

### **BAGIAN 3. PENILAIAN PEMBELAJARAN**

#### **SOFTWARE ANBUSO SEBAGAI ALAT ANALISIS BUTIR SOAL YANG PRAKTIS DAN APLIKATIF**

**Ali Muhson, Barkah Lestari, Supriyanto & Kiromim Baroroh**

*Fakultas Ekonomi, Universitas Negeri Yogyakarta*

*alimuchson@yahoo.com*

##### **Abstrak**

Penelitian ini bertujuan untuk mengembangkan, menguji kelayakan software AnBuso, dan mengidentifikasi kendala penggunaannya. Penelitian dan pengembangan ini menggunakan dokumentasi, kuesioner dan wawancara untuk mengumpulkan data. Data yang terkumpul dianalisis secara deskriptif. Penelitian ini menghasilkan software AnBuso dan buku panduan yang dapat dimanfaatkan guru dalam melakukan analisis butir soal secara praktis dan aplikatif. Software tersebut dinilai sangat layak oleh guru dilihat dari aspek kepraktisan dan kemudahan, kebermanfaatan, substansi isi, dan tampilan. Kendala yang dihadapi guru dalam menggunakan software ini terkait dengan lemahnya penguasaan guru terhadap program Microsoft Excel, kurang terbiasanya melakukan analisis butir soal, pemahaman konsep analisis butir soal yang terbatas, dan kendala teknis yang terdapat dalam software.

Kata Kunci: AnBuso, kelayakan, analisis butir soal

##### **PENDAHULUAN**

Pendidikan merupakan kunci utama kemajuan bangsa sehingga upaya peningkatan kualitas pendidikan harus terus dilakukan agar Indonesia mampu bersaing di kancah dunia global. Potret kualitas pendidikan di Indonesia masih memprihatinkan. UNESCO pada tahun 2011 melaporkan bahwa indeks *Education Development Index (EDI)* Indonesia belum beranjak dari kategori sedang (*medium*) dan berada di peringkat ke-57 dari 115 (UNESCO, 2011). Sementara itu *The United Nations Development Programme (UNDP)* tanggal 24 Juli 2014 melaporkan *Human Development Index (HDI)* Indonesia menempati peringkat 108 dari 187 negara, sementara Singapura di posisi 9, Malaysia (62), Thailand (89) (UNDP, 2014). Hal ini menunjukkan bahwa kondisi kualitas pendidikan di Indonesia masih perlu ditingkatkan agar tidak ketinggalan dengan negara lain.

Proses pembelajaran menjadi bagian yang penting dalam menentukan kualitas pendidikan secara keseluruhan. Hal ini mengindikasikan bahwa pendidik memiliki peran yang sangat sentral dalam meningkatkan kualitas pendidikan. Peran pendidik tersebut tidak hanya dalam merancang dan melaksanakan pembelajaran saja melainkan juga dalam melakukan asesmen proses dan hasil belajar. Asesmen merupakan komponen yang penting dalam pembelajaran (Russel & Airasian, 2012: 2), karena memiliki pengaruh yang kuat dalam meningkatkan proses pembelajaran (Raymond, et.al., 2012; Bers, 2008: 32) bahkan penggunaan prosedur asesmen yang benar dapat memberikan kontribusi langsung kepada peningkatan belajar peserta didik (Miller, Linn & Gronlund,

2009: 34). Oleh karena itu, guru harus mampu mengembangkan alat asesmen yang baik yang mampu memotret secara tepat kompetensi yang telah dicapai peserta didik.

Guna mengukur tingkat ketercapaian tujuan pendidikan perlu dikembangkan alat asesmen yang mampu mengungkap seluruh komponen yang ingin dicapai dalam pembelajaran. Asesmen merupakan kegiatan pengumpulan bukti-bukti tentang pembelajaran siswa sebagai informasi untuk pengambilan keputusan dalam pembelajaran (Stiggins & Chappius, 2012: 3). Oleh karena itu agar keputusan yang diambil tepat, asesmen harus memperhatikan keseluruhan aspek yang akan diukur agar mampu menggambarkan dengan tepat sasaran yang dituju.

Pemberlakuan kurikulum 2013 dalam implementasinya masih banyak mengalami masalah (Republika, 2014), misalnya timbul masalah sulitnya mengubah mindset guru (Metronews, 2014; Tempo, 2013). Hasil kajian juga menunjukkan bahwa sebanyak 87 persen guru masih kesulitan dalam memahami cara asesmen autentik (Susilowati, 2013). Satu hal yang membuat guru repot adalah sistem asesmen yang memiliki terlalu banyak aspek (Tempo, 2014). Hal ini mengindikasikan bahwa kemampuan guru untuk melakukan asesmen secara baik masih perlu ditingkatkan.

Asesmen adalah upaya sistematis dalam mengumpulkan, mengkaji, dan menggunakan informasi tentang program-program pendidikan yang dilakukan untuk tujuan meningkatkan dan mengembangkan pembelajaran (Banta, Palomba, & Kinzie, 2014: 2). Dengan demikian dalam asesmen terdapat proses pengumpulan informasi, pengkajian dan penggunaan informasi tersebut untuk membuat keputusan pembelajaran agar dapat meningkatkan proses pembelajaran.

Agar asesmen menghasilkan informasi yang tepat maka perlu dilakukan dengan baik dengan cara mengumpulkan bukti akurat terkait pencapaian hasil belajar siswa dan menjadikan proses asesmen kelas dan hasilnya bermanfaat bagi siswa, yaitu mampu meningkatkan motivasi dan prestasi belajarnya (Stiggins & Chappuis, 2012: 3). Dengan demikian, asesmen harus dapat menilai kemajuan belajar siswa. Informasi tersebut dapat digunakan sebagai dasar untuk mengambil suatu keputusan tentang status siswa dalam kelompoknya dan menentukan langkah-langkah selanjutnya. Oleh karena itu dalam melakukan asesmen hasil pembelajaran perlu dirancang langkah-langkahnya secara rinci agar mampu memotret kompetensi siswa secara tepat.

Asesmen membantu guru dalam memperjelas tujuan pembelajaran dan pencapaiannya, menciptakan pengalaman yang menerapkan pengetahuan dalam konteks kehidupan nyata, dan memberikan berbagai cara bagi siswa untuk menunjukkan kemampuan dan keterampilan mereka (Darling-Hammond, 2014: 54). Prosedur asesmen yang digunakan dengan benar dapat memberikan kontribusi langsung kepada peningkatan belajar siswa, yakni (1) mengklarifikasi sifat hasil belajar yang dimaksud, (2) menyiapkan tujuan jangka pendek agar terarah, (3) memberikan umpan balik terhadap kemajuan belajar, (4) memberikan informasi dalam mengatasi kesulitan belajar dan untuk memilih pengalaman belajar masa depan, dan (5) mengidentifikasi tujuan

pembelajaran berikutnya (Miller, Linn & Gronlund, 2009: 34). Prosedur tersebut merupakan langkah yang saling berkaitan dan menentukan langkah berikutnya.

Asesmen juga bertujuan menjaga keseimbangan kelas, merencanakan dan melaksanakan pembelajaran, menempatkan siswa, memberikan umpan balik dan penghargaan, mendiagnosis masalah siswa, dan menilai tingkat kemajuan akademik (Russell & Airasian, 2012: 5-8). Hal ini mengindikasikan bahwa melalui asesmen dapat ditentukan rancangan pembelajaran berikutnya dengan cara mendiagnosis masalah yang dihadapi siswa agar prestasi akademik siswa dapat berkembang secara optimal.

Pendapat lain juga menyatakan bahwa ada tiga tujuan utama dilakukannya asesmen. Pertama, untuk memantau kemajuan pendidikan atau perbaikan. Pendidik, pembuat kebijakan, orang tua dan masyarakat ingin tahu berapa banyak siswa mencapai standar kinerja yang ditentukan. Tujuan ini, sering disebut asesmen sumatif. Tujuan kedua adalah untuk memberikan umpan balik kepada guru dan siswa. Para guru dapat menggunakan umpan balik untuk memperbaiki proses pembelajaran, dan siswa dapat menggunakan umpan balik untuk memantau pembelajaran mereka sendiri. Tujuan ini, sering disebut asesmen formatif. Tujuan ketiga asesmen adalah untuk mendorong perubahan dalam praktek dan kebijakan untuk mencapai perubahan yang diinginkan. Tujuan ini, disebut asesmen akuntabilitas (National Research Council, 1999, 1-2). Dengan demikian asesmen dapat berfungsi untuk memantau kemajuan pembelajaran, memberikan informasi sebagai dasar pemberian umpan balik, dan melakukan perbaikan pembelajaran.

Dalam melakukan asesmen kelas tidak dapat dilakukan dengan mudah namun harus mendasarkan pada beberapa kriteria. Beberapa kriteria yang harus dipenuhi dalam melakukan asesmen kelas adalah validitas, reliabilitas, terfokus pada kompetensi, komprehensif, objektivitas, dan mendidik (Puskur, 2008). Pendapat lain juga menyatakan bahwa agar hasil asesmen dapat memberikan informasi yang tepat maka harus memenuhi validitas, reliabilitas, dan objektivitas (Anderson, 2003: 10; Kubiszyn & Borich, 2013: 326). Dengan demikian, validitas dan reliabilitas menjadi bagian yang penting dalam kegiatan asesmen agar informasi yang dihasilkan sesuai dengan yang diharapkan.

*Item analyses play a somewhat more important role in construct and predictive validation* (Nunnally & Bernstein, 1994: 304). Hal ini berarti analisis butir menjadi bagian yang penting dalam menjamin validitas butir soal. Tiga hal yang diperhatikan dalam melakukan analisis butir soal adalah tingkat kesukaran, daya beda dan distraktor. Tingkat kesukaran suatu butir soal merupakan salah satu parameter butir soal yang sangat berguna dalam penganalisisan suatu tes. Hal ini disebabkan karena dengan melihat parameter butir ini, akan diketahui seberapa baiknya kualitas suatu butir soal. Jika tingkat kesukaran mendekati 0, maka soal tersebut terlalu sukar, sedangkan jika tingkat kesukaran mendekati 1, maka soal tersebut terlalu mudah. Soal yang terlalu sukar dan terlalu mudah perlu dibuang karena butir tersebut tidak dapat membedakan kemampuan seorang siswa dengan siswa lainnya. Indeks kesukaran suatu butir yang baik terletak

dalam kategori sedang yakni pada interval 0,30 – 0,70 (Allen & Yen, 1979: 121; Kaplan & Saccuzzo, 2005: 170; Sudjana, 2011: 137). Pada interval ini, informasi tentang kemampuan siswa akan diperoleh secara maksimal.

Kriteria kedua yang perlu diperhatikan adalah daya beda butir soal. Daya beda butir merujuk pada kemampuan butir soal untuk membedakan peserta tes yang berkemampuan tinggi dan berkemampuan rendah. Untuk menentukan daya pembeda dapat menggunakan indeks diskriminasi, indeks korelasi biserial, indeks korelasi *point biserial*, dan indeks keselarasan. Indeks daya pembeda suatu butir yang kecil nilainya akan menyebabkan butir tersebut tidak dapat membedakan siswa yang kemampuannya tinggi dan siswa yang kemampuannya rendah. Jika nilai daya beda rendah menunjukkan adanya kemencengan distribusi skor dari populasi sehingga mengakibatkan validitas tes menjadi rendah. Indeks daya beda dikatakan baik jika lebih besar atau sama dengan 0,3 (Nunnally & Bernstein, 2009: 304; Kaplan & Saccuzzo, 2005: 176; Azwar, 2003: 153). Sementara itu koefisien antara 0,20 – 0,29 dianggap cukup baik (Alagumalai & Curtis, 2005: 8) dan koefisien di bawah 0,2 dianggap tidak baik sehingga perlu dibuang (Ebel & Frisbie, 1991: 232; Crocker & Algina, 2006: 315).

Khusus untuk tes objektif bentuk *multiple choice* perlu dilengkapi dengan beberapa alternatif jawaban, atau yang sering dikenal dengan istilah *option*. *Option* atau alternatif itu jumlahnya berkisar antara 3 sampai dengan 5 buah, dan dari kemungkinan-kemungkinan jawaban yang terpasang pada setiap butir *item* itu, salah satu di antaranya adalah merupakan jawaban betul (kunci jawaban) sedangkan yang lainnya salah. Alternatif jawaban salah itulah yang biasa dikenal dengan istilah *distractor* (pengecoh). Pada kenyataannya bisa terjadi alternatif yang diberikan pada butir tertentu sama sekali tidak dipilih oleh peserta tes. Hal ini berarti alternatif tersebut tidak mampu berfungsi sebagai pengecoh yang baik. Pengecoh dinyatakan telah dapat menjalankan fungsinya dengan baik apabila distraktor tersebut sekurang-kurangnya sudah dipilih oleh 5 % dari seluruh peserta tes.

Asesmen akan menjadi bermakna dalam proses pembelajaran manakala hasil asesmen tersebut dimanfaatkan dan ditindaklanjuti. Umpan balik asesmen merupakan bagian yang tidak terpisahkan dari sebuah proses pembelajaran dan asesmen itu sendiri. Umpan balik bukanlah hal yang asing dalam dunia pendidikan dan asesmen (Irons, 2008: 1). Umpan balik akan sangat bermakna jika dilakukan secara tepat (Brookhart, 2008: 2) karena dapat meningkatkan proses pembelajaran (Irons, 2008: 7). Hasil penelitian juga menunjukkan bahwa umpan balik berdampak positif terhadap hasil belajar (James & Folorunso, 2012; Delacruz, 2012). Artinya diperlukan kemampuan dan strategi khusus dalam memberikan asesmen yang baik serta *feedback* (umpan balik) yang tepat agar mampu meningkatkan kualitas pembelajaran. Oleh karena itu sudah selayaknya jika guru memiliki kompetensi yang memadai dalam melakukan asesmen dan umpan balik.

Umpan balik merupakan bukti yang menegaskan atas kebenaran suatu tindakan (Wiggins, 1993: 185). Umpan balik dikonseptualisasikan sebagai informasi yang diberikan oleh perantara (misalnya, guru, teman sebaya, buku, orang tua, diri,

pengalaman) mengenai aspek kinerja atau pemahaman seseorang (Hattie & Timperley 2007). Umpan balik formatif adalah setiap informasi, proses atau kegiatan yang memberi atau mempercepat belajar siswa berdasarkan komentar yang berkaitan dengan asesmen formatif dan kegiatan asesmen sumatif (Irons, 2008: 7). Umpan balik formatif biasanya disajikan sebagai informasi kepada peserta didik dalam menanggapi beberapa tindakan peserta didik. Bentuknya dapat berupa berbagai jenis (misalnya, verifikasi akurasi respon, penjelasan jawaban yang benar, petunjuk, pemberian contoh) (Shute 2007: i).

Pengertian tersebut menunjukkan bahwa *umpan balik* merupakan sesuatu yang diberikan untuk dapat mengkaji apa yang telah dilakukan. Umpan balik itu sendiri adalah salah satu upaya untuk mengobservasi siswa berkaitan dengan bagaimana mereka melakukan aktivitas serta apa yang harus dilakukan guru untuk meningkatkan kemampuan siswa.

Tujuan utama dari umpan balik formatif adalah untuk meningkatkan pengetahuan, keterampilan, dan pemahaman peserta didik (misalnya, pemecahan masalah) (Shute, 2007: 6) karena itu umpan balik harus bersifat interaktif, meningkatkan motivasi dan berupaya memecahkan masalah (Langer, 2011). Peterson & Irving (2008) dalam penelitiannya menjelaskan bahwa ternyata siswa berpandangan umpan balik dapat memberikan informasi dan balikan yang baik kepada mereka. Komentar yang diberikan dalam umpan balik formatif hanya dapat efektif jika siswa membaca dan memanfaatkannya (Higgins & Hartley, 2002). Umpan balik dapat digunakan untuk meningkatkan efektivitas pembelajaran di kelas (Hattie & Timperley, 2007). Umpan balik yang diberikan kepada peserta didik jika diberikan secara tepat akan membantu mereka meningkatkan kinerjanya, memberikan ide tentang bagaimana mereka berkembang, meningkatkan motivasi dan memberdayakan mereka sebagai peserta didik (Harvey, 2011: 20). Dengan demikian dapat dikatakan bahwa umpan balik memiliki peran yang penting dalam proses pembelajaran, artinya melalui umpan balik dapat mengarahkan proses pembelajaran sesuai dengan tujuan yang ingin dicapai.

Kajian di atas menunjukkan bahwa asesmen pembelajaran merupakan kegiatan yang penting dilakukan. Asesmen harus mampu mengukur secara tepat kompetensi peserta didik sehingga instrumen yang digunakan haruslah valid. Analisis butir soal memiliki peran penting untuk mengidentifikasi butir soal yang baik. Hasil asesmen juga memberikan informasi yang akurat tentang kemampuan peserta didik sehingga dapat diidentifikasi materi mana yang dianggap sulit, bahkan hasil analisis juga memberikan informasi tentang materi mana yang belum dikuasai oleh masing-masing peserta didik sehingga dapat dilakukan upaya perbaikan melalui kegiatan remedial. Guna memenuhi hal tersebut perlu dikembangkan sebuah software analisis soal yang praktis dan aplikatif sehingga dapat memotivasi guru senantiasa melakukan analisis butir soal.

Berbagai software analisis butir soal memang sudah banyak dikembangkan oleh para ahli namun belum dimanfaatkan secara optimal oleh guru. Hal itu terjadi karena sebagian besar software berbahasa asing sehingga sulit untuk memahami cara penggunaannya. Software tersebut juga cukup rumit untuk digunakan dan kurang praktis

dan aplikatif. Informasi yang diberikan dalam software tersebut juga ditampilkan dalam format yang sangat beragam sehingga mempersulit guru untuk menguasainya. Oleh karena itu perlu dikembangkan software analisis butir soal yang praktis dan aplikatif sesuai dengan kebutuhan guru di lapangan.

Guna memenuhi tuntutan tersebut, Muhson, dkk (2013) telah berhasil mengembangkan software yang diberi nama AnBuso (Analisis Butir Soal). Dalam software AnBuso tersebut dapat diketahui baik tidaknya soal yang dibuat guru, baik dari sisi daya beda, tingkat kesulitan, maupun efektivitas distraktornya. Di samping itu dalam software tersebut juga memberikan informasi tentang kemampuan seluruh siswa dan tingkat ketercapaian KKM (Kriteria Ketuntasan Minimal). Software ini juga dirancang untuk mampu mengidentifikasi dan mengelompokkan siswa yang masuk dalam program remedial berdasarkan materi yang belum dikuasai sehingga akan mempermudah guru dalam pelaksanaan program remedial. Semua hasil analisis tersebut sudah ditampilkan dan dapat dicetak dalam format laporan yang sangat mudah untuk dibaca dan ditafsirkan.

Hasil ujicoba terbatas ditemukan bahwa keberadaan software AnBuso disambut positif oleh guru sebagai alternatif untuk melakukan analisis butir soal. Bahkan beberapa guru yang telah menggunakan AnBuso merasa bahwa software ini lebih mudah digunakan, praktis, dan aplikatif sehingga mereka mengaku selalu menggunakan software AnBuso dibandingkan dengan software lain (Muhson, dkk, 2013).

Hasil temuan di atas menunjukkan bahwa software AnBuso yang dikembangkan pada tahap awal ini telah direspon positif oleh guru serta sangat bermanfaat dan siap untuk digunakan. Sebagai produk awal, perlu lebih dikembangkan dan disempurnakan lagi agar kelemahan dan kekurangan yang ada di software tersebut dapat diatasi. Oleh karena itu software tersebut perlu dikaji dan diujicoba lagi dengan melibatkan guru dan pengawas yang lebih banyak agar diperoleh masukan yang lebih komprehensif. Penelitian ini dimaksudkan untuk mengetahui kelayakan dan mengidentifikasi kendala yang dihadapi dalam menggunakan software AnBuso. Tujuan akhirnya adalah menghasilkan produk berupa software AnBuso yang siap untuk dipublish kepada khalayak sasaran.

## **METODE**

Penelitian ini menggunakan model *Research and Development (R & D)*. Prosedur pengembangan dilakukan dengan langkah perancangan dan pengembangan produk, validasi produk, uji coba produk, dan diseminasi produk. Kegiatan perancangan dan pengembangan produk sudah dilakukan sampai pada tahap ujicoba produk tetapi masih pada ujicoba terbatas. Penelitian ini berusaha untuk melanjutkan pengembangan produk dengan melakukan ujicoba produk pada khalayak yang lebih luas agar diperoleh informasi dan masukan yang lebih komprehensif untuk kepentingan penyempurnaan produk.

Penelitian ini melibatkan guru-guru dan pengawas sekolah di DIY. Responden yang dilibatkan 65 orang yang berasal dari lima kabupaten/kota di provinsi DIY. Penentuan responden dilakukan secara *purposive sampling* dengan mempertimbangkan

kemampuan guru dalam penguasaan komputer, khususnya program aplikasi Microsoft Excel.

Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi observasi, dokumentasi, angket, dan wawancara. Observasi digunakan untuk memperoleh data atau informasi tentang kemampuan guru dalam menggunakan program aplikasi yang telah dikembangkan. Hal ini diperlukan untuk diperoleh data tentang kemampuan guru dalam penggunaan software yang telah dikembangkan.

Teknik dokumentasi dalam penelitian ini digunakan untuk memperoleh data tentang dokumen tes yang digunakan guru dalam mengukur kemampuan peserta didik, baik tes formatif maupun tes sumatif. Dokumen tersebut dapat berupa soal-soal ujian dan ulangan harian, program remedial, dan hasil analisis butir soal yang selama ini digunakan guru.

Angket digunakan untuk mengungkap masukan-masukan yang diperlukan dari guru, pengawas, pejabat dinas pendidikan dan para pakar. Angket ini juga sekaligus digunakan untuk mengukur tingkat kelayakan dari software yang telah dikembangkan. Angket yang dikembangkan meliputi angket kelayakan software baik yang terkait dari sisi tampilan, substansi materi/isi, aspek kebermanfaatan, dan aspek kepraktisan dan kemudahan. Teknik terakhir yang digunakan adalah wawancara yang dilakukan kepada guru, pengawas, pejabat Dinas Pendidikan dan para pakar. Teknik ini digunakan untuk mengungkap berbagai kelebihan dan kelemahan dari software yang dikembangkan agar dapat dijadikan sebagai masukan untuk penyempurnaan.

Analisis data yang digunakan meliputi analisis deskriptif. Analisis yang dilakukan meliputi analisis kelayakan software baik dilihat dari sisi tampilan, substansi materi/isi, maupun kepraktisan dan kemudahan. Dalam melakukan analisis ini digunakan lima kategori seperti terlihat pada Tabel 1.

Tabel 1. Kategorisasi Penilaian Kelayakan Software

No	Skor	Kategori
1	Lebih dari $M + 1,8 SD$	Sangat layak
2	$M + 0,6 SD$ s.d. $M + 1,8 SD$	Layak
3	$M - 0,6 SD$ s.d. $M + 0,6 SD$	Cukup
4	$M - 1,8 SD$ s.d. $M - 0,6 SD$	Tidak layak
5	Kurang dari $M - 1,8 SD$	Sangat tidak layak

## **HASIL DAN PEMBAHASAN**

Untuk memperoleh gambaran tentang hasil pengembangan software dilakukan ujicoba luas yang melibatkan para user seperti guru, pengawas, dan pelaku pendidikan sebanyak 65 orang. Sebagian besar (72%) mereka mengajar di tingkat SLTA baik SMA, SMK maupun MA. Sebagian besar mereka 68% berasal dari sekolah negeri dan responden yang sudah PNS sebanyak 57%. Agar penelitian ini mampu memperoleh gambaran yang memadai maka guru-guru yang dilibatkan juga berasal dari berbagai

bidang studi, di antaranya Ekonomi, Akuntansi, Matematika, IPS, Bahasa Inggris, Bahasa Indonesia, Teknologi Informasi, Kimia, Fisika, Biologi, Geografi, Bahasa Arab, dan sebagainya.

Jika dilihat dari kemauan guru dalam melakukan analisis butir soal tampaknya masih memprihatinkan. Sebagian besar guru 57% memang sudah melakukan analisis butir soal namun masih bersifat kadang-kadang. Hanya 11% saja yang selalu melakukan analisis butir soal sedangkan yang tidak pernah melakukan analisis butir soal sebanyak 12%. Hal ini tentu menjadi penting untuk dikaji mengapa guru sebagai pelaku pendidikan memiliki kemauan yang rendah dalam melakukan analisis butir soal.

Pada umumnya guru hanya melakukan analisis butir soal jika memang dituntut oleh pengawas. Artinya kesadaran guru untuk melakukan analisis butir soal terhadap semua soal yang sudah diujikan kepada siswa masih kurang. Hal ini terjadi karena umumnya guru kurang menguasai software analisis butir soal yang sudah ada. Kalaupun menguasai tampaknya tidak mampu membangkitkan minat dan kemauan guru dalam melakukan analisis butir soal secara terus menerus. Tentu saja hal ini akan berdampak pada rendahnya kualitas butir soal yang dikembangkan guru karena tidak selalu dilakukan analisis.

Penelitian ini berhasil mengembangkan software AnBuso dan buku panduannya yang sudah diperbaiki sesuai masukan responden. Buku panduan dikembangkan untuk memudahkan pengguna dalam memanfaatkan software ini. Buku panduan ini sekaligus memberikan informasi tentang langkah-langkah dan cara menggunakan software ini sehingga diper memudahkan user untuk memanfaatkan software dalam melakukan analisis butir soal. Panduan ini berisi tentang pendahuluan, kerangka isi, data input, dan data laporan.

Beberapa perubahan yang penting yang dihasilkan adalah penyesuaian software ini dengan diberlakukannya kurikulum 2013 terutama yang terkait dengan masalah penilaian. Karena itu pada penelitian ini dilakukan revisi perbaikan yang meliputi perubahan tampilan, sheet **Input01**, sheet **Laporan Peserta**, sheet **Peserta Remedial**, dan perubahan formula.



Sebelum Perubahan

Sesudah Perubahan

Gambar 1. Tampilan Sheet Input01 Sebelum dan Sesudah Perubahan

Perubahan tampilan perlu dilakukan karena dianggap terlalu banyak variasi warna sehingga terlihat kurang menarik. Oleh karena itu dilakukan perubahan-perubahan sesuai masukan. Perubahan tampilan tidak hanya dilakukan pada sheet input (Gambar 1) tetapi juga dilakukan perubahan tampilan pada sheet laporan (Gambar 2).

Sebelum Perubahan

Sesudah Perubahan

Gambar 2. Tampilan Sheet Laporan Butir Sebelum dan Sesudah Perubahan

Akibat diberlakukannya kurikulum 2013, software AnBuso juga dilakukan penyesuaian agar software ini mampu mengakomodasikan kepentingan guru dalam membuat penilaian sesuai dengan kurikulum 2013. Beberapa perubahan yang dilakukan mencakup dalam hal penentuan skala penilaian. Pada software sebelumnya skala penilaian yang disediakan hanya 1-10 dan 1-100, sementara itu kurikulum 2013 menggunakan skala penilaian 1-4, karena itu dalam software ini dilakukan penyesuaian dengan menyediakan skala penilaian 1-4 (Gambar 3).

Keterangan	Kolom Pengisian	VALIDASI
Tempat Laporan	Yogyakarta	OK
Tanggal Laporan	20 Agustus 2014	OK
Skala Penilaian (4, 10 atau 100)	4	OK
Nilai KKM	2.66	OK

Jumlah dan Bobot Soal		
Jumlah soal pilihan ganda (Max 50)	20	OK
Jumlah soal essay (Max 10)	5	OK
Bobot soal pilihan ganda	60%	OK
Bobot soal essay	40%	OK

Gambar 3. Perubahan Skala Penilaian

Berdasarkan hasil ujicoba luas juga ditemukan beberapa kendala dalam menentukan bobot penilaian antara soal objektif dan soal essay. Karena itu software ini juga dilakukan perubahan dalam penentuan bobot tersebut dengan menyediakan kolom tersendiri untuk bobot soal objektif dan soal essay (perubahannya dapat dilihat pada Gambar 4). Dengan cara tersebut diharapkan guru atau user akan semakin praktis dalam menentukan bobot penilaiannya. Bahkan software ini juga memungkinkan untuk digunakan hanya untuk soal objektif saja atau untuk soal essay saja.

AnBuso 5.1 For Teacher © 2011-2014 by Ali Muhson		
(Hanya diperkenankan mengisi data atau menghapus tetapi tidak boleh memindah isi data atau menggunakan fasilitas Cut Paste)		
Keterangan	Kolom Pengisian	VALIDASI
Tempat Laporan	Yogyakarta	OK
Tanggal Laporan	20 Agustus 2014	OK
Skala Penilaian (4, 10 atau 100)	4	OK
Nilai KKM	2.66	OK
Jumlah dan Bobot Soal		
Jumlah soal pilihan ganda (Max 50)	20	OK
Jumlah soal essay (Max 10)	5	OK
Bobot soal pilihan ganda	60%	OK
Bobot soal essay	40%	OK
Data Soal Pilihan Ganda		
Jumlah Alternatif Jawaban (Max 5)	4	OK

Gambar 4. Penambahan Bobot Penilaian

Penyesuaian dengan kurikulum 2013 juga berdampak pada perubahan pada Sheet **Laporan Peserta**. Pada bagian ini dimunculkan hasil penilaian peserta tes menurut kurikulum 2013 lengkap dengan predikatnya (hasil perbaikannya dapat dilihat pada Gambar 5). Predikat penilaian dilakukan penyesuaian berdasarkan Permendikbud Nomor 52. Sementara itu pada bagian yang lain tidak mengalami perubahan karena sudah sesuai dengan yang diharapkan guru.

Pada dasarnya tujuan guru melakukan analisis butir soal di samping untuk mengetahui kualitas butir soal yang telah dibuat juga informasi hasil penilaian pesertanya dapat dimanfaatkan untuk melakukan rencana tindak lanjut baik untuk keperluan remedial maupun pengayaan. Oleh karena itu hasil laporan peserta haruslah mampu memberikan gambaran siapa saja peserta yang masuk pada kelompok pengayaan dan remedial. Pada pengembangan software ini sudah mampu mengantisipasi hal tersebut namun demikian pada versi sebelumnya hanya sebatas pengelompokan peserta remedial saja dan belum disediakan kolom untuk melakukan tindak lanjut.

DAFTAR NILAI SISWA											
Satuan Pendidikan		: SMA Negeri 80 Yogyakarta									
Nama Tes		: Sumatif									
Mata Pelajaran		: Ekonomi									
Kelas/Program		: XI/IPS									
Tanggal Tes		: 30 Agustus 2014									
Pokok Bahasan/Sub		:									
		<table border="1"> <tr> <td>KKM</td> </tr> <tr> <td>2.66</td> </tr> </table>								KKM	2.66
KKM											
2.66											
No	NAMA PESERTA	L/P	HASIL TES OBJEKTIF			NILAI TES ESSAY	NILAI AKHIR	PREDIKAT	KET		
			BENAR	SALAH	NILAI						
1	RIANDITA PRATIWI	P	19	1	3.80	3.51	3.68	A	Tuntas		
2	N VIAN BURHANUDIN L	L	13	7	2.60	2.77	2.67	B	Tuntas		
3	DWI CAHYO ADHI	L	12	8	2.40	3.14	2.70	B	Tuntas		
4	ISHMATUL FATHIYAH	P	12	8	2.40	3.32	2.77	B	Tuntas		
5	BAGUS PRATOMO ARIF AL HAKIM	L	11	9	2.20	2.89	2.48	B-	Belum tuntas		
6	R BRATA KUSUMASMARA	L	15	5	3.00	3.14	3.06	B+	Tuntas		
7	LILIANA ROMANTIKA SARI	P	12	8	2.40	3.02	2.65	B-	Belum tuntas		
8	RIYAN ABI SAYEKTI	P	18	2	3.60	2.52	3.17	B+	Tuntas		
9	DWI TITIK PUSPITOSARI	P	12	8	2.40	2.65	2.50	B-	Belum tuntas		
10	PIPIT HARYADI	L	11	9	2.20	2.58	2.35	B-	Belum tuntas		
11	AWIG JUWANTO	L	17	3	3.40	2.65	3.10	B+	Tuntas		
12	ELIN DAHRIAH	P	15	5	3.00	2.71	2.88	B	Tuntas		

Gambar 5. Perbaikan Sheet Laporan Peserta

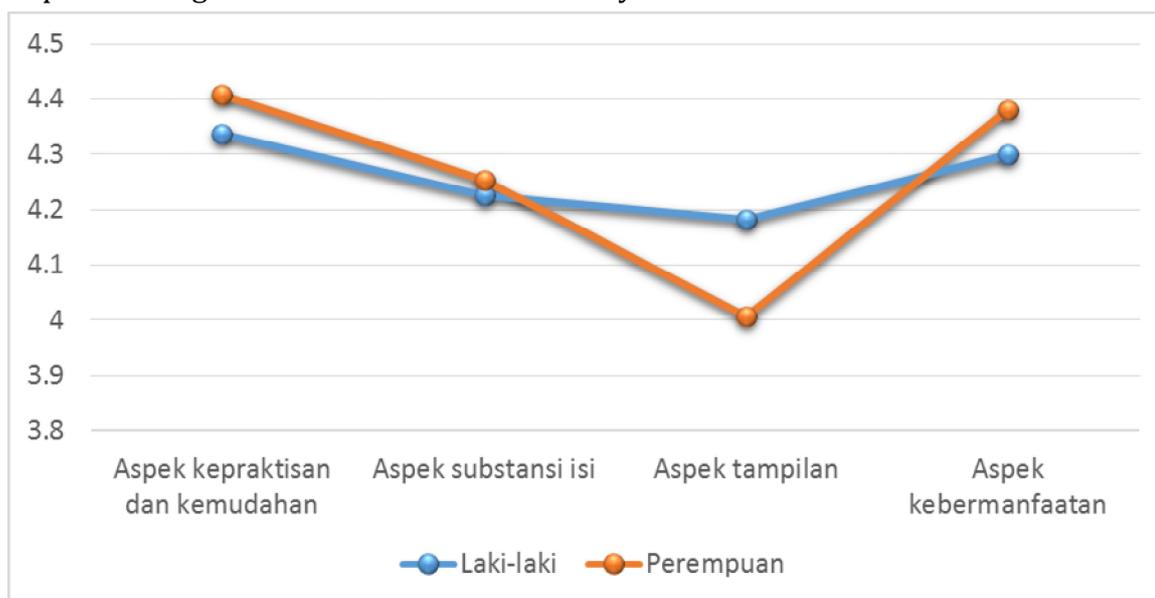
Guna memenuhi hal tersebut pada sheet **Peserta Remedial** dilakukan perbaikan yakni tidak hanya menemukan kelompok peserta remedial menurut kemampuan yang diukur melainkan disediakan kolom untuk pengisian jadwal kegiatannya. Hal itu diperlukan agar mampu mengakomodasikan kebutuhan guru dalam membuat jadwal remedial yang lebih praktis. Perbaikan pada sheet **Laporan Peserta Remedial** tersebut dapat dilihat pada Gambar 6.

PENGELOMPOKAN PESERTA REMIDIAL						
Satuan Pendidikan		: SMA Negeri 80 Yogyakarta				
Nama Tes		: Sumatif				
Mata Pelajaran		: Ekonomi				
Kelas/Program		: XI/IPS				
Tanggal Tes		: 30 Agustus 2014				
Pokok Bahasan/Sub		:				
No	Kompetensi Dasar	Peserta Remedial	Hari	Tgl	Jam	Tempat
1	Masalah Pokok Ekonomi	DWI CAHYO ADHI ; LILIANA ROMANTIKA SARI ; ELIN DAHRIAH; FITRIYANTI SYAH;				
2	Kelangkaan	Tidak Ada				
3	Sistem ekonomi	N VIAN BURHANUDIN L ; DWI CAHYO ADHI ; FITRIYANTI SYAH; TINA RETNANINGTYAS; NUR SABAR SETIANI;				
4	Perilaku Konsumen	N VIAN BURHANUDIN L ; DWI CAHYO ADHI ; ISHMATUL FATHIYAH; BAGUS PRATOMO ARIF AL HAKIM ; DWI TITIK PUSPITOSARI; FITRIYANTI SYAH; TINA RETNANINGTYAS; SITI SAONAH;				
5	Kebijakan Moneter	DWI CAHYO ADHI ; LILIANA ROMANTIKA SARI ; NING TYAS TRIASTUTI; NURLITA GADISA R;				
6	Elastisitas Permintaan	N VIAN BURHANUDIN L ; DWI CAHYO ADHI ; R BRATA				

Gambar 6. Perbaikan Sheet Laporan Peserta

Hasil uji kelayakan memperlihatkan bahwa software AnBuso yang dikembangkan dalam penelitian ini terbukti dinilai sangat layak. 51% responden menyatakan layak dan 46% menyatakan sangat layak sementara yang lainnya menyatakan cukup layak. Hal ini menunjukkan bahwa software yang dikembangkan ini memiliki kebermanfaatan yang tinggi dalam membantu guru untuk melakukan analisis butir soal.

Aspek kelayakan yang dinilai paling tinggi adalah aspek kepraktisan dan kemudahan, dan aspek kebermanfaatan. Sementara aspek yang dinilai paling rendah adalah aspek tampilan (Gambar 7). Dilihat dari jenis kelamin guru juga tampak tidak menunjukkan adanya perbedaan yang signifikan terkait dengan penilaian mereka terhadap kelayakan software. Software AnBuso dianggap praktis dan mudah untuk digunakan serta bermanfaat dalam membantu melakukan analisis butir soal. AnBuso dikembangkan dengan Microsoft Excel sehingga mempermudah guru dalam menggunakannya. Hasil analisis yang dihasilkan juga memberikan informasi yang lengkap. AnBuso tidak hanya mampu menganalisis butir soal objektif saja melainkan juga soal essay. Hasil analisis juga sudah dibuat dalam format laporan sehingga mempermudah guru dalam menafsirkan hasilnya.



Gambar 7. Hasil Uji Kelayakan Software

Aspek tampilan tampaknya dinilai paling rendah dibandingkan dengan yang lain. Hal ini terjadi karena adanya keterbatasan dalam mengatur tampilan karena program ini bukanlah program yang berdiri sendiri melainkan melekat dengan program Microsoft Excel. Akibatnya tampilan yang dihasilkan juga menyesuaikan dengan fitur yang tersedia dalam Microsoft Excel. Komponen yang dinilai rendah adalah kesesuaian pemilihan warna, tata letak dan topografi (pemilihan jenis font). Hal ini mengindikasikan bahwa software ini perlu perbaikan dari sisi tampilan. Variasi warna dan pemilihan font perlu dilakukan perubahan agar tampilannya menjadi lebih menarik. Bahkan bila perlu menggunakan program desain grafis dalam merancang tampilan.

Aspek substansi isi dari software dianggap sangat baik karena sesuai dengan kebutuhan guru. Software dinilai praktis untuk digunakan, menarik, inovatif, kreatif, interaktif dan unik. Informasi yang dihasilkan dari software ini sangat lengkap, tidak hanya menampilkan hasil analisis butir soal objektif dan essay melainkan juga menampilkan hasil pencapaian nilai dan KKM peserta didik. Bahkan dalam software ini dapat ditemukan materi-materi tertentu yang belum dikuasai oleh masing-masing peserta didik sehingga dapat memberikan informasi kepada guru dalam merancang program remedial dan pengayaan.

Dilihat dari aspek kepraktisan dan kemudahan dari software dianggap sangat baik. Software dinilai mudah digunakan, dipahami, dipelajari, dibaca dan ditafsirkan hasilnya. Untuk memanfaatkan software ini tidak perlu belajar program baru namun cukup menggunakan pengetahuan dan keterampilan dalam menggunakan Microsoft Excel. Oleh karena program ini umumnya sudah dikuasai guru maka software ini menjadi mudah untuk dimanfaatkan dan diaplikasikan.

Software juga dinilai memiliki manfaat yang tinggi oleh guru. Software yang dihasilkan dinilai sangat bermanfaat, aplikatif dan sesuai dengan kebutuhan guru. Hal ini menunjukkan bahwa keberadaan software ini memiliki kebermanfaatan yang tinggi dalam membantu guru untuk melakukan analisis butir soal. Hasil analisis yang ditampilkan dari software ini sangat sesuai dengan kebutuhan guru karena dapat dimanfaatkan untuk melakukan evaluasi terhadap instrumen dan soal yang dibuat guru dalam mengukur kompetensi peserta didik. Oleh karena tampilan hasil analisis sudah dibuat dalam format laporan maka hasil analisis ini juga dapat dipergunakan untuk keperluan membuat laporan administrasi guru.

Walaupun software AnBuso ini dinilai layak namun dalam kenyataannya ada beberapa kendala yang dihadapi guru dalam menggunakan software ini. Dari sisi manfaat yang dihasilkan dari software ini memang sangat baik namun masih ada beberapa guru yang kurang mahir dalam penguasaan komputer, khususnya program Microsoft Excel. Umumnya pengetahuan guru terhadap program ini sangatlah terbatas. Masih banyak menu dan fasilitas yang disediakan Microsoft Excel namun belum dimanfaatkan secara optimal. Oleh karena software ini terintegrasi dengan program Excel maka akibatnya guru kurang lancar dalam memanfaatkan software ini.

Kesediaan dan kebiasaan guru dalam melakukan analisis butir soal juga masih dalam kategori jarang. Hanya sedikit guru yang selalu melakukan analisis butir soal baik terhadap soal harian yang dibuatnya, soal semeseteran maupun soal ujian akhir. Karena kebiasaan mereka tersebut akibatnya guru kurang lancar dalam melakukan analisis butir soal karena memang belum terbiasa.

Pengetahuan dan penguasaan guru tentang konsep analisis butir soal juga masih terbatas. Sementara pengembangan software ini juga didasarkan pada konsep analisis butir soal terutama analisis klasik, akibatnya pemahaman guru terhadap angka-angka yang dihasilkan dari software masih kurang. Walaupun hasil analisis sudah dikemas

dalam bentuk laporan yang siap ditandatangani, namun masih perlu dijelaskan tentang arti dan makna dari hasil analisis tersebut.

## **SIMPULAN**

Penelitian ini berhasil mengembangkan software AnBuso yang dapat dimanfaatkan guru dalam melakukan analisis butir soal secara praktis dan aplikatif. Software ini dibuat dengan program Microsoft Excel yang di dalamnya terdapat sheet untuk input data, sheet data processing, sheet laporan hasil analisis dalam bentuk tabel dan gambar.

Software yang dihasilkan terbukti sangat layak oleh guru dilihat dari aspek kepraktisan dan kemudahan, aspek kebermanfaatan dan aspek substansi isi serta aspek tampilan. Walaupun demikian, ada beberapa kendala yang dihadapi guru dalam menggunakan software ini seperti lemahnya penguasaan guru terhadap program Microsoft Excel, kurang terbiasanya melakukan analisis butir soal, pemahaman konsep analisis butir soal yang terbatas, dan kendala teknis yang terdapat dalam software.

Software ini terbukti sangat layak dan sangat bermanfaat bagi guru karena itu perlu sosialisasi yang lebih luas tentang penggunaan software ini agar lebih dikenal oleh guru sehingga mampu meningkatkan kinerja guru dalam melakukan analisis butir soal. Pengembangan software ini masih perlu terus dilakukan agar mampu memenuhi kebutuhan guru dalam melakukan analisis butir soal.

## **DAFTAR PUSTAKA**

- Alagumalai, S. & Curtis, D.D. 2005. *Classical Test Theory*. In Alagumalai, S., et.al. (Eds.). *Applied Rasch Measurement: A Book of Exemplars*. Norwell, MA: Springer.
- Allen, M. J. & Yen, W. M. 1979. *Introduction to measurement theory*. Monterey, CA: Brooks/Cole Publishing Company.
- Anderson, L.W. 2003. *Classroom assessment: enhancing the quality of teacher decision making*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Azwar, S. 2003. *Sikap Manusia: Teori dan Pengukurannya, Edisi ke-2*. Yogyakarta: Pustaka Pelajar.
- Banta, T.W., Palomba, C.A., & Kinzie, J. 2014. *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Fransisco: Jossey-Bass.
- Bers, T.H. 2008. The role of institutional assessment in assessing student learning outcomes. *New Directions for Higher Education*, 141: 31-39.
- Brookhart, S.M. 2008. *How to give effective feedback to your students*. Virginia: Association for Supervision and Curriculum Development.
- Crocker, L & Algina, J. 2008. *Introduction to classical and modern test theory*. Ohio: Cengage Learning.
- Darling-Hammond, L. 2014. *Next generation assessment: Moving beyond the bubble test to support 21st century learning*. San Fransisco: Jossey-Bass.

- Ebel, R.L. & Frisbie, D.A. 1991. *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Harvey, L. 2011. The nexus of feedback and improvement. Dalam Nair, C.S. & Mertova, P. (eds.). *Student Feedback: The cornerstone to an effective quality assurance system in higher education*. New Delhi: Oxford Cambridge.
- Hattie J. & Timperley, H. 2007. The power of feedback. *Review of Educational Research*. 77(1): 81-112.
- Higgins, R., & Hartley, P. 2002. The conscientious consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1): 53-64.
- Irons, A. (2008). *Enhancing learning through formative assessment and feedback*. New York: Routledge.
- James, A.O. & Folorunso, A.M. 2012. Effect of feedback and remediation on students' achievement in junior secondary school mathematics. *International Education Studies*, 5(5): 153-162.
- Kaplan, R.M. & Saccuzzo, D.P. 2005. *Psychological Testing: Principles, Applications, and Issues, 6th edition*. Belmont, CA: Thomson Wadsworth.
- Kubiszyn, T., & Borich, G.D. 2013. *Educational testing and measurement: classroom application and practice. 10<sup>th</sup> edition*. Hoboken, NJ: John Wiley & Sons, Inc.
- Langer, P. 2011. The use of feedback in education: a complex instructional strategy *Psychological Reports*, 109(3): 775-784.
- Metronews. 2014. *Ini delapan masalah dalam implementasi kurikulum 2013*. (Online) (News.metronews.com), diakses 19 Oktober 2014.
- Miller, M.D., Linn, R.L., & Gronlund, N.E. 2009. *Measurement and assessment in teaching (tenth edition)*. New Jersey: Pearson Education Inc.
- Muhson, A., Lestari, B., Supriyanto, & Baroroh, K. 2013. *Pengembangan Software AnBuso Sebagai Solusi Alternatif Bagi Guru dalam Melakukan Analisis Butir Soal Secara Praktis dan Aplikatif*. Laporan Penelitian tidak dipublikasikan. Yogyakarta: LPPM UNY.
- National Research Council 1999. *The assessment of science meets the science of assessment. Board on Testing and Assessment Commission on Behavioral and Social Sciences and Education, National Research Council*. Washington, DC: National Academy Press.
- Nunnally, J.C. & Bernstein, I.H. 1994. *Psychometric Theory (Third Edition)*. New York: McGraw-Hill, Inc.
- Peterson, E.R., & Irving, S.E. 2008. Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, 18: 238-250.
- Puskur 2008. *Model Penilaian Kelas Kurikulum Berbasis Kompetensi*. Jakarta: Balitbang Depdiknas.
- Raymond, J.E., Homer, C.S.E., Smith, R. & Gray, J.E. 2012. Learning through authentic assessment: An evaluation of a new development in the undergraduate midwifery curriculum. *Nurse Education in Practice*, 30: 1-6.

- Republika. 2014. *Implementasi kurikulum 2013 masih dibayangi banyak masalah*. (Online) ([www.republika.co.id](http://www.republika.co.id)), diakses 18 November 2014.
- Russell, M.K. & Airasian, P.W. 2012. *Classroom assessment: concepts and applications* (7<sup>th</sup> edition). New York: McGraw-Hill.
- Shute, V.J. 2007. Focus on formative feedback. *Research Report*. Princeton, NJ: Educational Testing Service (ETS).
- Stiggins, R.J. & Chappuis, J. 2012. *An introduction to student involved assessment for learning*. Sixth edition. Boston: Pearson assessment training institute.
- Sudjana, N. 2011. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Susilowati. 2013. *Kurikulum 2013, 87 persen guru kesulitan cara penilaian*. (Online) (<http://unnes.ac.id>), diakses 18 November 2014.
- Tempo. 2013. *Problematika implementasi kurikulum 2013*, (Online) ([www.tempo.co](http://www.tempo.co)), diakses 10 Juli 2013.
- Tempo. 2014. *Kurikulum 2013, Apa Saja Kendalanya?*, (Online) ([www.tempo.co](http://www.tempo.co)), diakses 16 Agustus 2014.
- UNDP. 2014. *2014 human development report*. (Online) (<http://www.undp.org/content/undp/en/home/librarypage/hdr/2014-human-development-report.html>), diakses 5 Maret 2015.
- UNESCO. 2011. *Education For All Global Monitoring Report*. (Online) (<http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/statistics/efa-development-index/>), diakses 5 Maret 2015.
- Wiggins, G. P. 1993. *Assessing student performance: exploring the purpose and limits of testing*. San Francisco: Jossey-Bass Publishers.