

**KARAKTERISTIK SECARA KUALITATIF DAN KUANTITATIF
SOAL ULANGAN AKHIR SEMESTER GENAP BAHASA INDONESIA
KELAS XII SMA NEGERI DI KABUPATEN PURBALINGGA
TAHUN PELAJARAN 2010/2011**

SKRIPSI

Diajukan kepada Fakultas Bahasa dan Seni
Universitas Negeri Yogyakarta
untuk Memenuhi Sebagian Persyaratan
guna Memperoleh Gelar Sarjana Pendidikan



oleh

Rokhyati

NIM 07201241022

**PROGRAM STUDI PENDIDIKAN BAHASA DAN SASTRA INDONESIA
FAKULTAS BAHASA DAN SENI
UNIVERSITAS NEGERI YOGYAKARTA
2011**

PERSETUJUAN

Skripsi yang berjudul *Karakteristik Secara Kualitatif dan Kuantitatif Soal Ulangan Akhir Semester Genap Bahasa Indonesia Kelas XII SMA Negeri di Kabupaten Purbalingga Tahun Pelajaran 2010/2011* ini telah disetujui oleh pembimbing untuk diujikan.



Yogyakarta, 13 Juli 2011

Pembimbing I,

Prof. Dr. Pujiati Suyata

NIP 19420806 197803 2 001

Yogyakarta, 14 Juli 2011

Pembimbing II,

Dwi Hanti Rahayu, S. Pd.

NIP 19720229 200012 2 001

PENGESAHAN

Skripsi yang berjudul *Karakteristik Secara Kualitatif dan Kuantitatif Soal Ulangan Akhir Semester Genap Bahasa Indonesia Kelas XII SMA Negeri di Kabupaten Purbalingga Tahun Pelajaran 2010/2011* ini telah dipertahankan di depan Dewan Penguji pada 25 Juli 2011 dan dinyatakan lulus.

DEWAN PENGUJI

Nama	Jabatan	Tandatangan	Tanggal
Esti Swatika Sari, S.Pd., M.Hum.	Ketua Penguji		8 Agustus 2011
Dwi Hanti Rahayu, S.Pd.	Sekretaris Penguji		10 Agustus 2011
Dr. Anwar Efendi, M.Si.	Penguji I		8 Agustus 2011
Prof. Dr. Pujiati Suyata, M.Pd.	Penguji II		9 Agustus 2011

Yogyakarta, 10 Agustus 2011

Fakultas Bahasa dan Seni

Universitas Negeri Yogyakarta



Prof. Dr. Zamzani, M.Pd.

NIP 19550505 198011 1 001

PERNYATAAN

Yang bertanda tangan di bawah ini, saya

Nama : **Rokhyati**

NIM : 07201241022

Program Studi : Pendidikan Bahasa dan Sastra Indonesia

Fakultas : Bahasa dan Seni Universitas Negeri Yogyakarta

menyatakan bahwa karya ilmiah ini adalah hasil pekerjaan saya sendiri. Sepanjang pengetahuan saya, karya ilmiah ini tidak berisi materi yang ditulis oleh orang lain, kecuali bagian-bagian tertentu yang saya ambil sebagai acuan dengan mengikuti tata cara dan etika penulisan karya ilmiah yang lazim.

Apabila ternyata terbukti bahwa pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab saya.

Yogyakarta,

Penulis,

A handwritten signature in black ink, appearing to read 'Rokhyati', is shown within a light blue rectangular border.

Rokhyati

MOTTO

*Dia yang menciptakan segala sesuatu, lalu Dia menetapkan atasnya takdir
(ketetapan) yang sesempurna-sempurnanya. (Q.S. 25:2).*

Stay Hungry. Stay Foolish. (Steve Jobs)

*Anggaplah semua masalah adalah pelangi yang mewarnai hidup ini yang sesaat
kan memudar dan hilang. (Fajar Prianto)*

*Manusia diberi akal, pikiran, dan perasaan untuk memilih apa yang terbaik bagi
dirinya. (Pendapat pribadi)*

PERSEMBAHAN

Kupersembahkan karya sederhana ini kepada:

Allah S.W.T. dan Rasul Muhammad S.A.W. Alhamdulillahrabbi'l'amin....

Ibunda dan ayahanda tersayang, terima kasih atas doa dan pengertian yang mendalam, maafkan jika saya sering tak sejalan dengan keinginan kalian,

Kak Indarti dan keluarga, Kak Srinani dan keluarga, serta Dik Ningrum Irasari yang selalu memberikan senyum dan motivasinya saat saya menghadapi kesulitan dan kejenuhan.

Mas Undiono, terima kasih atas perhatian yang tiada henti, pengorbanan, dorongan, dan curahan kasih sayang sehingga saya tidak pernah putus asa menyelesaikan studi untuk menggapai cita-cita bersama.

Keluarga besar yang selalu menanyakan kapan saya selesai studi.

Teman-teman JPBSI khususnya kelas AB 2007, Kos Karangmalang A28a, KKN SMPN 15 Yk, facebook, dan LBPP LIA CV Class, yang selalu memberikan bantuan, solusi, dan komentar membangun dalam rangka perbaikan diri.

**KARAKTERISTIK SOAL ULANGAN AKHIR SEMESTER GENAP
MATA PELAJARAN BAHASA INDONESIA KELAS XII
SMA NEGERI DI KABUPATEN PURBALINGGA
TAHUN PELAJARAN 2010/2011**

oleh
ROKHYATI
07201241022

ABSTRAK

Penelitian ini bertujuan untuk mengetahui kualitas soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011 ditinjau dari karakteristik soal secara kualitatif dan karakteristik soal secara kuantitatif.

Sesuai dengan tujuan penelitian, jenis penelitian yang digunakan adalah penelitian evaluatif. Sampel penelitian ini adalah paket soal yang terdiri dari kisi-kisi soal, lembar jawaban siswa sejumlah 2147 lembar, dan kunci jawaban soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011. Data diperoleh dengan teknik dokumentasi. Data yang terkumpul dianalisis secara kualitatif dan kuantitatif. Analisis kualitatif dengan teknik *expert judgement* yang bertujuan untuk mengetahui validitas isi soal. Analisis kuantitatif dilakukan dengan bantuan komputer menggunakan program *MicroCat Iteman* versi 3.00 untuk mengetahui reliabilitas soal, indeks kesulitan, indeks daya beda, dan keefektifan distraktor.

Analisis data yang telah dilakukan diperoleh hasil (1) karakteristik soal secara kualitatif (validitas isi soal) tergolong sangat baik karena dari hasil telaah menunjukkan kesesuaian telaah soal sebesar 99,30%, (2) karakteristik soal secara kuantitatif; (a) reliabilitas soal tergolong kurang baik, koefisien *Alpha Cronbach* sebesar 0,646, dengan SEM sebesar 2,523; (b) analisis karakteristik butir soal (i) indeks kesulitan tiap-tiap butir soal tergolong kurang baik, hanya 30% butir soal yang baik, tetapi indeks rata-rata butir soal tergolong baik dengan *mean p* sebesar 0,736; (ii) daya beda soal diabaikan karena menggunakan kurikulum KTSP. Setiap butir soal dinyatakan baik asal indeks daya beda tidak bernilai negatif; (iii) keefektifan distraktor tergolong kurang baik, hanya 46,5% distraktor yang dapat berfungsi dengan baik.

Dengan demikian, secara umum karakteristik soal ini kurang baik karena hanya sebesar 30% butir soal yang diterima, sehingga perlu ada upaya perbaikan dan penyempurnaan.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah Tuhan Yang Maha Pemurah lagi Maha Penyayang. Berkat rahmat, hidayah, dan inayah-Nya penulis dapat menyelesaikan skripsi untuk memenuhi sebagian persyaratan guna memperoleh gelar sarjana.

Penulisan skripsi ini dapat terselesaikan karena bantuan dan bimbingan dari berbagai pihak. Untuk itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Zamzani, M. Pd. selaku Dekan Fakultas Bahasa dan Seni Universitas Negeri Yogyakarta beserta staf yang telah membantu dalam kelancaran penelitian ini.
2. Ibu Pangesti Wiedarti, Ph. D. selaku ketua jurusan Pendidikan Bahasa dan Sastra Indonesia yang telah memberikan ijin dalam penulisan skripsi ini.
3. Ibu Prof. Dr. Pujiati Suyata, M. Pd. selaku pembimbing I yang telah memberikan arahan dan masukan guna menyempurnakan proses penulisan skripsi ini.
4. Ibu Dwi Hanti Rahayu, S. Pd. selaku pembimbing II yang telah memberikan bimbingan dan arahan selama penulisan skripsi ini.
5. Ibu Siti Maslakhah, M. Hum. selaku pembimbing akademik atas bimbingan dan bantuannya.
6. Bapak Kastam Syamsi, M. Ed., Ibu Dra. Henifah, dan Bapak Kaozal Dadi Legawan, S. Pd. selaku *expert judgement* yang telah meluangkan waktu untuk membantu penelitian ini.
7. Semua pihak di jurusan Pendidikan Bahasa dan Sastra Indonesia atas bantuan dan penyediaan fasilitas dalam penulisan skripsi ini.
8. Kepala Dinas Pendidikan Purbalingga beserta staf karyawan atas bantuannya dalam pengambilan data penelitian.
9. Kepala sekolah SMA Negeri se-Purbalingga beserta guru dan staf karyawan atas ijin dan bantuannya dalam pengambilan data penelitian.

10. Kedua orang tuaku tercinta atas kasih sayang, doa, dan dukungan yang selalu diberikan sebagai penyemangat dalam menyelesaikan skripsi ini.
11. Teman-teman jurusan Pendidikan Bahasa dan Sastra Indonesia atas bantuan dan semangatnya.
12. Dewan penguji dan semua pihak yang telah membantu terselesaikannya skripsi ini dan tidak bisa saya sebutkan satu per satu

Penulis menyadari sepenuhnya bahwa laporan skripsi ini masih banyak kekurangannya. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan demi perbaikan laporan skripsi ini.

Yogyakarta,

A handwritten signature in black ink, appearing to be 'Rachmat', is written on a light blue rectangular background.

Penulis

DAFTAR ISI

Halaman

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN	iv
HALAMAN MOTTO	v
HALAMAN PERSEMBAHAN	vi
ABSTRAK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR BAGAN	xiv
DAFTAR LAMPIRAN	xv

BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Identifikasi Masalah	5
C. Batasan Masalah	6
D. Rumusan Masalah	7
E. Tujuan Penelitian	7
F. Manfaat Penelitian	8
G. Pembatasan Istilah	8
BAB II KAJIAN TEORI	9
A. Deskripsi Teoretis	10
1. Konsep Pengukuran, Penilaian, dan Evaluasi	10
2. Tujuan dan Fungsi Evaluasi	13
3. Kualitas Soal	14
4. Bentuk Soal	22
5. Analisis Butir Soal (Item Analysis)	26
6. Karakteristik Penilaian Pembelajaran Bahasa Indonesia	33

B. Kajian Penelitian yang Relevan	35
C. Kerangka Berpikir	37
D. Pertanyaan Penelitian	38
BAB III METODE PENELITIAN	39
A. Jenis Penelitian	39
B. Sampel Penelitian	39
C. Teknik Pengumpulan Data	40
D. Teknik Analisis Data	40
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	47
A. Deskripsi Hasil Penelitian	47
1. Hasil Analisis Kualitatif	47
2. Hasil Analisis Kuantitatif	52
a. Reliabilitas Soal	52
b. Analisis Butir Soal	52
1) Indeks Kesukaran	52
2) Indeks Daya Beda	53
3) Efektivitas Distraktor	54
3. Pemaknaan Hasil Analisis Butir Soal	56
4. Kunci Jawaban yang Perlu Dicek	57
B. Pembahasan Hasil Penelitian	58
1. Pembahasan Hasil Analisis Kualitatif	58
2. Pembahasan Hasil Analisis Kuantitatif	65
a. Pembahasan Pengujian Reliabilitas Soal	65
b. Pembahasan Analisis Butir Soal	65
1) Pembahasan Indeks Kesukaran	65
2) Pembahasan Indeks Daya Beda	66
3) Pembahasan Efektivitas Distraktor	67
3. Pembahasan Hasil Analisis Butir Soal	68
4. Pembahasan Kunci Jawaban yang Perlu Dicek	85
5. Relevansi Hasil Analisis Kualitatif dengan Kuantitatif	85

BAB V PENUTUP	88
A. Simpulan	88
B. Implikasi	89
C. Keterbatasan Penelitian	90
D. Saran	90
DAFTAR PUSTAKA	91
LAMPIRAN	93

DAFTAR TABEL

	Halaman
Tabel 1 : Jumlah Data Lembar Jawaban Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Purbalingga.....	40
Tabel 2 : Kriteria Kualitas Validitas Isi Soal	43
Tabel 3 : Klasifikasi Indeks Kesukaran	44
Tabel 4 : Klasifikasi Efektivitas Distraktor	46
Tabel 5 : Telaah Validitas Isi Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Kabupaten Purbalingga oleh Tiga Orang Penelaah.....	48
Tabel 6 : Rangkuman Hasil Telaah Validitas Isi Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Kabupaten Purbalingga oleh Tiga Orang Penelaah.....	51
Tabel 7 : Hasil Pengujian Tingkat Kesukaran Butir Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Kabupaten Purbalingga.....	53
Tabel 8 : Hasil Pengujian Daya beda Butir Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Purbalingga.....	53
Tabel 9 : Efektivitas Distraktor Butir Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Purbalingga.....	55

DAFTAR BAGAN

	Halaman
Bagan 1 : Alur Analisis Soal	38

DAFTAR LAMPIRAN

	Halaman
1. Daftar Nama Sekolah SMA di Kabupaten Purbalingga	91
2. Kisi-kisi Soal Ulangan Akhir Semester (UAS) Genap SMA kelas XII tahun pelajaran 2010/2011 di Purbalingga	92
3. Lembar Soal Ulangan Akhir Semester (UAS) Genap SMA kelas XII tahun pelajaran 2010/2011 di Purbalingga	97
4. Kunci Jawaban Soal Ulangan Akhir Semester (UAS) Genap SMA kelas XII tahun pelajaran 2010/2011 di Purbalingga	107
5. Lembar Jawaban Soal Ulangan Akhir Semester (UAS) Genap SMA kelas XII tahun pelajaran 2010/2011 di Purbalingga	108
6. Lembar Telaah Butir Soal	109
7. Hasil Analisis <i>Iteman</i>	127
8. Rangkuman Hasil Analisis Butir Soal yang Meliputi Indeks Kesulitan, Indeks Daya Beda, dan Efektivitas Distraktor	147
9. Surat Ijin Penelitian	154

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Keberhasilan suatu proses pembelajaran tidaklah terlepas dari sebuah proses penilaian hasil belajar yang handal. Pelaksanaan ujian yang baik sangat tergantung pada ketersediaan alat ukur yang berkualitas. Evaluasi hasil belajar merupakan bagian dari program pendidikan untuk mengetahui keberhasilan belajar siswa. Evaluasi hasil belajar juga dapat memberikan gambaran tentang kualitas pembelajaran dalam rangka perbaikan kualitas pendidikan secara umum.

Penilaian hasil belajar oleh pendidik dilakukan secara berkesinambungan. Penilaian ini bertujuan untuk memantau proses dan kemajuan belajar siswa serta meningkatkan efektivitas kegiatan pembelajaran. Hal ini dapat digunakan sebagai dasar pengambilan keputusan yang tepat dalam rangka peningkatan mutu pembelajaran. Penilaian ini diharapkan dapat berkontribusi terhadap keberhasilan evaluasi akhir yang dilakukan oleh pemerintah melalui ujian nasional sebagai instrumen pengendalian mutu pendidikan secara nasional.

Peraturan Menteri Pendidikan Nasional Nomor 20 tahun 2007 tentang standar penilaian pendidikan menyebutkan bahwa ulangan akhir semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan semua kompetensi dasar pada semester tersebut. Ulangan akhir semester di Sekolah Menengah Atas (SMA) dilaksanakan secara berkelanjutan dari kelas X sampai kelas XII.

Salah satu mata pelajaran yang diujikan dalam ujian nasional adalah Bahasa Indonesia. Bahasa adalah kemampuan dasar yang harus dimiliki oleh seorang manusia. Dengan bahasa, manusia dapat berkomunikasi untuk mengutarakan gagasannya. Pembelajaran Bahasa Indonesia diarahkan dalam rangka meningkatkan kemampuan peserta didik untuk berkomunikasi dalam bahasa Indonesia dengan baik dan benar (Suryaman, 2009: 5). Evaluasi proses belajar mengajar mata pelajaran Bahasa Indonesia merupakan sebuah keniscayaan untuk melihat sejauh mana kompetensi yang dimiliki siswa dalam mata pelajaran tersebut.

Kelas XII merupakan tahun terakhir proses pembelajaran di SMA dan akan dilakukan evaluasi hasil belajar oleh pemerintah melalui Ujian Nasional (UN). Hasil Ujian Nasional SMA/MA pada tahun 2010 menunjukkan kelulusan siswa hanya 89,61%, padahal di tahun 2009 kelulusan siswa mencapai 95,06% (www.bataviase.co.id). Merosotnya tingkat kelulusan siswa patut menjadi perhatian. Nilai mata pelajaran Bahasa Indonesia menempati posisi terendah di antara mata pelajaran yang diujikan di semua program. Rendahnya nilai bahasa Indonesia ini perlu diteliti agar dapat diketahui faktor penyebabnya.

KTSP pada pembelajaran Bahasa Indonesia menunjukkan kegiatan berbahasa mencakup aspek mendengar, berbicara, membaca, dan menulis. Tes pada mata pelajaran Bahasa Indonesia harus terdiri dari butir soal yang handal sehingga dapat mengukur kemampuan siswa yang terdiri dari keempat aspek tersebut.

Di Kabupaten Purbalingga, ulangan akhir semester kelas XII dilakukan secara bersama atas kesepakatan Kelompok Kerja Kepala Sekolah (KKKS). Perangkat tes dibuat oleh beberapa guru mata pelajaran sejenis melalui Musyawarah Guru Mata Pelajaran (MGMP). MGMP SMA di Kabupaten Purbalingga menyepakati bahwa

Ulangan Akhir Semester (UAS) Genap tahun pelajaran 2010/2011 juga berfungsi sebagai uji coba (*try out*) ujian nasional. Dengan demikian, Ulangan Akhir Semester (UAS) Genap tahun pelajaran 2010/2011 di Kabupaten Purbalingga berfungsi ganda, yaitu sebagai ulangan akhir semester yang merupakan bagian dari nilai rapor dan sebagai uji coba ujian nasional dalam rangka persiapan siswa menempuh ujian nasional.

Penilaian hasil belajar oleh pemerintah dilakukan dalam bentuk ujian nasional. Ujian nasional bertujuan menilai pencapaian kompetensi lulusan secara nasional. Hal ini sesuai dengan Peraturan Pemerintah Nomor 19 tahun 2005 tentang Standar Nasional Pendidikan pasal 66 ayat (1) yang menyebutkan bahwa penilaian hasil belajar oleh pemerintah bertujuan menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu dalam kelompok mata pelajaran ilmu pengetahuan teknologi dan dilakukan dalam bentuk ujian nasional.

Memperhatikan standar kelulusan yang ditentukan Badan Standardisasi Nasional Pendidikan (BSNP) pada Ujian Nasional (UN) setiap tahunnya terus meningkat, hal ini dirasa cukup berat. Oleh karena itu, sekolah selaku pengemban tugas mencerdaskan kehidupan bangsa melakukan berbagai upaya agar siswanya mampu mencapai kriteria kelulusan tersebut. Uji coba (*Try Out*) merupakan upaya yang dilakukan sekolah untuk mempersiapkan siswa dalam menempuh ujian nasional. Setiap tahunnya, uji coba ujian nasional telah dilakukan semua sekolah baik tingkat SMP/ sederajat maupun SMA/ sederajat.

Ulangan akhir semester yang juga sebagai uji coba ujian nasional mata pelajaran Bahasa Indonesia SMA di Purbalingga menggunakan teknik tes. Penggunaan instrumen tes yang kurang cermat akan memberikan hasil yang kurang

cermat pula. Teknik penyusunan tes yang tepat dapat diharapkan memberikan landasan yang kokoh untuk melakukan evaluasi yang tepat.

Perangkat soal yang digunakan harus memenuhi syarat-syarat pembuatan soal yang baik. Perangkat soal yang berfungsi mengukur kemampuan siswa terlebih dahulu dilakukan tahap penelaahan pada aspek materi, konstruksi, dan bahasanya. Soal dapat digunakan apabila soal tersebut memenuhi semua aspek telaah. Analisis empiris digunakan untuk mengetahui karakteristik setiap butir soal. Analisis empiris dilakukan setelah soal digunakan.

Hasil wawancara yang dilakukan pada bulan Februari 2011 kepada beberapa guru SMA di Kabupaten Purbalingga, diketahui bahwa guru belum secara maksimal mengembangkan kemampuannya dalam menyusun tes yang baik. Guru menyusun naskah soal mengambil dari buku pelajaran, buku bank soal, internet, dan soal-soal tahun sebelumnya. Setelah soal selesai dibuat, guru langsung menggunakannya untuk tes. Padahal, naskah soal untuk menguji kemampuan siswa tersebut karakteristiknya belum jelas karena belum dilakukan tahap analisis. Tahap analisis butir soal yang diabaikan dapat menyebabkan soal tes tidak mampu mengukur kemampuan siswa yang sesungguhnya.

Agar tes dapat memberikan gambaran sesungguhnya tentang pencapaian kompetensi yang diharapkan, perangkat tes harus memenuhi persyaratan alat ukur yang baik. Untuk mengetahui mutu seperangkat soal dalam berbagai aspek dan mengidentifikasi kekurangan-kekurangan dalam tes, maka perlu dilakukan analisis butir soal. Gambaran tentang baik tidaknya tes yang digunakan tercermin pada karakteristik perangkat tes itu sendiri, yaitu: validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas distraktor.

Uraian di atas memberikan gambaran betapa pentingnya usaha meningkatkan karakteristik soal tes melalui analisis butir soal, sehingga informasi yang didapatkan dari alat ukur yang baik betul-betul objektif dan tidak bias. Dengan demikian, karakteristik secara kualitatif dan kuantitatif soal ulangan akhir semester genap bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011 layak untuk dikaji dan diteliti.

B. Identifikasi Masalah

Identifikasi masalah berdasarkan latar belakang masalah sebagai berikut.

1. Validitas soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA di Kabupaten Purbalingga tahun pelajaran 2010/2011 belum diketahui.
2. Reliabilitas soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA di Kabupaten Purbalingga tahun pelajaran 2010/2011 belum diketahui.
3. Indeks kesulitan butir soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA di Kabupaten Purbalingga tahun pelajaran 2010/2011 belum diketahui.
4. Indeks daya beda butir soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA di Kabupaten Purbalingga tahun pelajaran 2010/2011 belum diketahui.
5. Keefektifan distraktor butir soal Ulangan Akhir Semester (UAS) Genap (UAS) mata pelajaran Bahasa Indonesia kelas XII SMA di Kabupaten Purbalingga tahun pelajaran 2010/2011 belum diketahui.

6. Kualitas soal Ulangan Akhir Semester (UAS) Genap (UAS) mata pelajaran Bahasa Indonesia kelas XII SMA di Kabupaten Purbalingga tahun pelajaran 2010/2011 belum diketahui.
7. Guru belum cukup baik dalam menyusun tes khususnya pada aspek perencanaan dan pengembangan.

C. Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut.

1. Karakteristik secara kualitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011.
2. Karakteristik secara kuantitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011. Secara rinci, analisis kuantitatif meliputi:
 - a. reliabilitas soal,
 - b. analisis butir soal, meliputi (1) indeks kesukaran butir soal, (2) indeks daya beda butir soal, dan (3) keefektifan distraktor.

D. Rumusan Masalah

Masalah-masalah yang diteliti dan dikaji dapat dirumuskan sebagai berikut.

1. Bagaimanakah karakteristik secara kualitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011?
2. Bagaimanakah karakteristik secara kuantitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011? Secara rinci, analisis kuantitatif meliputi:
 - a. reliabilitas soal,
 - b. analisis butir soal, meliputi (1) indeks kesukaran butir soal, (2) indeks daya beda butir soal, dan (3) keefektifan distraktor.

E. Tujuan Penelitian

Tujuan penelitian ini sebagai berikut.

1. Menjelaskan karakteristik secara kualitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011.
2. Menjelaskan karakteristik secara kuantitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011. Secara rinci, analisis kuantitatif meliputi:
 - a. reliabilitas soal,

- b. analisis butir soal, meliputi (1) indeks kesukaran butir soal, (2) indeks daya beda butir soal, dan (3) keefektifan distraktor.

F. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut.

1. Secara teoretis, hasil penelitian ini diharapkan dapat menambah khasanah ilmu pengetahuan di bidang evaluasi pembelajaran Bahasa Indonesia.
2. Secara praktis, hasil penelitian ini diharapkan dapat mengetahui kemampuan siswa SMA Negeri di Purbalingga dalam memahami dan menjawab soal Ulangan Akhir Semester (UAS) Genap sehingga dapat digunakan sebagai acuan ke arah perbaikan. Hasil penelitian ini juga dapat menjadi masukan kepada guru tentang prosedur penyusunan dan kriteria tes yang baik, serta alternatif metode analisis butir tes yang paling tepat dengan cara sederhana menggunakan teori tes klasik. Bagi sekolah, hasil penelitian ini dapat digunakan sebagai informasi menindaklanjuti pengambilan keputusan dalam penyusunan dan pengembangan tes, serta evaluasi pembelajaran.

G. Pembatasan Istilah

Penelitian ini berjudul *Karakteristik Secara Kualitatif dan Kuantitatif Soal Ulangan Akhir Semester Genap Bahasa Indonesia Kelas XII SMA Negeri di Kabupaten Purbalingga Tahun Pelajaran 2010/2011*. Pembatasan istilah pada penelitian ini adalah sebagai berikut.

1. Karakteristik Soal

Karakteristik soal adalah sifat khas, tetap, dan kekal yang dapat dijadikan ciri mengidentifikasi soal.

2. Analisis Kualitatif

Analisis kualitatif adalah analisis yang digunakan untuk mengetahui validitas isi soal.

3. Analisis Kuantitatif

Analisis kuantitatif adalah analisis yang dilakukan dengan cara klasik menggunakan program *MicroCat Iteman* yang mencakup informasi mengenai besarnya indeks reliabilitas dan analisis butir soal.

4. Ulangan Akhir Semester (UAS) Genap

Ulangan Akhir Semester (UAS) Genap adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester genap.

BAB II

KAJIAN TEORI

A. Deskripsi Teoretis

Deskripsi teoretis ini akan menguraikan tentang konsep pengukuran, penilaian, dan evaluasi; tujuan dan fungsi evaluasi; kualitas soal; bentuk soal; analisis butir soal (*item analysis*); dan karakteristik penilaian pembelajaran bahasa Indonesia.

1. Konsep Pengukuran, Penilaian, dan Evaluasi

Proses pengukuran, penilaian, evaluasi, dan pengujian merupakan suatu kegiatan atau proses yang bersifat hierarkis. Kegiatan dilakukan secara berurutan dan berjenjang, dimulai dari proses pengukuran kemudian penilaian dan terakhir evaluasi (Haryati, 2007: 14).

Menurut Arikunto (2006: 3), mengukur adalah membandingkan sesuatu dengan satu ukuran. Pengukuran bersifat kuantitatif. Djali dan Muljono (2008: 1) menyebutkan bahwa pengukuran yang dalam bahasa Inggris dikenal dengan istilah *measurement* merupakan suatu kegiatan yang dilakukan untuk mengukur dalam arti memberi angka terhadap sesuatu yang disebut objek pengukuran atau objek ukur.

Pendapat di atas menunjukkan bahwa inti dari pengukuran adalah suatu kegiatan memberikan angka. Angka di sini dimaknai sebagai skor yang merupakan langkah awal dalam proses evaluasi. Hal ini sejalan dengan pendapat Widyoko (2009: 2) yang menyebutkan bahwa esensi dari pengukuran adalah kuantifikasi atau penetapan angka tentang karakteristik atau keadaan individu menurut aturan-

aturan tertentu. Berbeda dengan pendapat Cangelosi (1995: 21) yang menjelaskan bahwa pengukuran adalah proses pengumpulan data melalui pengamatan empiris. Pendapat ini menyebutkan bahwa hasil pengukuran bukanlah berupa angka. Pengukuran hanyalah sebuah proses pengumpulan data tanpa tindak lanjut pemberian angka menurut aturan-aturan tertentu.

Menilai, menurut Arikunto (2006: 3) adalah mengambil suatu keputusan terhadap sesuatu dengan ukuran baik buruk. Penilaian bersifat kualitatif. Djali dan Muljono (2008: 2) menyebutkan bahwa penilaian yang dalam bahasa Inggris dikenal dengan istilah *assesment* berarti menilai sesuatu. Menilai itu sendiri berarti mengambil keputusan terhadap sesuatu yang mengacu pada ukuran tertentu, seperti menilai baik atau buruk, sehat atau sakit, pandai atau bodoh, tinggi atau rendah, dan sebagainya.

Pendapat di atas menunjukkan bahwa menilai adalah memberikan nilai terhadap sesuatu. Nilai dapat berupa angka dan non angka. Berbeda dengan pengukuran yang hasilnya berupa angka-angka, penilaian memberikan nilai berupa non angka, seperti baik atau buruk, sehat atau sakit, pandai atau bodoh, tinggi atau rendah, dan sebagainya. Seperti halnya dengan pendapat Cangelosi (1995: 21) yang menyebutkan bahwa penilaian adalah suatu keputusan tentang nilai. Ia menambahkan bahwa penilaian dipengaruhi oleh hasil pengukuran. Pendapat lain menyebutkan bahwa penilaian dapat diartikan sebagai kegiatan menafsirkan data hasil pengukuran berdasarkan kriteria maupun aturan-aturan tertentu (Widoyoko, 2009: 3). Jadi, hasil pengukuran sangat berperan penting dalam proses penilaian. Oleh karena itu, pengukuran harus dilakukan dengan tepat sehingga proses selanjutnya yaitu penilaian dapat dilakukan dengan tepat pula.

Mengadakan evaluasi meliputi dua langkah, yakni mengukur dan menilai (Arikunto, 2006: 3). Evaluasi dapat diartikan sebagai proses menilai sesuatu berdasarkan kriteria atau tujuan yang telah ditetapkan, yang selanjutnya diikuti dengan pengambilan keputusan atas objek yang dievaluasi (Djali dan Muljono, 2008: 1). Di sisi lain, Widoyoko (2009: 2) berpendapat bahwa evaluasi merupakan proses yang sistematis dan berkelanjutan untuk mengumpulkan, mendeskripsikan, menginterpretasikan dan menyajikan informasi tentang suatu program sehingga dapat digunakan sebagai dasar membuat keputusan, menyusun kebijakan maupun menyusun program selanjutnya. Sesuai dengan pengertian di atas maka setiap kegiatan evaluasi merupakan proses yang sengaja direncanakan untuk memperoleh informasi atau data. Data tersebut kemudian dianalisis untuk dijadikan dasar pengembangan suatu keputusan.

Pengertian tentang pengukuran, penilaian, dan evaluasi di atas menunjukkan bahwa pengukuran, penilaian, dan evaluasi merupakan tiga konsep yang berbeda. Namun, dalam praktiknya di dunia pendidikan, ketiga konsep tersebut sering dipraktikkan dalam satu rangkaian.

Menurut Arikunto (2006: 6), evaluasi pendidikan adalah kegiatan menilai yang terjadi dalam kegiatan pendidikan. Di bidang pendidikan, pengukuran memegang peranan yang sangat penting. Data hasil pengukuran dalam bidang pendidikan memiliki arti penting baik bagi sekolah atau lembaga pendidikan, guru, maupun bagi siswa dan orang tua siswa atau masyarakat (Djali dan Muljono, 2008: 5). Oleh karena itu, kegiatan pengukuran dan penilaian harus dilakukan dengan teliti dan tepat karena hasilnya memiliki arti penting dan berdampak luas.

2. Tujuan dan Fungsi Evaluasi

Menurut Daryanto (2008: 9), penilaian dalam dunia pendidikan, khususnya dunia persekolahan, mempunyai makna ditinjau dari berbagai segi.

a. Makna bagi siswa

Penilaian membuat siswa dapat mengetahui sejauh mana telah berhasil mengikuti pelajaran yang diberikan oleh guru. Hasilnya ada dua kemungkinan:

1) Memuaskan

Siswa akan mempunyai motivasi belajar lebih giat, agar mendapat hasil yang lebih memuaskan lagi. Namun sebaliknya, ada siswa yang sudah merasa puas dan kurang gigih belajar lagi.

2) Tidak memuaskan

Siswa akan berusaha agar lain kali tidak terulang lagi, lalu ia belajar giat. Namun sebaliknya, ada siswa menjadi putus asa dengan hasil kurang memuaskan yang telah diterimanya.

b. Makna bagi guru

1) Mengetahui siswa yang sudah berhak melanjutkan perjalanannya karena sudah berhasil menguasai bahan dan mengetahui siswa yang belum berhasil menguasai bahan. Guru dapat memusatkan perhatiannya pada siswa yang belum berhasil.

2) Mengetahui apakah materi yang diajarkan sudah tepat bagi siswa sehingga tidak perlu diadakan perubahan pengajaran di waktu yang akan datang.

3) Mengetahui apakah metode yang digunakan sudah tepat atau belum. Jika sebagian besar siswa memperoleh angka tidak baik pada penilaian yang diadakan, mungkin hal ini disebabkan oleh pendekatan atau metode yang kurang tepat. Maka guru harus mawas diri dan mencoba mencari metode lain dalam mengajar.

c. Makna bagi sekolah

1) Mengetahui apakah kondisi belajar di sekolah sudah sesuai dengan harapan atau belum. Hasil belajar merupakan cermin kualitas suatu sekolah.

2) Merupakan bahan pertimbangan bagi perencanaan sekolah di masa-masa yang akan datang.

3) Menjadi pedoman bagi sekolah apakah yang dilakukan oleh sekolah sudah memenuhi standar atau belum. Pemenuhan standar akan terlihat dari bagusnyanya angka-angka yang diperoleh siswa.

Purwanto (1992: 5) menyebutkan bahwa fungsi evaluasi dalam pendidikan dan pengajaran dapat dikelompokkan menjadi empat fungsi.

a. Mengetahui kemajuan dan perkembangan serta keberhasilan siswa setelah mengalami atau melakukan kegiatan belajar selama jangka waktu tertentu.

Hasil evaluasi selanjutnya dapat digunakan memperbaiki cara belajar siswa (fungsi formatif) dan mengisi rapor atau Surat Tanda Tamat Belajar yang berarti pula menentukan kenaikan kelas atau lulus-tidaknya seorang siswa (fungsi sumatif).

- b. Mengetahui tingkat keberhasilan program pengajaran. Fungsi ini berguna bagi supervisor untuk mengadakan perbaikan program beserta pelaksanaannya pada masa yang akan datang.
- c. Evaluasi berfungsi untuk kepentingan Bimbingan Konseling (BK).
- d. Evaluasi berfungsi untuk keperluan pengembangan dan perbaikan kurikulum sekolah yang bersangkutan. Meskipun pada umumnya di Indonesia kurikulum sekolah disusun secara nasional dan berlaku untuk semua sekolah yang sejenis dan setingkat, akan tetapi masukan dari para pelaksana kurikulum di lapangan berperan penting.

Pendapat kedua pakar di atas dapat disimpulkan bahwa evaluasi di bidang pendidikan sangat bermanfaat baik bagi siswa, guru, sekolah, maupun bagi supervisor. Kedua pendapat tersebut saling melengkapi dan menunjukkan betapa pentingnya evaluasi. Kegiatan evaluasi harus dilaksanakan dengan teliti karena hasilnya menyangkut orang banyak dan berdampak luas. Hasil evaluasi dapat memberikan informasi yang akurat dan objektif tentang suatu program, sehingga berguna dalam pengambilan keputusan atau kebijakan suatu program.

3. Kualitas Soal

Soal sebagai alat ukur dikatakan berkualitas apabila memenuhi syarat validitas dan reliabilitas. Soal bentuk tes, juga dilihat dari butir-butir tesnya (Suyata, 2009: 16). Menurut Purwanto (1992: 137), mengukur kesesuaian, efisiensi, dan kemantapan (*consistency*) suatu alat penilaian atau suatu tes dipergunakan bermacam-macam kualitas seperti validitas, keandalan, objektivitas, dan kepraktisan (*practicibility*). Lebih lengkap, Arikunto (2006: 57-62) menyebutkan bahwa sebuah tes yang dapat dikatakan baik sebagai alat pengukuran harus

memiliki persyaratan tes, yaitu: validitas, reliabilitas, objektivitas, praktikabilitas, dan ekonomis.

Pendapat para pakar di atas, dapat diambil kesimpulan bahwa kualitas soal yang baik atau berkualitas harus memiliki persyaratan tes berupa validitas, reliabilitas, objektivitas, praktikabilitas, dan ekonomis. Namun, validitas dan reliabilitas merupakan syarat terpenting dalam menentukan kualitas soal. Agar lebih jelas, berikut ini dijabarkan kelima syarat kualitas soal.

a. Validitas

Alat ukur dikatakan valid apabila alat ukur itu dapat dengan tepat mengukur apa yang hendak diukur. Sebuah data atau informasi dapat dikatakan valid apabila sesuai dengan keadaan senyatanya. Jika data yang dihasilkan dari sebuah instrumen valid, maka dapat dikatakan bahwa instrumen tersebut valid, karena dapat memberikan gambaran tentang data secara benar sesuai dengan kenyataan atau keadaan sesungguhnya (Arikunto, 2006: 58). Validitas sering diartikan kesahihan. Suatu alat ukur disebut memiliki validitas yang baik bilamana alat ukur tersebut isinya layak mengukur objek yang seharusnya diukur dan sesuai dengan kriteria tertentu (Thoha, 1990: 109).

Pendapat di atas dapat disimpulkan bahwa validitas adalah sifat yang sesuai dengan kenyataan, tidak dibuat-buat, misalnya validitas suatu tes. Tes dapat dikatakan memiliki validitas yang tinggi apabila mampu memberikan gambaran kemampuan siswa secara benar sesuai dengan kenyataan. Sebaliknya, tes dikatakan memiliki validitas yang rendah apabila tidak mampu menggambarkan kemampuan siswa yang sebenarnya.

Suatu teknik evaluasi dikatakan mempunyai validitas yang tinggi (disebut valid) jika teknik evaluasi atau tes itu dapat mengukur apa yang sebenarnya akan diukur (Purwanto, 1992: 137). Lebih dalam lagi, validitas suatu alat evaluasi bukanlah merupakan ciri yang absolut atau mutlak. Suatu tes dapat memiliki validitas yang bertingkat-tingkat: tinggi, sedang, rendah, bergantung pada tujuannya.

Menurut Arikunto (2006: 65), secara garis besar ada dua macam validitas, yaitu validitas logis dan validitas empiris. Macam validitas tersebut adalah sebagai berikut.

1) Validitas logis

Arikunto (2006: 65) menyebutkan bahwa validitas logis sebuah instrumen evaluasi menunjuk pada kondisi bagi sebuah instrumen yang memenuhi persyaratan valid berdasarkan hasil penalaran. Validitas logis dapat dicapai apabila instrumen yang disusun mengikuti ketentuan yang ada. Ada dua macam validitas logis yang dapat dicapai oleh sebuah instrumen, yaitu: validitas isi dan validitas konstruk (*construct validity*).

a) Validitas isi (*content validity*)

Validitas isi menunjuk suatu kondisi instrumen yang disusun berdasarkan isi materi pelajaran yang dievaluasi. Oleh karena materi yang diajarkan tertera dalam kurikulum maka validitas isi ini sering juga disebut validitas kurikuler (Arikunto, 2006: 65). Sama halnya dengan pendapat Purwanto (1992: 138) yang mengatakan bahwa tes dikatakan memiliki *content validity* apabila isi tes sesuai dengan sampel hasil belajar yang seharusnya dicapai menurut tujuan kurikulum.

b) Validitas konstruk (*construct validity*)

Menurut Arikunto (2006: 65), Validitas konstruk menunjuk kondisi instrumen yang disusun berdasarkan konstruk–aspek kejiwaan–yang seharusnya dievaluasi. Butir-butir soal tersebut dapat mengukur setiap aspek berpikir seperti yang disebutkan dalam Tujuan Instruksional Khusus (TIK). Purwanto (1992: 138) menambahkan, untuk menentukan adanya *construct validity*, hasil-hasil tes itu disesuaikan dengan tujuan atau ciri-ciri tingkah laku (domain) yang hendak diukur.

2) Validitas empiris

Arikunto (2006: 65) menyebutkan bahwa sebuah instrumen memiliki validitas empiris apabila sudah diuji dari pengalaman. Validitas empiris tidak dapat diperoleh hanya dengan menyusun instrumen berdasarkan ketentuan seperti halnya validitas logis, tetapi harus dibuktikan melalui pengalaman. Ada dua macam validitas empiris, yaitu validitas “ada sekarang” (*Concurrent validity*) dan validitas prediksi (*predictive validity*).

a) Validitas “ada sekarang” (*Concurrent validity*)

Validitas ini lebih umum disebut validitas empiris. Sebuah tes dikatakan memiliki validitas empiris jika hasilnya sesuai dengan pengalaman. Hal ini menunjukkan bahwa tes dipasangkan dengan hasil pengalaman, sehingga hasil tes merupakan sesuatu yang dibandingkan (Arikunto, 2006: 65). Di sisi lain, menurut Purwanto (1992: 138), tes dikatakan memiliki *concurrent validity* apabila hasil tes memiliki korelasi yang tinggi dengan hasil suatu alat ukur lain terhadap bidang yang sama pada waktu yang sama pula.

b) Validitas prediksi (*predictive validity*)

Menurut Arikunto (2006: 65), tes dikatakan memiliki validitas prediksi atau validitas ramalan apabila mempunyai kemampuan meramalkan apa yang akan terjadi pada masa yang akan datang. Purwanto (1992: 138) menambahkan, tes dikatakan memiliki *predictive validity* jika hasil koreksi tes itu meramalkan dengan tepat keberhasilan seseorang pada masa mendatang di dalam lapangan tertentu. Hal itu dapat dilihat dari korelasi koefisien antara hasil tes itu dengan hasil alat ukur lain pada masa mendatang.

Widoyoko (2009: 99) menambahkan bahwa validitas isi merupakan aspek terpenting dalam tes hasil belajar. Tingkat validitas isi tes dapat diketahui dengan penilaian ahli yang menguasai bidang studi tersebut. Jadi bersifat analisis kualitatif.

Pendapat para ahli di atas dapat disimpulkan macam-macam validitas, yaitu validitas isi (*content validity*), validitas konstruk (*construct validity*), validitas “ada sekarang” (*concurrent validity*), dan validitas prediksi (*predictive validity*). Validitas yang digunakan dalam penelitian ini adalah validitas logis dan validitas empiris. Validitas logis digunakan untuk melihat kesesuaian antara butir soal dengan materi pelajaran dalam kurikulum. Materi pelajaran selanjutnya dituangkan pada kisi-kisi soal. Butir soal memiliki validitas yang tinggi apabila dalam kisi-kisi tersebut memuat sampel materi yang mewakili seluruh pokok bahasan atau sub pokok bahasan yang telah diajarkan. Di sisi lain, validitas empiris digunakan setelah soal selesai diujikan. Validitas empiris digunakan untuk mengetahui reliabilitas soal dan karakteristik butir soal. Karakteristik butir soal meliputi indeks kesulitan, daya beda, dan efektivitas distraktor.

b. Reliabilitas

Reliabilitas berasal dari kata *reliability* berarti sejauh mana hasil suatu pengukuran dapat dipercaya (Djali dan Muljono, 2008: 55). Reliabilitas sering diartikan dengan keterandalan. Suatu tes memiliki keterandalan bilamana tes tersebut dipakai mengukur berulang-ulang hasilnya sama. Reliabilitas dapat pula diartikan dengan keajegan atau stabilitas (Thoha, 1990: 118). Dengan kata lain, tes dikatakan reliabel jika tes tersebut memberikan hasil yang tetap apabila diteskan berkali-kali. Reliabilitas adalah ketepatan atau keajegan alat dalam menilai, artinya ketepatan hasil manakala alat penilaian tersebut diberikan beberapa kali pada objek yang sama pada waktu yang berbeda.

Widoyoko (2009: 99) mengatakan bahwa kata reliabilitas (Inggris: *reliability*), berasal dari kata asal *reliable* yang artinya dapat dipercaya. Tes dikatakan dapat dipercaya jika memberikan hasil yang tetap atau ajeg apabila diteskan berkali-kali. Ajeg atau tetap tidak selalu harus sama, tetapi mengikuti perubahan secara ajeg. Pendapat tersebut sesuai dengan pendapat Jihad dan Haris (2008: 180) yang mengatakan bahwa reliabilitas soal merupakan ukuran yang menyangkut tingkat keajegan atau kekonsistenan suatu soal tes. Tingkat keajegan soal menggunakan perhitungan *Alpha Cronbach*.

Jadi, reliabilitas mengacu pada konsistensi pengukuran. Konsistensi pengukuran yaitu bagaimana suatu soal dalam beberapa kali pengukuran terhadap objek yang sama mempunyai hasil yang tetap atau ajeg. Karena mempunyai hasil yang tetap itulah membuat suatu pengukuran dapat dipercaya.

Djali dan Muljono (2008: 56) membedakan reliabilitas menjadi dua macam. Penjabaran kedua macam reliabilitas adalah sebagai berikut.

1) Reliabilitas konsistensi tanggapan

Reliabilitas konsistensi tanggapan responden mempersoalkan apakah tanggapan responden atau objek terhadap tes tersebut sudah baik atau konsisten.

2) Reliabilitas konsistensi gabungan item

Reliabilitas konsistensi gabungan item berkaitan dengan kemantapan atau konsistensi antara item-item suatu tes.

Faktor-faktor yang dapat memengaruhi reliabilitas soal menurut Purwanto (1992: 141) adalah sebagai berikut.

- 1) Luas-tidaknya *sampling* yang diambil. Makin luas *sampling*, berarti tes makin andal.
- 2) Perbedaan bakat dan kemampuan murid yang dites. Makin variabel kemampuan peserta tes, makin tinggi keandalan koefisien tes.
- 3) Suasana dan kondisi testing. Suasana ketika berlangsung testing, seperti tenang, gaduh, banyak gangguan, dapat memengaruhi hasil dan keandalan tes.

Di sisi lain, menurut <http://www.scribd.com>, secara rinci faktor yang memengaruhi reliabilitas skor tes adalah sebagai berikut.

- 1) Semakin banyak jumlah butir soal, semakin ajeg suatu tes
- 2) Semakin lama waktu tes semakin ajeg
- 3) Semakin sempit *range* kesukaran butir soal, semakin besar keajegan
- 4) Soal-soal yang saling berhubungan akan mengurangi keajegan
- 5) Semakin objektif pemberian skor, semakin besar keajegan
- 6) Ketidaktepatan pemberian skor
- 7) Menjawab besar soal dengan cara menebak
- 8) Semakin homogen materi semakin besar keajegan
- 9) Pengalaman peserta ujian
- 10) Salah penafsiran terhadap butir soal
- 11) Menjawab soal dengan buru-buru/cepat
- 12) Kesiapan mental peserta ujian
- 13) Adanya gangguan dalam pelaksanaan tes
- 14) Jarak antara tes pertama dengan tes kedua
- 15) Mencontek dalam mengerjakan tes
- 16) Posisi individu dalam belajar
- 17) Kondisi fisik peserta ujian

Uraian faktor yang memengaruhi reliabilitas di atas menunjukkan bahwa untuk mendapatkan tes yang memiliki reliabilitas tinggi tidaklah mudah. Faktor yang dapat memengaruhi reliabilitas soal meliputi faktor internal pada diri peserta tes, dan faktor eksternal seperti situasi tempat tes.

Ukuran besarnya koefisien reliabilitas soal yang dikemukakan oleh para ahli berada pada rentang angka -1,0 sampai +1,0. Koefisien reliabilitas yang mendekati angka 1,0 merupakan indeks reliabilitas tinggi. Indeks reliabilitas tinggi mengandung pengertian bahwa alat penilaian memiliki tingkat keajegan atau ketetapan (reliabel) (Sudjana, 2009:17). Menurut Suyata (2009: 15), indeks reliabilitas yang baik adalah 0,90 atau lebih, tetapi instrumen buatan guru indeks reliabilitas 0,70 sudah dianggap cukup baik.

Soal Ulangan Akhir Semester (UAS) Genap SMA di Kabupaten Purbalingga dibuat oleh tim Musyawarah Guru Mata Pelajaran (MGMP) sehingga koefisien reliabilitas yang relevan adalah sebesar $\geq 0,90$. Apabila koefisien reliabilitas kurang dari 0,90 menunjukkan bahwa soal belum cukup baik.

c. Objektivitas

Objektivitas suatu tes ditentukan oleh tingkat atau kualitas kesamaan skor-skor yang diperoleh dengan tes tersebut meskipun hasil tes tersebut dinilai oleh beberapa orang penilai (Purwanto, 1992:141). Widoyoko (2009: 100) menjelaskan bahwa objektivitas berarti tidak adanya unsur pribadi yang memengaruhinya.

Kualitas objektivitas dalam suatu tes dapat dibedakan menjadi tiga tingkatan. Ketiga tingkatan tersebut adalah sebagai berikut.

- 1) Objektivitas tinggi, ialah jika hasil-hasil tes itu menunjukkan tingkat kesamaan yang tinggi. Contohnya: tes yang distandardisasi.

- 2) Objektivitas sedang, ialah seperti tes yang sudah distandardisasi, tetapi pandangan subjektif skor masih mungkin muncul dalam penilaian dan interpretasinya.
- 3) Objektivitas fleksibel, ialah seperti beberapa jenis tes yang digunakan oleh LBP (Lembaga Bimbingan dan Penyuluhan) untuk keperluan *Counseling* (Purwanto, 1992:141).

d. Kepraktisan

Suatu tes dikatakan mempunyai kepraktisan yang baik jika kemungkinan menggunakan tes itu besar (Purwanto, 1992:141). Menurut Widoyoko (2009: 99), tes yang praktis adalah tes yang:

- a. mudah dilaksanakan, artinya tidak menuntut peralatan yang banyak dan memberi kebebasan kepada siswa mengerjakan terlebih dahulu bagian yang dianggap mudah oleh siswa;
- b. mudah pemeriksaannya, artinya bahwa tes itu dilengkapi dengan kunci jawaban maupun pedoman skoringnya. Soal bentuk objektif, pemeriksaannya akan lebih mudah dilakukan jika dikerjakan oleh siswa dalam lembar jawaban;
- c. dilengkapi dengan petunjuk-petunjuk sehingga dapat diberikan oleh orang lain.

Pendapat di atas menunjukkan bahwa tes yang baik adalah tes yang mudah dipakai, dilakukan, dan diteliti. Tes yang baik juga memiliki ketentuan yang memberikan arah atau bimbingan bagaimana suatu tes harus dilakukan.

d. Ekonomis

Ekonomis adalah sifat hati-hati terhadap pengeluaran dan penggunaan baik berupa uang, barang, bahasa, maupun waktu. Menurut Arikunto (2006: 63), ekonomis berarti pelaksanaan tes tidak membutuhkan biaya yang mahal, tenaga yang banyak, dan waktu yang lama.

4. Bentuk Soal

Bentuk Soal yang digunakan di lembaga pendidikan dapat berupa tes dan nontes. Secara umum tes diartikan sebagai alat mengukur pengetahuan atau

penguasaan objek ukur terhadap seperangkat konten dan materi tertentu (Djali dan Muljono, 2008: 6). Bentuk tes terdiri dari tes objektif dan tes nonobjektif. Pada tes objektif, siapa saja yang memeriksa lembar jawaban tes akan menghasilkan skor yang sama. Tes nonobjektif sistem penskorannya dipengaruhi oleh subjektivitas pemberi skor (Depdiknas, 2003: 27-28), sedangkan nontes dapat berbentuk pedoman observasi, pedoman wawancara, angket (kuisisioner), dan pemeriksaan dokumen (Djali dan Muljono, 2008: 16-23).

Bentuk tes yang digunakan di lembaga pendidikan dilihat dari segi sistem penskorannya dapat dikategorikan menjadi dua, yaitu tes objektif dan tes subjektif.

a. Tes objektif

Menurut Widoyoko (2009: 46), tes objektif adalah bentuk tes yang mengandung kemungkinan jawaban atau respons yang harus dipilih oleh peserta tes.

- 1) Kelebihan tes objektif
 - a) Lebih representatif mewakili isi dan luas bahan.
 - b) Lebih mudah dan cepat cara memeriksanya karena dapat menggunakan kunci jawaban, bahkan dapat menggunakan alat-alat kemajuan teknologi misalnya mesin *scanner*.
 - c) Pemeriksaannya dapat diserahkan orang lain.
 - d) Dalam pemeriksaan maupun penskoran, tidak ada unsur subjektif yang memengaruhi, baik dari segi guru maupun siswa.
- 2) Kelemahan tes objektif
 - a) Membutuhkan persiapan yang lebih sulit dari pada tes esai karena butir soal atau item tesnya banyak dan harus teliti untuk menghindari kelemahan-kelemahan yang lain.
 - b) Butir-butir soal cenderung hanya mengungkap ingatan dan pengenalan kembali (*recalling*) saja, dan sukar mengukur kemampuan berpikir yang tinggi seperti sintesis maupun kreativitas.
 - c) Banyak kesempatan bagi siswa untuk spekulasi atau untung-untungan (*guessing*) dalam menjawab soal tes.
 - d) Kerja sama antar siswa pada waktu mengerjakan soal tes lebih terbuka.
- 3) Cara mengatasi kelemahan
 - a) Kesulitan menyusun tes objektif dapat diatasi dengan jalan banyak berlatih menyusun soal tes secara terus-menerus, sehingga semakin lama semakin terampil.
 - b) Menggunakan tabel spesifikasi

- c) Menggunakan norma (standar) penilaian yang memperhitungkan faktor tebakan (*guessing*) yang bersifat spekulatif.

Menurut Widoyoko (2009: 50-60), secara umum ada tiga tipe tes objektif, yaitu benar salah (*true false*), menjodohkan (*matching*), dan pilihan ganda (*multiple choice*). Arikunto (2006: 175) menambahkan bahwa tes isian (*completion test*) sebagai tipe tes objektif. Kajian teori ini hanya akan menjelaskan tipe tes pilihan ganda (*multiple choice*) karena Ulangan Akhir Semester (UAS) Genap kelas XII SMA Negeri di Kabupaten Purbalingga menggunakan tipe tes ini.

Tes pilihan ganda (*multiple choice test*) adalah tes di mana setiap butir soalnya memiliki jumlah alternatif jawaban lebih dari satu. Menurut Shirran (2008: 93), soal pilihan ganda memiliki dua bagian, yaitu soal atau masalah (*stem*) dan tiga atau empat jawaban pilihan (alternatif). Soal pilihan ganda mengharuskan peserta memilih alternatif yang paling sesuai untuk melengkapi pernyataan atau menjawab soal. Jawaban yang kemungkinan tidak betul disebut distraktor karena mereka didisain untuk mengalihkan perhatian dari jawaban yang betul itu. Pada umumnya jumlah alternatif jawaban berkisar antara 2 (dua) atau 5 (lima). Jumlah alternatif tidak boleh terlalu banyak.

Hal yang paling sulit dilakukan dalam menulis soal bentuk pilihan ganda adalah menuliskan pengecohnya. Pengecoh yang baik adalah pengecoh yang tingkat kerumitan atau tingkat kesederhanaan, serta panjang pendeknya relatif sama dengan kunci jawaban. Oleh karena itu, penulisan soal bentuk pilihan ganda perlu mengikuti langkah-langkah berikut. Langkah pertama adalah menuliskan pokok soalnya, langkah kedua menuliskan kunci jawabannya, langkah ketiga menuliskan pengecohnya (BSNP, hlm. 19).

Menurut Widoyoko (2009: 52-53), kelebihan tes pilihan ganda adalah dapat digunakan mengukur segala level tujuan pembelajaran; waktu mengerjakan sangat minimal; penskoran dapat dilakukan secara objektif; tipe butir soal dapat disusun sedemikian rupa sehingga menuntut kemampuan peserta tes membedakan berbagai tingkatan kebenaran sekaligus. Kelebihan lain yang dimiliki tes bentuk ini adalah jumlah pilihan yang disediakan melebihi dua sehingga mengurangi keinginan peserta tes untuk menebak; memungkinkan dilakukan analisis butir soal secara baik; tingkat kesukaran butir soal dapat diatur dengan hanya mengubah tingkat homogenitas alternatif jawaban; dan informasi yang diberikan lebih kaya.

Meskipun banyak kelebihan, tes bentuk ini juga memiliki kekurangan. Kekurangan tes bentuk pilihan ganda antara lain: lebih sulit dalam penyusunan butir soal; kecenderungan penyusunan soal hanya mengukur aspek ingatan; serta pengaruh kebiasaan peserta tes terhadap tes bentuk pilihan ganda (*testwise*) terhadap hasil peserta.

Pendapat di atas menunjukkan tes pilihan ganda memiliki banyak kelebihan dibanding bentuk tes lain. Namun, tes jenis ini juga memiliki kekurangan. Penyusunan butir soal yang sulit dapat diminimalisir dengan berlatih menyusun soal secara berkesinambungan.

b. Tes Subjektif

Arikunto (2006: 162) menyebutkan bahwa tes subjektif, pada umumnya berbentuk uraian (esai), yaitu sejenis tes kemajuan belajar yang memerlukan jawaban bersifat pembahasan. Soal bentuk esai berkisar 5-10 soal yang dikerjakan dalam waktu 90-120 menit. Soal ini menuntut siswa dapat mengorganisir, menginterpretasi, menghubungkan pengertian-pengertian yang telah dimiliki.

Ditambahkan oleh Arikunto (2006: 163), kelebihan tes subjektif antara lain: mudah disiapkan dan disusun; tidak memberi banyak kesempatan berspekulasi; mendorong siswa mengemukakan pendapat dalam bentuk kalimat yang bagus; memberi kesempatan siswa mengutarakan maksudnya dengan gaya bahasa sendiri; dan mengetahui sejauh mana siswa mendalami sesuatu masalah yang diteskan. Tes subjektif juga memiliki beberapa keburukan antara lain: kadar validitas dan reliabilitas rendah; kurang representatif dalam mewakili seluruh *scope* bahan pelajaran yang diteskan; pemeriksaannya lebih sulit; dan waktu pengoreksian lama.

Pendapat di atas menunjukkan bahwa bentuk tes subjektif menuntut kemampuan siswa menyampaikan, memilih, menyusun, dan memadukan gagasan atau ide yang telah dimilikinya dengan menggunakan kata-katanya sendiri. Namun, teknik penskoran dalam tes ini cenderung dipengaruhi subjektivitas penilai.

Menurut Widoyoko (2009: 79), berdasarkan tingkat kebebasan peserta tes menjawab soal tes uraian, secara umum tes uraian dapat dibagi menjadi dua bentuk. Bentuk soal tes uraian yaitu: tes uraian bebas atau uraian terbuka (*extended response*) dan tes uraian terbatas (*restricted response*).

5. Analisis Butir Soal (*Item Analysis*)

Analisis soal antara lain bertujuan untuk mengidentifikasi soal-soal yang baik, kurang baik, dan soal yang tidak baik. Dengan mengetahui soal-soal yang tidak baik itu selanjutnya kita dapat mencari kemungkinan sebab-sebab mengapa item itu tidak baik. Menurut Purwanto (1992: 118-119), dengan membuat analisis soal, dapat diketahui hal penting yang diperoleh dari tiap soal, yaitu: sampai di

mana tingkat atau taraf kesukaran soal; apakah soal mempunyai daya pembeda; dan apakah semua alternatif jawaban (*options*) menarik jawaban-jawaban.

a. Taraf kesukaran (*difficulty level*) soal

Arikunto (2006: 207) menjelaskan bahwa soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa belajar. Sebaliknya soal yang terlalu sukar menyebabkan siswa putus asa karena diluar jangkauannya.

Bilangan yang menunjukkan sukar dan mudahnya soal disebut indeks kesukaran (*difficulty indeks*). Besarnya indeks kesukaran antara 0,00 sampai 1,0. Soal dengan indeks kesukaran 0,0 menunjukkan bahwa soal tersebut terlalu sukar, dan indeks 1,0 menunjukkan bahwa soal terlalu mudah.

Rumus mencari indeks kesukaran

$$p = \frac{B}{JS}$$

Di mana:

p = indeks kesukaran

B = banyaknya siswa yang menjawab soal dengan betul

JS = Jumlah siswa peserta tes

Menurut ketentuan, indeks kesukaran diklasifikasikan sebagai berikut.

– Soal dengan p 1,00 sampai p 0,30 adalah soal sukar.

– Soal dengan p 0,30 sampai p 0,70 adalah soal sedang.

– Soal dengan p 0,70 sampai p 1,00 adalah soal mudah.

Soal-soal yang dianggap baik adalah soal-soal sedang dengan p 0,30 sampai p 0,70 (Arikunto, 2006: 208-210).

Lain halnya dengan pendapat Nurgiyantoro (2010: 11) bahwa indeks tingkat kesukaran yang diterima berkisar antara 0,20-0,80. Di sisi lain, menurut Fernandes (1984: 26), *“Typicaly the purpose of a national achievement test is to differentiate students at all levels of achievement, and item having difficulty indices ranging from about 75% to 25% are most useful.”* Pendapat tersebut menjelaskan bahwa

besarnya tingkat kesukaran yang diterima menyatakan butir soal yang baik adalah antara 0,25 sampai 0,75.

Penelitian ini menggunakan pendapat fernandes yang menyebutkan bahwa indeks kesulitan yang baik adalah antara 0,25 sampai 0,75. Rentang tersebut menunjukkan soal yang baik tidak terlalu mudah dan tidak terlalu sukar.

Bila suatu butir soal termasuk dalam kategori sukar, maka prediksi terhadap informasi ini adalah seperti berikut.

- 1) Butir soal itu “mungkin” salah kunci jawaban
- 2) Butir soal itu mempunyai 2 atau lebih jawaban yang benar
- 3) Materi yang ditanyakan belum diajarkan atau belum tuntas pembelajarannya, sehingga kompetensi minimum yang harus dikuasai siswa belum tercapai
- 4) Materi yang diukur tidak cocok ditanyakan dengan menggunakan bentuk soal yang diberikan
- 5) Pernyataan atau kalimat soal terlalu panjang (<http://www.scribd.com>).

b. Daya pembeda (*discriminating power*)

Menurut Arikunto (2006: 211), daya pembeda soal adalah kemampuan sesuatu soal membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang bodoh (berkemampuan rendah). Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi (D) yang berkisar antara 0,00 sampai 1,0. Pada indeks diskriminasi ada tanda negatif (-) yang digunakan jika sesuatu soal “terbalik” menunjukkan kualitas *testee*, yaitu anak pandai disebut bodoh dan anak bodoh disebut pandai.

Soal yang dapat dijawab benar oleh siswa pandai maupun siswa bodoh, soal tersebut tidak baik karena tidak mempunyai daya pembeda. Demikian pula jika semua siswa pandai dan bodoh tidak dapat menjawab dengan benar. Soal yang baik adalah soal yang dapat dijawab benar oleh siswa-siswa yang pandai saja.

Klasifikasi daya pembeda

$D = 0,00 - 0,20$ = tidak baik (*poor*)

$D = 0,20 - 0,40$ = cukup (*satisfactory*)

$D = 0,40 - 0,70$ = baik (*good*)

$D = 0,70 - 1,00$ = baik sekali (*excellent*)

D = negatif, semuanya tidak baik, jadi semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja.

Lain halnya dengan Depdiknas (2003:56) yang menyatakan bahwa daya beda dinyatakan baik untuk butir soal acuan norma jika minimum besarnya 0,3. Pada

butir soal acuan kriteria, jika seluruh siswa sudah berhasil menguasai indikator dari suatu kompetensi dasar, maka indeks daya beda akan sebesar 0. Namun butir ini tetap dinyatakan baik dan tetap dapat dipakai menunjukkan efektivitas proses pembelajaran manakala seluruh siswa sebelum mengalami proses pembelajaran tidak dapat mengerjakan butir soal yang bersangkutan.

Pendapat di atas menunjukkan adanya perbedaan dalam klasifikasi daya beda yang dinyatakan baik. Peneliti lebih mengacu pada klasifikasi daya beda dari Depdiknas yang menyatakan bahwa daya beda yang baik adalah $\geq 0,3$ untuk butir soal acuan norma.

Apabila suatu butir soal tidak dapat membedakan kedua kemampuan siswa itu, maka butir soal itu dapat dicurigai “kemungkinannya” sebagai berikut.

- 1) Kunci jawaban butir soal itu tidak tepat.
- 2) Butir soal itu memiliki 2 atau lebih kunci jawaban yang benar
- 3) Kompetensi yang diukur tidak jelas
- 4) Pengecoh tidak berfungsi
- 5) Materi yang ditanyakan terlalu sulit, sehingga banyak siswa yang menebak (<http://www.scribd.com>).

Soal UAS Bahasa Indonesia SMA Negeri di Purbalingga tahun ajaran 2010/2011 disusun berdasarkan Kurikulum Tingkat Satuan Pendidikan (KTSP), sistem penilaian pun harus mengarah pada penilaian berbasis KTSP. Acuan yang dipakai dalam menafsirkan hasil tes adalah acuan kriteria. Pemakaian indeks daya beda pada acuan kriteria tidak dipakai. Indeks daya beda pada dasarnya adalah perbandingan antara banyaknya anggota kelompok yang berhasil (kelompok atas) dan banyaknya anggota kelompok yang gagal (kelompok bawah). Hal tersebut hanya dipakai jika penilaiannya adalah penilaian norma. *Point Biser* dalam *print out Iteman* menunjukkan indeks daya beda tiap butir soal (Riestalinda, 2007: 25).

c. Pola jawaban soal

Arikunto (2006: 219-220) menjelaskan bahwa yang dimaksud pola jawaban soal adalah distribusi *testee* dalam hal menentukan pilihan jawaban pada soal bentuk pilihan ganda. Pola jawaban soal diperoleh dengan menghitung banyaknya *testee* yang memilih pilihan jawaban a, b, c, d, atau e dan yang tidak memilih pilihan manapun (*blanko*). *Testee* yang tidak memilih pilihan jawaban manapun disebut omit (O).

Pola jawaban soal dapat menentukan pengecoh (*distractor*) yang berfungsi dengan baik dan yang tidak berfungsi dengan baik. Pengecoh yang tidak dipilih sama sekali oleh *testee* menunjukkan bahwa pengecoh itu tidak baik. Sebaliknya distraktor dapat dikatakan berfungsi dengan baik apabila distraktor tersebut dipilih oleh banyak *testee*.

Pola jawaban soal digunakan mengetahui taraf kesukaran soal, daya pembeda soal, dan baik tidaknya distraktor. Suatu distraktor berfungsi baik jika paling sedikit dipilih oleh 5% pengikut tes (Arikunto, 2006: 220). Berbeda dengan Fernandes (1984: 29) yang menyatakan bahwa “*Here a rule of thumb is that an alternative should be eliminated or revised unless one of every 50 student (or 2%) select it.*” Pendapat tersebut menyebutkan bahwa soal tidak perlu dibuang atau direvisi apabila telah dipilih oleh 2% peserta tes. Penelitian ini menggunakan pendapat fernandes yang menyebutkan bahwa distraktor telah berfungsi baik jika telah dipilih oleh 2% peserta tes.

Analisis distraktor hanya dapat diterapkan pada bentuk tes pilihan ganda (tes objektif). Suatu butir soal yang baik adalah butir soal yang kunci jawabannya

positif dan distraktornya negatif. Efektivitas distraktor butir soal ditunjukkan pada kolom *Proportion Endorsing* dalam hasil *print out Iteman* (Riestalinda, 2007: 27).

Selanjutnya, Suyata (2009: 16) menjelaskan bahwa analisis butir soal dilakukan secara (1) kualitatif atau teoretik dan (2) kuantitatif atau empiris.

a. Analisis kualitatif

Depdiknas (2003: 53) menjelaskan bahwa analisis butir soal secara kualitatif atau telaah butir dilakukan sebelum uji coba, yakni dengan mencermati butir-butir soal yang telah disusun dilihat dari kesesuaian dengan kompetensi dasar dan indikator yang diukur serta pemenuhan persyaratan baik dari aspek materi, konstruksi, dan bahasa. Hal yang sama dijelaskan oleh Suyata (2009: 16) bahwa analisis kualitatif dilakukan sebelum instrumen diujikan.

Pendapat di atas menunjukkan bahwa analisis kualitatif digunakan mengetahui validitas isi soal. Validitas ini digunakan untuk melihat kesesuaian antara butir soal dengan materi pelajaran dalam kurikulum. Butir soal akan dianalisis apakah telah memenuhi persyaratan atau belum. Persyaratan butir soal meliputi aspek materi, konstruksi, dan bahasa.

b. Analisis kuantitatif

Analisis secara kuantitatif dilakukan setelah instrumen diujikan (Suyata, 2009: 16). Depdiknas (2003: 53) menyebutkan bahwa analisis butir soal secara kuantitatif ada dua cara, yaitu analisis cara klasik/tradisional dan analisis cara modern dengan mendasarkan pada *item response theory* (IRT). Suyata (2009: 16) menambahkan, pada analisis secara klasik, tingkat pencapaian suatu kompetensi dasar adalah proporsi jumlah peserta tes yang menjawab benar terhadap indikator

kompetensi dasar yang bersangkutan, yaitu perbandingan antara jumlah peserta tes yang menjawab benar dengan jumlah peserta tes seluruhnya.

Depdiknas (2003: 53) menjelaskan bahwa analisis butir soal secara klasik dibedakan menjadi dua macam berdasarkan tujuan penilaian yang dilakukan. Butir soal harus memenuhi standar butir soal acuan kriteria jika menggunakan pendekatan acuan kriteria (*criterion referenced test*). Jika menggunakan pendekatan penilaian acuan norma, maka butir soal harus memenuhi standar sebagai butir soal acuan norma (*norm referenced test*).

Beberapa formula dalam analisis butir tes acuan kriteria dan acuan norma adalah sama tetapi penafsirannya berbeda. Langkah-langkah analisis dengan acuan norma tidak dipakai lagi menganalisis butir soal untuk mengukur indikator-indikator ketercapaian penguasaan kemampuan dasar pada kurikulum berbasis kompetensi (Depdiknas, 2003: 54).

Tujuan penilaian acuan kriteria adalah mengetahui kemampuan seseorang menurut kriteria tertentu. Karakteristik utama butir soal acuan kriteria tercermin dari besarnya harga indeks sensitivitas yang menunjukkan efektivitas proses pembelajaran (Depdiknas, 2003: 55).

Indeks sensitivitas butir soal memiliki interval dari -1 sampai dengan +1. Menurut Depdiknas (2003: 55), indeks sensitivitas suatu butir soal (I_s) penilaian formatif:

$$I_s = \frac{R_A - R_B}{T}$$

Di mana

I_s = Indeks sensitivitas butir soal

R_A = Banyaknya siswa yang berhasil mengerjakan suatu butir tes soal sesudah proses pembelajaran

R_B = Banyaknya siswa yang berhasil mengerjakan suatu butir tes soal sebelum proses pembelajaran

T = Banyaknya siswa yang mengikuti penilaian

Jika tidak diadakan tes awal, maka R_B tidak diketahui. Oleh karena itu, dapat dilihat dari besarnya tingkat pencapaiannya berdasarkan hasil tes akhir (*posttest*) atau besarnya R_A . Dengan kata lain, sebagai indikator kualitas butir soal hanya tergantung pada indeks kesukaran butir.

Penelitian ini menggunakan analisis cara klasik/tradisional dengan menggunakan acuan kriteria. Analisis cara klasik ini menggunakan program *MicroCat Iteman*. Hasil pengolahan data dengan *Iteman* mencakup informasi mengenai besarnya indeks reliabilitas dan analisis butir soal. Analisis butir soal meliputi indeks kesukaran butir, indeks daya beda, dan efektivitas distraktor. Penafsiran dalam penelitian ini menggunakan acuan kriteria sehingga daya beda tidak berfungsi atau diabaikan.

5. Karakteristik Penilaian Pembelajaran Bahasa Indonesia

Penilaian pembelajaran bahasa Indonesia adalah proses mendapatkan informasi, mempertimbangkannya, serta memutuskan mengenai prestasi atau kinerja siswa di dalam mata pelajaran Bahasa Indonesia. Hasil penilaian dapat digunakan untuk melakukan evaluasi terhadap ketuntasan belajar peserta didik dan efektivitas proses pembelajaran (Suryaman, 2009:197).

Penilaian pembelajaran bahasa Indonesia harus memperhatikan hakikat dan fungsi bahasa. Pada hakikatnya, bahasa merupakan hasil budaya manusia yang juga berfungsi sebagai sarana komunikasi. Penilaian pembelajaran bahasa tidak mengarah pada sistem bahasa, melainkan pada bagaimana menggunakan bahasa secara benar sesuai dengan sistem itu. Secara pragmatis, bahasa lebih merupakan satu bentuk kinerja dan performansi pada suatu sistem ilmu. Pandangan ini mengarahkan penilaian pembelajaran bahasa haruslah menekankan pada fungsi bahasa sebagai alat komunikasi pada sebagai suatu sistem bahasa (Suyata, 2009: 6).

Suyata (2009: 7) menyebutkan bahwa standar kompetensi penilaian pembelajaran bahasa meliputi (1) kompetensi mendengarkan, (2) kompetensi berbicara, (3) kompetensi membaca, (4) kompetensi menulis, serta (5) kompetensi pragmatik, yaitu kemampuan menggunakan bahasa dengan memperhatikan konteks, situasi, dan kondisi. Kemampuan bersastra tetap perlu diberikan pada siswa dengan cara diintegrasikan dalam kelima kemampuan yang ada di atas.

Pendapat di atas menunjukkan bahwa standar kompetensi mata pelajaran Bahasa Indonesia di Sekolah Menengah Atas (SMA) terdiri dari dua aspek, yaitu aspek kemampuan berbahasa dan aspek kemampuan bersastra. Kedua aspek tersebut terbagi atas beberapa aspek, yaitu mendengarkan, berbicara, membaca, dan menulis.

Komponen atau unsur bahasa yang ditekankan adalah meliputi hal-hal yang menjadi cakupan pembelajaran bahasa. Cakupan pembelajaran bahasa meliputi kompetensi bahasa (kompetensi linguistik), kompetensi berbahasa (kompetensi komunikatif), dan kompetensi bersastra (Nurgiyantoro, 2010: 280-284).

a) Tes kompetensi bahasa

Tes yang menyangkut kompetensi kebahasaan secara garis besar dapat dikelompokkan menjadi tes struktur dan kosakata (dengan tanpa mengabaikan sistem fonologi). Struktur dan kosakata perlu mendapat perhatian secara khusus, walaupun secara umum pembelajaran dan tes bahasa lebih ditekankan pada fungsi komunikatif bahasa. Tes kompetensi bahasa terdiri dari tes struktur gramatikal, dan tes kosakata.

b) Tes kompetensi berbahasa

Tinggi rendahnya kompetensi kebahasaan seseorang pada umumnya tercermin dari kemampuan berbahasanya. Kompetensi berbahasa dibedakan menjadi tes kompetensi aktif reseptif dan tes kompetensi aktif produktif.

c) Tes kompetensi bersastra

Pengetahuan tentang sastra mencakup bahan yang bersifat teoretis dan historis. Pentingnya pengetahuan tentang sastra tersebut karena ia merupakan alat bantu mengapresiasi karya sastra. Sesuai dengan peranannya yang

sebagai “alat bantu”, tes pengetahuan tentang sastra harus bukan merupakan prioritas. Tes kesastraan harus diprioritaskan pada usaha mengungkap kompetensi mengapresiasi sastra peserta didik atau kompetensi bersastra, dan secara langsung berhubungan dengan berbagai karya sastra.

Dengan kata lain, pembelajaran Bahasa Indonesia menekankan pencapaian kemampuan berbahasa secara benar dan bukan pencapaian kemampuan tentang sistem bahasa. Di sisi lain, pembelajaran sastra lebih ditekankan pada pencapaian kemampuan apresiasi sastra pada pengetahuan sastra sebagai ilmu. Hal tersebut juga berdampak pada sistem penilaiannya.

Fokus penilaian pembelajaran Bahasa Indonesia adalah keberhasilan belajar peserta didik dalam mencapai standar kompetensi berbahasa dan bersastra Indonesia (Suryaman, 2009:197). Kompetensi yang harus dicapai dalam mata pelajaran Bahasa Indonesia secara jelas telah termuat dalam standar kompetensi, yang kemudian dijabarkan ke dalam kemampuan dasar dan indikator. Indikator digunakan sebagai dasar menyusun soal ujian untuk mengukur tingkat pencapaian kemampuan yang dimaksud.

B. Kajian Penelitian yang Relevan

Penelitian yang dilakukan oleh Indriyati (2006) dengan judul *Analisis Soal Tes Uji Standar Mutu Mata Pelajaran Bahasa dan Sastra Indonesia Kelas X Tahun Ajaran 2004/2005 Kabupaten Sleman*, secara keseluruhan alat ukur tersebut masih kurang bagus. Validitas isi soal tes tersebut tidak dapat dilaporkan karena kisi-kisi soal tidak berhasil ditemukan. Indeks reliabilitas soal kurang tinggi dengan angka 0,657. Dengan pendekatan penilaian KBK, 22 butir soal (51,7%) ditolak atau gugur dengan IDB diabaikan. Dengan pendekatan penilaian non-KBK, 13 butir soal (21,7%) ditolak atau gugur dengan IDB butir yang baik sebesar 30% dan rerata

daya beda soal 0,201 yang berarti kurang baik. Tingkat kesukaran soal tergolong cukup baik dan distraktor cukup baik (77,5%).

Riestalinda (2007) telah mengadakan penelitian yang berjudul *Analisis Soal Tes Penjajagan Hasil Belajar Siswa (TPHBS) Mata pelajaran Bahasa Indonesia SMA Tahun Ajaran 2006/2007 Se-Kabupaten Sleman* disimpulkan bahwa: dengan menggunakan teori tes klasik menunjukkan (1) apabila soal yang direvisi menjadi baik, validitas isi soal tergolong baik karena 35 butir soal (76,1%) dinyatakan diterima dan 11 butir soal (23,9%) dinyatakan ditolak. (2) Reliabilitas soal tergolong cukup karena koefisien *Alpha* sebesar 0,815; (3) sesuai dengan nilai *Mean P* sebesar 0,671, indeks kesukaran butir soal tergolong sedang karena 28 butir soal (56%) indeks kesulitannya baik dan 22 butir soal (44%) indeks kesukarannya tidak baik; (4) efektivitas distraktor tergolong sedang karena 26 butir soal (52%) memiliki distraktor yang efektif dan 23 butir soal (46%) distraktornya tidak efektif. Dengan demikian, secara umum perangkat soal tergolong cukup berkualitas.

Penelitian lain dilakukan oleh Isti Akhiromah (2011) dengan judul *Karakteristik Secara Teoretis dan Empiris Butir Soal Ulangan Umum Bersama Semester Genap Bahasa Indonesia SMP Kabupaten Gunung Kidul Tahun Pelajaran 2009/2010*. Hasil yang diperoleh menyebutkan (1) karakteristik soal secara teoretis (validitas isi) tergolong baik karena 94% soal sesuai kisi-kisi, (2) karakteristik soal secara empiris adalah reliabilitas soal tergolong kurang baik, koefisien *Alpha* sebesar 0,510, dengan SEM sebesar 3,085; karakteristik butir soal yaitu: indeks kesulitan tergolong sedang, terdapat 60% butir soal yang baik, daya beda soal baik yaitu 92% butir soal baik, keefektifan distraktor baik, 80% butir soal baik dan memiliki distraktor yang efektif. Dengan demikian, secara umum

karakteristik soal ini belum baik karena hanya 60% diterima sehingga perlu ada upaya perbaikan dalam penyempurnaan.

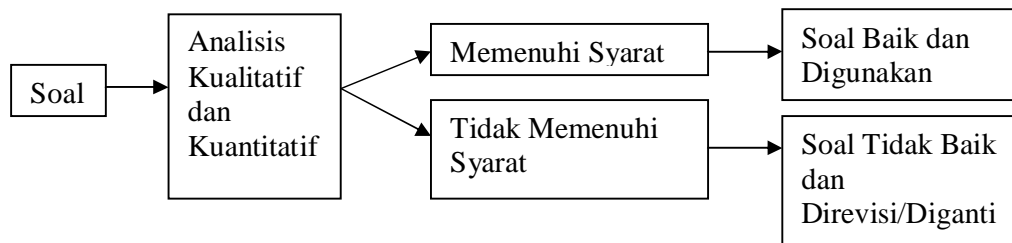
Penelitian ini ada relevansinya dengan tiga penelitian di atas karena sama-sama meneliti tentang karakteristik butir soal. Karakteristik butir soal ini diteliti secara kualitatif dan kuantitatif. Analisis kualitatif dilakukan untuk mengetahui validitas isi soal. Analisis kuantitatif dilakukan melalui program *MicroCat Iteman* untuk mengetahui besarnya indeks reliabilitas dan analisis butir soal. Analisis butir soal meliputi indeks kesukaran butir, indeks daya beda, dan efektivitas distraktor.

C. Kerangka Berpikir

Evaluasi hasil belajar digunakan untuk mengetahui pencapaian keefektifan tujuan pembelajaran. Soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011 memiliki fungsi yang strategis terhadap prediksi hasil ujian nasional. Oleh karena itu, sangat ideal apabila soal tersebut memenuhi persyaratan karakteristik soal yang baik.

Analisis soal digunakan untuk mengetahui karakteristik soal apakah sudah berfungsi dengan baik atau belum dan sesuaikah dengan teori pengukuran. Secara teoretis untuk mengetahui tes yang baik dapat dilakukan dengan analisis kualitatif yaitu dilihat dari segi materi, konstruksi, dan bahasa. Secara empiris dilakukan analisis kuantitatif dengan pendekatan teori tes klasik. Analisis dengan teori tes klasik menggunakan paket program *MicroCat Iteman* versi 3.0, tujuannya mengetahui indeks reliabilitas, tingkat kesukaran, daya beda, dan efektivitas distraktor apakah berfungsi dengan baik atau belum.

Penelitian ini dapat mengidentifikasi soal yang memiliki karakter baik dan tidak baik sesuai dengan kaidah yang telah ditentukan. Soal yang baik dapat digunakan pada tes berikutnya atau disimpan di bank soal, sedangkan soal yang tidak memenuhi syarat harus direvisi atau diganti. Hal ini, dapat diperjelas dengan bagan di bawah ini.



Gambar 1. **Alur Analisis Soal**

D. Pertanyaan Penelitian

Pertanyaan penelitian sebagai berikut.

1. Bagaimanakah karakteristik secara kualitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011?
2. Bagaimanakah karakteristik secara kuantitatif soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011? Secara rinci, analisis kuantitatif meliputi:
 - a. reliabilitas soal,
 - b. analisis butir soal, meliputi (1) indeks kesukaran butir soal, (2) indeks daya beda butir soal, dan (3) keefektifan distraktor.

BAB III

METODE PENELITIAN

A. Jenis penelitian

Penelitian ini merupakan penelitian evaluatif, yaitu mengevaluasi instrumen yang sudah ada berupa soal Ulangan Akhir Semester (UAS) Genap. Pada umumnya, penelitian evaluatif dimaksudkan untuk mengetahui akhir dari sebuah program kebijakan, yaitu mengetahui hasil akhir dari adanya kebijakan, dalam rangka menentukan rekomendasi atas kebijakan yang lalu, yang pada tujuan akhirnya adalah menentukan kebijakan selanjutnya (Arikunto dan Jabar, 2008: 7). Hasil analisis menggambarkan tentang karakteristik perangkat tes mata pelajaran Bahasa Indonesia pada Ulangan Akhir Semester (UAS) Genap kelas XII SMA Negeri tahun pelajaran 2010/2011 di Purbalingga. Karakteristik butir dilihat dari telaah butir soal dan teori tes klasik.

B. Sampel Penelitian

Sampel penelitian ini adalah paket soal yang terdiri dari kisi-kisi soal, seluruh lembar jawaban siswa, dan kunci jawaban soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011. Kisi-kisi soal digunakan untuk mencari validitas isi, sedangkan lembar jawaban siswa dianalisis guna memperoleh indeks reliabilitas soal, indeks tingkat kesukaran butir soal, indeks daya beda butir soal, dan efektivitas distraktor.

Lembar jawaban siswa berjumlah 2147 lembar, dihimpun dari sepuluh SMA Negeri di Kabupaten Purbalingga. Uraian data lembar jawaban sebagai berikut.

Tabel 1: Jumlah Data Lembar Jawaban Ulangan Akhir Semester (UAS) Genap SMA kelas XII tahun pelajaran 2010/2011 di Purbalingga

No	Nama Sekolah	Jumlah Data
1	SMA N 1 Bukateja	271
2	SMA N 1 Kejobong	150
3	SMA N 1 Purbalingga	353
4	SMA N 2 Purbalingga	252
5	SMA N 1 Kutasari	199
6	SMA N 1 Bobotsari	309
7	SMA N 1 Rembang	233
8	SMA N 1 Karangreja	139
9	SMA N 1 Padamara	149
10	SMA N 1 Kemangkon	92
	Total Data	2147

C. Teknik Pengumpulan Data

Data dikumpulkan dengan teknik dokumentasi. Data penelitian yang dianalisis berupa paket soal yang terdiri : (1) lembar jawaban siswa, (2) lembar soal, (3) lembar kunci jawaban, dan (4) kisi-kisi soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Purbalingga tahun pelajaran 2010/2011.

Data lembar jawaban siswa dihimpun oleh peneliti dari sepuluh SMA Negeri di Kabupaten Purbalingga. Data soal, kunci jawaban, dan kisi-kisi diperoleh dari ketua Musyawarah Guru Mata Pelajaran (MGMP) Bahasa Indonesia SMA Kabupaten Purbalingga.

D. Teknik Analisis Data

Analisis data dengan pendekatan kualitatif dan kuantitatif. Analisis secara kualitatif dilakukan untuk mengetahui validitas isi soal melalui penelaahan butir

soal. Telaah butir soal dilakukan dari aspek materi, aspek konstruksi, dan aspek bahasa.

Pendekatan kuantitatif dilakukan melalui analisis empiris terhadap butir soal yang direspon oleh peserta tes. Penelitian ini menggunakan analisis butir soal berdasarkan teori tes klasik, dengan program *MicroCat Iteman* versi 3.00. *Iteman* merupakan program komputer yang digunakan untuk menganalisis butir soal secara klasik. Program ini termasuk satu paket program dalam *MicroCat* yang dikembangkan oleh *Assessment Systems Corporation* pada tahun 1982 dan mengalami revisi pada tahun 1984, 1986, 1988, dan 1993. (<http://www.scribd.com>). Analisis kualitatif dilakukan untuk mengetahui reliabilitas dan analisis butir soal yang mencakup indeks kesulitan, indeks daya beda, serta efektivitas distraktor. Kedua analisis ini merupakan upaya mengetahui kualitas soal yang baik.

1. Analisis Kualitatif

Analisis kualitatif dilakukan untuk mengetahui validitas isi soal. Analisis ini dilakukan untuk mengetahui kesesuaian butir soal ditinjau dari aspek materi, konstruksi, dan bahasa. Telaah terhadap butir soal yang disusun pada dasarnya adalah penilaian (*judgement*) pada setiap butir soal dengan kriteria telaah dari aspek materi, konstruksi, dan bahasa. Penelaahan dilakukan oleh 2 orang guru yang mempunyai kompetensi mata pelajaran yang diujikan dan tergabung dalam Musyawarah Guru Mata Pelajaran (MGMP), serta satu orang dosen yang mempunyai kompetensi bidang studi yang diujikan. Tiga penelaah tersebut adalah sebagai berikut.

1. Nama penelaah : Dra. Henifah
Tempat tugas : SMA Negeri 1 Rembang
Bidang Keahlian : Guru Bahasa Indonesia
2. Nama penelaah : Kaozal Dadi Legawan, S.Pd.
Tempat tugas : SMA Negeri 1 Kejobong
Bidang Keahlian : Guru Bahasa Indonesia
3. Nama penelaah : Kastam Syamsi, M. Ed.
Tempat tugas : Universitas Negeri Yogyakarta
Bidang Keahlian : Dosen Pengajaran Bahasa

Pedoman penelaahan butir soal yang digunakan untuk menelaah butir soal mengacu pada pedoman yang telah ditetapkan oleh Departemen Pendidikan Nasional (2003: 78-79) yaitu:

- a. Aspek Materi
 - 1) Butir soal sesuai dengan indikator
 - 2) Hanya ada satu kunci atau jawaban yang benar
 - 3) Isi materi sesuai dengan tujuan pengukuran
 - 4) Isi materi sesuai dengan jenjang, jenis sekolah dan tingkat kelas
 - 5) Pilihan benar-benar berfungsi, jika pilihan merupakan hasil perhitungan, maka pengecoh berupa pilihan yang salah rumus/salah hitung
- b. Aspek Konstruksi
 - 6) Pokok soal (stem) dirumuskan dengan jelas
 - 7) Rumusan soal dan pilihan dirumuskan dengan tegas
 - 8) Pokok soal tidak memberi petunjuk/mengarah kepada pilihan jawaban yang benar
 - 9) Pokok soal tidak mengandung pernyataan negatif ganda
 - 10) Bila terpaksa menggunakan kata negatif, maka harus digaris bawahi atau dicetak lain
 - 11) Pilihan jawaban homogen
 - 12) Hindari adanya alternatif jawaban: “seluruh jawaban di atas benar” atau “tak satu jawaban di atas yang benar” dan yang sejenisnya
 - 13) Panjang alternatif/pilihan jawaban relatif sama, jangan ada yang sangat panjang dan ada yang sangat pendek
 - 14) Pilihan jawaban dalam bentuk angka/waktu diurutkan
 - 15) Wacana, gambar atau grafik benar-benar berfungsi
 - 16) Antar butir tidak bergantung satu sama lain

c. Aspek Bahasa

- 17) Rumusan kalimat komunikatif
- 18) Kalimat menggunakan bahasa yang baik dan benar sesuai dengan jenis bahasanya
- 19) Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian
- 20) Menggunakan bahasa/kata yang umum (bukan bahasa lokal)
- 21) Rumusan soal tidak menyinggung perasaan siswa

Hasil telaah tersebut, dirangkum dan ditentukan tingkat karakteristik perangkat tes yang memenuhi kriteria sebagai berikut.

a. Menghitung persentase validitas dalam butir soal

$$\text{Rumus: } \frac{\text{Jumlah aspek yang sesuai}}{\text{Jumlah seluruh aspek telaah}} \times 100\%$$

b. Menentukan kualitas validitas isi tes. Penentuan tinggi rendahnya kualitas validitas isi tes ditentukan dengan menggunakan kriteria sebagai berikut.

Tabel 2: **Kriteria Kualitas Validitas Isi Soal**

No.	Persentase butir soal	Kriteria
1	90% sampai dengan 100%	Sangat baik
2	80% sampai dengan 89%	Baik
3	70% sampai dengan 79%	Cukup baik
4	60% sampai dengan 69%	Sedang
5	$\leq 59\%$	Kurang baik

2. Analisis Kuantitatif

Respon siswa terhadap perangkat tes dianalisis dengan menggunakan paket program *MicroCat Iteman* versi 3.00. Hasil pengolahan data dengan *Iteman* mencakup informasi mengenai besarnya indeks reliabilitas dan analisis butir soal.

Analisis butir soal meliputi indeks kesukaran butir, indeks daya beda, dan efektivitas distraktor.

a. Reliabilitas

Indeks reliabilitas tes dilihat pada nilai *Alpha Cronbach* pada *Output Iteman*, semakin besar nilainya menunjukkan semakin kecil tingkat kesalahan pengukuran. Semakin tinggi koefisien reliabilitas (mendekati 1), maka semakin tinggi pula keajegan atau ketepatannya (Depdiknas, 2004: 31). Reliabilitas soal yang relevan dengan penelitian ini adalah sebesar $\geq 0,90$.

b. Analisis Butir Soal

1) Indeks Kesukaran Butir Soal

Indeks kesukaran butir soal (p) adalah proporsi peserta tes yang menjawab benar butir soal. Indeks kesukaran butir soal dalam program *Iteman* dinyatakan dalam *prop. Correct*. Besarnya tingkat kesukaran yang diterima untuk menyatakan butir soal yang baik adalah antara 0,25 sampai 0,75. Artinya butir dinyatakan gugur apabila memiliki tingkat kesukaran tinggi dan apabila butir tersebut memiliki tingkat kesukaran rendah. Butir soal yang telah dianalisis indeks kesukaraan butir soalnya, selanjutnya dibuat persentase indeks kesukaran keseluruhan butir soal. Indeks kesukaran dapat dinyatakan baik, sedang, atau rendah sesuai dengan klasifikasi sebagai berikut.

Tabel 3: **Klasifikasi Indeks Kesukaran**

No.	Persentase Indeks Kesukaran	Klasifikasi
1.	80% sampai dengan 100%	Baik
2.	60% sampai dengan 79%	Sedang
3.	< 60%	Kurang

2) Indeks Daya Beda

Daya pembeda butir soal adalah besarnya korelasi antara rata-rata skor peserta yang menjawab benar pada butir soal dengan rata-rata skor total. Menentukan daya beda butir dapat dilihat dari besar nilai korelasi biseral (*Point Biser*). Butir soal yang memiliki indeks daya beda $\geq 0,30$ dinyatakan baik dan butir soal yang indeks daya beda $< 0,30$ dinyatakan tidak baik. Hal ini hanya berlaku pada butir soal yang menggunakan acuan norma dalam menafsirkan hasil tes.

SMA Negeri di Kabupaten Purbalingga menggunakan Kurikulum Tingkat Satuan Pendidikan (KTSP), sehingga penafsirannya menggunakan acuan kriteria. Indeks daya beda pada penafsiran acuan kriteria tidak berfungsi atau diabaikan. Indeks daya beda hanya akan dideskripsikan sebagai hasil analisis *Iteman*. Setiap butir soal dinyatakan baik asal indeks daya beda tidak bernilai negatif.

3) Efektivitas Distraktor

Proportion Endorsing pada *print out* hasil *Iteman* menunjukkan efektivitas distraktor. Distribusi jawaban (distraktor) yang berfungsi dapat ditentukan dengan kriteria sebagai berikut.

- a) Distraktor dikatakan berfungsi dengan baik atau efektif jika direspon oleh peserta tes minimal 2% pada pengecoh dan *point biser* pada alternatif jawaban negatif kecuali kunci jawaban.
- b) Distraktor belum berfungsi dengan baik atau tidak efektif, apabila respon siswa memilih pengecoh kurang dari 2% dan angka *point biser* pada alternatif jawaban bernilai positif selain kunci jawaban.
- c) Distraktor dinyatakan gugur apabila tidak ada peserta tes yang memilih distraktor tersebut.

Butir soal yang telah dianalisis keefektifan distraktornya, selanjutnya dibuat persentase efektivitas distraktor secara keseluruhan. Efektivitas distraktor dapat dinyatakan baik, sedang, atau rendah sesuai dengan klasifikasi sebagai berikut.

Tabel 4: **Klasifikasi Efektivitas Distraktor**

No.	Persentase Efektivitas Distraktor	Klasifikasi
1.	80% sampai dengan 100%	Baik
2.	60% sampai dengan 79%	Sedang
3.	< 60%	Kurang

3. Kriteria Kualitas Perangkat Soal

Kualitas perangkat soal dapat dilihat dari hasil analisis secara kualitatif dan kuantitatif. Analisis kualitatif digunakan untuk melihat validitas isi soal. Analisis kuantitatif digunakan untuk mengetahui reliabilitas soal dan analisis butir soal. Pemaknaan terhadap hasil analisis butir soal dilakukan dengan mempertimbangkan acuan kriteria, yaitu dengan memperhitungkan indeks kesulitan dan efektivitas distraktor, sedangkan daya beda butir soal diabaikan.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

A. Deskripsi Hasil Penelitian

Bab ini mendeskripsikan hasil penelitian yang telah dilakukan. Analisis karakteristik soal Ulangan Akhir Semester (UAS) Genap mata pelajaran Bahasa Indonesia kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011 dilakukan dengan cara kualitatif dan kuantitatif. Analisis kualitatif dikerjakan untuk mendapatkan validitas isi soal. Analisis kuantitatif dilaksanakan untuk mendapatkan indeks reliabilitas soal, indeks kesukaran, indeks daya beda, dan efektivitas distraktor. Analisis kuantitatif dilakukan melalui program *MicroCat Iteman* versi 3.00.

1. Hasil Analisis Kualitatif

Analisis kualitatif dilakukan untuk mengetahui validitas isi soal ditinjau dari aspek materi, konstruksi, dan bahasa. Tiga orang ahli yang telah menelaah butir soal, memperoleh hasil sebagai berikut.

Tabel 5: Telaah Validitas Isi Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Kabupaten Purbalingga oleh Tiga Orang Penelaah

No. Soal	Telaah validitas isi soal			
	A	B	C	Rata-rata
1	100%	100%	100%	100%
2	100%	100%	100%	100%
3	100%	100%	100%	100%
4	100%	100%	100%	100%
5	100%	100%	100%	100%
6	100%	100%	100%	100%
7	100%	100%	100%	100%
8	94,74%	94,74%	94,74%	94,74%
9	100%	100%	100%	100%
10	100%	94,74%	100%	98,25%
11	94,74%	89,47%	100%	94,74%
12	100%	100%	100%	100%
13	94,74%	94,74%	94,74%	94,74%
14	100%	100%	100%	100%
15	100%	100%	100%	100%
16	100%	100%	100%	100%
17	100%	100%	100%	100%
18	100%	100%	100%	100%
19	100%	100%	100%	100%
20	100%	100%	100%	100%
21	100%	100%	100%	100%
22	100%	100%	100%	100%
23	100%	100%	100%	100%
24	100%	94,74%	100%	98,25%
25	100%	94,74%	100%	98,25%
26	100%	100%	100%	100%
27	100%	100%	100%	100%
28	100%	100%	100%	100%
29	100%	100%	100%	100%
30	100%	100%	100%	100%
31	100%	100%	100%	100%
32	100%	100%	94,74%	98,25%
33	100%	100%	100%	100%
34	100%	100%	100%	100%
35	100%	100%	100%	100%
36	100%	100%	100%	100%
37	95%	95%	95%	95%
38	100%	100%	100%	100%
39	100%	100%	100%	100%
40	100%	94,74%	100%	98,25%
41	94,74%	100%	94,74%	96,49%

42	94,74%	100%	100%	98,25%
43	100%	100%	100%	100%
44	100%	100%	100%	100%
45	100%	100%	100%	100%
46	100%	100%	100%	100%
47	100%	100%	100%	100%
48	100%	100%	100%	100%
49	100%	100%	100%	100%
50	100%	100%	100%	100%
$\Sigma A=99,37\%$		$\Sigma B=99,06\%$	$\Sigma C=99,48\%$	$\Sigma ABC=99,30\%$

Penelaah memberikan telaahnya pada tempat dan waktu yang berbeda. Hasil telaah dari masing-masing penelaah memberikan hasil berbeda. Hasil berbeda tersebut terdapat pada nomor 10, 11, 24, 25, 32, 40, 41, dan 42. Hasil telaah oleh tiga orang penelaah disajikan pada tabel sebagai berikut.

Tabel 6: Rangkuman Hasil Telaah Validitas Isi Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Kabupaten Purbalingga oleh Tiga Orang Penelaah

Penelaah	Persentase syarat validitas	Hasil telaah validitas isi soal			Ket.
		Jumlah butir soal	Persentase butir soal	Rata-rata kesesuaian	
A	100%	44	88%	99,37%	Sangat baik
	95%	1	2%		
	94,74%	5	10%		
B	100%	42	84%	99,06%	Sangat baik
	95%	1	2%		
	94,74%	6	12%		
	89,47%	1	2%		
C	100%	45	90%	99,48%	Sangat baik
	95%	1	2%		
	94,74%	4	8%		
ΣABC				99,30%	Sangat baik

2. Hasil Analisis Kuantitatif

a. Reliabilitas Soal

Hasil analisis lewat program *Iteman* menunjukkan bahwa koefisien *Alpha* = 0.646. Hal ini menunjukkan bahwa perangkat tes tersebut memiliki reliabilitas yang kurang karena masih dibawah 0,90. Rendahnya indeks reliabilitas menunjukkan bahwa perangkat tes tersebut kurang reliabel sehingga hasilnya belum bisa dengan baik menggambarkan kemampuan sebenarnya yang dimiliki peserta tes. Tidak reliabelnya perangkat tes menyebabkan tingginya kesalahan pengukuran. Kesalahan pengukuran dapat disebabkan oleh banyak faktor seperti lingkungan ketika proses UAS yang kurang mendukung, kondisi psikologis maupun biologis peserta tes atau terjadinya kecurangan ketika pelaksanaan ujian seperti saling kerjasama antara peserta atau adanya bocoran kunci jawaban.

b. Analisis Butir Soal

1) Indeks Kesukaran

Pengujian tingkat kesukaran butir soal dilakukan dengan menghitung banyaknya siswa yang menjawab benar melalui program *Iteman*. Hasil pengujian menunjukkan tingkat kesukaran tiap-tiap butir soal tergolong mudah. Namun, dilihat dari nilai *Mean p* pada hasil *Iteman* menunjukkan angka sebesar 0,736, artinya indeks kesukaran rata-rata butir soal tergolong baik. Besarnya tingkat kesukaran yang diterima untuk menyatakan butir soal yang baik adalah antara 0,25 sampai 0,75. Dilihat dari indeks tingkat kesukaran, butir yang termasuk sulit ada 3 butir soal, sedang ada 15 butir soal, dan mudah ada 32 butir soal dengan perbandingan 6% sulit, 30% sedang, dan 64% mudah. “Sedang” di sini memiliki

arti bahwa soal tersebut baik, yaitu butir soal itu tidak terlalu mudah dan tidak terlalu sulit. Hasil pengujian tingkat kesukaran tersebut adalah sebagai berikut.

Tabel 7: Hasil Pengujian Tingkat Kesukaran Butir Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Kabupaten Purbalingga

Tingkat kesukaran	Butir soal	Jumlah	Persentase
Sulit (... < 0,25)	21, 42, 47	3	6%
Sedang/baik (0,25 – 0,75)	6, 7, 11, 20, 24, 25, 27, 29, 30, 36, 38, 39, 45, 48, 50	15	30%
Mudah (... > 0,75)	1, 2, 3, 4, 5, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 22, 23, 26, 28, 31, 32, 33, 34, 35, 37, 40, 41, 43, 44, 46, 49	32	64%

2) Indeks Daya Beda

Pengujian daya beda dapat dilihat pada *point biser* dalam hasil *Iteman*. Dilihat dari indeks daya beda, butir yang baik ($\geq 0,30$) ada 11 butir soal (22%). Butir soal yang tidak baik ada 39 butir soal (78%). Nilai *point biser* dalam hasil *Iteman* tidak ada yang bernilai negatif. Hasil pengujian daya beda sebagai berikut.

Tabel 8: Hasil Pengujian Daya beda Butir Soal Ulangan Akhir Semester (UAS) Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Purbalingga

Tingkat kesukaran	Butir soal	Jumlah soal	Persentase
Baik ($\geq 0,30$)	3, 6, 7, 25, 29, 30, 34, 44, 45, 46, 50	11	22%
Tidak baik ($< 0,30$)	1, 2, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 31, 32, 33, 35, 36, 37, 38, 39, 40, 41, 42, 43, 47, 48, 49	39	78%

3) Efektivitas Distraktor

Soal Ulangan Akhir Semester (UAS) Genap berjumlah 50 butir soal berbentuk objektif dengan 5 alternatif jawaban. Tiap soal memiliki satu jawaban dan empat distraktor atau pengecoh. Total distraktor dalam Ulangan Akhir Semester (UAS) Genap ini ada 200 distraktor. Distraktor dapat dipakai apabila jumlah pemilihnya lebih dari 2%. Jika distraktor tersebut pemilihnya kurang dari 2% maka perlu direvisi atau diperbaiki. Distraktor akan digugurkan apabila tidak ada (0%) peserta tes yang memilih distraktor tersebut.

Terdapat distraktor yang tidak dipilih oleh peserta tes (0%), yaitu distraktor D pada soal nomor 2, distraktor C pada soal nomor 14, distraktor A pada soal nomor 28, distraktor D pada soal nomor 37, distraktor E pada soal nomor 43, dan distraktor A pada soal nomor 44. Rangkuman analisis efektivitas distraktor selanjutnya disajikan dalam tabel di bawah ini.

**Tabel 9: Efektivitas Distraktor Butir Soal Ulangan Akhir Semester (UAS)
Genap SMA Kelas XII Tahun Pelajaran 2010/2011 di Purbalingga**

No soal	Kunci jawaban	Efektivitas distraktor			Keterangan
		Berfungsi ($\geq 2\%$)	Tidak berfungsi ($\leq 2\%$)	Gugur (0%)	
1	A	B	C, D, E	-	Tidak efektif pada distraktor C, D, E
2	B	A	C, E	D	Tidak efektif pada distraktor C, E Distraktor D gugur
3	E	D	A, B, C	-	Tidak efektif pada distraktor A, B, C
4	E	B, C	A, D	-	Tidak efektif pada distraktor A, D
5	B	-	A, C, D, E	-	Tidak efektif pada distraktor A, C, D, E
6	B	A, C, D, E	-	-	Efektif
7	D	B, C, E	A	-	Tidak efektif pada distraktor A
8	D	A, C	B, E	-	Tidak efektif pada distraktor B, E
9	C	A, D	B, E	-	Tidak efektif pada distraktor B, E
10	C	A	B, D, E	-	Tidak efektif pada distraktor B, D, E
11	A	B, E	C, D	-	Tidak efektif pada distraktor C, D
12	A	C	B, D, E	-	Tidak efektif pada distraktor B, D, E
13	E	C	A, B, D	-	Tidak efektif pada distraktor A, B, D
14	E	D	A, B	C	Tidak efektif pada distraktor A, B Distraktor C gugur
15	B	A, C, E	D	-	Tidak efektif pada distraktor D
16	B	A, C	D, E	-	Tidak efektif pada distraktor D, E
17	C	B	A, D, E	-	Tidak efektif pada distraktor A, D, E
18	C	A, D, E	B	-	Tidak efektif pada distraktor B
19	D	-	A, B, C, E	-	Tidak efektif pada distraktor A, B, C, E
20	D	A, B, C, E	-	-	Efektif
21	E	B	A, C, D	-	Tidak efektif pada distraktor A, C, D
22	E	B, C	A, D	-	Tidak efektif pada distraktor A, D
23	D	B	A, C, E	-	Tidak efektif pada distraktor A, C, E
24	C	A, B, D, E	-	-	Efektif
25	B	A, C, D, E	-	-	Efektif
26	A	B, C	D, E	-	Tidak efektif pada distraktor D, E
27	B	D	A, C, E	-	Tidak efektif pada distraktor A, C, E
28	C	B	D, E	A	Tidak efektif pada distraktor D, E Distraktor A gugur
29	D	C	A, B, E	-	Tidak efektif pada distraktor A, B, E
30	E	A, B, C, D	-	-	Efektif
31	E	-	A, B, C, D	-	Tidak efektif pada distraktor A, B, C, D
32	D	E	A, B, C	-	Tidak efektif pada distraktor A, B, C

33	C	A	B, D, E	-	Tidak efektif pada distraktor B, D, E
34	B	A, C	D, E	-	Tidak efektif pada distraktor C, D, E
35	A	-	B, C, D, E	-	Tidak efektif pada distraktor B, C, D, E
36	A	D	B, C, E	-	Tidak efektif pada distraktor B, C, E
37	B	E	A, C	D	Tidak efektif pada distraktor A, C Distraktor D gugur
38	C	D, E	A, B	-	Tidak efektif pada distraktor A, B
39	D	A, B, C	E	-	Tidak efektif pada distraktor E
40	E	A, C	B, D	-	Tidak efektif pada distraktor B, D
41	A	D, E	B, C	-	Tidak efektif pada distraktor B, C
42	B	D, E	A, C	-	Tidak efektif pada distraktor A, C
43	C	A, D	B, E	E	Tidak efektif pada distraktor B, E Distraktor E gugur
44	D	C, E	A, B	A	Tidak efektif pada distraktor A, B Distraktor A gugur
45	E	A, B, C, D	-	-	Efektif
46	A	B, E	C, D	-	Tidak efektif pada distraktor C, D
47	B	A, C, E	D	-	Tidak efektif pada distraktor D
48	C	B, D, E	A	-	Tidak efektif pada distraktor A
49	D	-	A, B, C, E	-	Tidak efektif pada distraktor A, B, C, E
50	E	A, B, C, D	-	-	Efektif

Penelitian ini menunjukkan dari 200 distraktor, ada 93 distraktor (46,5%) yang dapat dipakai atau efektif, 101 (50,5%) distraktor yang perlu diperbaiki, dan 6 (3%) distraktor digugurkan. Sebanyak 5 butir (10%) memiliki distraktor yang semuanya tidak baik, 7 (14%) butir yang semua distraktornya berfungsi atau efektif, serta 38 (76%) butir yang distraktornya ada yang berfungsi dan ada yang tidak berfungsi.

3. Pemaknaan Hasil Analisis Butir Soal

Setelah perhitungan statistik berupa indeks kesulitan, indeks daya beda, dan efektivitas distraktor setiap butir soal selesai dilakukan, langkah selanjutnya adalah peneliti melakukan pemaknaan terhadap hasil analisis butir soal. Pemaknaan

terhadap hasil analisis butir soal ini dilakukan dengan mempertimbangkan acuan yang digunakan dalam pemaknaan hasil tes, yaitu acuan kriteria. Parameter yang digunakan dalam penelitian ini adalah indeks kesulitan dan efektivitas distraktor.

Hasil analisis butir soal, dari segi indeks kesulitan butir soal, sebanyak 15 butir soal (30%) memiliki indeks kesulitan yang baik dan 35 butir soal (70%) memiliki indeks kesulitan yang tidak baik, baik itu terlalu mudah atau terlalu sukar. Efektivitas distraktor yang dapat dipakai sebanyak 86 distraktor, 108 distraktor yang perlu diperbaiki, dan 6 distraktor digugurkan.

4. Kunci Jawaban yang Perlu Dicek

Hasil analisis menggunakan *Iteman* menunjukkan bahwa satu butir soal (2%) yang perlu dicek kuncinya. Hal ini ditandai dengan munculnya tanda *check the key* pada hasil *print out* analisis kuantitatif melalui program *Iteman*. Tanda *check the key* terdapat pada butir soal nomor 4.

B. Pembahasan Hasil Penelitian

Hasil penelitian yang telah dijabarkan pada bagian sebelumnya, selanjutnya akan dibahas pada bagian ini. Berikut pembahasan hasil analisisnya.

1. Pembahasan Hasil Analisis Kualitatif

Penelaah memberikan telaahnya pada tempat dan waktu yang berbeda. Hasil telaah dari masing-masing penelaah memberikan hasil berbeda. Hasil berbeda tersebut terdapat pada nomor 10, 11, 24, 25, 32, 40, 41, dan 42. Perbedaan penelaah dapat disebabkan oleh berbagai faktor. Hasil telaah dari masing-masing penelaah dapat dilihat pada lampiran.

Penelaah A (Dra. Henifah) memberikan hasil telaah soalnya dengan rata-rata 99,37%. Sebanyak 44 butir soal (88%) telah 100% memenuhi syarat validitas isi soal, 1 butir soal (2%) memenuhi syarat validitas sebesar 95%, dan 5 butir soal (10%) memenuhi syarat validitas sebesar 94,74%. Kriteria kualitas validitas isi tes yang telah ditentukan dalam bab sebelumnya, menunjukkan bahwa butir tes ini tergolong sangat baik karena memiliki kesesuaian telaah soal lebih dari 90% yaitu, 99,37%.

Penelaah B (Kaozal Dedi Legawan, S.Pd.) memberikan hasil telaah soalnya dengan rata-rata 99,06%. Sebanyak 42 butir soal (84%) telah 100% memenuhi syarat validitas isi soal, 1 butir soal (2%) memenuhi syarat sebesar 89,47%, 6 butir soal (12%) memenuhi syarat sebesar 94,74%, dan 1 butir soal (2%) memenuhi syarat sebesar 95%. Kriteria kualitas validitas isi tes yang telah ditentukan, menunjukkan bahwa butir tes ini tergolong sangat baik karena memiliki kesesuaian telaah soal lebih dari 90% yaitu, 99,06%.

Penelaah C (Kastam Syamsi, M.Ed.) memberikan hasil telaah dengan rata-rata 99,48%. Sebanyak 45 butir soal (90%) telah 100% memenuhi syarat validitas isi soal, 1 soal (2%) memenuhi syarat sebesar 95%, dan 4 soal (8%) memenuhi syarat sebesar 94,74%. Kriteria kualitas validitas isi tes yang ditentukan menunjukkan bahwa butir tes ini tergolong sangat baik karena memiliki kesesuaian telaah soal lebih dari 90% yaitu, 99,48%.

Validitas isi soal UAS genap kelas XII SMA Negeri di Purbalingga secara keseluruhan menunjukkan kesesuaian telaah soal sebesar 99,30%. Sebanyak 39 butir soal (78%) telah 100% memenuhi syarat validitas isi soal, 6 soal (12%) memenuhi syarat validitas sebesar 98,25%, 1 soal (2%) memenuhi syarat validitas sebesar 96,49%, 1 soal (2%) memenuhi syarat validitas sebesar 95%, 3 soal (6%) memenuhi syarat validitas sebesar 94,74%. Kriteria kualitas validitas isi tes yang telah ditentukan, menunjukkan bahwa butir tes ini tergolong sangat baik karena hasil telaah yang dilakukan oleh tiga penelaah menunjukkan rata-rata lebih dari 90%, yaitu sebesar 99,30%.

Kualitas soal pilihan ganda dapat lebih ditingkatkan apabila penulisannya juga mengikuti kaidah penulisan soal yang telah ditetapkan. Kaidah-kaidah penulisan soal merupakan petunjuk atau pedoman yang perlu diikuti oleh penulis soal agar soal yang dihasilkan memiliki kualitas yang baik. Soal yang kualitasnya baik adalah soal yang mampu menyaring informasi yang diperlukan dan berfungsi secara optimal. Kaidah penulisan soal pilihan ganda mencakup tiga aspek yaitu materi, konstruksi, dan bahasa.

Butir soal dikatakan baik apabila tiga segi, yaitu materi, konstruksi, dan bahasa baik. Butir soal yang 100% sesuai dari ketiga penelaah adalah butir soal

nomor 1, 2, 3, 4, 5, 6, 7, 9, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 38, 39, 43, 44, 45, 46, 47, 48, 49, dan 50. Sebagai contoh, soal nomor: 1, 2, dan 3 seperti kutipan di bawah ini.

Cermati kutipan teks berikut untuk soal 1-4!

Tata niaga elpiji yang dilakukan pemerintah perlu diperjelas (2) Elpiji yang sudah menjadi komoditas yang menguasai hajat hidup orang banyak selayaknya diatur penjualannya sebaga mana bahan bakar minyak. (3) Harga elpiji yang selama ini ditetapkan oleh Pertamina memunculkan monopoli. (4) Seharusnya, elpiji yang mengatur adalah pemerintah, bukan 'korporasi. (5) Apalagi dengan adanya program konversi minyak tanah ke elpiji seperti sekarang, pemakaiannya sudah sedemikian luas (6) Terhitung mulai 1 Juli, PT Pertamina menaikkan harga elpiji kemasan 12 kilogram dari Rp 4.250,00 per kg menjadi Rp 5.250,00 per kg. (7) Dengan kenaikan 23 persen itu, harga elpiji tabung 12 kg naik dari RP 53.000,00 per tabung menjadi Rp 63.000,00 per tabung. (8) Harga tersebut berlaku untuk agen dalam radius 60 kilometer dari instalasi pengisian elpiji Pertamina. (9) Di luar jangkauan itu, agen menambahkan biaya angkutan sesuai aturan Menteri Perhubungan.

1. Kalimat utama paragraf tersebut terdapat pada nomor
 - a. (1)
 - b. (2)
 - c. (3)
 - d. (4)
 - e. (5)
2. Ide pokok paragraf tersebut adalah
 - a. harga jual elpiji
 - b. tata niaga elpiji
 - c. perpindahan ke elpiji
 - d. kemasan baru elpiji
 - e. instalasi pengisian elpiji
3. Kalimat fakta dalam kutipan tersebut terdapat pada nomor
 - a. (1), (3), (4), dan (5)
 - b. (3), (4), (5), dan (6)
 - c. (4), (5), (6), dan (7)
 - d. (5), (6), (7), dan (8)
 - e. (6), (7), (8), dan (9)

Segi materi, soal di atas termasuk soal yang baik karena soal-soal tersebut memenuhi tuntutan indikator yang tertera pada kisi-kisi. Indikator nomor 1 dan 2

adalah “disajikan sebuah paragraf yang di dalamnya terdapat kalimat utama dan salah satu kalimat penjelasnya tidak padu, siswa dapat menentukan: kalimat utama paragraf, ide pokok paragraf”. Indikator soal nomor 3 adalah “disajikan sebuah paragraf yang di dalamnya terdapat fakta/opini, siswa dapat menentukan fakta/opini dalam paragraf.” Soal di atas juga memiliki satu kunci jawaban yang benar. Kunci jawaban soal nomor 1 adalah A, nomor 2 adalah B, dan nomor 3 adalah E.

Segi konstruksi, soal-soal tersebut juga tergolong baik. Semua aspek soal yang terdapat pada lembar telaah telah sesuai. Aspek telaah butir soal tersebut antara lain: pokok soal dirumuskan dengan jelas, panjang alternatif/pilihan jawaban relatif sama, pilihan jawaban dalam bentuk angka/waktu diurutkan, dan wacana, gambar, atau grafik benar-benar berfungsi.

Segi bahasa, semua aspek telaah juga sesuai. Hal ini berarti soal tersebut tergolong baik. Aspek telaah butir soal dari segi bahasa antara lain: rumusan kalimat komunikatif, kalimat menggunakan bahasa yang baik dan benar, serta rumusan kalimat tidak menimbulkan penafsiran ganda.

Pembahasan tersebut dapat disimpulkan bahwa soal nomor 1, 2, dan 3 telah memenuhi kriteria dan tergolong soal yang baik. Soal di atas sesuai dengan aspek telaah soal segi materi, konstruksi, dan bahasa.

Butir soal yang kriteria penelaahan tidak 100% terpenuhi terdapat pada aspek konstruksi, dan bahasa. Aspek konstruksi yang tidak terpenuhi yaitu kriteria penelaahan nomor 13: panjang alternatif/pilihan jawaban relatif sama, jangan ada yang sangat panjang dan ada yang sangat pendek, dan kriteria penelaahan nomor 14: pilihan jawaban dalam bentuk angka/waktu diurutkan. Kriteria aspek bahasa

yang tidak terpenuhi adalah kriteria penelaahan nomor 18: kalimat menggunakan bahasa yang baik dan benar sesuai dengan jenis bahasanya.

Butir soal yang tidak 100% sesuai dengan aspek telaah adalah butir soal nomor 8, 10, 11, 13, 24, 25, 32, 37, 40, 41, 42. Sebagai contoh butir soal nomor 8, kriteria penelaahan yang tidak terpenuhi yaitu kriteria penelaahan nomor 18: kalimat menggunakan bahasa yang baik dan benar sesuai dengan jenis bahasanya. Kesalahan tersebut sebagaimana terdapat pada kutipan di bawah ini.

8. Mengapa negara-negara yang memiliki kekayaan tambang mineral sering menunjukkan harus menghadapi kenyataan bahwa harta itu lebih merupakan kutukan daripada berkah?

Jawaban pertanyaan tersebut sesuai isi teks adalah

- a. Kejayaan dan kekayaan negara akan berlimpah ruah dari penghasilan minyak yang merupakan komoditas beraroma KKN.
- b. Karena akan menghadapi kenyataan yang menyakitkan, yaitu menjadi salah satu negara importir minyak terbesar di dunia.
- c. Karena melimpahnya minyak di suatu negara masanya tidak akan lama bertahan, seperti di sejumlah negara di Afrika.
- d. Karena sering menjadi rebutan kelompok yang berkonflik atau perang saudara, korupsi, dan pengelolaan yang salah urus.
- e. Akan menjadi pengeksport serta penikmat booming harga minyak yang akan menjadikan negara kaya dan maju.

Kutipan di atas menggunakan kata 'adalah' sebanyak dua kali berturut-turut. Kalimat tersebut menjadi tidak baik karena menggunakan kata yang tidak diperlukan. Soal tersebut perlu direvisi dengan menghilangkan kata adalah.

Butir soal nomor 11 tidak memenuhi kriteria penelaahan nomor 18 yaitu kalimat menggunakan bahasa yang baik dan benar sesuai dengan jenis bahasanya. Kesalahan tersebut sebagaimana terdapat pada kutipan di bawah ini.

11. Simpulan data tabel tersebut adalah ...

- a. Pengangguran di Kalimantan Tengah pada Agustus 2006-Februari 2007 jumlahnya lebih dari separo jumlah pengangguran di Indonesia.
- b. Tabel menggambarkan jumlah angkatan kerja penduduk yang bekerja dan pengangguran pada tahun 2006-2007.

- c. Angka pengangguran di Indonesia dari Kalimantan Tengah besarnya hampir sama pada tahun 2006-2007.
- d. Penduduk Indonesia yang bekerja semakin berkurang jumlahnya dibandingkan dengan penduduk yang bekerja pada bulan Agustus 2006.
- e. Di Kalimantan Tengah angkatan kerja pada Agustus 2006-Februari 2007 jumlahnya lebih meningkat pada bulan Agustus 2006.

Kutipan nomor 11 di atas menggunakan kata yang tidak baku pada alternatif/pilihan jawaban A yaitu kata *separo*. Soal di atas memerlukan adanya revisi. Revisi tersebut terdapat pada kata yang tidak baku, yaitu *separo* diubah menjadi *separuh*. Butir soal nomor 13, tidak memenuhi kriteria penelaahan nomor 13 yaitu panjang alternatif/pilihan jawaban relatif sama. Hal ini dapat dilihat pada kutipan soal nomor 13 sebagai berikut.

13. Nilai moral baik yang terkandung dalam kutipan tersebut adalah ...
- a. Perbuatan menyembah kepada orang yang lebih tua merupakan suatu perbuatan yang tidak manusiawi.
 - b. Memperlakukan orang sebagai inang dan pengasuh yang harus tunduk perintah.
 - c. Perbuatan menyesali nasib buruk yang menimpa diri secara berlebihan.
 - d. Menyindir seseorang yang terlalu berlebihan sehingga membuat orang menangis.
 - e. Kita perlu meminta maaf atas kesalahan yang telah dilakukan.

Kutipan soal nomor 13 di atas menunjukkan bahwa panjang alternatif/pilihan jawaban ada yang terlalu panjang, yaitu alternatif A. Pilihan jawaban A terdiri atas 13 kata, B. 10 kata, C. 9 kata, D. 9 kata, dan E. 9 kata. Alternatif A yang terlalu panjang membuat pilihan jawaban panjangnya relatif tidak sama sehingga soal perlu direvisi.

Butir soal nomor 37, kriteria penelaahan yang tidak terpenuhi yaitu kriteria penelaahan nomor 14: pilihan jawaban dalam bentuk angka/waktu diurutkan. Hal ini dapat dilihat pada kutipan soal nomor 37 sebagai berikut.

37. Susunan kalimat-kalimat tersebut agar menjadi paragraf yang padu adalah ...

- a. (4), (3), (2), (1), (5)
- b. (3), (4), (1), (2), (5)
- c. (4), (1), (3), (2), (5)
- d. (1), (2), (3), (4), (5)
- e. (3), (4), (2), (5), (1)

Pada kutipan nomor 37 dapat diketahui bahwa pilihan jawaban dalam bentuk angka tidak diurutkan. Seharusnya pilihan jawaban tersebut diurutkan menjadi:

- a. (1), (2), (3), (4), (5)
- b. (3), (4), (1), (2), (5)
- c. (3), (4), (2), (5), (1)
- d. (4), (1), (3), (2), (5)
- e. (4), (3), (2), (1), (5)

2. Pembahasan Hasil Analisis Kuantitatif

Pembahasan hasil analisis kuantitatif meliputi hasil pengujian reliabilitas tes, pembahasan hasil analisis butir soal yang meliputi pembahasan hasil pengujian indeks tingkat kesukaran butir, daya beda butir soal, efektivitas distraktor, dan kunci jawaban yang perlu dicek.

a. Pembahasan Pengujian Reliabilitas Soal

Reliabilitas soal penelitian ini dapat dilihat dari nilai koefisien *Alpha* yang diperoleh dari program *Iteman* yaitu 0,646. Tes yang bersifat seragam dan buatan tim, koefisien *Alpha* sebesar 0,646 kurang tinggi karena ketentuan reliabilitas yang baik untuk tes buatan tim setidaknya mencapai angka 0,90.

Tinggi rendahnya indeks reliabilitas tes dipengaruhi oleh indeks kesalahan pengukuran. Semakin tinggi indeks reliabilitas akan semakin kecil indeks kesalahan pengukuran (*Standard Error of Measurement* atau SEM). SEM dalam hasil *print out Iteman* menunjukkan angka sebesar 2,523. Indeks kesalahan pengukuran dapat diperkecil dengan meningkatkan indeks reliabilitas soal. Hal ini dapat dilakukan dengan memperhatikan faktor-faktor yang dapat memengaruhi reliabilitas soal. Salah satu cara yang dapat dilakukan adalah menambah panjang tes yang digunakan atau menambah waktu mengerjakan tes.

b. Pembahasan Analisis Butir Soal

1) Pembahasan Indeks Kesukaran

Hasil pengujian tingkat kesukaran soal, dapat disimpulkan bahwa sebagian besar butir soal dinyatakan belum baik. Namun, dilihat dari nilai *Mean p* pada hasil *Iteman* menunjukkan angka sebesar 0,736, artinya indeks kesukaran rata-rata butir soal tergolong baik. Butir soal yang masuk dalam kategori sukar dan mudah

menunjukkan bahwa dari segi materi, butir soal tersebut belum representatif atau belum mewakili materi yang telah diajarkan.

Butir soal yang dinyatakan tidak baik hendaknya ditindaklanjuti oleh pembuat soal. Butir soal yang masuk dalam kategori sulit diteliti ulang sehingga dapat diketahui faktor yang menyebabkan butir soal yang bersangkutan sulit dijawab oleh peserta tes. Demikian juga dengan butir soal yang berada pada kategori mudah. Setelah dilakukan perbaikan, butir-butir soal tersebut dapat dikeluarkan kembali pada tes berikutnya atau disimpan di bank soal.

Butir soal yang sukar dijawab oleh peserta tes dapat disebabkan oleh beberapa hal. Pengecekan terhadap kunci jawaban perlu dilakukan. Apabila kunci jawaban sudah benar, sulitnya soal dimungkinkan karena materi yang ditanyakan belum diajarkan atau belum tuntas pembelajarannya, sehingga kompetensi minimum yang harus dikuasai siswa belum tercapai. Faktor lain yang berpengaruh adalah materi yang diukur tidak cocok ditanyakan dengan menggunakan bentuk soal yang diberikan dan pernyataan atau kalimat soal terlalu panjang.

2) Pembahasan Indeks Daya Beda

Hasil analisis butir soal melalui program *Iteman*, menunjukkan bahwa indeks daya beda butir yang baik ($\geq 0,30$) ada 11 butir soal (22%). Butir soal yang memiliki indeks daya beda yang tidak baik ada 39 butir soal (78%). Dengan demikian, dapat disimpulkan bahwa soal memiliki indeks daya beda yang kurang baik. Hal ini sesuai dengan nilai *Mean Item Tot* sebesar 0.230 yang menunjukkan indeks daya beda berada pada kategori kurang baik.

Butir soal yang memiliki indeks daya beda tidak baik ditindaklanjuti dengan perbaikan dan penyempurnaan pada pokok soalnya atau dibuang. Butir soal yang

memiliki indeks daya beda yang baik dapat digunakan pada tes berikutnya. Butir soal yang tidak mampu membedakan siswa pandai dan tidak pandai dimungkinkan karena kompetensi yang diukur tidak jelas dan pengecohnya tidak berfungsi. Materi yang ditanyakan terlalu sulit sehingga banyak siswa yang menebak juga dapat memengaruhi butir soal sehingga tidak dapat membedakan kedua kemampuan siswa tersebut.

SMA Negeri di Purbalingga telah menggunakan Kurikulum Tingkat Satuan Pendidikan (KTSP) sehingga indeks daya beda dapat diabaikan. Setiap butir soal dikatakan baik asal indeks daya beda tidak bernilai negatif. Butir soal yang indeks daya bedanya tidak baik tetap dinyatakan baik dan dapat dipakai menunjukkan efektivitas proses pembelajaran.

3) Pembahasan Efektivitas Distraktor

Penelitian ini, dari 200 distraktor, terdapat 93 distraktor (46,5%) yang dapat dipakai atau efektif. Distraktor yang kurang efektif ada 101 distraktor (50,5%). Distraktor yang digugurkan karena tidak ada peserta tes yang memilih (0%) ada 6 distraktor (3%).

Distraktor yang sudah baik dapat langsung diterima tanpa melakukan revisi. Distraktor yang baik adalah distraktor yang dipilih lebih dari 2% peserta tes dan memiliki *point biser* distraktor negatif. Distraktor direvisi apabila distraktor tersebut belum efektif, yaitu apabila pemilihnya kurang dari 2% dan atau memiliki *point biser* distraktor positif. Distraktor ditolak apabila tidak berfungsi karena dipilih oleh 0% peserta tes dan atau *point biser* distraktor positif.

Efektivitas distraktor dapat dipengaruhi oleh homogenitas distraktor dan panjang pendeknya distraktor. Distraktor yang homogen membuat siswa merasa

bingung memilih alternatif jawaban yang benar. Panjang pendeknya distraktor juga membuat siswa bingung menentukan jawaban yang benar.

3. Pembahasan Hasil Analisis Butir Soal

Pemaknaan terhadap hasil analisis butir soal dilakukan dengan mempertimbangkan acuan yang digunakan, yaitu acuan kriteria. Dalam penerapan acuan kriteria, parameter yang digunakan adalah indeks kesulitan dan efektivitas distraktor. Hasil analisis butir soal menunjukkan bahwa indeks kesulitan butir soal tergolong tidak baik karena dari 50 butir soal, hanya 15 soal (30%) yang memiliki indeks kesulitan baik. Efektivitas distraktor juga kurang baik karena hanya 46,5% distraktor saja yang efektif. Hal tersebut menunjukkan kurang dari setengah distraktor yang berfungsi.

Butir soal nomor 1 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,948. Hal ini berarti soal tergolong mudah karena sebanyak 94,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 3,1% peserta merespon distraktor B, 1,5% merespon distraktor C, 0,3% peserta merespon distraktor D dan E. Distraktor B merupakan distraktor yang efektif karena minimal telah dipilih oleh 2% peserta tes. Namun, tidak demikian dengan distraktor C, D, dan E yang pemilihnya kurang dari 2% sehingga ketiga distraktor tersebut tidak efektif. Distraktor D juga memiliki nilai *point biser* yang positif sehingga tidak efektif.

Butir soal nomor 2 tergolong mudah karena sebanyak 82% peserta tes merespon kunci jawaban. Hal ini dapat dilihat dari indeks kesulitannya sebesar 0,820. Ditinjau dari distribusi jawabannya, sebanyak 17% peserta merespon distraktor A, 0,9% merespon distraktor C, 0% peserta merespon distraktor D dan

0,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A merupakan distraktor yang efektif karena minimal telah dipilih oleh 2% peserta tes. Namun, tidak demikian dengan distraktor D dan E yang pemilihnya kurang dari 2% sehingga kedua distraktor tersebut tidak efektif. Tidak ada peserta tes yang memilih distraktor C sehingga distraktor tersebut digugurkan.

Butir soal nomor 3 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,791. Hal ini berarti soal tergolong mudah karena sebanyak 79,1% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,3% peserta merespon distraktor A, 0,6% merespon distraktor B, 0,9% peserta merespon distraktor C dan 19% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor D merupakan distraktor yang efektif karena minimal telah dipilih oleh 2% peserta tes. Namun, tidak demikian dengan distraktor A, B, dan C yang pemilihnya kurang dari 2% sehingga ketiga distraktor tersebut tidak efektif.

Butir soal nomor 4 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,840. Hal ini berarti soal tergolong mudah karena sebanyak 84% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 4,0% peserta merespon distraktor A, 7,9% merespon distraktor B, 2,7% peserta merespon distraktor C dan 1,4% peserta merespon distraktor D. *Point biser* pada distraktor A bernilai positif. Distraktor B, dan C merupakan distraktor yang efektif karena minimal telah dipilih oleh 2% peserta tes dan nilai *point biser* negatif. Namun, tidak demikian dengan distraktor D yang pemilihnya kurang dari 0,2% sehingga distraktor tersebut tidak efektif. Tanda positif pada distraktor A menunjukkan

bahwa peserta tes yang kemampuannya rendah cenderung dapat menjawab benar. Oleh karena itu, distraktor A tidak efektif.

Butir soal nomor 5 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,972. Hal ini berarti soal tergolong mudah karena sebanyak 97,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,1% peserta merespon distraktor A, 0,4% merespon distraktor C, 0,6% peserta merespon distraktor D dan 1,7% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, C, D, dan E merupakan distraktor yang tidak efektif karena pemilihnya kurang dari 2%.

Butir soal nomor 6 memiliki indeks kesulitan baik, yaitu sebesar 0,592. Hal ini berarti tingkat kesukaran soal tergolong sedang karena sebanyak 59,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 9,9% peserta merespon distraktor A, 16,9% merespon distraktor C, 10,2% peserta merespon distraktor D dan 3,7% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, C, D, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

Butir soal nomor 7 memiliki indeks kesulitan baik, yaitu sebesar 0,614. Hal ini berarti tingkat kesukaran soal tergolong sedang karena sebanyak 61,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1% peserta merespon distraktor A, 11,6% merespon distraktor B, 2,7% peserta merespon distraktor C dan 23,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B, C, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 8 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,825. Hal ini berarti soal tergolong mudah karena sebanyak 82,5% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 11,1% peserta merespon distraktor A, 1,8% merespon distraktor B, 3% peserta merespon distraktor C dan 1,4% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A dan C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 9 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,834. Hal ini berarti soal tergolong mudah karena sebanyak 83,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 10,5% peserta merespon distraktor A, 0,8% merespon distraktor B, 3,4% peserta merespon distraktor D dan 1,9% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A dan D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 10 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,878. Hal ini berarti soal tergolong mudah karena sebanyak 87,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 8,0% peserta merespon distraktor A, 1% merespon distraktor B, 1,4% peserta merespon distraktor D dan 1,6% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B, D, dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 11 memiliki indeks kesulitan baik, yaitu sebesar 0,294. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 29,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 57,3% peserta merespon distraktor B, 3,9% merespon distraktor C, 1,1% peserta merespon distraktor D dan 8,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B, C, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 12 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,928. Hal ini berarti soal tergolong mudah karena sebanyak 92,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1,6% peserta merespon distraktor B, 4,2% merespon distraktor C, 0,2% peserta merespon distraktor D dan 1,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B, D, dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 13 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,942. Hal ini berarti soal tergolong mudah karena sebanyak 94,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,7% peserta merespon distraktor A dan B, 2,7% merespon distraktor C, dan 1,5% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, B, dan D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 14 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,948. Hal ini berarti soal tergolong mudah karena sebanyak 94,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,3% peserta merespon distraktor A dan B, 0% merespon distraktor C, dan 4,5% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A dan B tidak efektif karena direspon kurang dari 2% peserta tes. Tidak ada peserta tes yang memilih distraktor C sehingga distraktor tersebut digugurkan.

Butir soal nomor 15 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,786. Hal ini berarti soal tergolong mudah karena 78,6% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 10,3% peserta merespon distraktor A, 4,2% merespon distraktor C, 1,4% peserta merespon distraktor D dan 5,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, C, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 16 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,919. Hal ini berarti soal tergolong mudah karena 91,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 5,3% peserta merespon distraktor A, 2% merespon distraktor C, 0,5% peserta merespon distraktor D dan 0,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A dan C merupakan distraktor yang efektif

karena pemilihnya lebih dari 2%. Di sisi lain, distraktor D dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 17 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,937. Hal ini berarti soal tergolong mudah karena 92,7% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1,7% peserta merespon distraktor A, 2,5% merespon distraktor B, 0,5% peserta merespon distraktor D dan 1,5% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, D, dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 18 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,882. Hal ini berarti soal tergolong mudah karena 88,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 2,8% peserta merespon distraktor A, 1,8% merespon distraktor B, 4,2% peserta merespon distraktor D dan 2,8% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, D, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 19 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,987. Hal ini berarti soal tergolong mudah karena 98,7% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,1% peserta merespon distraktor A dan B, 0,7% merespon distraktor C, dan 0,3% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif.

Distraktor A, B, C, dan E merupakan distraktor yang tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 20 memiliki indeks kesulitan baik, yaitu sebesar 0,544. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 54,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 8,3% peserta merespon distraktor A, 16,8% merespon distraktor B, 6,3% peserta merespon distraktor C dan 14,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, C, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

Butir soal nomor 21 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,205. Hal ini berarti soal tergolong sulit karena hanya 20,5% peserta tes yang merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 57,8% peserta merespon distraktor A, 20,7% merespon distraktor B, 0,8% peserta merespon distraktor C dan 0,2% peserta merespon distraktor E. *Point biser* pada distraktor A dan C bernilai positif. Distraktor B merupakan distraktor yang efektif karena pemilihnya lebih dari 2% dan nilai *point bisernya* negatif. D tidak efektif karena direspon kurang dari 2% peserta tes. Distraktor C tidak efektif karena direspon kurang dari 2% peserta tes dan *point bisernya* bernilai positif. Distraktor A juga tidak efektif karena nilai *point bisernya* positif.

Butir soal nomor 22 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,899. Hal ini berarti soal tergolong mudah karena 89,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,7% peserta merespon distraktor A, 2,7% merespon distraktor B, 6,4% peserta merespon distraktor C dan 0,3% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor

tersebut bertanda negatif. Distraktor B dan C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A dan D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 23 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,954. Hal ini berarti soal tergolong mudah karena 95,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,3% peserta merespon distraktor A, C dan E, 3,7% merespon distraktor B. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, C dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 24 memiliki indeks kesulitan baik, yaitu sebesar 0,448. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 44,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 19,3% peserta merespon distraktor A, 19,6% merespon distraktor B, 9,8% peserta merespon distraktor D dan 6,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, D, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

Butir soal nomor 25 memiliki indeks kesulitan baik, yaitu sebesar 0,422. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 42,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 11,8% peserta merespon distraktor A, 21,3% merespon distraktor C, 11% peserta merespon distraktor D dan 13,7% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, C D, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

Butir soal nomor 26 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,897. Hal ini berarti soal tergolong mudah karena 89,7% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 2,5% peserta merespon distraktor B, 5,9% merespon distraktor C, 0,4% peserta merespon distraktor D dan 1,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B dan C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor D dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 27 memiliki indeks kesulitan baik, yaitu sebesar 0,672. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 67,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,2% peserta merespon distraktor A, 0,7% merespon distraktor C, 31,1% peserta merespon distraktor D dan 0,7% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, C, dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 28 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,924. Hal ini berarti soal tergolong mudah karena 92,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0% peserta merespon distraktor A, 6,8% merespon distraktor B, 0,2% peserta merespon distraktor D dan 0,5% peserta merespon distraktor E. Distraktor B merupakan distraktor yang efektif karena pemilihnya lebih dari 2% dan *point biser*-nya bernilai negatif. Di sisi lain, distraktor D dan E tidak efektif karena direspon kurang dari 2% peserta tes. Tidak ada peserta tes yang memilih distraktor A sehingga distraktor tersebut digugurkan.

Butir soal nomor 29 memiliki indeks kesulitan baik, yaitu sebesar 0,626. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 62,6% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,2% peserta merespon distraktor A, 0,4% merespon distraktor B, 36,7% peserta merespon distraktor C dan 0,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, B, dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 30 memiliki indeks kesulitan baik, yaitu sebesar 0,449. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 44,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 18,9% peserta merespon distraktor A, 27,4% merespon distraktor B, 5,9% peserta merespon distraktor C dan 2,8% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, C, dan D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

Butir soal nomor 31 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,984. Hal ini berarti soal tergolong mudah karena 98,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,2% peserta merespon distraktor A dan B, 0,5% merespon distraktor C, dan 0,7% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, C, dan D merupakan distraktor yang tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 32 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,929. Hal ini berarti soal tergolong mudah karena 92,9% peserta tes merespon kunci

jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1,0% peserta merespon distraktor A, 1,5% merespon distraktor B, 1,4% peserta merespon distraktor C dan 3,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, B, dan C tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 33 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,918. Hal ini berarti soal tergolong mudah karena 91,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 5,8% peserta merespon distraktor A, 0,8% merespon distraktor B, 0,9% peserta merespon distraktor D dan 0,6% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A, B, dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 34 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,886. Hal ini berarti soal tergolong mudah karena 88,6% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 3,1% peserta merespon distraktor A, 6,2% merespon distraktor C, 1,1% peserta merespon distraktor D dan 1,0% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A dan C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor D dan E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 35 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,985. Hal ini berarti soal tergolong mudah karena 98,5% peserta tes merespon kunci

jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,1% peserta merespon distraktor B, 0,8% merespon distraktor C, 0,4% peserta merespon distraktor D dan 0,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B, C, D, dan E merupakan distraktor yang tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 36 memiliki indeks kesulitan baik, yaitu sebesar 0,330. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 33% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,1% peserta merespon distraktor B, 0,4% merespon distraktor C, 64,9% peserta merespon distraktor D dan 1,4% peserta merespon distraktor E. Distraktor D merupakan distraktor yang efektif karena pemilihnya lebih dari 2% dan nilai *point biser*-nya bertanda negatif. Di sisi lain, distraktor B, C, dan E tidak efektif karena direspon kurang dari 2% peserta tes. Selain itu, pada distraktor C, *point biser*-nya juga bertanda positif.

Butir soal nomor 37 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,968. Hal ini berarti soal tergolong mudah karena 96,8% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,6% peserta merespon distraktor A, 0,4% merespon distraktor C, 0% peserta merespon distraktor D dan 2,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A dan C tidak efektif karena direspon kurang dari 2% peserta tes. Tidak ada peserta tes yang memilih distraktor D sehingga distraktor tersebut digugurkan.

Butir soal nomor 38 memiliki indeks kesulitan baik, yaitu sebesar 0,289. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 28,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1,9% peserta merespon distraktor A, 1,1% merespon distraktor B, 4,1% peserta merespon distraktor D dan 63,7% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor D dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A dan B tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 39 memiliki indeks kesulitan baik, yaitu sebesar 0,509. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 50,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 2,7% peserta merespon distraktor A, 4,8% merespon distraktor B, 40,7% peserta merespon distraktor C dan 0,8% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, dan C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor E tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 40 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,880. Hal ini berarti soal tergolong mudah karena 88% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 3,7% peserta merespon distraktor A, 1,2% merespon distraktor B, 5,6% peserta merespon distraktor C dan 1,4% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A dan C merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B dan D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 41 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,774.

Hal ini berarti soal tergolong mudah karena 77,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1,5% peserta merespon distraktor B, 0,2% merespon distraktor C, 4,7% peserta merespon distraktor D dan 16,1% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor D dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B dan C tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 42 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,119.

Hal ini berarti soal tergolong sulit karena hanya 11,9% peserta tes yang merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 17,7% peserta merespon distraktor A, 42,1% merespon distraktor C, 9,4% peserta merespon distraktor D dan 18,9% peserta merespon distraktor E. Distraktor D dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2% dan nilai *point biser*-nya negatif. Distraktor A dan C meskipun dipilih lebih dari 2% peserta tes, tetap tidak efektif karena nilai *point biser*-nya positif.

Butir soal nomor 43 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,912.

Hal ini berarti soal tergolong mudah karena 91,2% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 4,9% peserta merespon distraktor A, 0,7% merespon distraktor B, 3,1% peserta merespon distraktor D dan 0% peserta merespon distraktor E. Distraktor A dan D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor B tidak efektif karena direspon kurang dari 2% peserta tes. Tidak ada peserta tes yang memilih distraktor E sehingga distraktor tersebut digugurkan.

Butir soal nomor 44 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,769. Hal ini berarti soal tergolong mudah karena 76,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0% peserta merespon distraktor A, 1,4% merespon distraktor B, 17,4% peserta merespon distraktor C dan 4,2% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor C dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, B tidak efektif karena direspon kurang dari 2% peserta tes. Tidak ada peserta tes yang memilih distraktor A sehingga distraktor tersebut digugurkan.

Butir soal nomor 45 memiliki indeks kesulitan baik, yaitu sebesar 0,699. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 69,9% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 5,4% peserta merespon distraktor A, 9,6% merespon distraktor B, 7,3% peserta merespon distraktor C dan D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, C, dan D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

Butir soal nomor 46 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,905. Hal ini berarti soal tergolong mudah karena 90,5% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 2,8% peserta merespon distraktor B, 1,2% merespon distraktor C, 0,2% peserta merespon distraktor D dan 5,3% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor C dan D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 47 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,146. Hal ini berarti soal tergolong sulit karena hanya 28,2% peserta tes yang merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 76,3% peserta merespon distraktor A, 3,3% merespon distraktor C, 0,8% peserta merespon distraktor D dan 4,8% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, C, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor D tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 48 memiliki indeks kesulitan baik, yaitu sebesar 0,354. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 35,4% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 1,6% peserta merespon distraktor A, 33,8% merespon distraktor B, 17,4% peserta merespon distraktor D dan 11,4% peserta merespon distraktor E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor B, D, dan E merupakan distraktor yang efektif karena pemilihnya lebih dari 2%. Di sisi lain, distraktor A tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 49 memiliki indeks kesulitan tidak baik, yaitu sebesar 0,985. Hal ini berarti soal tergolong mudah karena 98,5% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 0,1% peserta merespon distraktor A, 0,3% merespon distraktor B, 0,5% peserta merespon distraktor C dan E. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, C, dan E merupakan distraktor yang tidak efektif karena direspon kurang dari 2% peserta tes.

Butir soal nomor 50 memiliki indeks kesulitan baik, yaitu sebesar 0,683. Hal ini berarti tingkat kesukaran soal tergolong baik karena sebanyak 68,3% peserta tes merespon kunci jawaban. Ditinjau dari distribusi jawabannya, sebanyak 10,9% peserta merespon distraktor A, 6,5% merespon distraktor B, 5,8% peserta merespon distraktor C dan 8,4% peserta merespon distraktor D. Nilai *point biser* pada keempat distraktor tersebut bertanda negatif. Distraktor A, B, C, dan D merupakan distraktor yang efektif karena pemilihnya lebih dari 2%.

4. Pembahasan Kunci Jawaban yang Perlu Dicek

Hasil *print out Iteman* menunjukkan butir soal nomor 4 terdapat peringatan "*check the key*". Peringatan tersebut menandakan bahwa kunci jawaban perlu dicek apakah sudah benar atau belum. Jika kunci jawaban sudah benar, maka adanya tanda tersebut disebabkan oleh berbagai faktor. Faktor yang dapat berpengaruh antara lain karena distraktor pada soal nomor 4 sangat homogen dengan kunci jawaban sehingga banyak siswa yang terkecoh. Faktor lainnya adalah disebabkan kurang mengertinya siswa dengan pertanyaan yang terdapat dalam soal nomor 4.

5. Relevansi Hasil Analisis Kualitatif dengan Kuantitatif

Dilihat dari karakteristik secara kualitatif (validitas isi) soal UAS genap kelas XII SMA Negeri di Purbalingga tergolong sangat baik. Kualitas validitas isi tes menunjukkan validitas sebesar 99,30%. Validitas soal yang sangat baik ini menunjukkan bahwa guru sudah mampu mengembangkan soal berdasarkan kisi-kisi. Kisi-kisi tersebut memuat sampel materi yang mewakili seluruh pokok bahasan atau sub pokok bahasan yang telah diajarkan. Namun dilihat dari karakteristik secara kuantitatif, soal UAS genap kelas XII SMA Negeri di

Purbalingga tergolong kurang baik. Reliabilitas soal tergolong kurang baik karena koefisien *Alpha Cronbach* hanya sebesar 0,646, dengan SEM (*Standard Error of Measurement*) yaitu 2,523. Reliabilitas soal yang baik adalah apabila nilai *Alpha Cronbach* $\geq 0,90$. Dilihat dari karakteristik butir soal, indeks kesulitan butir soal tergolong kurang baik karena hanya terdapat 15 butir soal (30%) yang baik, tetapi indeks rata-rata butir soal tergolong baik karena *Mean p* sebesar 0,736. Daya beda tergolong kurang baik karena hanya 11 butir soal (22%) yang baik. Di sisi lain, distraktor soal tergolong kurang baik karena hanya 46,5% distraktor yang dapat berfungsi dengan baik. Kurang baiknya indeks reliabilitas, indeks kesulitan, indeks daya beda, dan efektivitas distraktor dapat terjadi karena berbagai faktor. Faktor tersebut dapat berasal dari diri siswa, guru, dari maupun dari situasi tempat tes dilakukan.

Reliabilitas yang kurang baik ini dapat disebabkan karena *sampling* tes yang terlalu sempit, rendahnya kemampuan peserta tes, serta keadaan tes yang tidak kondusif. Di sisi lain, kurang baiknya indeks kesulitan ini disebabkan karena banyak butir soal yang tergolong mudah. Mudahnya butir soal dapat dikarenakan materi yang ditanyakan telah diajarkan. Pembuat soal menjabarkan materi kurang bagus juga dapat mempengaruhi indeks kesukaran butir soal. Sementara itu, indeks daya beda tidak dipakai karena penelitian ini menggunakan acuan kriteria. Namun, daya beda soal juga tergolong kurang baik. Hal tersebut menunjukkan butir soal belum dapat membedakan antara kelompok yang berhasil (kelompok atas) dan kelompok yang gagal (kelompok bawah). Butir soal tidak dapat membedakan kedua kemampuan tersebut kemungkinan karena kompetensi yang diukur tidak

jelas. Efektivitas distraktor juga kurang baik, artinya distraktor tersebut belum menjalankan fungsinya dengan baik.

Penjelasan di atas menunjukkan bahwa hasil analisis kualitatif dan kuantitatif tidak berbanding lurus. Analisis kualitatif tergolong sangat baik, sedangkan analisis kuantitatif tergolong kurang baik. Hubungan yang berbanding terbalik ini disebabkan oleh banyak faktor. Faktor tersebut dapat berupa faktor internal dari diri peserta tes seperti peserta kurang fokus dalam mengerjakan soal. Faktor lain yang dapat berpengaruh adalah faktor eksternal seperti situasi ujian yang kurang optimal.

BAB V

PENUTUP

A. Simpulan

Penelitian ini menghasilkan kesimpulan sebagai berikut:

1. Guru menyusun soal sangat baik. Hasil telaah secara kualitatif (validitas isi) menunjukkan soal sangat baik dengan kesesuaian telaah soal sebesar 99,30%.
2. Karakteristik soal ditinjau dari analisis kuantitatif tergolong kurang baik. Secara rinci, analisis kuantitatif meliputi:
 - a. ketepatan atau keajegan soal kurang baik yang berarti pengukuran kurang dapat dipercaya. Hal ini sesuai dengan reliabilitas soal yang tergolong kurang baik karena berdasarkan hasil analisis *Iteman* didapat koefisien *Alpha Cronbach* sebesar 0,646, dengan SEM (*Standard Error of Measurement*) yaitu 2,523.
 - b. karakteristik butir soal:
 - 1) butir soal tergolong mudah. Indeks kesulitan tiap-tiap butir soal tergolong kurang baik karena hanya terdapat 15 butir soal (30%) yang baik, tetapi indeks rata-rata butir soal tergolong baik karena *Mean p* pada hasil *Iteman* menunjukkan angka sebesar 0,736;
 - 2) daya beda butir soal diabaikan karena menggunakan Kurikulum Tingkat Satuan Pendidikan (KTSP). Setiap butir soal dinyatakan baik asal indeks daya beda tidak bernilai negatif.

- 3) pola jawaban soal kurang berfungsi dengan baik. Hal ini sesuai dengan distraktor yang tergolong kurang baik karena hanya 46,5% distraktor yang berfungsi dengan baik.

Secara umum kualitas soal Ulangan Akhir Semester (UAS) Genap kelas XII SMA Negeri di Kabupaten Purbalingga tahun pelajaran 2010/2011 tergolong kurang baik. Oleh karena itu, perlu adanya upaya perbaikan dan penyempurnaan soal.

B. Implikasi

Hasil penelitian ini menunjukkan bahwa soal Ulangan Akhir Semester (UAS) Genap kelas XII memiliki kualitas yang kurang baik. Dengan kata lain, soal yang digunakan kurang dapat mencerminkan mutu pembelajaran. Kualitas soal yang belum baik ini kemungkinan disebabkan guru belum secara maksimal mengembangkan kemampuannya dalam menyusun tes yang baik. Guru belum cukup baik dalam menyusun tes khususnya pada aspek perencanaan dan pengembangan.

Butir soal yang telah dianalisis secara kualitatif dan kuantitatif serta dinyatakan diterima dapat dipergunakan dalam meningkatkan penyusunan atau pembuatan soal di daerah Purbalingga. Sementara itu, butir soal yang direvisi hendaknya ditindaklanjuti dengan perbaikan dan penyempurnaan. Hal ini dilakukan sebagai langkah awal dalam upaya meningkatkan kualitas soal Ulangan Akhir Semester (UAS) Genap kelas XII di Kabupaten Purbalingga.

C. Keterbatasan Penelitian

Keterbatasan penelitian yang diungkapkan dalam penelitian ini berkaitan dengan hal-hal yang tidak dapat diperkirakan, dilacak, atau di luar jangkauan peneliti. Penelitian ini hanya terbatas pada SMA Negeri di Kabupaten Purbalingga. Banyaknya butir soal yang mudah berdasarkan analisis teori tes klasik serta reliabilitas tes yang rendah dapat disebabkan oleh beberapa hal seperti faktor lingkungan ketika dilaksanakan ujian, perangkat tes itu sendiri serta keadaan peserta. Untuk mengungkap variabel mana yang memberikan pengaruh paling signifikan terhadap mudahnya butir soal dan rendahnya indeks reliabilitas perlu dilaksanakan penelitian secara menyeluruh.

D. Saran

Saran yang dapat dikemukakan sebagai berikut.

1. Guru-guru SMA perlu memperhatikan berbagai faktor dalam upaya mencapai karakteristik soal yang baik. Faktor-faktor tersebut meliputi kegiatan pembelajaran, pembuatan soal, analisis soal, penilaian, kondisi siswa, dan situasi serta kondisi tes.
2. Dinas Pendidikan Nasional Kabupaten Purbalingga diharapkan dapat mensosialisasikan hasil penelitian ini dalam upaya perbaikan, peningkatan, dan penyempurnaan kualitas pembelajaran dan evaluasi di Purbalingga.
3. Terbuka kemungkinan diadakan penelitian sejenis dan berkelanjutan dengan objek yang lebih luas dan menggunakan teori tes modern agar hasil yang diperoleh dapat menjadi gambaran secara umum karakteristik soal yang baik.

DAFTAR PUSTAKA

- Akhiromah, Isti. 2011. *Karakteristik Secara Teoretis dan Empiris Butir Soal Ulangan Umum Bersama Semester Genap Bahasa Indonesia SMP Kabupaten Gunungkidul Tahun Pelajaran 2009/2010*. Skripsi S1. Yogyakarta: Jurusan Pendidikan Bahasa dan Sastra Indonesia, FBS Universitas Negeri Yogyakarta.
- Arikunto, Suharsimi dan Cepi Safrudin Abdul Jabar. 2008. *Evaluasi Program Pendidikan Pedoman Teoretis Praktis Bagi Mahasiswa dan Praktisi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2006. *Dasar-dasar Evaluasi Pendidikan (Edisi Revisi)*. Jakarta: PT Bumi Aksara.
- BSNP. *Materi Bimbingan Teknis KTSP dan Soal Terstandar 2010 Panduan Penulisan Butir Soal*. Direktorat Pembinaan SMP, Ditjen Manajemen Pendidikan Dasar dan Menengah, Kementerian Pendidikan Nasional.
- Cangelosi, James S. 1995. *Merancang Tes untuk Menilai Prestasi Siswa*. Bandung: ITB.
- Daryanto, H.M. 2008. *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta.
- Depdiknas. 2003. *Kurikulum 2004 Sekolah Menengah Pertama (SMP)*. Jakarta.
- Djali, H. Dan Pudji Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: Gramedia.
- Fernandez, H. J. X. 1984. *Testing and Measurement*. Jakarta: National Education Planing, Evaluation and Curriculum Development.
- Haryati, Mimin. 2007. *Model dan Teknik Penilaian pada Tingkat Satuan Pendidikan*. Jakarta: Gaung Persada Press.
- Indriyati. 2006. *Analisis Soal Tes Uji Standar Mutu Mata Pelajaran Bahasa dan Sastra Indonesia SMA Kelas X Tahun Ajaran 2004/2005 Kabupaten Sleman*. Skripsi S1. Yogyakarta: Jurusan Pendidikan Bahasa dan Sastra Indonesia, FBS Universitas Negeri Yogyakarta.
- Jihad, Asep dan Abdul Haris. 2008. *Evaluasi Pembelajaran*. Yogyakarta: Multi Presindo.
- Lampiran Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 Tanggal 11 Juni 2007 tentang Standar Penilaian Pendidikan.

Noname. 2011. "Analisis Butir Soal dengan Komputer", <http://www.scribd.com/doc/53350731/5/V-ANALISIS-BUTIR-SOAL-DENGAN-KOMPUTER>. Diunduh tanggal 03 Mei 2011 pukul 19.29 WIB.

_____. 2011. "Sulitkah Pelajaran Bahasa Indonesia?" <http://www.bataviase.co.id/node/197980/> Sulitkah Pelajaran Bahasa Indonesia? Diunduh tanggal 27 Juli 2011 pukul 16.21 WIB.

Nurdiyanto, Burhan. 2010. *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*. Yogyakarta: BPFE-Yogyakarta

Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan.

Purwanto, M. Ngalm. 1992. *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.

Riestalinda. 2007. *Analisis Soal Tes Penjajagan Hasil Belajar Siswa (TPHBS) Mata pelajaran Bahasa Indonesia SMA Tahun Ajaran 2006/2007 Se-Kabupaten Sleman*. Skripsi S1. Yogyakarta: Jurusan Pendidikan Bahasa dan Sastra Indonesia, FBS Universitas Negeri Yogyakarta.

Shirran, Alex. 2008. *Evaluating Students*. Jakarta: PT Grasindo.

Sudjana, Nana. 2009. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya.

Suryaman, Maman. 2009. *Panduan Pendidikan dalam Pembelajaran Bahasa Indonesia SMP/MTS*. Pusat Perbukuan Departemen Pendidikan Nasional.

Suyata, Pujiati. 2009. *Pengembangan Sistem Penilaian Berbasis Kompetensi yang Menyatu pada Pembelajaran Bahasa Indonesia*. Yogyakarta: Universitas Negeri Yogyakarta.

Thoha, M. Chabib. 1990. *Teknik Evaluasi Pendidikan*. Jakarta: Rajawali.

Widoyoko, Eko Putro. 2009. *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Pelajar.

LAMPIRAN

Daftar Nama SMA Negeri di Kabupaten Purbalingga

No.	Nama Sekolah	Alamat
1.	SMA N 1 Bukateja	JL. Raya Purwandaru Bukateja
2.	SMA N 1 Kejobong	JL. Raya Gumiwang
3.	SMA N 1 Purbalingga	JL. Letjend M.T. Haryono
4.	SMA N 2 Purbalingga	JL. Raya Pucung Rumbak
5.	SMA N 1 Kutasari	JL. Raya Tobong-Kutasari
6.	SMA N 1 Bobotsari	Majapura
7.	SMA N 1 Rembang	JL. Monumen Jenderal Sudirman
8.	SMA N 1 Karangreja	JL. Raya Karangreja
9.	SMA N 1 Padamara	JL. Raya Desa Padamara
10.	SMA N 1 Kemangkon	JL. Raya Panican

No Soal	Indeks Kesulitan/IK (Prop. Correct)		Daya Beda/DB (Point Biser)		Pengecoh/ distraktor (Prop. Endorsing)			Interpretasi
1	0.948	Jelek	0.095	Jelek	A 0.948 0.198 0.095 *	A.	B. Baik C. Jelek D. Jelek E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
2	0.820	Jelek	0.219	Baik	A 0.170 -0.297 -0.200 B 0.820 0.321 0.219 * C 0.009 -0.345 -0.088 D 0.000 -9.000 -9.000 E 0.002 -0.170 -0.023 Other 0.000 -9.000 -9.000	A. Baik B.	C. Jelek D. Jelek E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
3	0.791	Jelek	0.305	Baik	A 0.003 -0.174 -0.030 B 0.006 -0.201 -0.044 C 0.009 -0.465 -0.121 D 0.190 -0.391 -0.270 E 0.791 0.432 0.305 * Other 0.000 -0.845 -0.065	A. Jelek B. Jelek C. Jelek D. Baik E.		IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
4	0.840	Jelek	0.074	Jelek	A 0.040 0.176 0.077 ? B 0.079 -0.128 -0.070 CHECK THE KEY E was specified, A works better C 0.027 -0.147 -0.057 D 0.014 -0.380 -0.117 E 0.840 0.111 0.074 * Other 0.000 -0.185 -0.014	A. Jelek B. Baik C. Baik D. Jelek E.		IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
5	0.972	Jelek	0.208	Baik	A 0.001 -0.404 -0.042 B 0.972 0.535 0.208 * C 0.004 -0.399 -0.073 D 0.006 -0.493 -0.106 E 0.017 -0.451 -0.146 Other 0.001 -0.613 -0.063	A. Jelek B. C. Jelek D. Jelek E. Jelek		IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
6	0.592	Baik	0.315	Baik	A 0.099 -0.296 -0.172 B 0.592 0.399 0.315 * C 0.169 -0.130 -0.087 D 0.102 -0.318 -0.187 E 0.037 -0.153 -0.066 Other 0.001 -0.578 -0.060	A. Baik B. C. Baik D. Baik E. Baik		IK = Baik DB = Baik Pengecoh= Baik (soal dipakai)
7	0.614	Baik	0.403	Baik	A 0.010 -0.445 -0.118	A. Jelek		IK = Baik

					B 0.116 -0.196 -0.119 C 0.027 -0.413 -0.159 D 0.614 0.513 0.403 * E 0.231 -0.393 -0.284 Other 0.002 -0.133 -0.018	B. Baik C. Baik D. E. Baik	DB = Baik Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh B, C, E)
8	0.825	Jelek	0.275	Baik	A 0.111 -0.265 -0.160 B 0.018 -0.203 -0.068 C 0.030 -0.398 -0.159 D 0.825 0.406 0.275 * E 0.014 -0.482 -0.148 Other 0.001 -0.106 -0.013	A. Baik B. Jelek C. Baik D. E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
9	0.834	Jelek	0.270	Baik	A 0.105 -0.233 -0.139 B 0.008 -0.430 -0.108 C 0.834 0.403 0.270 * D 0.034 -0.511 -0.212 E 0.019 -0.224 -0.076 Other 0.000 0.079 0.006	A. Baik B. Jelek C. D. Baik E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
10	0.878	Jelek	0.183	Jelek	A 0.080 -0.131 -0.072 B 0.010 -0.609 -0.162 C 0.878 0.296 0.183 * D 0.014 -0.444 -0.135 E 0.016 -0.200 -0.064 Other 0.001 -0.082 -0.010	A. Baik B. Jelek C. D. Jelek E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
11	0.294	Baik	0.256	Baik	A 0.294 0.339 0.256 * B 0.573 -0.174 -0.138 C 0.039 -0.172 -0.075 D 0.011 -0.347 -0.095 E 0.082 -0.162 -0.089 Other 0.001 0.086 0.011	A. B. Baik C. Baik D. Jelek E. Baik	IK = Baik DB = Baik Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh D)
12	0.928	Jelek	0.104	Jelek	A 0.928 0.196 0.104 * B 0.016 -0.216 -0.069 C 0.042 -0.101 -0.045 D 0.002 -0.347 -0.052 E 0.011 -0.227 -0.062 Other 0.000 -9.000 -9.000	A. B. Jelek C. Baik D. Jelek E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
13	0.942	Jelek	0.236	Baik	A 0.007 -0.554 -0.133 B 0.007 -0.294 -0.069 C 0.027 -0.292 -0.113 D 0.015 -0.504 -0.158 E 0.942 0.476 0.236 * Other 0.000 0.079 0.006	A. Jelek B. Jelek C. Baik D. Jelek E.	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
14	0.942	Jelek	0.144	Jelek	A 0.003 -0.854 -0.148	A. Jelek	IK = Jelek

					B 0.003 -0.285 -0.049 C 0.000 -0.053 -0.004 D 0.045 -0.206 -0.094 E 0.948 0.301 0.144 * Other 0.000 -0.779 -0.060	B. Jelek C. Jelek D. Baik E.	DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
15	0.786	Jelek	0.193	Jelek	A 0.103 -0.238 -0.140 B 0.786 0.272 0.193 * C 0.042 -0.199 -0.089 D 0.014 -0.149 -0.046 E 0.052 -0.120 -0.058 Other 0.002 -0.041 -0.006	A. Baik B. C. Baik D. Jelek E. Baik	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
16	0.919	Jelek	0.221	Baik	A 0.053 -0.353 -0.170 B 0.919 0.403 0.221 * C 0.020 -0.263 -0.092 D 0.005 -0.567 -0.113 E 0.002 -0.181 -0.027 Other 0.000 -9.000 -9.000	A. Baik B. C. Baik D. Jelek E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
17	0.937	Jelek	0.168	Jelek	A 0.017 -0.169 -0.055 B 0.025 -0.339 -0.127 C 0.937 0.329 0.168 * D 0.005 -0.296 -0.061 E 0.015 -0.221 -0.069 Other 0.001 -0.334 -0.035	A. Jelek B. Baik C. D. Jelek E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
18	0.882	Jelek	0.234	Baik	A 0.028 -0.301 -0.118 B 0.018 -0.301 -0.101 C 0.882 0.381 0.234 * D 0.042 -0.248 -0.111 E 0.028 -0.305 -0.119 Other 0.002 -0.115 -0.016	A. Baik B. Jelek C. D. Baik E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
19	0.987	Jelek	0.116	Jelek	A 0.001 -0.509 -0.053 B 0.001 -0.923 -0.113 C 0.007 -0.292 -0.070 D 0.987 0.397 0.116 * E 0.003 -0.126 -0.020 Other 0.000 -9.000 -9.000	A. Jelek B. Jelek C. Jelek D. E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
20	0.544	Baik	0.280	Baik	A 0.083 -0.172 -0.095 B 0.168 -0.150 -0.100 C 0.063 -0.218 -0.111 D 0.544 0.352 0.280 * E 0.141 -0.211 -0.135 Other 0.000 -0.779 -0.060	A. Baik B. Baik C. Baik D. E. Jelek	IK = Baik DB = Baik Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh E)
21	0.205	Jelek	0.089	Jelek	A 0.578 0.035 0.028 B 0.207 -0.174 -0.123	A. Baik B. Baik	IK = Jelek

					C 0.008 0.052 0.013 D 0.002 -0.152 -0.021 E 0.205 0.127 0.089 * Other 0.000 -0.119 -0.009	C. Jelek D. Jelek E.	DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
22	0.899	Jelek	0.239	Baik	A 0.007 -0.232 -0.054 B 0.027 -0.395 -0.151 C 0.064 -0.330 -0.169 D 0.003 -0.185 -0.032 E 0.899 0.408 0.239 * Other 0.000 -9.000 -9.000	A. Jelek B. Baik C. Baik D. Jelek E. E.	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
23	0.954	Jelek	0.208	Baik	A 0.003 -0.430 -0.074 B 0.037 -0.443 -0.189 C 0.003 -0.177 -0.029 D 0.954 0.452 0.208 * E 0.003 -0.228 -0.037 Other 0.000 -9.000 -9.000	A. Jelek B. Baik C. Jelek D. E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
24	0.448	Baik	0.275	Baik	A 0.193 -0.217 -0.150 B 0.196 -0.054 -0.038 C 0.448 0.345 0.275 * D 0.098 -0.060 -0.035 E 0.061 -0.408 -0.206 Other 0.004 -0.215 -0.041	A. Baik B. Baik C. D. Baik E. Baik	IK = Baik DB = Baik Pengecoh= Baik (soal dipakai)
25	0.422	Baik	0.397	Baik	A 0.118 -0.245 -0.150 B 0.422 0.502 0.397 * C 0.213 -0.322 -0.228 D 0.110 -0.140 -0.084 E 0.137 -0.120 -0.076 Other 0.001 -0.613 -0.063	A. Baik B. C. Baik D. Baik E. Baik	IK = Baik DB = Baik Pengecoh= Baik (soal dipakai)
26	0.897	Jelek	0.183	Jelek	A 0.897 0.310 0.183 * B 0.025 -0.279 -0.105 C 0.059 -0.222 -0.111 D 0.004 -0.260 -0.050 E 0.012 -0.251 -0.072 Other 0.003 -0.228 -0.037	A. B. Baik C. Baik D. Jelek E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
27	0.672	Baik	0.179	Jelek	A 0.002 -1.000 -0.141 B 0.672 0.233 0.179 * C 0.007 -0.316 -0.074 D 0.311 -0.183 -0.139 E 0.007 -0.367 -0.086 Other 0.001 -0.056 -0.006	A. Jelek B. C. Jelek D. Baik E. Jelek	IK = Baik DB = Jelek Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh A, C, E)
28	0.924	Jelek	0.203	Baik	A 0.000 0.013 0.001 B 0.068 -0.346 -0.180	A. Jelek B. Baik	IK = Jelek

					C 0.924 0.375 0.203 * D 0.002 -0.410 -0.057 E 0.005 -0.385 -0.080 Other 0.001 -0.125 -0.013	C. D. Jelek E. Jelek	DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
29	0.626	Baik	0.360	Baik	A 0.002 -0.468 -0.071 B 0.004 -0.103 -0.019 C 0.367 -0.442 -0.345 D 0.626 0.460 0.360 * E 0.002 -0.577 -0.080 Other 0.000 -9.000 -9.000	A. Jelek B. Jelek C. Baik D. E. Jelek	IK = Baik DB = Baik Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh A, B, E)
30	0.449	Baik	0.347	Baik	A 0.189 -0.168 -0.116 B 0.274 -0.260 -0.194 C 0.059 -0.267 -0.133 D 0.028 -0.139 -0.054 E 0.449 0.436 0.347 * Other 0.002 -0.041 -0.006	A. Baik B. Baik C. Baik D. Baik E. Other	IK = Baik DB = Baik Pengecoh= Baik (soal dipakai)
31	0.984	Jelek	0.121	Jelek	A 0.002 -0.059 -0.008 B 0.002 -0.688 -0.095 C 0.005 -0.170 -0.035 D 0.007 -0.405 -0.097 E 0.984 0.376 0.121 * Other 0.000 -9.000 -9.000	A. Jelek B. Jelek C. Jelek D. Jelek E. Other	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
32	0.929	Jelek	0.252	Baik	A 0.010 -0.438 -0.119 B 0.015 -0.351 -0.110 C 0.014 -0.259 -0.079 D 0.929 0.477 0.252 * E 0.031 -0.433 -0.174 Other 0.001 -0.021 -0.002	A. Jelek B. Jelek C. Jelek D. E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
33	0.918	Jelek	0.211	Baik	A 0.058 -0.313 -0.155 B 0.008 -0.286 -0.070 C 0.918 0.383 0.211 * D 0.009 -0.509 -0.133 E 0.006 -0.142 -0.030 Other 0.000 -0.053 -0.004	A. Baik B. Jelek C. D. Jelek E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
34	0.886	Jelek	0.332	Baik	A 0.031 -0.461 -0.186 B 0.886 0.547 0.332 * C 0.062 -0.494 -0.251 D 0.011 -0.165 -0.045 E 0.010 -0.298 -0.079 Other 0.000 -9.000 -9.000	A. Baik B. C. Baik D. Jelek E. Jelek	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
35	0.985	Jelek	0.160	Jelek	A 0.985 0.510 0.160 * B 0.001 -0.250 -0.031	A. B. Jelek	IK = Jelek

					C 0.008 -0.484 -0.119 D 0.004 -0.370 -0.067 E 0.002 -0.544 -0.082 Other 0.000 -9.000 -9.000	C. Jelek D. Jelek E. Jelek	DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
36	0.330	Baik	0.126	Jelek	A 0.330 0.164 0.126 * B 0.001 -0.610 -0.075 C 0.004 0.051 0.010 D 0.649 -0.113 -0.088 E 0.014 -0.332 -0.102 Other 0.001 -0.961 -0.099	A. B. Jelek C. Jelek D. Baik E. Jelek	IK = Baik DB = Jelek Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh B, C, E)
37	0.968	Jelek	0.126	Jelek	A 0.006 -0.362 -0.078 B 0.968 0.310 0.126 * C 0.004 -0.429 -0.082 D 0.000 -9.000 -9.000 E 0.021 -0.214 -0.076 Other 0.000 0.013 0.001	A. Jelek B. C. Jelek D. Jelek E. Baik	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
38	0.289	Baik	0.140	Jelek	A 0.019 -0.316 -0.107 B 0.011 -0.394 -0.108 C 0.289 0.186 0.140 * D 0.041 -0.359 -0.159 E 0.637 -0.007 -0.005 Other 0.004 -0.340 -0.062	A. Jelek B. Jelek C. D. Baik E. Baik	IK = Baik DB = Jelek Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh A, B)
39	0.509	Baik	0.286	Baik	A 0.027 -0.318 -0.123 B 0.048 -0.215 -0.100 C 0.407 -0.238 -0.188 D 0.509 0.358 0.286 * E 0.008 -0.382 -0.096 Other 0.000 -9.000 -9.000	A. Baik B. Baik C. Baik D. E. Jelek	IK = Baik DB = Baik Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh E)
40	0.880	Jelek	0.250	Baik	A 0.037 -0.193 -0.083 B 0.012 -0.405 -0.115 C 0.056 -0.315 -0.155 D 0.014 -0.457 -0.139 E 0.880 0.406 0.250 * Other 0.001 -0.439 -0.045	A. Baik B. Jelek C. Baik D. Jelek E.	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
41	0.774	Jelek	0.262	Baik	A 0.774 0.364 0.262 * B 0.015 -0.268 -0.084 C 0.002 -0.393 -0.059 D 0.047 -0.348 -0.161 E 0.161 -0.254 -0.169 Other 0.000 -9.000 -9.000	A. B. Jelek C. Jelek D. Baik E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)

42	0.119	Jelek	0.237	Baik	A 0.177 0.021 0.015 B 0.119 0.385 0.237 * C 0.421 0.005 0.004 D 0.094 -0.052 -0.030 E 0.189 -0.279 -0.193 Other 0.000 -9.000 -9.000	A. Jelek B. C. Jelek D. Baik E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
43	0.912	Jelek	0.194	Jelek	A 0.049 -0.278 -0.131 B 0.007 -0.159 -0.038 C 0.912 0.343 0.194 * D 0.031 -0.336 -0.136 E 0.000 0.145 0.011 Other 0.000 -9.000 -9.000	A. Baik B. Jelek C. D. Baik E. Jelek	IK = Jelek DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
44	0.769	Jelek	0.348	Baik	A 0.000 -0.449 -0.035 B 0.014 -0.429 -0.130 C 0.174 -0.404 -0.274 D 0.769 0.482 0.348 * E 0.042 -0.301 -0.134 Other 0.000 0.013 0.001	A. Jelek B. Jelek C. Baik D. E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
45	0.699	Baik	0.326	Baik	A 0.054 -0.208 -0.101 B 0.096 -0.374 -0.216 C 0.073 -0.215 -0.114 D 0.073 -0.222 -0.119 E 0.699 0.429 0.326 * Other 0.004 -0.180 -0.034	A. Baik B. Baik C. Baik D. Baik E.	IK = Baik DB = Baik Pengecoh= revisi (soal dipakai)
46	0.905	Jelek	0.356	Baik	A 0.905 0.616 0.356 * B 0.028 -0.367 -0.143 C 0.012 -0.560 -0.159 D 0.002 -0.106 -0.016 E 0.053 -0.580 -0.279 Other 0.001 -0.265 -0.027	A. B. Baik C. Jelek D. Jelek E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
47	0.146	Jelek	0.282	Baik	A 0.763 -0.219 -0.159 B 0.146 0.434 0.282 * C 0.033 -0.195 -0.080 D 0.008 -0.103 -0.026 E 0.048 -0.151 -0.070 Other 0.002 -0.115 -0.016	A. Baik B. C. Baik D. Jelek E. Baik	IK = Jelek DB = Baik Pengecoh= revisi (soal digugurkan karena IK jelek)
48	0.354	Baik	0.260	Baik	A 0.016 -0.365 -0.117 B 0.338 -0.095 -0.073 C 0.354 0.334 0.260 * D 0.174 -0.174 -0.118 E 0.114 -0.143 -0.087 Other 0.004 -0.221 -0.040	A. Jelek B. Baik C. D. Baik E. Baik	IK = Baik DB = Baik Pengecoh= revisi (soal dipakai dengan revisi pada pengecoh A)
49	0.985	Jelek	0.118	Jelek	A 0.001 -0.334 -0.035	A. Jelek	IK = Jelek

					B 0.003 -0.386 -0.067 C 0.005 -0.192 -0.040 D 0.985 0.381 0.118 * E 0.005 -0.415 -0.086 Other 0.000 -0.317 -0.024	B. Jelek C. Jelek D. E. Jelek	DB = Jelek Pengecoh= revisi (soal digugurkan karena IK jelek)
50	0.683	Baik	0.338	Baik	A 0.109 -0.364 -0.218 B 0.065 -0.292 -0.150 C 0.058 -0.119 -0.059 D 0.084 -0.247 -0.138 E 0.683 0.441 0.338 * Other 0.001 -0.082 -0.010	A. Baik B. Baik C. Baik D. Baik E.	IK = Baik DB = Baik Pengecoh= Baik (soal dipakai)