN

Penelitian ini dilaksanakan dengan metode deskriptif untuk mengetahui hasil uji kriteria instrumen penilaian hasil belajar kimia kelas XI. Metode ini dimulai dengan mengumpulkan data, menganalisis data dan menginterpretasikannya (Suryana, 2010). Populasi dalam penelitian ini adalah instrumen penilaian hasil belajar kimia kelas XI di SMA/MA Negeri di Kabupaten Semarang. Teknik pengambilan sampel yang digunakan adalah *purposive sampling*, yaitu pengambilan sampel sumber data dengan pertimbangan tertentu, berkaitan dengan seseorang yang mempunyai informasi yang diperlukan sehingga memudahkan peneliti menjelajahi objek yang diteliti (Arifianti, 2013). Adapun sampel yang diambil adalah instrumen penilaian hasil belajar kimia kelas XI di SMA Negeri 1 Ambarawa dan MA Negeri Suruh.

Metode pengambilan data dalam penelitian ini menggunakan metode dokumentasi. Data-data yang diambil mencakup lembar kisi-kisi, soal UAS kimia kelas XI semester gasal tahun ajaran 2013/2014 yang terdiri atas 40 soal tipe pilihan ganda dan lembar jawaban siswa. Lembar jawaban yang dianalisis dalam uji kriteria instrumen penilaian hasil belajar kimia ini adalah

lembar jawaban siswa kelas XI IPA 4 SMA Negeri 1 Ambarawa dan XI IPA 1 MA Negeri Suruh. Dengan demikian, materi penelitian adalah seluruh materi kimia kelas XI IPA yang diujikan dalam soal UAS di kedua sekolah tersebut. Adapun variabel yang diteliti dalam uji kriteria instrumen penilaian hasil belajar ini adalah validitas butir, indeks kesukaran, jenjang soal, efektifitas *distractor* dan reliabilitas instrumen.

Data-data penelitian yang sudah dikumpulkan selanjutnya digunakan untuk uji kriteria instrumen penilaian hasil belajar, yaitu dengan menganalisis butir soal. Analisis butir soal yang dilaksanakan dalam penelitian ini mencakup beberapa hal, diantaranya adalah analisis validitas butir, indeks kesukaran, jenjang soal, efektifitas *distractor* dan reliabilitas instrument. Analisis jenjang soal dilaksanakan berdasarkan

ranah kognitif dalam taksonomi *Bloom* versi terbaru.

HASIL DAN PEMBAHASAN

Hasil uji validitas butir soal UAS mata pelajaran kimia di SMA Negeri 1 Ambarawa menunjukkan bahwa dari 40 butir soal yang diujikan, ternyata terdapat 27 butir soal yang valid dan 13 butir soal yang tidak valid. Soal yang tergolong valid, diantaranya adalah soal-soal dengan nomor 2, 5, 6, 7, 8, 9, 14, 15, 16, 18, 20, 22, 23, 24, 25, 27, 28, 31, 33, 34, 35, 36, 37, 38, 39, 40. Sedangkan soal yang tidak valid, diantaranya adalah soal-soal dengan nomor 1, 3, 4, 10, 11, 12, 13, 17, 19, 21, 26, 29, 30, 32. Adapun hasil uji validitas butir soal di SMA Negeri 1 Ambarawa tersebut dapat dilihat pada Tabel 1.

Tabel 1. Hasil uji validitas butir soal di salah satu SMA Negeri di Ambarawa

No	Hasil Uji	Kriteria	Butir Soal	No Soal
1	$t_{hitung} > t_{tabel}$	Valid	27	2, 5, 6, 7, 8, 9, 13, 14, 15, 16, 18, 20, 22, 23, 24, 25, 27, 28, 31, 33, 34, 35, 36, 37, 38, 39, 40
2	$t_{hitung} < t_{tabel}$	Tidak valid	13	1, 3, 4, 10, 11, 12, 17, 19, 21, 26, 29, 30, 32
Jumlah			40	40

Hasil uji validitas terhadap 40 butir soal UAS kimia kelas XI tahun ajaran 2013/2014 di MA Negeri Suruh menunjukkan bahwa 28 butir soal yang valid dan 12 lainnya tidak valid. Butir soal yang valid, yaitu soal nomor 1, 3, 5, 6, 7, 8, 10, 11, 12, 14, 16, 17, 19, 22, 23, 25, 26, 28, 29, 31, 32, 33, 34, 35,

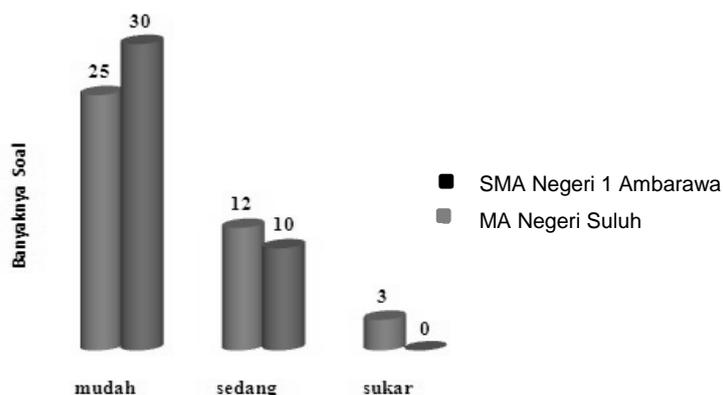
36, 37, 38, 40, sedangkan 12 butir soal lainnya yang tidak valid, yakni soal nomor 2, 4, 9, 13, 15, 18, 20, 21, 24, 27, 30, 39. Adapun hasil uji validitas butir soal di MA Negeri Suruh tersebut dapat dilihat pada Tabel 2.

Tabel 2. Hasil uji validitas butir soal di MA Negeri Suruh

No	Hasil Uji	Kriteria	Butir Soal	No Soal
1	$t_{hitung} > t_{kritis}$	Valid	28	1, 3, 5, 6, 7, 8, 10, 11, 12, 14, 16, 17, 19, 22, 23, 25, 26, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 40
2	$t_{hitung} < t_{kritis}$	Tidak valid	12	2, 4, 9, 13, 15, 18, 20, 21, 24, 27, 30, 39
	Jumlah		40	40

Berdasarkan Tabel 1 dan Tabel 2, dapat dilihat bahwa jumlah soal yang valid pada kedua sekolah melebihi separuh dari total jumlah soal UAS yang diujikan. Dengan demikian, sebagian besar soal sudah dapat menjalankan fungsi dan pengukurannya dengan baik. Hal ini dikarenakan persyaratan tes yang paling utama adalah valid sehingga soal dapat digunakan untuk mengukur hasil belajar siswa dengan tepat (Nuswawati, *et al.*, 2010). Ketika suatu tes tidak memiliki validitas yang baik, ada dua hal yang mungkin akan terjadi, yaitu: (1) siswa tidak dapat menunjukkan kemampuan mereka yang sebenarnya dikarenakan tidak ada soal yang menguji kemampuan tersebut dan (2) adanya pertanyaan-pertanyaan yang tidak berhubungan sehingga menyebabkan siswa tidak dapat menjawab dengan benar (Jandaghi dan Fatemeh, 2008). Namun demikian, ternyata validitas butir saja belum bisa digunakan untuk menentukan kualitas suatu soal. Oleh karenanya perlu diadakan analisis lain seperti analisis indeks kesukaran, jenjang soal, efektifitas *distractor* dan reliabilitas instrumen.

Hasil uji analisis indeks kesukaran soal UAS di SMA Negeri 1 Ambarawa menunjukkan bahwa dari 40 butir soal yang diujikan, terdapat 25 butir soal yang tergolong mudah, yakni soal nomor 2, 3, 4, 6, 12, 13, 14, 15, 16, 17, 18, 19, 22, 23, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 40. Adapun 12 soal tergolong sedang, yakni soal nomor 1, 5, 7, 8, 10, 11, 20, 24, 25, 26, 27, 39 dan 3 soal lainnya tergolong sukar, yakni soal nomor 9, 21, 35. Sedangkan hasil analisis indeks kesukaran soal UAS di MA Negeri Suruh menunjukkan hasil yang berbeda. Terdapat 30 butir soal mudah, dengan nomor soal 1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 35, 37, 38, 40 dan 10 butir soal sedang, dengan nomor soal 2, 4, 11, 14, 23, 24, 33, 34, 36, 39. Berdasarkan hasil tersebut, terlihat bahwa sebagian besar soal UAS di kedua sekolah tergolong mudah. Sementara soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar (Suharsimi, 2009). Adapun hasil uji analisis indeks kesukaran butir soal UAS SMA Negeri 1 Ambarawa dan MA Negeri Suruh dapat dilihat pada Gambar 1.

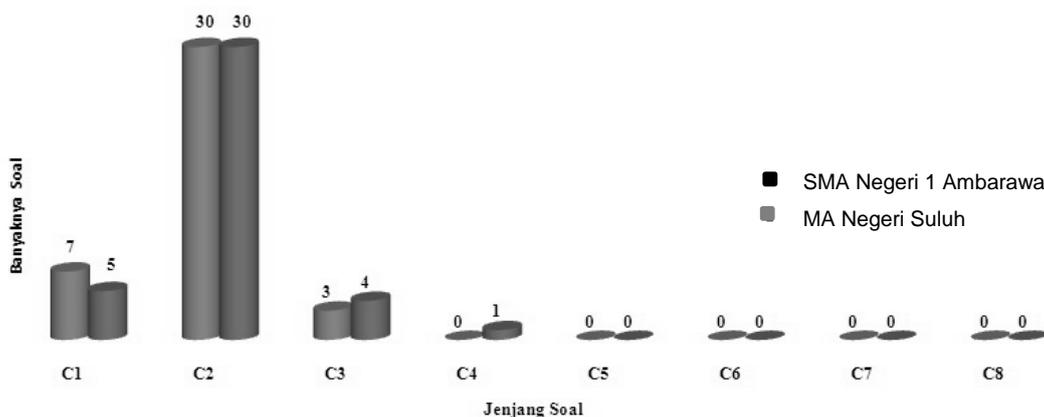


Gambar 1. Hasil analisis indeks kesukaran

Hasil analisis proporsi jenjang soal UAS di SMA Negeri 1 Ambarawa menunjukkan bahwa ada 7 butir soal yang termasuk jenjang C1/pengetahuan, yakni soal nomor 1, 11, 21, 22, 30, 31, 32. Sementara sebagian besar soal, yakni 30 dari 40 butir soal UAS sekolah tersebut ternyata termasuk dalam jenjang C2/ pemahaman, yakni soal nomor 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23, 24, 26, 28, 29, 33, 34, 35, 36, 37, 38, 39, 40. Adapun 3 soal lainnya yang termasuk dalam jenjang C3/ aplikasi adalah soal nomor 4, 25, 27.

Hasil analisis proporsi jenjang soal terhadap 40 butir soal UAS MA Negeri Suruh menunjukkan bahwa terdapat

jenjang soal C1/pengetahuan sampai dengan C4/ analisis di dalamnya. Terdapat 5 butir soal yang termasuk dalam jenjang C1/pengetahuan, yakni soal nomor 1, 2, 25, 29, 32. Selain itu, 30 butir soal termasuk dalam jenjang C2/ pengetahuan, yakni soal nomor 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 30, 31, 33, 34, 35, 36, 37, 40. Sedangkan untuk jenjang C3/ aplikasi ada 4 butir soal, yakni soal nomor 4, 28, 38, 39 dan untuk jenjang C4/ analisis hanya ada 1 butir soal, yakni soal nomor 27. Adapun hasil analisis proporsi jenjang soal UAS SMA Negeri 1 Ambarawa dan MA Negeri Suruh dapat dilihat pada Gambar 2.



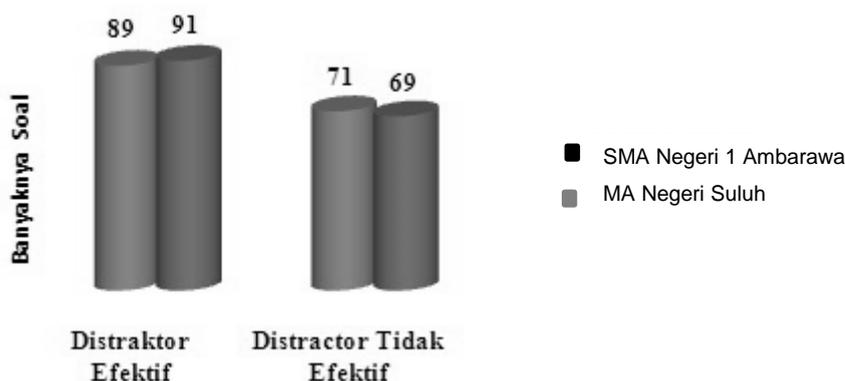
Gambar 2. Hasil analisis proporsi jenjang soal

Berdasarkan proporsi jenjang soal pada Gambar 2, terlihat bahwa sebagian besar soal UAS termasuk dalam jenjang C2. Hasil analisis tersebut sesuai dengan hasil yang memberikan informasi bahwa semua siswa Indonesia ternyata hanya menguasai pelajaran sampai level 3 saja. Padahal, jenjang C2 bukan termasuk dalam kategori soal berpikir tingkat tinggi dan tidak melatih keterampilan berpikir siswa (Lissa, *et al.*, 2012). Yang termasuk dalam soal berpikir tingkat tinggi adalah soal-soal analisis, sintesis, dan evaluasi. Hal ini berarti bahwa dari seluruh soal UAS kimia kelas XI semester gasal tahun ajaran 2013/2014 di SMA Negeri 1 Ambarawa dan MA Negeri Suruh hanya ada 1 soal saja yang termasuk dalam soal berpikir tingkat tinggi, yaitu soal dengan jenjang C4/ analisis.

Hasil analisis *distractor* soal UAS kimia kelas XI di SMA Negeri 1 Ambarawa menunjukkan bahwa 89 dari 160 butir *distractor* termasuk dalam kriteria efektif,

sedangkan 71 butir yang lainnya dinyatakan tidak efektif. Oleh karena jumlah keseluruhan *testee* di kelas XI IPA 4 SMA Negeri 1 Ambarawa ada 40 orang siswa, maka *distractor* dinyatakan efektif atau dapat menjalankan fungsinya dengan baik apabila sekurang-kurangnya dipilih oleh 2 orang *testee*. *Distractor* efektif apabila sekurang-kurangnya dipilih oleh 5% dari seluruh peserta.

Hasil analisis *distractor* soal UAS kimia kelas XI di MA Negeri Suruh menunjukkan bahwa sebanyak 91 dari 160 butir *distractor* termasuk dalam kriteria efektif, sedangkan 69 butir yang lainnya dinyatakan tidak efektif. Dalam hal ini, *testee* pada kelas XI IPA 1 MA Negeri Suruh berjumlah 21 orang siswa. Oleh karenanya, *distractor* sudah bisa disebut efektif apabila sekurang-kurangnya dipilih oleh 1 orang *testee*. Adapun hasil uji efektifitas *distractor* soal UAS SMA Negeri 1 Ambarawa dan MA Negeri Suruh dapat dilihat pada Gambar 3.



Gambar 3. Hasil uji efektifitas *distractor*

Berdasarkan hasil uji efektifitas *distractor* pada Gambar 3, sebanyak 89 butir *distractor* soal UAS kimia kelas XI SMA

Negeri 1 Ambarawa dan 91 butir *distractor* soal UAS kimia di MA Negeri Suruh yang termasuk kategori *distractor* efektif telah

Ana Yustika, dkk, Uji Kriteria Instrumen

menjalankan fungsinya dengan baik, yaitu dapat mengecoh *testee*, khususnya yang berkemampuan rendah sehingga memilih *distractor* sebagai jawaban yang benar (Widyantoro, *et al.*, 2009).

Berdasarkan hasil uji reliabilitas yang telah dilaksanakan, diperoleh koefisien reliabilitas instrumen untuk SMA Negeri 1 Ambarawa dan MA Negeri Suruh berturut-turut sebesar 0,70 dan 0,81. Ini artinya bahwa paket soal UAS tersebut reliabel, karena menurut Suparji (2010), tes dikatakan reliabel jika koefisien reliabilitasnya lebih dari 0,70. Dengan demikian, kapanpun soal UAS SMA Negeri 1 Ambarawa dan MA Negeri Suruh tersebut digunakan akan memberikan hasil ukur yang sama (Djanuarsih, 2012).

Uji kriteria instrumen penilaian hasil belajar kimia yang dilaksanakan di SMA Negeri 1 Ambarawa dan MA Negeri Suruh ini didukung pula dengan adanya 2 macam data angket, yaitu data angket respon guru terhadap prinsip penilaian hasil belajar siswa dan angket keterbacaan soal. Pengadaan angket ini bertujuan untuk mengetahui adakah pengaruh respon guru dan tingkat keterbacaan soal terkait hasil uji kriteria instrumen penilaian hasil belajar yang telah dilaksanakan. Berdasarkan hasil analisis, ternyata dalam pembuatan soal UAS, guru pengampu mata pelajaran kimia kelas XI IPA 4 SMA di Negeri 1 Ambarawa dan kelas XI IPA 1 di MA Negeri Suruh tidak melakukan uji coba soal, analisis butir soal, ataupun perhitungan reliabilitas instrumen tes. Guru enggan melaksanakan analisis butir soal tes karena: (1) tidak ada tuntutan

sehingga dapat diabaikan; (2) tidak memiliki waktu luang karena jam mengajarnya penuh sebab menganalisis butir soal tes memerlukan waktu ekstra cukup banyak; (3) belum mengetahui manfaat dari menganalisis butir soal-soal tes, karena soal yang hampir sama untuk siswa tahun lalu hasilnya bagus tetapi untuk siswa tahun sekarang kurang memuaskan; (4) tidak mengetahui cara-cara menganalisis butir soal-soal tes; dan (5) menganggap bahwa soal yang telah dianalisis dan digunakan kembali untuk tes tidak bermanfaat, karena sudah menjadi kebiasaan bahwa soal tes dibagikan kepada siswa dan siswa belajar dari soal tersebut (Widodo, 2010).

Angket tingkat keterbacaan soal yang diisi oleh seluruh siswa kelas XI IPA 4 SMA Negeri 1 Ambarawa dan kelas XI IPA 1 MA Negeri Suruh memberikan hasil yang baik. Dalam hal ini, tingkat keterbacaan soal UAS di kedua sekolah tergolong bagus, karena bahasa yang digunakan jelas dan mudah dipahami. Akan tetapi, penggunaan bentuk negatif (seperti kecuali dan bukan) pada kedua paket soal UAS sama-sama tidak ditandai dengan cetak miring. Sebagaimana yang disampaikan oleh Suharsimi (2009), salah satu kriteria soal bentuk pilihan ganda yang baik adalah penggunaan tanda cetak miring pada bentuk-bentuk negatif tersebut sehingga tidak membingungkan siswa. Berdasarkan hal-hal tersebut, dapat dikatakan bahwa tingkat keterbacaan soal tidak terlalu berpengaruh terhadap hasil analisis butir soal yang dilaksanakan.

SIMPULAN

Berdasarkan hasil uji kriteria instrumen penilaian hasil belajar di SMA Negeri 1 Ambarawa, diketahui bahwa sebanyak 27 butir soal tergolong valid dengan jenjang soal C1/ pengetahuan sampai dengan C3/ aplikasi. Dari total 160 butir *distractor* yang digunakan, 89 butir diantaranya tergolong efektif. Instrumen tergolong reliabel karena memiliki koefisien reliabilitas 0,70. Sedangkan untuk hasil uji kriteria instrumen penilaian hasil belajar di MA Negeri Suruh, sebanyak 28 butir soal tergolong valid dengan jenjang soal C1/ pengetahuan sampai dengan C4/ analisis. *Distractor* yang tergolong efektif berjumlah 91 butir. Instrumen tergolong reliabel karena memiliki koefisien reliabilitas 0,81.

DAFTAR PUSTAKA

- Arifianti, R., 2013, Analisis Kualitas Produk Sepatu Tomkins, *Jurnal Dinamika Manajemen*, Vol 1, no 4, Hal:46-58.
- Djanuarsih, E., 2012, Validitas dan Reliabilitas Butir Soal, *Jurnal Dinas Pendidikan Kota Surabaya*, Vol 1, No 1, Hal: 1-12.
- Jandaghi, G. dan Fatemeh, S., 2008, Rate of Validity, Reliability, and Difficulty Indices for Teacher-Designed Exam Questions in First Year High School, *International Journal of Human Sciences*, Vol 2, No 5, Hal:1-6.
- Khan, M.U.Z. dan Badr, M.A., 2011, Evaluation of Modified Essay Questions and Multiple Choice Questions as a Tool For Assessing the Cognitive Skills of Undergraduate Medical Students, *International Journal of Health Sciences, Qassim University*, Vol 1, No 5, Hal:39-43.
- Lissa, Andreas, P.B.P., dan Dyah, R.I., 2012, Pengembangan Instrumen Penilaian Keterampilan Berpikir Tingkat Tinggi Materi Sistem Respirasi dan Ekskresi, *Jurnal Ilmu Kependidikan*, Vol 1, No 41, Hal:27-32.
- Nuswowati, M., Binadja, A., Soeprodjo, dan Khida, E.N.I., 2010, Pengaruh Validitas dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Kimia terhadap Pencapaian Kompetensi, *Jurnal Inovasi Pendidikan Kimia*, Vol 1, No 4, Hal:566-573.
- Patel, K.A. dan Neeraj, R. M., 2013, Itemized Analysis of Questions of Multiple Choice Question (MCQ) Exam, *International Journal of Scientific Research*, Vol 2, No 2, Hal:279-280.
- Purwati, A. dan Irni, W., 2009, Studi Kualitas Soal Ujian Akhir Sekolah Berstandar Nasional Mata Pelajaran Matematika Provinsi DKI Jakarta wilayah Jakarta Timur tahun pelajaran 2007/2008, *Jurnal Lingkar Mutu Pendidikan*, Vol 2, No 2, Hal:128-136.
- Ratnaningsih, D.J., 2011, Analisis Butir Soal Pilihan Ganda Ujian Akhir Semester Mahasiswa di Universitas Terbuka dengan Pendekatan Teori Tes Klasik, *Jurnal Pendidikan Terbuka dan Jarak Jauh*, Vol 2, No 12, Hal:92-99.
- Suharsimi, A., 2009, *Dasar-dasar Evaluasi Pendidikan*, Edisi revisi, cetakan 9, Jakarta: Bumi Aksara.
- Suparji, 2010, Kualitas Butir Soal Buatan Guru-Guru SMP Mata Pelajaran Matematika dan IPA di Kabupaten Sumenep, *Jurnal Pendidikan Dasar*, Vol 1, No 11, Hal:48-52.
- Suryana, 2010, *Metodologi Penelitian Model Praktis Penelitian Kuantitatif dan Kualitatif*, Bandung: Universitas Pendidikan Indonesia.
- Widodo, 2010, Analisis Butir Soal Tes, *Jurnal Pendidikan Penabur*, Vol 9, No 14, Hal:58-67.
- Widyantoro, D., Boenasir, dan Karsono, 2009, Pengembangan Soal Tes Pilihan Ganda Kompetensi Sistem

Ana Yustika, dkk, Uji Kriteria Instrumen

Starter dan Pengisian Program Keahlian Teknik Mekanik Otomotif Kelas XII, *Jurnal PTM*, Vol 1, No 9, Hal:14-21.

Zaman, A., Asaf, N., Fayyaz, A. F., Muhammad, A. D., dan Alamgir, 2010,

1339

Analysis of Multiple Choice Items and the Effect of Items' Sequencing on Difficulty Level in the Test of Mathematics, *European Journal of Social Sciences*, Vol 1, No 17, Hal:61-67.