

PEMETAAN KUALITAS EMPIRIK SOAL UJIAN AKHIR SEMESTER PADA MATA PELAJARAN BAHASA INDONESIA SMA DI KABUPATEN KLATEN

Yulinda Erma Suryani

Universitas Widya Dharma Klaten

Jln. Ki Hajar Dewantara, Karanganom, Macanan, Karanganom, Klaten, Jawa Tengah 57438, Indonesia

Email: yulinda@unwidha.ac.id

Abstrak

Tujuan penelitian ini adalah untuk mengetahui kualitas empirik soal Mata Pelajaran Bahasa Indonesia yang digunakan dalam pelaksanaan ujian akhir semester SMA di Kabupaten Klaten dalam empat tahun terakhir. Jenis penelitian ini adalah deskriptif kuantitatif. Metode pengumpulan data yang digunakan dalam penelitian ini adalah dokumentasi yang mengutip respon jawaban 1125 peserta didik pada ujian akhir semester tingkat SMA di Kabupaten Klaten pada Mata Pelajaran Bahasa Indonesia kelas X pada Tahun Pelajaran 2011/2012 sampai Tahun Pelajaran 2014/2015. Analisis kualitas empirik soal ujian akhir semester meliputi: tingkat kesukaran, daya pembeda, efektivitas distraktor, validitas dan reliabilitas. Berdasarkan hasil analisis data dapat diketahui bahwa tingkat kesukaran butir soal ujian akhir semester di Kabupaten Klaten pada Tahun Pelajaran 2011/2012 kategori mudah sebanyak 64%, sedang 24 % dan sukar 12%. Pada Tahun Pelajaran 2012/2013 78% soal mudah, 12% sedang 10% sukar 10%. Pada tahun pelajaran 2013/2014 12% soal mudah, 14% sedang dan 74% sukar. Pada tahun pelajaran 2014/2015 64% mudah, 16% sedang dan 28% sukar. Daya pembeda soal ujian akhir semester pada tahun pelajaran 2011/2012 yang memenuhi kriteria sebanyak 62%, pada Tahun Pelajaran 2012/2013 sebanyak 76%, pada tahun pelajaran 2013/2014 sebanyak 62%, pada Tahun Pelajaran 2014/2015 sebanyak 64%. Distraktor pada soal ujian akhir semester Tahun Pelajaran 2011/2012 dapat berfungsi dengan baik sebanyak 86%, pada Tahun Pelajaran 2012/2013 sebanyak 90%, pada Tahun Pelajaran 2013/2014 sebanyak 76%, pada Tahun Pelajaran 2014/2015 sebanyak 90%. Hasil analisis validitas empirik pada soal ujian akhir semester tahun pelajaran 2011/2012 diperoleh 36% butir yang valid dengan nilai koefisien reliabilitas sebesar 0,232. Pada soal ujian akhir semester tahun pelajaran 2012/2013, 56% butir soal valid dengan nilai koefisien reliabilitas 0,428. Koefisien reliabilitas soal ujian akhir semester pada tahun pelajaran 2013/2014 diperoleh sebesar 0,238 dengan butir soal yang valid 40%. Validitas empirik pada soal ujian akhir semester tahun pelajaran 2014/2015 sebesar 50% butir soal valid dengan nilai koefisien reliabilitas sebesar 0,603.

Kata kunci: *pemetaan, kualitas, ujian semester, bahasa Indonesia*

MAPPING OF APPING OF THE EMPIRICAL QUALITY OF INDONESIAN LANGUAGE SUBJECT FINAL EXAM FOR SENIOR HIGH SCHOOL IN KLATEN REGENCY

Abstract

This research aims at determining the empirical quality of Indonesian Language subject test questions used for Final Exam for senior high school in Klaten Regency within the last four years. This research uses a descriptive quantitative method. The data for this research were collected by means of documentation. Documentation that cites the responses of the students of the final exam of high school semester in Klaten Regency on the Indonesian Language Subject of class X in the 2011/2012 Year Learning Year until the 2014/2015 Lesson Year which amounts to 1125 students. Empirical quality analysis of the final exam of the semester include: item difficulty, discriminating power, distractor effectiveness, validity and reliability. Based on the results of data analysis can be seen that the level of difficulty item semester final exam in Klaten regency in the academic year 2011/2012 easy categories as much as 64%, while 24% and 12% difficult. In school year 2012/2013 78% easy problem, 12% medium 10% difficult 10%. In the academic year 2013/2014 12% easy questions, 14% medium and 74% difficult. In academic year 2014/2015 64% easy, 16% medium and 28% difficult. The discriminating power of final exam of semester in academic year 2011/2012 that have qualified as much as 62%, Lessons 2012/2013 as much as 76%, 62/2014 years of education as much as 62%, Lesson 2014/2015 as much as 64%. Distractor at the end of semester exam in the academic year 2011/2012 to function properly as much as 86%, in the Academic Year 2012/2013, as much as 90%, in the Academic Year 2013/2014 by 76%, in the Academic Year 2014/2015 by 90%. The results of empirical validity analysis on the final exam of the academic year 2011/2012 obtained 36% valid grain with the value of reliability coefficient of 0.232. On the matter of the final exam of the year 2012/2013 academic year, 56% items are valid with the value of reliability coefficient of 0.428. Reliability coefficient about final exam of semester in academic year 2013/2014 obtained equal to 0,238 with item 40% valid matter. Empirical validity in semester test case final year 2014/2015 amounted to 50% items valid with reliability coefficient value of 0.603

Keywords: *mapping, quality, semester test, Indonesian language*

Permalink/DOI: <http://dx.doi.org/10.21831/jpep.v21i2.10725>

Pendahuluan

Desentralisasi pendidikan secara resmi dimulai dengan diberlakukannya Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Menurut Umiarso & Gojali (2010) desentralisasi pendidikan muncul dan berkembang sebagai bagian dari agenda besar global tentang demokratisasi dan desentralisasi pemerintahan dalam rangka mewujudkan tata pemerintahan yang lebih baik (*good governance*). Desentralisasi dalam penyelenggaraan pendidikan disebut dengan manajemen berbasis sekolah (MBS). Konsep dasar manajemen berbasis sekolah adalah pengelolaan peningkatan mutu pendidikan yang dilakukan sekolah secara mandiri dengan melibatkan semua pihak yang terkait dengan pendidikan yang biasa disebut dengan otonomi pendidikan atau sekolah. Esensi MBS adalah otonomi sekolah dan pengambilan keputusan partisipatif untuk mencapai sasaran mutu pendidikan di sekolah. Sehingga dalam proses perencanaan, pelaksanaan dan evaluasi mutu pendidikan di sekolah mampu melibatkan *stakeholder* sekolah.

Evaluasi mutu pendidikan di sekolah dilakukan berdasarkan Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tanggal 11 Juni 2007 tentang Standar Penilaian Pendidikan. Dalam peraturan tersebut disebutkan bahwa standar penilaian pendidikan adalah standar nasional pendidikan yang berkaitan dengan mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik, dan penilaian pendidikan adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik (Menteri Pendidikan Nasional, 2007). Menurut Mardapi (1999) Penilaian merupakan proses penggambaran, memperoleh dan memberikan informasi yang berguna sebagai alternatif pengambilan keputusan. Suatu proses penilaian, ada yang didasarkan atas hasil pengukuran yang bersifat kuantitatif dan ada pula yang kualitatif yang didasarkan atas non-pengukuran. Ciri penting dalam pengukuran adalah adanya proses perbandingan. Oleh karena itu, pengukuran, penilaian, dan evaluasi merupakan

kegiatan yang berkaitan dan berurutan. Secara umum penilaian atau evaluasi adalah suatu sistem sistematis untuk mengetahui tingkat keberhasilan dan efisiensi suatu program. Evaluasi dalam sistem pendidikan adalah salah satu kegiatan yang sangat penting dilaksanakan secara teratur pada periode-periode tertentu, antara lain untuk memantau kualitas mutu pendidikan dan membantu proses belajar mengajar (PBM) di kelas, karena itu diperlukan alat ukur. Pemahaman siswa terhadap materi pelajaran dapat diketahui seseorang berdasarkan hasil evaluasi, sehingga informasi yang bermakna dapat diperoleh dalam pengambilan keputusan. Oleh karena itu dalam praktiknya masalah pengukuran mempunyai kedudukan yang sangat penting dalam proses evaluasi. Baik buruknya hasil evaluasi tergantung pada hasil pengukuran.

Umar (1991) mengemukakan jika kualitas kegiatan evaluasi cukup baik maka data yang dihasilkan dapat digunakan untuk mengambil keputusan mengenai peserta didik dan untuk penyempurnaan kurikulum. Sebaliknya jika evaluasi rendah mutunya, akan menyesatkan para pengambil keputusan atau kebijakan. Hal ini berarti bahwa evaluasi bertujuan untuk memperbaiki proses belajar siswa serta dapat memberi umpan balik yang dapat dipergunakan sebagai dasar untuk perencanaan di masa yang akan datang. Agar hasil ujian semester ini dapat menggambarkan keadaan yang sesungguhnya, maka perangkat soal yang digunakan harus memenuhi segala aspek persyaratan bagisebuah alat ukur yang baik dan teruji dalam berbagai aspek. Perangkat soal yang digunakan dalam ujian semester dibuat oleh masing-masing guru di masing-masing sekolah sesuai dengan bidang studi yang diajarkan. Menurut Masrun (1978) gambaran tentang baiknya suatu alat ukur tercermin pada karakteristik dari perangkat soal itu sendiri. Karakteristik adalah ciri-ciri yang melekat atau dimiliki oleh suatu perangkat soal yaitu validitas, reliabilitas, tingkat kesukaran, dan daya beda.

Dengan adanya kebijakan otonomi pendidikan maka pelaksanaan pendidikan

menjadi hak dan wewenang daerah setempat. perangkat dapat diketahui dari hasil bukti empiris. Perbaikan mutu pendidikan dapat dilakukan melalui sistem evaluasi. Perbaikan sistem evaluasi dapat dimulai dari proses penyelenggaraan ujian semester baik itu semester I maupun semester II. Evaluasi keberhasilan belajar peserta didik dalam bidang studi, baik yang diselenggarakan pada setiap semester maupun akhir program pendidikan perlu direncanakan sebaik-baiknya. Perangkat tes untuk kepentingan evaluasi tersebut, pada masa yang akan datang sistem dan penyelenggaraannya harus didasarkan pada tes standar sehingga hasilnya dapat dipertanggungjawabkan. Melalui penelitian ini akan dilakukan pemetaan mutu pendidikan di Kabupaten Klaten, karena sistem pengajaran dan sistem evaluasi memiliki keterkaitan dalam rangka memantau mutu pendidikan. Untuk dapat meningkatkan mutu pendidikan harus dimulai dengan memperbaiki sistem pengajarannya, tetapi untuk mendapatkan informasi tentang efektivitas dan efisiensi sistem pengajaran yang telah dilakukan diperlukan sistem evaluasi yang baik. Kedua hal tersebut harus merupakan suatu sistem yang harus sejalan dalam skala yang lebih besar, yaitu sebagai bagian dari sistem pendidikan. Tujuan penelitian ini adalah untuk mengetahui kualitas empirik soal Mata Pelajaran Bahasa Indonesia yang digunakan dalam pelaksanaan ujian semester di Kabupaten Klaten dalam empat tahun terakhir. Melalui hasil penelitian ini Guru-guru Mata Pelajaran Bahasa Indonesia di Kabupaten Klaten memperoleh gambaran mengenai kualitas soal yang telah digunakan dalam proses evaluasi yang dilaksanakan setiap akhir semester.

Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode penelitian deskriptif kuantitatif. Metode pengumpulan data yang digunakan dalam penelitian ini adalah dokumentasi dengan menggunakan respon jawaban 1125 siswa pada soal ulangan akhir semester SMA di Kabupaten Klaten pada Mata Pelajaran Bahasa Indonesia kelas X

mulai Tahun Pelajaran 2011/2012 sampai Tahun Pelajaran 2014/2015. Soal ulangan akhir semester mata pelajaran Bahasa Indonesia kelas X dengan bentuk soal pilihan ganda, yang berjumlah 50 butir soal.

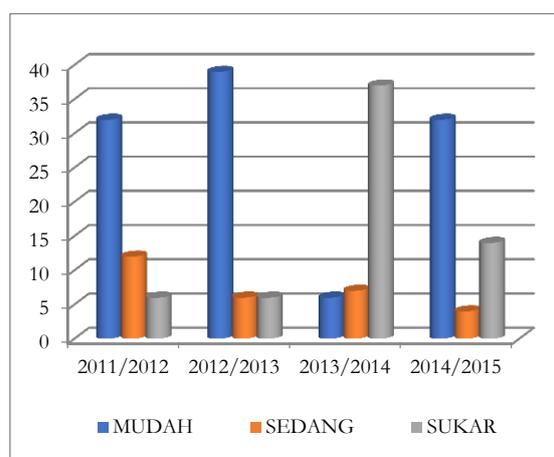
Untuk mengetahui kualitas empirik soal ujian semester pada Mata Pelajaran Bahasa Indonesia mulai Tahun Pelajaran 2011/2012 sampai Tahun Pelajaran 2014/2015 di Kabupaten Klaten, dilakukan tahapan analisis terhadap data yang diperoleh, yaitu dengan melakukan analisis kualitas soal ujian semester Tingkat SMA di Kabupaten Klaten, dilakukan dengan menggunakan pendekatan teori tes klasik. Teori tes klasik digunakan untuk mengetahui tingkat kesukaran, daya beda soal dan efektivitas distraktor. Analisis kualitas soal secara klasik dilakukan dengan menggunakan program IteMan. Biasanya tingkat kesukaran hasil analisis dengan menggunakan program IteMan dikategorikan menjadi tiga bagian, yaitu: (1) Mudah jika nilai $propcorrect > 0,70$; (2) Sedang jika $0,30 \leq \text{nilai } propcorrect \leq 0,70$; dan (3) sukar jika nilai $propcorrect < 0,30$ (Hayat & Suprananto, 1999). Pengkategorian hasil analisis daya pembeda dengan menggunakan program IteMan menggunakan kriteria berikut: (1) bagus sekali jika nilai Biser $> 0,40$; (2) lumayan bagus tapi mungkin masih perlu peningkatan jika $0,30 < \text{nilai Biser} < 0,39$; (3) belum memuaskan, perlu diperbaiki jika $0,20 < \text{nilai Biser} < 0,29$; dan (4) jelek dan harus dibuang jika nilai Biser $< 0,20$ (Azwar, 2016). Hasil analisis distraktor pada program IteMan dapat dilihat pada nilai Biser dan *Point Biser* yang dihasilkan, jika nilai Biser dan *Point Biser* negatif dan tinggi maka distraktor dikategorikan dapat berfungsi dengan baik. Validitas butir soal dianalisis dengan menggunakan perhitungan koefisien korelasi antara skor butir soal dengan total soal (Pradana, Parno, & Handayanto, 2017). Reliabilitas butir soal dianalisis dengan menggunakan perhitungan koefisien Cronbach Alpha (Djaali & Muljono, 2008). Butir soal ujian akhir semester SMA pada mata pelajaran Bahasa Indonesia kelas X dinyatakan valid apabila nilai Biser setiap butir soal $> 0,30$ dan distraktor berfungsi

dengan baik. Estimasi reliabilitas instrumen menggunakan konsistensi internal dengan formula *alpha cronbach*. Prosedur yang dilakukan dalam menganalisis data adalah: (1) melakukan koordinasi dengan dinas Pendidikan Kabupaten Klaten mengenai data hasil ujian akhir semester di kabupaten Klaten, (2) melakukan koordinasi dengan percetakan CV. Saudara yang bekerjasama dengan Dinas Pendidikan Kabupaten Klaten dalam penyelenggaraan ujian akhir semester, dalam pengadaan lembar jawaban komputer sekaligus melakukan *scanning* terhadap lembar jawaban komputer yang telah dikerjakan oleh siswa, (3) data hasil *scanning* ujian akhir semester siswa yang ada di percetakan CV. Saudara tersimpan dalam format ASCII, (4) mengkonversi data dengan format ASCII menjadi Excel, (5) melakukan pembersihan data, yang disesuaikan dengan format file data program Iteman, (6) mengkonversi data Excel ke dalam format ASCII kembali untuk keperluan analisis data, (7) menyiapkan baris kontrol yang dibutuhkan pada program Iteman, (8) melakukan analisis data, (9) melakukan interpretasi terhadap hasil analisis data.

Hasil Penelitian dan Pembahasan

Tingkat Kesukaran

Berdasarkan hasil analisis data, tingkat kesukaran butir soal ujian semester di Kabupaten Klaten dapat diilustrasikan pada Gambar 1.



Gambar 1. Grafik Tingkat Kesukaran Butir Soal

Pada Gambar 1 tampak bahwa tingkat kesukaran butir soal ujian semester dalam empat tahun terakhir sangat bervariasi. Pada tahun pelajaran 2011/2012 tingkat kesukaran butir soal yang termasuk dalam kategori Mudah sebanyak 32 butir soal atau 65%, yang termasuk dalam kategori Sedang sebanyak 12 butir soal atau 24% dan yang termasuk dalam kategori Sukar sebanyak 6 butir soal atau 12%. Pada tahun pelajaran 2012/2013 tingkat kesukaran butir soal yang termasuk dalam kategori Mudah sebanyak 39 butir soal atau 78%, yang termasuk dalam kategori Sedang sebanyak 6 butir soal atau 12% dan yang termasuk dalam kategori Sukar sebanyak 6 butir soal atau 10%. Pada tahun pelajaran 2013/2014 tingkat kesukaran butir soal yang termasuk dalam kategori Mudah sebanyak 6 butir soal atau 12%, yang termasuk dalam kategori sedang sebanyak 7 butir soal atau 14% dan yang termasuk dalam kategori Sukar sebanyak 37 butir soal atau 74%. Pada tahun pelajaran 2014/2015 tingkat kesukaran butir soal yang termasuk dalam kategori Mudah sebanyak 32 butir soal atau 64%, yang termasuk dalam kategori Sedang sebanyak 4 butir soal atau 16% dan yang termasuk dalam kategori Sukar sebanyak 14 butir soal atau 28%.

Tingkat kesukaran merupakan rasio antara penjawab soal dengan benar dan banyaknya peserta ujian. Secara teoritik dikatakan bahwa tingkat kesukaran merupakan probabilitas empirik untuk lulus pada butir soal tertentu bagi peserta ujian tertentu. Dalam teori skor-murni klasik, tingkat kesukaran adalah parameter yang mendeskripsikan seberapa sukarlah bagi sekelompok peserta ujian yang dites untuk memberikan jawaban yang benar terhadap suatu butir soal (Azwar, 2016). Menurut Sudjana (2004), sebaiknya paket soal yang diberikan kepada siswa memiliki keseimbangan antara mudah, cukup, sukar dengan perbandingan 3:4:3 atau 3:5:2. Apabila soal ujian akhir semester berjumlah 50 maka seharusnya perbandingan antara soal mudah:cukup:sukar adalah 15:20:15 atau 11:28:11. Perbandingan hasil analisis tingkat kesukaran butir soal ujian akhir semester SMA pada mata pelajaran Bahasa

Indonesia kelas X di Kabupaten Klaten dapat dilihat pada Tabel 1.

Tabel 1. Perbandingan Hasil Analisis Tingkat Kesukaran Butir Soal

| No | Tahun Pelajaran | Perbandingan |
|----|-----------------|-----------------|
| 1 | 2011/2012 | 6,4 : 2,4 : 1,2 |
| 2 | 2012/2013 | 7,8 : 1,2 : 1,2 |
| 3 | 2013/2014 | 1,2 : 1,4 : 7,4 |
| 4 | 2014/2015 | 6,4 : 0,8 : 2,8 |

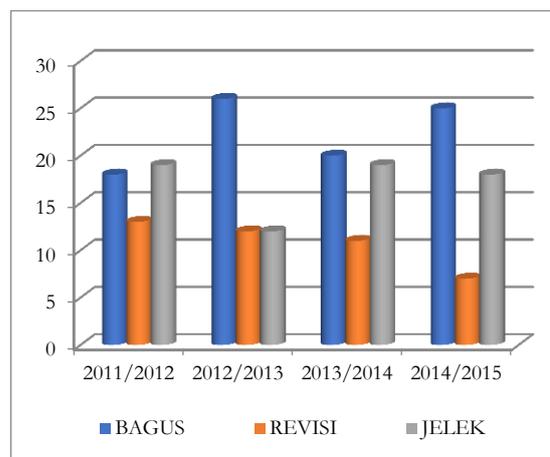
Berdasarkan Tabel 1 dapat dilihat bahwa tingkat kesukaran butir soal ujian akhir semester SMA pada mata pelajaran Bahasa Indonesia kelas X belum memenuhi unsur keseimbangan. Pada Tahun Pelajaran 2011/2012 soal didominasi oleh tingkat kesukaran mudah, demikian pula untuk Tahun pelajaran 2013/2014 dan Tahun Pelajaran 2014/2015. Namun, pada Tahun Pelajaran 2013/2014 soal didominasi oleh tingkat kesukaran sukar. Butir soal yang baik adalah yang tingkat kesulitannya cukup, tidak terlalu mudah atau terlalu sukar. Apabila butir soal yang terlalu mudah atau sukar maka butir soal tersebut tidak dapat mencerminkan capaian hasil pembelajaran yang dilakukan. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang peserta didik untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan peserta didik menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena diluar jangkauannya (Arikunto, 2007).

Daya Pembeda

Berdasarkan hasil analisis data, daya pembeda soal ujian semester di Kabupaten Klaten dapat diilustrasikan pada Gambar 2.

Pada Gambar 2 tampak bahwa daya pembeda butir soal ujian semester dalam empat tahun terakhir sangat bervariasi. Pada tahun pelajaran 2011/2012 daya pembeda butir soal yang termasuk dalam kategori Bagus sebanyak 18 butir soal atau 36%, yang termasuk dalam kategori Revisi sebanyak 13

butir soal atau 26% dan yang termasuk dalam kategori Jelek sebanyak 19 butir soal atau 38%. Pada tahun pelajaran 2012/2013 daya pembeda butir soal yang termasuk dalam kategori Bagus sebanyak 26 butir soal atau 52%, yang termasuk dalam kategori Revisi sebanyak 12 butir soal atau 24% dan yang termasuk dalam kategori Jelek sebanyak 12 butir soal atau 24%. Pada tahun pelajaran 2013/2014 daya pembeda butir soal yang termasuk dalam kategori Bagus sebanyak 20 butir soal atau 40%, yang termasuk dalam kategori Revisi sebanyak 11 butir soal atau 22% dan yang termasuk dalam kategori Jelek sebanyak 19 butir soal atau 38%. Pada tahun pelajaran 2014/2015 daya pembeda butir soal yang termasuk dalam kategori Bagus sebanyak 25 butir soal atau 50%, yang termasuk dalam kategori Revisi sebanyak 7 butir soal atau 14% dan yang termasuk dalam kategori Jelek sebanyak 18 butir soal atau 36%.



Gambar 2. Grafik Daya Pembeda Butir Soal

Daya pembeda merupakan kemampuan suatu butir soal untuk membedakan individu yang satu dari yang lainnya berdasarkan atribut yang diukur oleh tes (Azwar, 2016). Pada tes prestasi belajar Bahasa Indonesia daya pembeda adalah kemampuan butir soal dalam membedakan individu berdasarkan hasil belajar siswa dalam bidang Bahasa Indonesia. Daya pembeda merupakan suatu pernyataan tentang seberapa besar daya sebuah butir soal dapat membe-

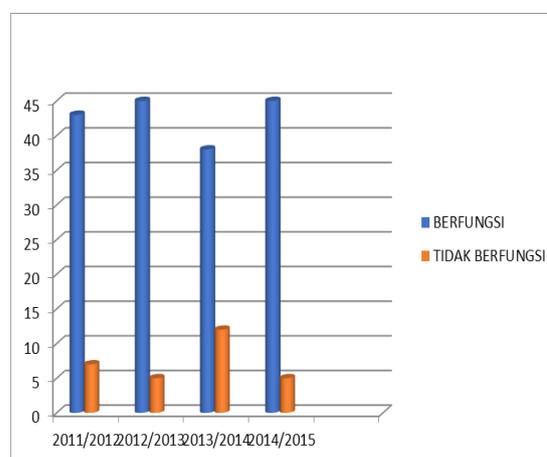
dakan kemampuan antara peserta kelompok tinggi dan kelompok rendah. Secara teoritis peserta ujian akhir semester kelompok tinggi harus bisa menjawab dengan benar butir-butir soal yang dikerjakan secara lebih banyak daripada jawaban benar kelompok rendah. Apabila jumlah jawaban benar peserta kelompok rendah lebih banyak daripada kelompok tinggi, maka butir tersebut dinyatakan tidak baik. Sebuah butir soal yang baik adalah yang mempunyai daya beda untuk membedakan kemampuan antara peserta tes ujian akhir semester pada kedua kelompok (Nurgiyantoro, 2010). Siswa yang tergolong pada kelompok tinggi memiliki peluang yang lebih besar untuk menjawab dengan benar suatu butir soal dibandingkan dengan siswa pada kelompok rendah.

Secara prinsip daya pembeda dicerminkan oleh perbedaan jawaban terhadap butir soal diantara kelompok subjek yang cerdas dan yang tidak cerdas. Besar kecilnya nilai indeks daya pembeda butir menunjukkan tinggi rendahnya daya sebuah butir soal untuk dapat membedakan kemampuan peserta ujian akhir semester. Semakin tinggi indeks yang dimiliki oleh sebuah butir soal, maka akan semakin baik butir soal tersebut karena mempunyai daya untuk membedakan kemampuan peserta ujian akhir semester. Daya pembeda soal ujian akhir semester SMA pada Mata Pelajaran Bahasa Indonesia di Kabupaten Klaten dalam empat tahun bersifat fluktuatif. Soal ujian akhir semester pada tahun pelajaran 2011/2012 hanya 36% butir soal yang bisa membedakan kemampuan peserta antar kelompok tinggi dengan kelompok rendah. Soal ujian akhir semester pada tahun pelajaran 2012/2013 yang bisa membedakan kemampuan peserta antara kelompok tinggi dengan kelompok rendah meningkat menjadi 52%. Soal ujian akhir semester pada tahun pelajaran 2013/2014 turun menjadi 40% butir soal yang bisa membedakan kemampuan peserta antar kelompok tinggi dengan kelompok rendah. Soal ujian akhir semester pada tahun pelajaran 2014/2015 naik kembali menjadi 50% butir soal yang bisa membedakan kemampuan peserta antar kelompok tinggi de-

ngan kelompok rendah. Penentuan kelompok tinggi dan kelompok rendah dengan mengambil 27% siswa yang memperoleh skor tertinggi sebagai kelompok tinggi dan 27% siswa yang memperoleh skor terendah sebagai kelompok rendah, 46% siswa ditengah distribusi tidak diikutkan dalam analisis. Menurut Kelly, pada kondisi normal, titik optimum di mana dua kondisi seimbang dicapai pada 27% kelompok tinggi dan bawah (Anastasi & Urbina, 1997).

Distraktor

Berdasarkan hasil analisis data, efektivitas distraktor soal ujian semester di Kabupaten Klaten disajikan pada Gambar 3.



Gambar 3. Grafik Distraktor

Pada Gambar 3 tampak bahwa Distraktor soal ujian semester dalam empat tahun terakhir sangat bervariasi. Pada tahun pelajaran 2011/2012 Distraktor yang Berfungsi dengan Baik sebanyak 43 butir soal atau 86%, yang Tidak berfungsi sebanyak 7 butir soal atau 14%. Pada tahun pelajaran 2012/2013 Distraktor yang Berfungsi dengan Baik sebanyak 45 butir soal atau 90%, yang Tidak berfungsi sebanyak 5 butir soal atau 10%. Pada tahun pelajaran 2013/2014 Distraktor yang Berfungsi dengan Baik sebanyak 38 butir soal atau 76%, yang Tidak berfungsi sebanyak 12 butir soal atau 24%. Pada tahun pelajaran 2014/2015 Distraktor yang Berfungsi dengan Baik sebanyak 45 butir soal atau 90%, yang Tidak berfungsi sebanyak 5 butir soal atau 10%.

Distraktor bukan sekedar pelengkap pilihan jawaban. Pengecoh diadakan dengan tujuan untuk menyesatkan siswa agar tidak memilih kunci jawaban. Agar dapat melakukan fungsinya dengan baik maka distraktor harus dibuat semirip mungkin dengan kunci jawaban (Purwanto, 2009). Agar semua pilihan jawaban efektif, penyusunan pilihan jawaban yang salah harus dilakukan sedemikian rupa sehingga tidak terlalu mencolok terlihat sebagai jawaban yang salah. Pilihan jawaban salah yang baik adalah yang serupa tapi tidak sama dengan pilihan jawaban benar sehingga mempunyai peluang untuk dipilih oleh peserta ujian akhir semester yang tidak hati-hati (Nurgiyantoro, 2010).

Item yang baik tidak saja memiliki daya pembeda yang tinggi dan tingkat kesukaran yang sesuai tapi juga memiliki distraktor yang efektif. Fungsi distraktor berlawanan dengan fungsi daya pembeda. Daya pembeda ditunjukkan dengan besarnya proporsi peserta ujian akhir semester pada kelompok tinggi yang dapat menjawab dengan benar dibandingkan dengan proporsi peserta ujian akhir semester pada kelompok rendah. Fungsi distraktor justru sebaliknya, distraktor akan berfungsi efektif jika besarnya proporsi peserta ujian akhir semester pada kelompok rendah yang terjebak oleh distraktor dibandingkan dengan proporsi peserta ujian akhir semester pada kelompok tinggi (Azwar, 2016). Idealnya distraktor harus dipilih hanya oleh peserta ujian akhir semester yang termasuk pada kelompok bawah, sedangkan peserta ujian pada kelompok atas tidak ada yang memilih distraktor. Pada soal ujian akhir semester pada tahun pelajaran 2011/2012 ada 14% distraktor yang tidak berfungsi, tahun pelajaran 2012/2013 ada 10%, tahun pelajaran 2013/2014 ada 24% dan pada tahun pelajaran 2014/2015 ada 10% distraktor yang tidak berfungsi. Terdapat beberapa faktor yang mempengaruhi berfungsi tidaknya pengecoh yaitu jika soal terlalu mudah, pokok soal memberi petunjuk pada kunci jawaban dan siswa sudah mengetahui materi yang akan ditanyakan terlalu mudah. Pengecoh dikatakan berfungsi jika dipilih oleh seba-

gian besar siswa yang berkemampuan rendah dan dipilih minimal 5% dari seluruh peserta tes yang berkemampuan tinggi. Jika pengecoh lebih banyak dipilih oleh siswa yang berkemampuan tinggi maka dapat dikatakan pengecoh tersebut menyesatkan. Apabila pengecoh dipilih secara merata oleh peserta tes, maka pengecoh tersebut berfungsi (Suaedi, 2016)

Validitas dan Reliabilitas

Berdasarkan hasil analisis validitas dapat diketahui bahwa pada soal ujian akhir semester tahun pelajaran 2011/2012 ada 18 atau 36% butir soal yang tidak valid, tahun pelajaran 2012/2013 ada 26 atau 52%, tahun pelajaran 2013/2014 ada 20 atau 40% dan tahun 2014/2015 ada 25 atau 50%. Estimasi Reliabilitas dilakukan dengan menggunakan formula *alpha cronbach*. Hasil estimasi reliabilitas soal ujian akhir semester tahun 2011/2012 yang diperoleh sebesar 0,232. Analisis estimasi reliabilitas soal ujian akhir semester tahun 2012/2013 diperoleh koefisien reliabilitas sebesar 0,428. Hasil estimasi reliabilitas soal ujian akhir semester tahun 2013/2014 yang diperoleh sebesar 0,238. Hasil estimasi reliabilitas soal ujian akhir semester tahun 2014/2015 yang diperoleh sebesar 0,603. Validitas dan reliabilitas adalah dua hal penting dalam pengukuran. Reliabilitas akan mempengaruhi atau bahkan menentukan validitas walaupun tidak semua yang reliabel pasti valid. Proses validasi sangat penting dilakukan untuk setiap tes kemampuan kognitif. Validitas adalah karakteristik terpenting dalam pengukuran yang mengacu kepada akurasi dan kecermatan fungsi ukur tes yang bersangkutan. Melakukan validasi tes adalah mencari bukti empiris bahwa hasil ukur dari tes tersebut memang memberikan informasi yang akurat dan cermat mengenai atribut yang diukur, tanpa dicemari oleh informasi yang tidak relevan (Azwar, 2016). Reliabilitas mengacu pada konsistensi pengukuran yang berarti bahwa perbedaan skor yang diperoleh dalam pengukuran memang mencerminkan adanya perbedaan kemampuan yang sesungguhnya, bukan perbedaan yang disebabkan

oleh adanya eror pengukuran. Estimasi reliabilitas artinya mencari bukti empiris bahwa hasil ukur dari tes memang memberikan variasi perbedaan yang konsisten, bukan perbedaan yang terjadi secara random atau secara kebetulan.

Wells & Wollack (2003) mengatakan bahwa tes yang dibuat oleh guru untuk keperluan *class-room test* harus memiliki koefisien reliabilitas paling tidak 0,700 atau lebih. Hasil estimasi reliabilitas soal ujian akhir semester SMA pada mata pelajaran Bahasa Indonesia kelas X di Kabupaten Klaten dari tahun pelajaran 2011/2012 sampai tahun pelajaran 2014/2015 belum ada yang mencapai 0,700. Hal ini berarti Soal ujian akhir semester SMA pada mata pelajaran Bahasa Indonesia kelas X di Kabupaten Klaten belum memenuhi kriteria koefisien konsistensi internal. Koefisien reliabilitas soal akhir ujian semester SMA belum dapat dikatakan memuaskan, artinya hasil ukur dari soal ujian akhir semester belum banyak menampilkan variasi skor yang sesungguhnya karena masih banyak dipengaruhi oleh variasi eror pengukuran. Koefisien reliabilitas soal ujian akhir semester pada tahun pelajaran 2011/2012 sebesar 0,232 menunjukkan bahwa perbedaan (variasi) yang tampak pada skor ujian akhir semester mencerminkan

23,3% dari perbedaan pada skor peserta tes yang sesungguhnya. Hal itu dapat dikatakan bahwa 76,7% dari perbedaan skor yang tampak adalah akibat adanya variasi eror pengukuran. Koefisien reliabilitas soal ujian akhir semester tahun pelajaran 2012/2013 sebesar 0,428 berarti bahwa 37,2% dari variasi skor tes Cuma menampilkan variasi eror.

Pemetaan Kualitas Butir Soal Ujian Semester Pada Mata pelajaran Bahasa Indonesia di Kabupaten Klaten

Berdasarkan hasil analisis data, Tingkat Kesukaran, Daya Pembeda dan efektivitas Distraktor, Validitas dan Reliabilitas soal ujian semester di Kabupaten Klaten dapat dilihat pada Tabel 2.

Berdasarkan Tabel 2 dapat dilihat soal ujian akhir semester tahun pelajaran 2011/2012 memiliki tingkat kesukaran butir soal mudah sebanyak 64%, sedang 24% dan sukar 12%. Daya pembeda bagus 36%, Revisi 26% dan jelek 38%. Distraktor berfungsi dengan baik sebanyak 86% dan yang tidak berfungsi sebanyak 14%. Berdasarkan hasil analisis validitas empirik diperoleh 36% butir yang valid dengan nilai koefisien reliabilitas sebesar 0,232.

Tabel 2. Pemetaan Kualitas Butir Soal Mata Pelajaran Bahasa Indonesia dalam Empat Tahun Terakhir

| Karakteristik Instrumen | | Tahun Pelajaran | | | |
|-----------------------------|---------------|-----------------|---------------|---------------|---------------|
| | | 2011/ 2012 | 2012/ 2013 | 2013/ 2014 | 2014/ 2015 |
| Tingkat Kesukaran | Mudah | 64% | 78% | 12% | 64% |
| | Sedang | 24% | 12% | 14% | 16% |
| | Sukar | 12% | 10% | 74% | 28% |
| Daya Pembeda | Bagus | 36% | 52% | 40% | 50% |
| | Revisi | 26% | 24% | 22% | 14% |
| | Jelek | 38% | 24% | 38% | 36% |
| distraktor | Berfungsi | 86% | 90% | 76% | 90% |
| | Tdk Berfungsi | 14% | 10% | 24% | 10% |
| Validitas Empirik terpenuhi | | 36% | 52% | 40% | 50% |
| Reliabilitas | | 0,232 | 0,428 | 0,238 | 0,603 |

Hasil analisis soal ujian akhir semester tahun pelajaran 2012/2013 menunjukkan bahwa tingkat kesukaran butir soal mudah sebanyak 78%, sedang 12% dan sukar 10%. Daya pembeda soal bagus sebanyak 52%, Revisi 24% dan jelek 24%. Distraktor berfungsi dengan baik sebanyak 90% dan yang tidak berfungsi sebanyak 10%. Dari 50 butir soal ada 56% butir soal yang valid dengan nilai koefisien reliabilitas 0,428.

Tingkat kesukaran soal ujian akhir semester tahun pelajaran 2013/2014 yang termasuk dalam kategori mudah sebanyak 12%, sedang 14% dan sukar 74%. Daya pembeda bagus sebanyak 40%, Revisi 22% dan jelek 38%. Distraktor dapat berfungsi dengan baik sebanyak 76% dan yang tidak berfungsi sebanyak 24%. Koefisien reliabilitas yang diperoleh sebesar 0,238 dengan butir soal yang valid sebanyak 40%.

Hasil analisis kualitas butir soal tahun pelajaran 2014/2015 menunjukkan bahwa tingkat kesukaran butir soal yang termasuk dalam kategori mudah sebanyak 64%, sedang 16% dan sukar 28%. Daya pembeda soal bagus sebanyak 50%, revisi 14% dan jelek 36%. Distraktor berfungsi dengan baik sebanyak 90% dan yang tidak berfungsi sebanyak 10%. 50% butir soal valid dengan nilai koefisien reliabilitas sebesar 0,603.

Hasil analisis Tingkat kesukaran soal ujian akhir semester SMA pada mata pelajaran Bahasa Indonesia kelas X di Kabupaten Klaten pada tahun pelajaran 2011/2012 sampai 2014/2015 belum memenuhi unsur keseimbangan antara butir soal yang mudah, sedang dan sukar seperti yang dikatakan Sudjana (2004), sebaiknya paket soal yang diberikan kepada peserta didik memiliki keseimbangan antara mudah:cukup:sukar dengan perbandingan 3:4:3 atau 3:5:2. Temuan dalam penelitian ini sejalan dengan temuan penelitian Sulistiawan (2016) dan Suaedi (2016), tingkat kesukaran butir soal buatan guru belum memenuhi keseimbangan antara butir soal mudah, sedang, dan sukar. Keseimbangan komposisi butir soal mudah, sedang, sukar sangat penting untuk diperhatikan. Apabila soal ujian akhir semester yang diberikan lebih banyak soal yang sukar maka

tidak seorang siswapun yang dapat menjawabnya dengan benar, yang akan berakibat pada skor atau nilai yang diperoleh siswa tidak akan bagus. Sesungguhnya tingkat kesukaran merupakan probabilitas empirik untuk lulus atau berhasil dalam mengerjakan soal. Demikian pula halnya dengan daya pembeda soal, masih banyak soal ujian akhir semester yang dibuat guru untuk penilaian kelas yang belum memiliki daya pembeda bagus. Jika soal ujian akhir semester yang diberikan memiliki daya pembeda yang belum bagus, maka butir soal tersebut tidak bisa membedakan antara siswa pada kelompok tinggi (kemampuan tinggi) dengan siswa pada kelompok rendah (kemampuan rendah). Distraktor pada soal ujian akhir semester pada tahun pelajaran 2011/2012 sampai tahun pelajaran 2014/2015 sudah bisa mencapai 90% yang dapat berfungsi dengan bagus walaupun masih ada yang belum berfungsi. Hasil analisis validitas dan reliabilitas yang diperoleh belum menunjukkan hasil yang bagus. Soal ujian akhir semester pada Mata Pelajaran Bahasa Indonesia kelas X yang berjumlah 50 butir soal setiap tahunnya, validitas tertinggi baru mencapai 52 % yaitu soal ujian akhir semester tahun pelajaran 2012/2013. Demikian pula dengan nilai koefisien reliabilitas, soal ujian akhir semester tahun pelajaran 2011/2012 sampai 2014/2015 belum ada yang memenuhi kriteria sebagaimana yang disyaratkan Wells & Wollack (2003) bahwa koefisien reliabilitas untuk penilaian kelas harus memiliki koefisien reliabilitas lebih dari 0,700.

Simpulan

Tingkat kesukaran butir soal ujian akhir semester di Kabupaten Klaten pada tahun pelajaran 2011/2012 memiliki tingkat kesukaran butir soal mudah sebanyak 64%, sedang 24 % dan sukar 12%. Pada tahun pelajaran 2012/2013 menunjukkan bahwa tingkat kesukaran butir soal mudah sebanyak 78%, sedang 12% dan sukar 10%. Pada tahun pelajaran 2013/2014 yang termasuk dalam kategori mudah sebanyak 12%, sedang 14% dan sukar 74%. Pada tahun pelajaran 2014/2015 menunjukkan bahwa

tingkat kesukaran butir soal yang termasuk dalam kategori mudah sebanyak 64%, sedang 16% dan sukar 28%.

Daya pembeda soal ujian akhir semester pada tahun pelajaran 2011/2012 yang termasuk dalam kategori bagus sebanyak 36%, Revisi 26% dan jelek 38%. Pada tahun pelajaran 2012/2013, daya pembeda soal bagus sebanyak 52%, Revisi 24% dan jelek 24%. Pada tahun pelajaran 2013/2014 daya pembeda bagus sebanyak 40%, Revisi 22% dan jelek 38%. Daya pembeda soal pada tahun pelajaran 2014/2015 termasuk dalam kategori bagus sebanyak 50%, revisi 14% dan jelek 36%.

Distraktor pada soal ujian akhir semester tahun pelajaran 2011/2012 dapat berfungsi dengan baik sebanyak 86% dan yang tidak berfungsi sebanyak 14%. Distraktor berfungsi dengan baik sebanyak 90% dan yang tidak berfungsi sebanyak 10% pada tahun pelajaran 2012/2013. Pada tahun pelajaran 2013/2014 distraktor dapat berfungsi dengan baik sebanyak 76% dan yang tidak berfungsi sebanyak 24%. Distraktor pada tahun pelajaran 2014/2015 dapat berfungsi dengan baik sebanyak 90% dan yang tidak berfungsi sebanyak 10%.

Hasil analisis validitas empirik pada soal ujian akhir semester tahun pelajaran 2011/2012 diperoleh 36% butir yang valid dengan nilai koefisien reliabilitas sebesar 0,232. Pada soal ujian akhir semester tahun pelajaran 2012/2013, dari 50 butir soal ada 56% butir soal yang valid dengan nilai koefisien reliabilitas 0,428. Koefisien reliabilitas soal ujian akhir semester pada tahun pelajaran 2013/2014 diperoleh sebesar 0,238 dengan butir soal yang valid sebanyak 40%. Validitas empirik pada soal ujian akhir semester tahun pelajaran 2014/2015 diperoleh 50% butir soal valid dengan nilai koefisien reliabilitas sebesar 0,603.

Berdasarkan analisis tingkat kesukran, daya pembeda dan efektifitas distraktor maka dapat disimpulkan bahwa kualitas soal ujian akhir semester pada tahun pelajaran 2011/2012 yang memenuhi kriteria sebanyak 62%, pada Tahun Pelajaran 2012/2013 sebanyak 76%, pada tahun pelajaran

2013/2014 sebanyak 62%, pada Tahun Pelajaran 2014/2015 sebanyak 64%.

Kualitas butir soal ujian semester yang digunakan dalam ujian semester di Kabupaten Klaten dalam empat tahun terakhir rata-rata termasuk dalam kategori mudah. Komposisi tingkat kesukran butir soal belum memenuhi unsur perimbangan antara soal yang mudah, sedang dan sukar. Oleh karena itu, disarankan bagi guru-guru yang menulis soal dan pihak Dinas Pendidikan dan Kebudayaan Kabupaten Kepada agar memperhatikan kualitas butir soal yang akan digunakan untuk ujian semester. Agar nanti skor yang diperoleh siswa mencerminkan kemampuan siswa sesungguhnya. Dalam arti bahwa siswa mendapatkan skor 90 pada mata pelajaran Bahasa Indonesia memang merupakan kemampuan (*trait*), bukan karena soal-soal yang diberikan dalam ujian semester termasuk dalam kategori mudah. Sehingga nilai yang diperoleh siswa di Wilayah Kabupaten Klaten dapat dibandingkan dengan wilayah lainnya.

Daftar Pustaka

- Anastasi, A. & Urbina, S. (1979). *Psychological testing*. (4th ed.). New York: Macmillan Publishing Co. Inc.
- Arikunto, S. (2007). *Dasar-dasar evaluasi pendidikan*. Edisi Revisi. Cetakan ketujuh. Jakarta: Bumi Aksara.
- Azwar, S. (2016). *Konstruksi tes kemampuan kognitif*. Yogyakarta: Pustaka Pelajar.
- Djaali & Muljono, P. (2008). *Pengukuran dalam bidang pendidikan*. Jakarta: Grasindo.
- Hayat, B., & Suprananto. (1999). *Analisis butir soal secara klasik dengan menggunakan program iteman. Pengelolaan pengujian bagi guru mata pelajaran*. Jakarta: Direktorat Pendidikan Menengah Umum, Direktorat Jendral Pendidikan Dasar dan Menengah, Depdikbud.
- Mardapi, D. (1999). Evaluasi Penyelenggaraan ebtanas. *Laporan penelitian*. Yogyakarta: Lembaga

- Penelitian IKIP Yogyakarta.
- Masrun. (1978). *Pengukuran dalam pendidikan*. Yogyakarta: Fakultas Psikologi Universitas Gadjah Mada Yogyakarta.
- Menteri Pendidikan Nasional. Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan (2007). Jakarta
- Nurgiyantoro, B. (2010). *Penilaian pembelajaran bahasa berbasis kompetensi*. Yogyakarta: BPFE.
- Pradana, S. D. S., Parno, P., & Handayanto, S. K. (2017). Pengembangan tes kemampuan berpikir kritis pada materi Optik Geometri untuk mahasiswa Fisika. *Jurnal Penelitian dan Evaluasi Pendidikan*, 21(1), 51-64. doi:<http://dx.doi.org/10.21831/pep.v21i1.13139>
- Purwanto. (2009). *Evaluasi hasil belajar*. Yogyakarta: Pustaka Pelajar.
- Suaedi, M. (2016). Analisis terhadap kualitas butir soal buatan guru IPA kelas VIII MTsN Slawi tahun pelajaran 2015/2016 *Jurnal Kamil*, 3.
- Sudjana, N. (2004). *Penilaian hasil dan proses Belajar mengajar*. Bandung: Remaja Rosdakarya.
- Sulistiawan, C. H. (2016). Kualitas soal ujian sekolah matematika program IPA dan kontribusinya terhadap hasil ujian nasional. *Jurnal Penelitian dan Evaluasi Pendidikan*, 20(1). Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/7516>.
- Umar, J. (1991). *Pengantar Penilaian Pendidikan (Makalah)*. Jakarta: PUSISJIAN, Balitbang, Depdikbud
- Umiarso & Gojali, I. (2010). *Manajemen mutu sekolah di era otonomi pendidikan*. Yogyakarta: IRCiSoD.
- Wells, C. S. & Wollack, J. A. (2003). *An instructor's guide to understanding test reliability, paper, testing & evaluation Services*. Madison: University of Wisconsin.