



## KUALITAS SOAL BUATAN GURU DAN DINAS PENDIDIKAN BIDANG STUDI KIMIA

YP Nugraheni , AT Widodo, W Sugiyo

Jurusan Kimia FMIPA Universitas Negeri Semarang

Gedung D6 Kampus Sekaran Gunungpati Telp. 8508112 Semarang 50229

### Info Artikel

#### Sejarah Artikel:

Diterima 10 Januari 2013

Disetujui 10 Februari 2013

Dipublikasikan April 2013

#### Keywords:

analysis of a test

quality of test

final semester test

### Abstrak

Berlakunya kurikulum KTSP mengakibatkan adanya variasi dalam pembuatan soal tes. Soal pada beberapa kabupaten dibuat oleh guru, sedangkan beberapa kabupaten yang lain disusun oleh Dinas Pendidikan (MGMP). Dalam fungsinya sebagai alat untuk mengetahui ketercapaian kompetensi siswa, suatu soal harus dapat dikatakan sebagai soal yang baik dengan memenuhi kriteria validitas, reliabilitas, tingkat kesukaran, daya beda, objektifitas, praktikabilitas, dan keekonomisan. Penelitian ini bertujuan untuk mengetahui kualitas soal ulangan akhir semester ganjil buatan guru dan dinas pendidikan bidang studi kimia. Metode yang digunakan dalam penelitian ini adalah metode dokumentasi. Sampel jawaban siswa dianalisis untuk mengetahui kandungan kualitas soal yang dikerjakan siswa tersebut. Hasil penelitian menunjukkan bahwa soal buatan guru dan dinas pendidikan belum sepenuhnya memenuhi kriteria parameter kualitas soal . Kedua jenis soal ini belum dapat dikatakan sebagai soal yang baik untuk dapat digunakan sebagai soal ulangan umum semester. Berdasarkan hasil tersebut disimpulkan bahwa soal Ulangan Akhir Semester Ganjil buatan Guru di Kabupaten Semarang dan Dinas Pendidikan Kabupaten Pemalang untuk bidang studi kimia tahun 2011/2012 belum memenuhi kriteria soal yang baik.

### Abstract

Curriculum enactment of SBC lead to variations in the manufacturing of test questions. Problem in some districts created by the teacher, while some other districts compiled by the Department of Education (MGMP). In its function as a tool to determine student achievement of competence, the test must be regarded as a good problem to meet the criteria of validity, reliability, level of difficulty, different power, objectivity, praktikabilitas, and economics. This study aims to determine the quality of the final semester of odd-test questions and the teacher made the education field of study of chemistry. The method used in this study is the method of documentation. Samples were analyzed to determine the students' responses about the quality of the content of student work. The results showed that the question of teacher-made and not yet fully meet the education criteria about the quality parameters. Both of these types of problems can not be said as a test of good to be used as a general test questions semester. Based on these results concluded that the question of artificial Odd Semester End of Deuteronomy in the District of Semarang and Teacher Education Office Pemalang district to subject areas of chemistry in 2011/2012 do not meet the criteria of a good question.

© 2013 Universitas Negeri Semarang

Alamat korespondensi:

Email: ana\_gembil@yahoo.com

HP:085729112045

ISSN NO 2252-6609

## Pendahuluan

Kegiatan pembelajaran tidak bisa lepas dari aktivitas evaluasi sebagai tindakan untuk mengetahui tercapainya tujuan pembelajaran. Salah satu jenis evaluasi yang cukup sering didengar adalah evaluasi hasil belajar untuk mengukur domain kognitif siswa. Tes Hasil Belajar yang paling umum dan dominan digunakan adalah Ulangan Akhir Semester dan Ujian Nasional.

Tes sebagai salah satu alat evaluasi memiliki kriteria untuk dapat dinyatakan sebagai tes yang baik. Kriteria tes tersebut adalah tes harus memiliki validitas isi, validitas butir, reliabilitas 0,7; daya beda 0,20; tingkat kesukaran 0,30–0,70; objektif; praktis; dan ekonomis.

Tujuan melakukan penilaian atau evaluasi dalam kegiatan pembelajaran adalah untuk mendapatkan informasi yang akurat mengenai tingkat penguasaan siswa terhadap kompetensidasar. Sehingga dapat dilakukan tindak lanjut khususnya umpan balik mengenai kelemahan kelemahan siswa untuk diperbaiki bagi kemajuan siswa. Penilaian atau evaluasi sangat penting bagi siswa, karena sebagai alat untuk mengetahui tingkat keberhasilan belajar atau penguasaan kompetensi. Kemungkinan bisa terjadi perbedaan antara hasil penilaian yang diharapkan dengan kenyataannya. Guru telah melaksanakan pembelajaran dengan baik dan melihat respon siswa telah menguasai kompetensi dasar. Guru dan siswa sama-sama berharap hasil penilaian baik, tetapi ternyata tidak baik. Masalah tersebut mungkin bukan cara mengajar guru yang salah, bukan cara belajar siswa yang salah, kemungkinan alat evaluasinya atau soal-soal tes yang bermasalah (Widodo, 2010). Untuk itu diperlukan analisis terhadap soal yang digunakan dalam evaluasi untuk mengetahui kualitas soal tersebut.

Setelah berlakunya KTSP, suatu sekolah memiliki otonomi penuh terhadap pembelajaran peserta didiknya, begitupun pada evaluasinya. Pada beberapa daerah, tes yang digunakan juga diserahkan kepada sekolah masing-masing. Artinya tes tersebut dibuat oleh guru dari sekolah tersebut dan diujikan di sekolah itu saja, sehingga tiap sekolah di daerah tersebut akan memiliki tes yang berbeda. Beberapa daerah yang lain menggunakan tes yang dibuat MGMP. Pada daerah tersebut, tes yang digunakan masing-masing sekolah akan sama.

Masalah dalam penelitian ini adalah fakta baik pada tes yang dibuat oleh guru mata

pelajaran maupun MGMP ternyata tidak melalui uji coba soal. Tidak adanya uji coba soal ini tentunya membuat soal tes tersebut tidak diketahui sudah atau belum memenuhi kriteria tes yang baik. Selain itu, penelitian yang dilakukan oleh Khifda Efti Nelly Ifada (2009) dalam skripsinya yang berjudul 'Analisis Validitas dan Reliabilitas Butir Soal Ujian Akhir Semester Bidang Study Kimia Kelas X SMA N 1 Pati Tahun Ajaran 2007/2008 Terhadap Pencapaian Kompetensi', menyatakan bahwa tidak semua soal yang dibuat oleh guru mata pelajaran memenuhi persyaratan sebagai soal yang baik. Bahkan untuk soal Ujian Nasional, tidak semua soalnya memenuhi kriteria kualitas soal yang baik (Ridhadani, 2007).

Hal inilah yang mendasari dilakukannya penelitian terhadap soal yang dibuat oleh guru dan dinas pendidikan untuk mengetahui kualitas soal yang digunakan dalam Ujian Akhir Semester pada mata pelajaran kimia.

Soal tes yang digunakan dalam penelitian ini adalah soal Ulangan Umum Semester Ganjil kelas X dan XI tahun Pelajaran 2011/2012. Penelitian ini bertujuan untuk mengetahui kualitas soal yang dibuat oleh Guru dan Dinas Pendidikan (MGMP) pada Bidang Studi Kimia

## Metode Penelitian

Metode penelitian adalah cara yang digunakan dalam mengumpulkan data penelitian (Suharsimi, 2010). Penelitian ini termasuk penelitian deskriptif dengan metode dokumentasi. Metode dokumentasi yaitu suatu metode penelitian yang mencari data mengenai hal-hal atau variabel yang berupa isi catatan, transkrip, buku, surat kabar, majalah, prasasti, notulen rapat, legger, agenda, dan sebagainya (Suharsimi: 2010). Sampel dalam penelitian ini adalah soal Ulangan Umum Semester Ganjil dari SMA N 1 Bergas dan SMA N 2 Ungaran untuk soal buatan guru, sedangkan soal buatan Dinas Pendidikan (MGMP) diambil dari soal Ulangan Umum Semester Ganjil pada SMA N 1 Randudongkal dan SMA N 1 Bantarbolang. Sekolah tersebut diambil sebagai sampel secara acak. Masing-masing sampel mewakili populasi yang berbeda. Sampel soal dari guru mewakili populasi dari soal Ulangan Umum Semester Ganjil mata pelajaran kimia untuk kabupaten Ungaran, sedangkan sampel soal dari Dinas Pendidikan mewakili populasi pada soal

Ulangan Umum Semester Ganjil mata pelajaran Kimia untuk kabupaten Pemalang.

Variabel dalam penelitian ini adalah kualitas soal dengan parameter validitas, reliabilitas, tingkat kesukaran, daya beda, objektifitas, praktikabilitas, dan keekonomisan.

Analisis dalam penelitian ini menggunakan rumus-rumus stastistika dan beberapa analisis diskriptif menggunakan lembar pengamatan. Analisis yang dilakukan meliputi analisis validitas isi, analisis validitas butir, analisis reliabilitas, analisis tingkat kesukaran, analisis daya beda, analisis objektifitas, analisis praktikabilitas, dan analisis keekonomisan.

Analisis terhadap validitas isi dilakukan dengan mencocokkan materi dengan indikator dan tujuan di silabus. Analisis terhadap validitas butir dilakukan dengan menggunakan rumus korelasi point biserial yang dilanjutkan dengan t hitung untuk soal pilihan ganda dan rumus korelasi produk moment untuk soal uraian. Sedangkan analisis terhadap reliabilitas menggunakan rumus KR-21 untuk soal pilihan ganda dan koefisien alpha-kronbach untuk soal uraian (Suharsimi, 2009). Tingkat kesukaran yang merupakan angka yang menunjukkan

proporsi peserta didik yang menjawab betul suatu soal, dicari dengan menghitung perbandingan jumlah siswa yang menjawab benar terhadap jumlah total siswa. Daya beda ditentukan dengan menghitung selisih antara proporsi jawaban benar kelompok atas dengan proporsi jawaban benar kelompok bawah. Sedangkan pada analisis terhadap objektifitas, praktikabilitas, dan keekonomisan digunakan analisis deskriptif dengan lembar pengamatan.

#### Hasil dan Pembahasan

Penelitian ini melakukan analisis terhadap soal ulangan umum semester ganjil kelas X dan XI. Analisis yang dilakukan meliputi analisis terhadap validitas isi, validitas butir, reliabilitas, tingkat kesukaran, daya beda, objektifitas, praktikabilitas, dan keekonomisan.

Validitas isi menyatakan kesesuaian butir soal terhadap indikator dan tujuan yang harus diwakili oleh butir soal tersebut. Analisis terhadap validitas isi dilakukan dengan mencocokkan tujuan dalam butir soal dengan tujuan dan indikator yang tercantum dalam silabus. Hasil analisis validitas isi dicantumkan dalam Tabel 1.

Tabel 1. Hasil analisis validitas isi

KRITERIA	SOAL			
	UNG 1	UNG 2	PMLG 1	PMLG 2
<b>A. Soal kelas X</b>				
Ketercakupan indikator	66,66%	55,55%	92,59%	92,59%
Ketercakupan tujuan	46,55%	37,93%	54,45%	54,45%
Soal yang tidak mencakup indikator maupun tujuan	-	-	-	-
<b>B. Soal Kelas XI</b>				
Ketercakupan indikator	73,08%	26,92%	92,30%	92,30%
Ketercakupan tujuan	45,31%	18,75%	56,25%	56,25%
Soal yang tidak mencakup indikator maupun tujuan	-	1 soal	-	-

Berdasarkan hasil analisis tersebut terlihat bahwa hanya soal dari MGMP yang memenuhi kriteria validitas isi yaitu mencakup sekurangnya 75% dari indikator.

Analisis validitas butir diperlukan untuk mengetahui jumlah soal valid dalam satu paket soal. Hasil analisis validitas butir diperlihatkan dalam Tabel 2 dan 3.

Tabel 2 dan 3 memperlihatkan tak ada satupun soal yang memenuhi kriteria untuk dikatakan merupakan soal yang valid secara

keseluruhan (memenuhi 75% butir soal valid).

Reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang sudah ditetapkan. Hasil analisis reliabilitas soal diperlihatkan pada Tabel 4

Selain validitas dan reliabilitas, Soal yang baik harus memiliki perbandingan tingkat kesukaran yang proporsional. Hasil analisis terhadap tingkat kesukaran dicantumkan dalam Tabel 5

Tabel 2. Hasil analisis validitas butir soal kelas X

KRITERIA	SOAL			
	UNG 1	UNG 2	PMLNG 1	PMLNG 2
Jumlah soal total	40	25	40	40
Jumlah soal valid	20	13	19	13
Jumlah soal tidak valid	20	12	21	27
Presentase soal valid (%)	50,0	52,0	47,5	32,5

Tabel 3. Hasil analisis validitas butir soal kelas XI

KRITERIA	SOAL			
	UNG 1	UNG 2	PMLNG 1	PMLNG 2
Jumlah total soal	40	29	40	40
Jumlah soal valid	20	15	23	21
Jumlah soal tidak valid	20	14	17	19
Presentase soal valid (%)	50,0	51,7	57,5	52,5

Tabel 4. Hasil analisis reliabilitas

NO	JENIS SOAL	TIPE SOAL	KRITERIA		KELAS XI	
			Skor	Kriteria	Skor	Kriteria
1	Ung 1	Pilihan Ganda	0,67	Reliabel	0,66	Reliabel
2	Ung 2	Pilihan Ganda	0,46	Tdk reliabel	0,60	Tdk Reliabel
		Uraian	0,39	Reliabel	0,45	Reliabel
3	Pmlg 1	Pilihan Ganda	0,41	Tidak reliabel	0,69	Reliabel
4	Pmlg 2	Pilihan Ganda	-0,11	Tidak reliabel	0,48	Tidak reliabel

Tabel 5. Hasil analisis tingkat kesukaran

KRITERIA	UNG 1		UNG 2		PMLNG 1		PMLNG 2	
	jumlah	Persen (%)						
<b>A. Soal Kelas X</b>								
Soal mudah	6	15,0	4	16,0	9	22,5	6	15,0
Soal sedang	23	57,5	17	68,0	20	50,0	16	40,0
Soal sulit	10	25,0	4	16,0	11	27,5	18	45,0
Soal sangat sulit	1	2,5	0	0,0	0	0,0	0	0,0
<i>Jumlah</i>	<i>40</i>	<i>100</i>	<i>25</i>	<i>100</i>	<i>40</i>	<i>100</i>	<i>40</i>	<i>100</i>
<b>B. Soal Kelas XI</b>								
Soal mudah	11	27,5	3	10,3	5	12,5	5	12,5
Soal sedang	19	47,5	22	75,8	25	62,5	27	67,5
Soal sulit	9	22,5	4	13,7	10	25	8	20,0
Soal sangat sulit	1	2,5	0	0,0	0	0,0	0	0,0
<i>Jumlah</i>	<i>40</i>	<i>100</i>	<i>29</i>	<i>100</i>	<i>40</i>	<i>100</i>	<i>40</i>	<i>100%</i>

Berdasarkan hasil pada Tabel 5, soal yang memiliki perbandingan soal mudah:sedang:sulit yang paling proporsional adalah soal kelas XI dari kode Ung 1.

Analisis yang berikutnya adalah analisis daya beda butir soal yang diperlukan untuk menyatakan kemampuan soal membedakan peserta didik pintar dan kurang pintar. Soal yang baik mampu memberikan hasil yang berbeda pada peserta didik pintar dan kurang pintar. Hasil analisis daya beda terangkum dalam Tabel 6

Selain berkaitan dengan butir soal, penelitian ini juga melakukan analisis terhadap objektifitas, praktikabilitas, dan keekonomisan soal. Hasil analisis objektifitas diperlihatkan pada Tabel 7

Kepraktisan, atau kemudahan soal termasuk salah satu syarat soal yang baik. Kemudahan ini meliputi dalam persiapan, penggunaan, pengolahan dan penafsiran, serta pengadministrasianya (Arifin, 2011). Hasil analisis praktikabilitas soal dalam penelitian ini dicantumkan dalam Tabel 8.

Tabel 6. Hasil analisis daya beda

KRITERIA DAYA BEDA	SOAL							
	UNG 1		UNG 2		PMLNG 1		PMLNG 2	
	Jumlah	Persen (%)						
<b>A. Soal Kelas X</b>								
Negatif	3	7,5	0	0,0	2	5,0	7	17,5
Jelek	11	27,5	10	40,0	23	57,5	20	50,0
Cukup	15	37,5	9	36,0	8	20,0	10	25,0
Baik	10	25,0	6	24,0	7	17,5	3	7,5
Sangat Baik	1	2,5	0	0,0	0	0,0	0	0,0
<i>Jumlah</i>	<i>40</i>	<i>100</i>	<i>25</i>	<i>100</i>	<i>40</i>	<i>100</i>	<i>40</i>	<i>100</i>
Soal memenuhi kriteria	26	65,0	15	60,0	15	37,5	13	32,5
<b>B. Soal Kelas XI</b>								
Negatif	4	10,0	1	3,4	4	10,0	6	15,0
Jelek	14	35,0	10	34,5	13	32,5	11	27,5
Cukup	9	22,5	7	24,1	13	32,5	13	32,5
Baik	13	32,5	11	37,9	10	25,0	9	22,5
Sangat Baik	0	0,0	0	0,0	0	0,0	1	2,5
<i>Jumlah</i>	<i>40</i>	<i>100</i>	<i>29</i>	<i>100</i>	<i>40</i>	<i>100</i>	<i>40</i>	<i>100</i>
Soal memenuhi kriteria	22	55,0	18	62,1	23	57,5	23	57,5

Tabel 7. Hasil analisis objektifitas

NO	SEKOLAH	KRITERIA	
		KELAS X	KELAS XI
1	UNG 1	Objektif	Objektif
2	UNG 2	Objektif	Objektif
3	PMLG 1	Objektif	Objektif
4	PMLG 2	Objektif	Objektif

Tabel 8. Hasil analisis praktikabilitas

NO	SEKOLAH	KELAS X		KELAS XI	
		Nilai	Kriteria	Nilai	Kriteria
1	UNG 1	65	Praktis	65	Praktis
2	UNG 2	50	Cukup praktis	55	Cukup praktis
3	PMLG 1	60	Cukup praktis	65	Praktis
4	PMLG 2	60	Cukup praktis	65	Praktis

Selain itu, syarat yang tak kalah penting adalah keekonomisan. Suatu tes dikatakan ekonomis bila dalam pelaksanaan tes tersebut tidak membutuhkan biaya yang mahal, tenaga yang banyak, dan waktu yang lama (Suharsimi,

2009). Hasil analisis keekonomisan soal dalam penelitian ini dirangkum dalam Tabel 9.

Hasil analisis secara keseluruhan dapat diringkas dalam Tabel 10.

Tabel 9. Hasil analisis keekonomisan

NO	SEKOLAH	KELAS X		KELAS XI	
		Nilai	Kriteria	Nilai	Kriteria
1	UNG 1	75,0	Ekonomis	68,75	Ekonomis
2	UNG 2	87,5	Sangat ekonomis	81,25	Sangat ekonomis
3	PMLG 1	75,0	Ekonomis	75,0	Ekonomis
4	PMLG 2	75,0	Ekonomis	75,0	Ekonomis

Tabel 10. Hasil analisis keterpenuhan kriteria

ASPEK	UNG 1		UNG 2		PMLG 1		PMLG 2	
	M	TM	M	TM	M	TM	M	TM
<b>A. Soal Kelas X</b>								
Validitas isi		✓		✓	✓		✓	
Validitas butir		✓		✓		✓		✓
Reliabilitas soal pilihan ganda	✓			✓		✓		✓
Reliabilitas soal uraian	-	-	✓		-	-	-	-
Tingkat kesukaran		✓		✓		✓		✓
Daya beda		✓		✓		✓		✓
Objektifitas	✓		✓		✓		✓	
Praktikabilitas	✓			✓		✓		✓
Ekonomis	✓		✓		✓		✓	
<b>B. Soal Kelas XI</b>								
Validitas isi		✓		✓	✓		✓	
Validitas butir		✓		✓		✓		✓
Reliabilitas soal pilihan ganda	✓			✓	✓			✓
Reliabilitas soal uraian	-	-	✓		-	-	-	-
Tingkat kesukaran		✓	✓			✓		✓
Daya beda		✓		✓		✓		✓
Objektifitas	✓		✓		✓		✓	
Praktikabilitas	✓			✓	✓		✓	
Ekonomis	✓		✓		✓		✓	

Penelitian ini dilakukan untuk mengetahui kualitas soal Ulangan Akhir Semester ganjil kelas X dan XI IPA mata pelajaran kimia yang dibuat oleh guru mata pelajaran dan dinas pendidikan.

Berdasarkan hasil analisis validitas isi seperti yang terlihat dalam Tabel 1, soal yang memenuhi validitas isi hanyalah soal dari MGMP. Perbedaan hasil validitas isi ini dikarenakan pada soal yang disusun oleh tim MGMP, materi yang masuk dalam soal didiskusikan terlebih dahulu. Selain itu, karena soal digunakan untuk satu daerah maka tidak ada penekanan materi dalam penyusunan soalnya. Penyusunan soal buatan guru mata pelajaran cenderung berdasarkan apa yang selama ini diajarkan kepada peserta didik. Soal juga hanya digunakan oleh satu sekolah saja, sehingga guru bebas menyusun soal dengan materi yang ada tanpa proporsi yang jelas. Hal ini tentu saja membuat kualitas soal dari segi isinya kurang bagus, karena soal tersebut tidak mampu mengukur semua kompetensi yang seharusnya dikuasai peserta didik. Dengan soal yang ketercakupan tujuan dan indikator yang cukup rendah, kemampuan peserta didik dalam memahami pelajaran pun tak dapat terukur dengan pasti.

Hasil analisis juga menunjukkan bahwa walaupun sama-sama disusun oleh guru mata pelajaran, soal dengan kode sekolah Ung 1 dan Ung 2 memiliki hasil ketercakupan yang

berbeda. Selain itu, soal dari sekolah dengan kode Ung 2 untuk kelas XI dan kelas X memiliki ketercakupan yang jauh berbeda. Perbedaan ini disebabkan penyusun antara soal kelas X dan XI adalah guru yang berbeda dan tidak melalui diskusi antar guru terlebih dahulu. Tidak adanya komunikasi antar guru penyusun soal dalam merumuskan soal yang ada tentu cukup merugikan peserta didik, terutama ketika ada kelas atau peserta didik yang tidak diajar oleh guru yang membuat soal untuk jenjang kelasnya.

Analisis yang berikutnya adalah analisis validitas butir soal untuk mengetahui kevalidan suatu butir soal sebagai alat ukur. Seperti yang terlihat pada Tabel 4 mengenai hasil analisis validitas butir soal, validitas butir soal tertinggi dimiliki oleh soal dengan kode Ung 2, kemudian Ung 1, Pmlg 1 dan Pmlg 2 untuk kelas X. Sementara kelas XI urutannya adalah Ung 2, Pmlg 1, Ung 1 dan Pmlg 2. Data ini menunjukkan bahwa soal yang disusun oleh guru mata pelajaran (Ung 1 dan 2) memiliki validitas yang lebih baik dibanding soal yang disusun tim MGMP. Namun pada dasarnya kedua soal tersebut masih belum valid butir soalnya. Lima puluh persen dari jumlah soal merupakan soal yang tidak valid baik itu dari soal buatan guru mata pelajaran maupun soal buatan MGMP. Sementara, soal dapat dikatakan valid jika minimal 75% dari butir soal tersebut valid. Hal ini tentunya membuat kedua

jenis soal ini belum mampu mengukur kompetensi yang seharusnya dimiliki peserta didik secara maksimal.

Data yang diperoleh juga menunjukkan bahwa untuk soal yang sama, ketika diujikan di sekolah yang berbeda memiliki harga validitas yang berbeda pula. Dalam penelitian ini validitas untuk kode Pmlg 1 lebih tinggi dibanding validitas pada kode Pmlg 2. Seharusnya jumlah butir soal valid pada kedua sekolah sama atau paling tidak mendekati sama. Tidak adanya uji coba soal yang seharusnya dilakukan sebelum soal digunakan menjadi faktor penting dalam ketidaksamaan jumlah butir soal valid ini. Soal yang tidak diuji coba terlebih dahulu tidak dapat memberikan gambaran menyeluruh tentang kualitasnya terutama dalam validitas butir. Hal ini seharusnya menjadi sebuah pertimbangan penting bagi penyusunan soal terutama oleh tim MGMP untuk dapat menyusun soal yang valid untuk semua sekolah yang menggunakan soal tersebut..

Hasil uji terhadap reliabilitas soal juga menunjukkan hasil yang bervariasi. Soal yang disusun oleh guru mata pelajaran mendapatkan hasil yang reliabel kecuali untuk soal pilihan ganda untuk soal kelas X dan XI pada sekolah dengan kode Ung 2. Sementara untuk soal yang disusun tim MGMP, untuk soal kelas X tidak reliabel pada kedua sekolah sampel sementara soal kelas XI mendapatkan hasil yang berbeda pada kedua sekolah tersebut.

Suatu tes dikatakan reliabel apabila beberapa kali pengujian menunjukkan hasil yang relatif sama (Suharsimi, 2009). Salah satu faktor yang mempengaruhi reliabilitas adalah panjang tes yang dalam penelitian ini adalah jumlah soal (Ikhwanuddin, 2011). Instrumen yang memiliki lebih banyak soal cenderung memiliki harga reliabilitas yang lebih tinggi. Hal ini yang mengakibatkan soal pilihan ganda untuk sekolah dengan kode Ung 2 tidak reliabel. Soal pilihan ganda dengan kode Ung 2 berjumlah 20 soal untuk kelas X dan 25 soal untuk kelas XI. Jumlah ini relatif sedikit bila dibanding jumlah soal kode yang lain yang mencapai 40 soal.

Faktor lain yang mempengaruhi reliabilitas soal adalah kesesuaian tingkat kesukaran (Ikhwanuddin, 2011). Kesesuaian tingkat kesukaran inilah yang mengakibatkan pada soal kelas XI dari MGMP (Pmlg 1 dan Pmlg 2) terdapat perbedaan hasil reliabilitas . Persebaran tingkat kesukaran soal pada kedua

sekolah tersebut tidak sama. Akibatnya soal yang menurut sekolah dengan kode Pmlg 2 sukar belum tentu sulit bagi peserta didik di sekolah dengan kode Pmlg 1. Hal yang dapat dilakukan untuk menaikkan harga reliabilitas pada soal ini adalah dengan membuat soal tersebut berada pada kurva normal tingkat kesukaran untuk sekolah dengan kode Pmlg 1 maupun Pmlg 2.

Hasil reliabilitas yang baik ditunjukkan oleh soal buatan guru dengan kode sekolah Ung 1. Sementara reliabilitas terendah ditunjukkan oleh soal MGMP kelas X yang tidak reliabel.

Analisis tingkat kesukaran pada masing-masing soal juga menunjukkan hasil yang berbeda -beda. Untuk soal kelas X, proporsi tingkat kesukaran yang paling tidak proporsional adalah di sekolah dengan kode Pmlg 2. Sementara soal dengan proporsi yang paling seimbang adalah Ung 2. Pada soal kelas XI, proporsi tingkat kesukaran paling buruk dimiliki soal dengan kode Ung 1. Sementara proporsi paling seimbang adalah Ung 2. Dari hasil ini terlihat bahwa soal yang disusun oleh guru mata pelajaran memiliki proporsi tingkat kesukaran yang lebih seimbang. Hal ini dikarenakan untuk soal ini, pada saat penyusunannya guru benar- benar paham kemampuan peserta didik yang akan mengerjakan soal. Dengan begitu, guru dapat memperkirakan proporsi soal yang seimbang antara soal mudah, sedang dan sulit bagi peserta didiknya.

Pada soal yang disusun tim MGMP terdapat perbedaan hasil analisis tingkat kesukaran untuk 2 sekolah yang berbeda. Hal ini kembali lagi dikarenakan perbedaan kualitas dari kedua sekolah yang menggunakan soal tersebut. Soal yang sulit untuk peserta didik pada sekolah dengan kode Pmlg 2 bisa saja menjadi soal yang sedang bagi peserta didik pada sekolah dengan kode Pmlg 1. Hal ini perlu menjadi perhatian bagi tim MGMP untuk dapat membuat soal dengan taraf kesukaran yang hampir sama bagi sekolah penggunanya.

Hampir senada dengan tingkat kesukaran, hasil analisis daya beda untuk soal kelas X menunjukkan bahwa soal buatan guru memiliki daya beda yang lebih tinggi. Hal ini dikarenakan guru lebih paham pada kondisi dan kemampuan peserta didik peserta evaluasi. Akibatnya, guru dapat lebih memperkirakan bentuk soal yang mampu dijawab dengan baik oleh peserta didik yang pintar dan dijawab kurang tepat oleh peserta didik yang kurang

pintar.

Hasil yang berbeda diperlihatkan analisis daya beda soal kelas XI. Hasil analisis menunjukkan bahwa keempat soal memiliki persentase pemenuhan daya beda yang hampir sama. Sedikit sekali perbedaan antara soal buatan guru dan buatan dinas pendidikan. Fakta ini tentu saja adalah sebuah hal yang baik. Kenyataan ini memberikan masukan bahwa kedua soal ini sama-sama memerlukan revisi yang sejenis untuk dapat memenuhi kelayakan daya beda soal.

Walaupun hasil analisis daya beda menunjukkan perbedaan yang cukup besar, namun kedua jenis soal ini secara umum belum memenuhi kriteria daya beda yang baik. Suatu soal dikatakan memenuhi kriteria daya beda jika 75% dari butir soal tersebut memiliki daya beda

0,2. Hasil tertinggi daya beda 0,2 pada analisis ini hanyalah 65%, yang menegaskan bahwa semua soal dalam analisis ini belum memenuhi kriteria daya beda soal.

Keterpenuhan syarat yang sama dimiliki setiap soal untuk kriteria objektifitas. Walaupun soal dengan kode Ung 2 sedikit berbeda karena memiliki soal uraian namun soal ini juga merupakan soal yang objektif sama seperti soal yang lain. Objektifitas soal berkaitan dengan ada tidaknya penilaian pribadi guru dalam memberikan nilai kepada peserta didik. Mayoritas soal yang digunakan dalam evaluasi memang memenuhi syarat objektifitas terutama soal yang berbentuk pilihan ganda. Begitupun untuk soal uraian. Faktor subjektif dalam penilaian biasanya memang tidak muncul dari penilaian pada ulangan umum saja melainkan pemberian nilai akhir atau raport. Sehingga baik itu soal yang disusun guru mata pelajaran maupun soal yang disusun MGMP, semuanya memenuhi syarat objektivitas.

Analisis terhadap praktikabilitas soal menunjukkan hasil yang relatif sama untuk semua soal. Skor praktikabilitas terendah dimiliki oleh soal dengan kode Ung 2. Soal ini memiliki praktikabilitas terendah karena tersusun dari soal pilihan ganda dan uraian sementara soal lain hanya pilihan ganda saja. Keberadaan soal uraian inilah yang mengurangi kepraktisan suatu soal. Soal uraian memerlukan jawaban yang lebih panjang dan jelas, waktu koreksi yang lebih lama dan panduan penilaian yang lebih kompleks. Hal inilah yang mengakibatkan soal uraian menjadi tidak

sepraktis soal pilihan ganda.

Sementara untuk analisis terhadap keekonomisan soal, hampir diperoleh nilai yang sama untuk semua jenis soal. Soal yang paling ekonomis adalah soal dengan kode Ung 2 dan paling mahal adalah soal kelas XI kode Ung 1, Soal dengan kode Ung 2 memiliki harga keekonomisan tertinggi karena jumlah soalnya paling sedikit sehingga kertas yang digunakan tidak banyak.

Hasil analisis menunjukkan bahwa tidak ada soal yang lebih baik pada semua kriteria. Pada beberapa kriteria terlihat soal buatan guru mata pelajaran lebih baik dari soal buatan tim MGMP. Namun, pada kriteria yang lain terlihat sebaliknya. Hal ini menunjukkan bahwa soal buatan guru mata pelajaran dan dinas pendidikan masih harus banyak dibenahi lagi untuk dapat disebut sebagai soal yang baik.

Hasil analisis ini dapat digunakan sebagai dasar revisi bagi guru mata pelajaran maupun dinas pendidikan dalam melakukan revisi terhadap soal-soal yang digunakan dalam setiap evaluasi

## Simpulan

Berdasarkan hasil penelitian, maka dapat diambil simpulan bahwa soal Ulangan Umum Semester Ganjil buatan guru dan Dinas Pendidikan belum memenuhi persyaratan sebagai soal yang baik dan layak digunakan.

## Daftar Pustaka

- Arifin, Z. 2011. Evaluasi Pembelajaran. Bandung: PT Remaja Rosdakarya
- Ifada, KFN. 2009. Analisis validitas dan reliabilitas butir soal ujian akhir semester bidang study kimia kelas X SMA N 1 Pati tahun ajaran 2007/2008. Skripsi. Semarang: FMIPA UNNES
- Ikhwanuddin, M. 2011. Makalah reliabilitas. Diunduh di mikhwanuddin.wordpress.com tanggal 30 Juni 2012
- Ridhahani. 2007. Analisis butir soal bahasa Indonesia dalam ujian akhir nasional (UAN) pada madrasah aliyah negeri di kalimantan selatan. Ittihad Jurnal Kopertis Wilayah XI Kalimantan, 5(7) :1-15
- Suharsimi, A. 2009. Dasar-dasar evaluasi pendidikan. Jakarta: Bumi Aksara.
- Suharsimi, A. 2010. Prosedur penelitian suatu pendekatan praktik. Jakarta: Rineka Cipta.
- Widodo. 2010. Analisis butir soal. Jurnal pendidikan penabur. 9(14): 58-67