

KELAYAKAN ANBUSO SEBAGAI *SOFTWARE* ANALISIS BUTIR SOAL BAGI GURU

Ali Muhson, Barkah Lestari, Supriyanto, dan Kiromim Baroroh

Fakultas Ekonomi Universitas Negeri Yogyakarta

email: alimuhson@uny.ac.id

Abstrak

Penelitian ini bertujuan untuk menguji kelayakan dan mengidentifikasi kendala yang dihadapi dalam menggunakan *software* AnBuso. Penelitian ini menggunakan metode R&D. Data dikumpulkan melalui dokumentasi, kuesioner, dan wawancara. Data dianalisis secara deskriptif. Penelitian ini menghasilkan produk berupa *software* AnBuso yang dapat dimanfaatkan guru dalam melakukan analisis butir soal secara praktis dan aplikatif. Aspek yang dinilai sangat layak adalah kepraktisan dan kemudahan, substansi isi, dan kebermanfaatannya. Aspek yang dinilai layak adalah tampilan. Kendala yang dihadapi dalam menggunakan *software* adalah penguasaan guru terhadap program *Microsoft Excel* masih kurang. Guru juga kurang terbiasa melakukan analisis butir soal karena pemahaman terhadap konsep analisis butir soal yang masih terbatas.

Kata kunci: analisis butir soal, AnBuso, kelayakan *software*

FEASIBILITY OF ANBUSO AS ITEM-ANALYSIS SOFTWARE FOR TEACHER

Abstract

This study was aimed at testing the feasibility and identifying the obstacles encountered in using AnBuso software. This study used Research and Development method. The data were collected through documentation, questionnaires, and interviews. The data were analyzed descriptively. The final product of this study is AnBuso software that could help teachers analyzing the test items in a practical way. The software is considered very worthy in the aspect of practicality, simplicity, usefulness, and substance while the display is considered worthy. This study also finds some constraints such as low teacher's performance in using Microsoft Excel program, and limited understanding of item analysis concept.

Keywords: AnBuso, item analysis, software feasibility

PENDAHULUAN

Kunci utama peningkatan kualitas pendidikan terletak pada proses pembelajaran. Asesmen sebagai bagian yang tidak terpisahkan dari proses pembelajaran tentu tidak dapat kita abaikan begitu saja. Asesmen merupakan komponen yang penting dalam pembelajaran (Russel & Airasian, 2012: 2). Asesmen memiliki pengaruh yang kuat dalam meningkatkan proses pembelajaran (Raymond, *et al.*, 2012; Bers, 2008: 32). Oleh karena itu, guru

harus memiliki kemampuan yang memadai dalam melakukan asesmen.

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 menghasilkan informasi kemajuan belajar

siswa untuk digunakan sebagai dasar pengambilan keputusan.

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 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). 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.

Dalam melakukan asesmen perlu memperhatikan beberapa prinsip umum. Prinsip umum tersebut di antaranya adalah (1) menentukan secara jelas bahwa hal yang akan dinilai memiliki prioritas dalam proses asesmen, (2) prosedur asesmen yang dipilih harus relevan dengan karakteristik

atau kinerja yang akan diukur, (3) asesmen yang komprehensif memerlukan berbagai prosedur, (4) penggunaan prosedur asesmen yang tepat memerlukan kesadaran akan keterbatasan prosedur tersebut, dan (5) asesmen adalah alat untuk mencapai tujuan, bukan tujuan itu sendiri (Miller, Linn, & Gronlund, 2009: 29-31). Dengan demikian, asesmen harus dilakukan secara jelas, relevan, komprehensif, serta menggunakan prosedur dan alat yang tepat.

Pendapat lain juga mengemukakan ada sepuluh prinsip dalam asesmen. Kesepuluh prinsip tersebut adalah (1) asesmen harus meningkatkan proses pembelajaran; (2) metode asesmen harus memfasilitasi dan melaporkan kemajuan semua tujuan pembelajaran; (3) prosedur asesmen harus mencakup proses jelas untuk memastikan diperoleh informasi yang valid dan andal; (4) asesmen harus menawarkan pemahaman masyarakat akan tujuan pembelajaran yang relevan dengan kehidupan sekarang dan masa depan siswa; (5) asesmen hasil belajar harus diperlakukan sebagai perkiraan, tunduk pada kesalahan yang tidak mungkin dihindari; (6) asesmen harus menjadi bagian dari proses pembelajaran yang memungkinkan siswa untuk memahami tujuan pembelajaran dan kualitas prestasi belajar; (7) metode asesmen harus melibatkan siswa secara aktif dalam pembelajaran dan asesmen; (8) asesmen harus mengaktifkan dan memotivasi siswa untuk menampilkan apa yang bisa mereka lakukan; (9) asesmen harus menggabungkan informasi dari berbagai sumber, termasuk asesmen diri, untuk menginformasikan keputusan tentang proses belajar belajar dan prestasi siswa; dan (10) metode asesmen harus memenuhi standar yang mencerminkan konsensus yang luas pada kualitas di semua tingkatan mulai dari kelas sampai pada kebijakan nasional (Harlen, 2010: 30-31). Pendapat tersebut

menunjukkan bahwa dalam melakukan asesmen harus mampu mengungkap proses, dilakukan dengan jelas, dilaporkan hasilnya, melibatkan peserta didik secara aktif, membangkitkan motivasi peserta didik, menggunakan berbagai sumber, dan memenuhi standar yang telah digariskan.

Asesmen akan menghasilkan informasi yang tepat jika instrumen yang digunakan untuk melakukan pengukuran memenuhi beberapa kriteria seperti validitas, reliabilitas, dan objektivitas (Anderson, 2003: 10; Kubiszyn & Borich, 2013: 326). Analisis butir menjadi bagian yang penting dalam menjamin validitas butir soal (Nunnally & Bernstein, 1994: 304). Dengan demikian guru tidak hanya dituntut untuk mampu menguasai konsep analisis butir soal tetapi juga melakukan analisis butir soal dalam praktik asesmen.

Pada kenyataannya dalam praktik asesmen, guru belum sepenuhnya melakukan analisis butir soal. Hasil studi Muhson, dkk, (2013) menunjukkan bahwa kemauan dan kemampuan guru untuk melakukan analisis butir soal masih rendah. Hal ini dapat terjadi karena kurangnya pengetahuan dan kemampuan guru dalam melakukan analisis butir soal baik secara manual maupun menggunakan *software* analisis butir. Hal inilah yang mengakibatkan kualitas soal yang dikembangkan guru untuk mengukur kemampuan peserta didik menjadi kurang memenuhi standar.

Agar mampu menghasilkan soal yang baik perlu dilakukan kegiatan analisis butir soal. Analisis butir soal merupakan upaya menguji kualitas soal untuk menentukan butir mana yang perlu dipertahankan, dibuang atau direvisi. Analisis ini memberikan informasi tentang kualitas butir soal dilihat dari tingkat kesukaran, daya beda dan efektivitas pengecoh.

Tingkat kesukaran suatu butir soal merupakan salah satu parameter butir

soal yang sangat berguna dalam analisis tes karena memberikan informasi tentang sukar mudahnya butir soal. Soal dikatakan sukar jika tingkat kesukaran mendekati 0 dan soal dianggap mudah jika semakin tingkat kesukaran mendekati 1. Soal yang terlalu sukar atau terlalu mudah tidak dapat membedakan kemampuan seorang siswa dengan siswa lainnya sehingga butir soal tersebut perlu dibuang atau direvisi. Soal yang baik adalah soal yang tingkat kesukarannya sedang yakni memiliki indeks kesukaran antara 0,30-0,70 (Allen & Yen, 1979: 121; Kaplan & Saccuzzo, 2005: 170).

Daya beda butir soal merupakan kemampuan butir untuk membedakan peserta tes yang berkemampuan tinggi dan rendah. Indeks daya beda yang rendah menunjukkan adanya kemencengan distribusi skor dari populasi sehingga mengakibatkan validitas tes menjadi rendah. Soal dikatakan baik jika memiliki indeks daya beda lebih besar atau sama dengan 0,3 (Nunnally & Bernstein, 1994: 304; Kaplan & Saccuzzo, 2005: 176). 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). Berdasarkan pendapat tersebut dapat dikatakan bahwa butir soal dianggap cukup baik jika memiliki indeks daya beda melebihi atau sama dengan 0,2.

Berbagai pendapat di atas menunjukkan bahwa asesmen pembelajaran merupakan kegiatan yang penting dilakukan. Asesmen harus mampu mengukur secara tepat kompetensi peserta didik sehingga instrumen yang digunakan harus 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. Hasil analisis juga memberikan informasi tentang materi mana yang belum dikuasai oleh masing-masing peserta didik sehingga dapat dilakukan upaya perbaikan melalui kegiatan remedial. Oleh karena itu perlu dikembangkan *software* penilaian berbasis komputer yang praktis dan aplikatif yang mampu mengakomodasikan berbagai keperluan di atas.

Program komputer mampu memberikan penilaian terhadap kemampuan siswa, baik dalam pengujian maupun pelaksanaan proses pembelajaran secara otomatis (Mardapi, Haryanto, & Hadi, 2012). Hal ini menunjukkan bahwa penilaian berbasis komputer akan mampu menghasilkan informasi yang akurat dan cepat tentang kemampuan siswa serta dapat segera dilakukan perbaikan terhadap proses pembelajaran. Penilaian berbasis komputer tampaknya merupakan alternatif yang tepat karena dapat menghasilkan informasi yang cepat dan akurat. Guna memenuhi hal tersebut perlu dikembangkan sebuah *software* analisis butir 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.

Penelitian Muhson, dkk. (2013) telah berhasil mengembangkan *software* yang diberi nama AnBuso (Analisis Butir Soal). Penelitian tersebut menghasilkan juga *flowchart diagram* dan buku panduan penggunaan AnBuso. *Flowchart* tersebut menggambarkan proses pengembangan *software* yang dimulai dari input data, proses analisis, dan hasil analisis. Sementara itu buku panduan yang dihasilkan sangat berguna bagi pengguna sebagai acuan dalam menggunakan *software* AnBuso. Semua petunjuk penggunaan *software* dan hasil analisisnya dituangkan secara lengkap dalam panduan tersebut.

Fitur yang ditawarkan AnBuso juga sudah cukup lengkap. Dalam *software* AnBuso tersebut dapat diketahui baik tidaknya soal yang dibuat guru, baik dari sisi daya beda, tingkat kesulitan, maupun efektivitas distraktornya. Disamping 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.

Sebagai produk awal, *software* ini perlu lebih dikembangkan dan disempurnakan lagi agar kelemahan dan kekurangan yang ada di *software* tersebut dapat diatasi. Untuk itu studi kelayakan perlu dilakukan untuk menilai tampilan, keefektifan, kemudahan, dan kebermanfaatan *software* tersebut.

Untuk menilai kualitas sebuah *software* dapat dilihat dari tiga komponen, yaitu spesifikasi, desain dan model, dan implementasi (Kim, Jerath, & Sheldon,

2003). Sementara itu, empat dimensi kualitas yang perlu diperhatikan adalah spesifikasi, desain, pengembangan, dan kesesuaian (Chemuturi, 2011: 25). Kedua pendapat tersebut menyebutkan bahwa sebuah software harus memiliki spesifikasi tertentu baik dilihat dari fungsi, kapasitas, tujuan penggunaan, keandalan, keselamatan dan keamanan. Artinya sebuah *software* haruslah memiliki fungsi dan manfaat yang jelas.

Desain dan model yang dikembangkan dalam sebuah *software* juga perlu dilihat karena akan menentukan kemenarikannya. Sementara itu dari sisi implementasi juga perlu dinilai kepraktisan dan kemudahan dalam menggunakan *software*. Kesesuaian *software* juga merupakan bagian yang penting karena *software* yang baik mestinya dikembangkan sesuai dengan kebutuhan dan keinginan penggunanya. Dengan demikian pengujian kelayakan software dalam penelitian ini didasarkan pada tampilan (desain), substansi isi (spesifikasi), kepraktisan/kemudahan, dan kebermanfaatannya.

Mengingat pentingnya penyempurnaan *software* maka 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 dipublikasikan kepada khalayak sasaran.

METODE

Penelitian ini menggunakan model *Research and Development (R&D)*. Prosedur pengembangan dilakukan dengan langkah perancangan dan pengembangan produk, validasi produk, dan uji coba produk (Gambar 1). Pada tahap perancangan dan

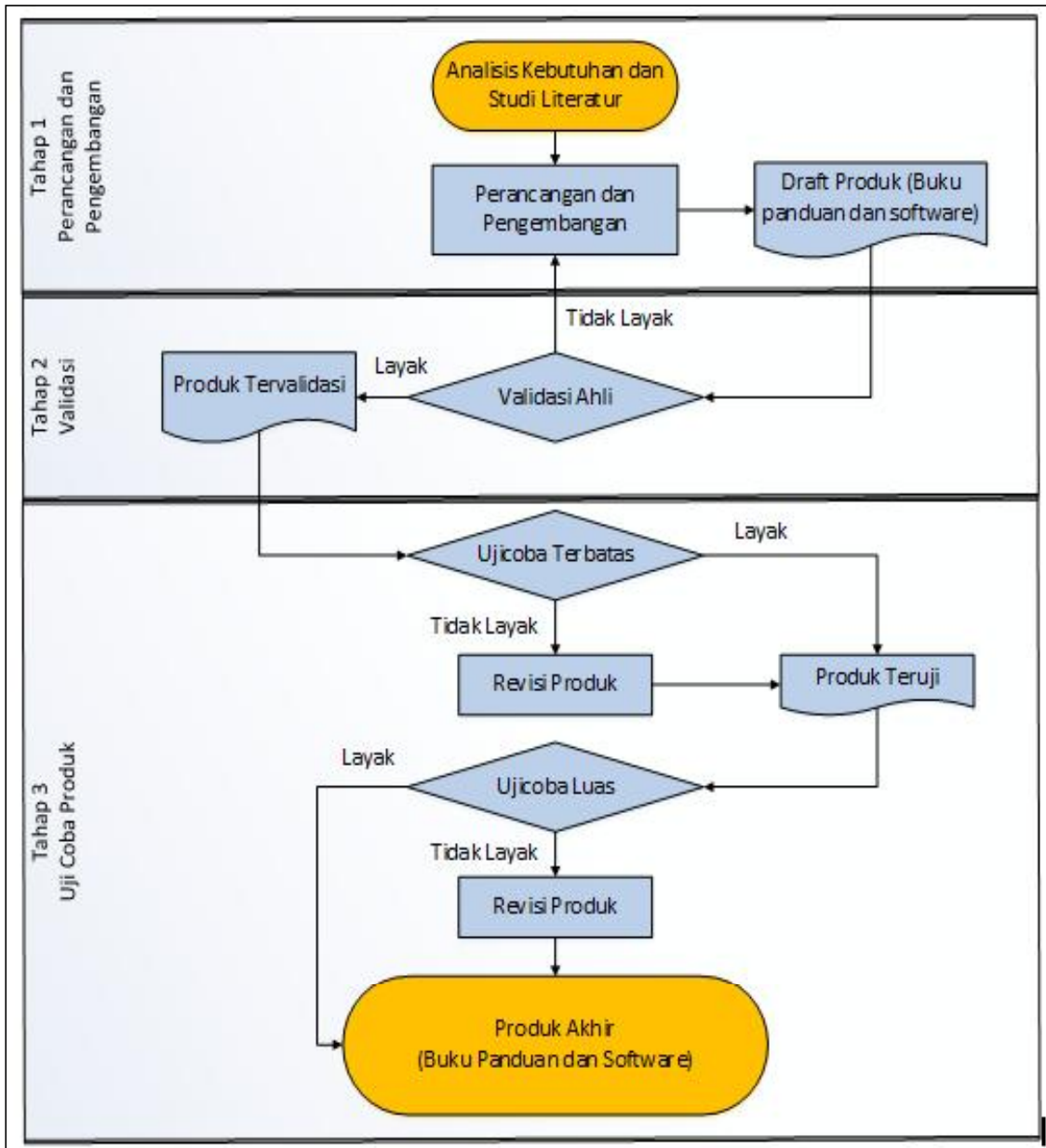
pengembangan produk dilakukan studi pendahuluan untuk melihat kemampuan dan kemauan guru dalam melakukan analisis butir soal sekaligus mengidentifikasi kebutuhan. Langkah selanjutnya adalah merancang draft model *software* dan buku panduannya kemudian dilakukan validasi ahli dan dilakukan revisi sehingga diperoleh produk yang sudah tervalidasi.

Validasi dilakukan dengan metode Delphi dan melibatkan beberapa pakar yakni pakar pendidikan, pengukuran dan pemrograman. Hasil validasi menunjukkan bahwa semua instrumen, draft model *software* dan buku panduan yang dihasilkan secara substansi dinyatakan layak walaupun ada beberapa saran untuk dilakukan perbaikan.

Terkait dengan instrumen dilakukan perubahan dengan lebih menegaskan lagi kalimat yang digunakan agar tidak memiliki salah arti. Sementara itu perubahan dalam buku panduan terkait dengan sistematika dan tata penyajiannya sehingga lebih mudah untuk dipahami pembaca. Adapun perbaikan yang dilakukan terhadap *software* bukan terletak pada hasil analisisnya melainkan pada penyajian laporan hasil analisis agar dibuat lebih menarik.

Langkah berikutnya adalah melakukan ujicoba produk yang dilakukan sebanyak dua kali, yakni ujicoba terbatas dan ujicoba luas untuk mendapatkan informasi dan masukan yang lebih komprehensif dari pengguna. Hasil ujicoba tersebut kemudian dilakukan perbaikan untuk mendapatkan produk akhir berupa *software* AnBuso dan buku panduan yang siap untuk disebarluaskan.

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



Gambar 1. Prosedur Pengembangan

komputer, khususnya program aplikasi *Microsoft Excel*.

Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi observasi, angket, dan wawancara. Observasi digunakan untuk memperoleh data atau informasi tentang kemampuan guru dalam menggunakan program aplikasi yang telah dikembangkan. 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/kemudahan, dan kebermanfaatan. 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 isi, kepraktisan/kemudahan, dan kebermanfaatan. Dalam melakukan analisis digunakan lima kategori seperti terlihat pada Tabel 1.

HASIL PENELITIAN DAN PEMBAHASAN

Untuk memperoleh gambaran tentang hasil pengembangan *software* dilakukan ujicoba luas yang melibatkan para pengguna 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 me-

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 |

manfaatkan *software* ini. Buku panduan ini sekaligus memberikan informasi tentang langkah-langkah dan cara menggunakan *software* ini sehingga memudahkan *user* untuk memanfaatkan *software* dalam melakukan analisis butir soal. Panduan ini berisi tentang pendahuluan, kerangka isi, data *input*, dan data laporan.

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. *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 esai. Hasil analisis juga sudah dibuat dalam format laporan sehingga mempermudah guru dalam menafsirkan hasilnya.

Dilihat dari aspek tampilan (Tabel 2) menunjukkan bahwa *software* yang dikembangkan dinilai layak oleh responden. Sebagian besar responden 57% menyatakan layak dan 35% menilai sangat layak dan selebihnya 8% menilai cukup layak. Tidak ada satu pun responden yang menilai tidak layak. Walaupun demikian aspek ini dinilai paling rendah 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*.

Tabel 2. Kelayakan Aspek Tampilan

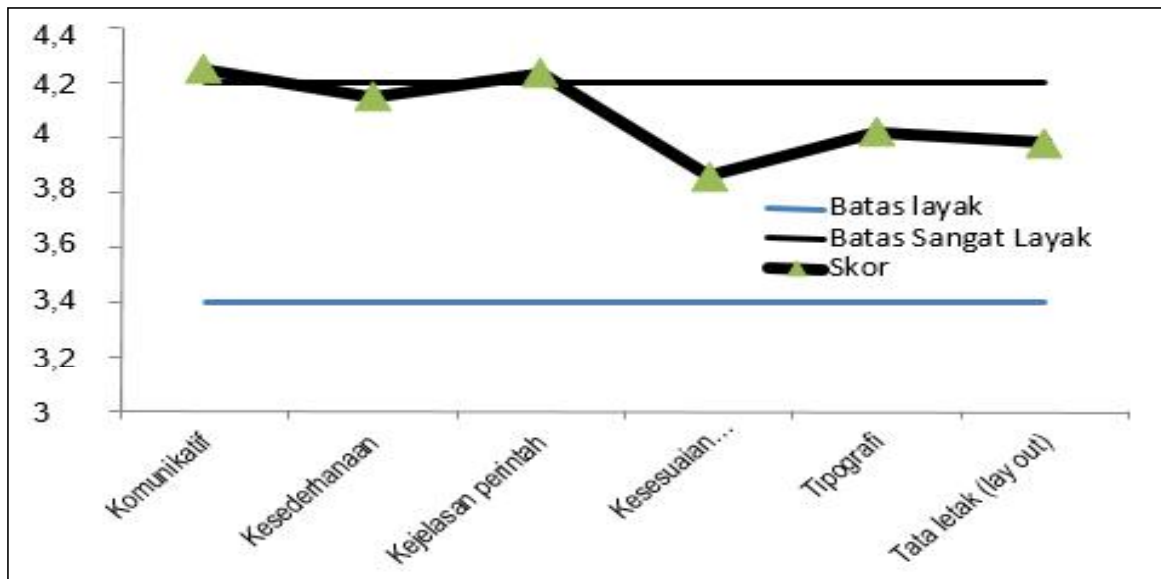
| No | Kategori Penilaian | Frekuensi | Persentase |
|-------|--------------------|-----------|------------|
| 1 | Cukup layak | 5 | 7,7 |
| 2 | Layak | 37 | 56,9 |
| 3 | Sangat layak | 23 | 35,4 |
| Total | | 65 | 100,0 |

Jika dilihat dari per komponen dari aspek tampilan yang dinilai paling baik adalah komunikatif, kejelasan perintah, dan kesederhanaan (Gambar 2). Sementara itu 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.

Untuk aspek substansi isi, *software* ini dinilai sangat layak (Tabel 3). Sebagian besar responden (52%) menganggap bahwa substansi isi dari *software* ini sangat layak, sedangkan yang menilai layak 43%, cukup layak 5% dan tidak ada satu pun yang

Tabel 3. Kelayakan Aspek Substansi

| No | Kategori Penilaian | Frekuensi | Persentase |
|-------|--------------------|-----------|------------|
| 1 | Cukup layak | 3 | 4,6 |
| 2 | Layak | 28 | 43,1 |
| 3 | Sangat layak | 34 | 52,3 |
| Total | | 65 | 100,0 |



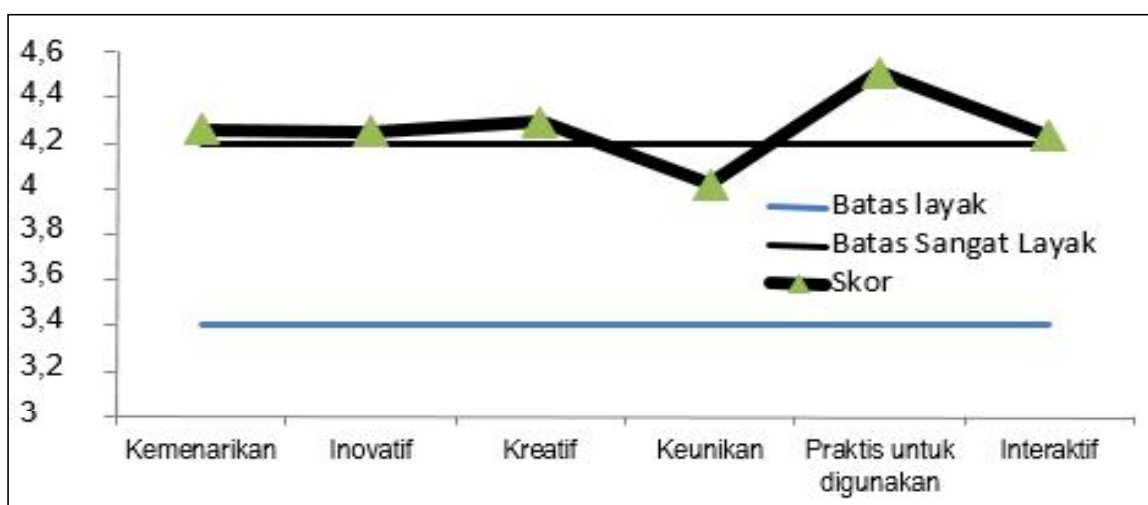
Gambar 2. Hasil Penilaian Tampilan *Software*

menilai tidak layak. Hal ini menunjukkan bahwa dilihat dari aspek substansi isi dari *software* dianggap sangat baik karena sesuai dengan kebutuhan guru.

Informasi yang dihasilkan dari *software* ini sangat lengkap, tidak hanya menampilkan hasil analisis butir soal objektif dan esai 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.

Jika dilihat per komponen dalam aspek substansi isi *software* seperti terlihat pada Gambar 3 tampak bahwa hampir semua komponen dinilai sangat layak. Komponen yang dinilai sangat layak adalah



Gambar 3. Hasil Penilaian Substansi Isi *Software*

kepraktisan untuk digunakan, kemenarikan, inovatif, kreatif dan interaktif. Sementara itu komponen keunikan nilai layak.

Aspek kepraktisan dan kemudahan *software* disajikan pada Table 4. *Software* ini dinilai sangat layak dari segi kepraktisan dan kemudahan. Sebagian besar responden (55%) menganggap bahwa substansi isi dari *software* ini sangat layak, sedangkan yang menilai layak 40%, cukup layak 5% dan tidak ada satu pun yang menilai tidak layak. Hal ini menunjukkan bahwa dilihat dari aspek kepraktisan dan kemudahan dari *software* dianggap sangat baik.

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.

Jika dilihat per komponen dalam aspek kepraktisan dan kemudahan *software* seperti terlihat pada Gambar 4 tampak bahwa semua komponen dinilai sangat layak. Dengan demikian *software* ini dianggap

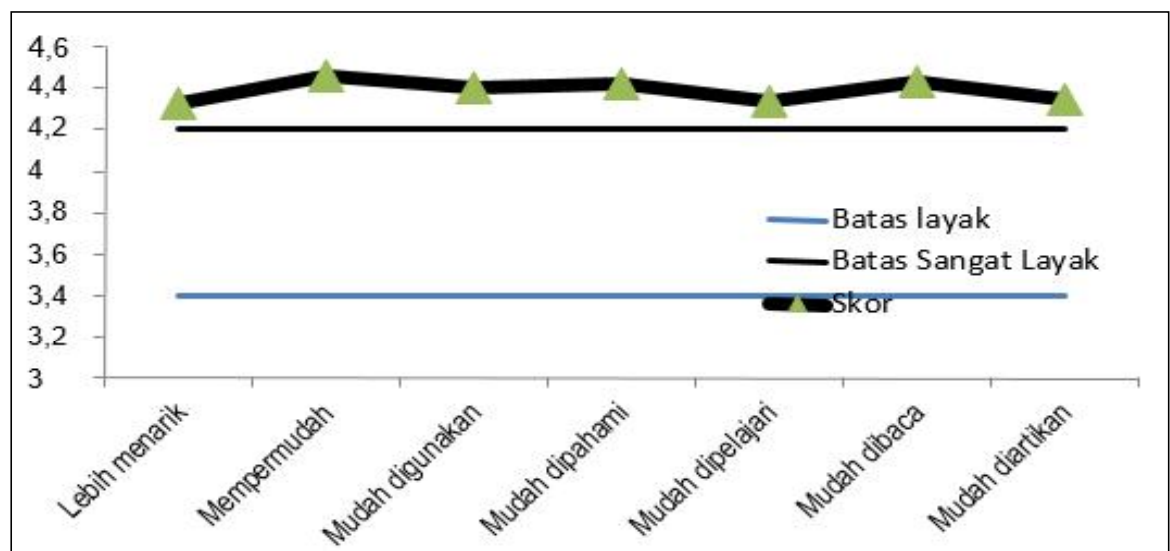
praktis dan mudah untuk digunakan guru dalam melakukan analisis butir soal.

Tabel 4. Kelayakan Aspek Kepraktisan dan Kemudahan

| No | Kategori Penilaian | Frekuensi | Persentase |
|-------|--------------------|-----------|------------|
| 1 | Cukup layak | 3 | 4,6 |
| 2 | Layak | 26 | 40,0 |
| 3 | Sangat layak | 36 | 55,4 |
| Total | | 65 | 100,0 |

Software yang baik seharusnya memiliki nilai manfaat yang tinggi bagi penggunaannya. Dilihat dari aspek kebermanfaatan menunjukkan bahwa *software* yang dikembangkan dinilai sangat layak oleh responden (Tabel 5). Sebagian besar responden 54% menilai sangat layak dan 42% menilai layak dan selebihnya 5% menilai cukup layak. Tidak ada satu pun responden yang menilai tidak layak. Hal ini menunjukkan bahwa *software* ini memiliki nilai manfaat yang tinggi bagi guru.

Hasil analisis yang ditampilkan dari *software* ini sangat sesuai



Gambar 4. Hasil Penilaian Kepraktisan dan Kemudahan *Software*

Tabel 5. Kelayakan Aspek Kebermanfaatan

| No | Kategori Penilaian | Frekuensi | Persentase |
|-------|--------------------|-----------|------------|
| 1 | Cukup layak | 3 | 4,6 |
| 2 | Layak | 27 | 41,5 |
| 3 | Sangat layak | 35 | 53,8 |
| Total | | 65 | 100,0 |

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 guru untuk keperluan membuat laporan administrasi.

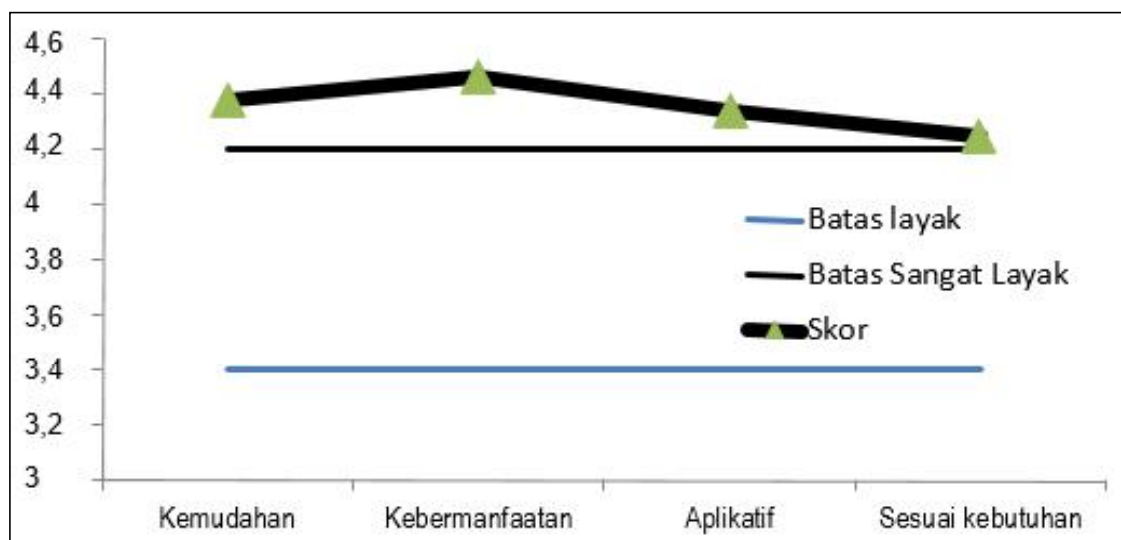
Jika dilihat dari per komponen dalam aspek kebermanfaatan *software* memperlihatkan bahwa semua komponen dinilai sangat layak oleh responden (Gambar 5). Komponen yang dinilai paling unggul adalah komponen kegunaan sementara komponen yang dinilai rendah adalah kesesuaian dengan kebutuhan. Hal ini menunjukkan bahwa keberadaan

software ini memiliki kebermanfaatan yang tinggi dalam membantu guru untuk melakukan analisis butir soal.

Walaupun *software* AnBuso ini dinilai layak, namun masih ada beberapa kendala yang dihadapi guru dalam menggunakan *software* ini. *Pertama*, masih ada beberapa guru yang kurang mahir dalam penguasaan komputer, khususnya program *Microsoft Excel*. Umumnya pengetahuan guru terhadap program ini sangatlah terbatas sehingga masih banyak menu dan fasilitas belum dimanfaatkan secara optimal. Oleh karena *software* ini terintegrasi dengan program *Excel* maka guru kurang lancar dalam memanfaatkan *software* ini.

Kedua, 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 semesteran maupun soal ujian akhir. Oleh karena kebiasaan tersebut, maka guru menjadi kurang lancar dalam melakukan analisis butir soal.

Ketiga, pengetahuan dan penguasaan guru tentang konsep analisis butir soal juga



Gambar 5. Hasil Penilaian Kebermanfaatan *Software*

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.

Selain kendala tersebut, masih ada beberapa tuntutan guru yang belum mampu diakomodasikan dalam *software* ini. *Pertama*, *software* ini belum merupakan program aplikasi yang berdiri sendiri yang cara menjalankannya harus melalui instalasi, namun *software* ini masih menjadi bagian integral dalam program *Microsoft Excel*. Akibatnya *software* ini tidak akan dapat dijalankan jika tidak menggunakan program *Excel*.

Kedua, keterbatasan dalam hal jumlah peserta tes/siswa. Jumlah maksimal yang bisa dianalisis hanya sebanyak 50 siswa sehingga tidak dapat digunakan untuk analisis yang peserta tesnya melebihi 50. *Ketiga*, keterbatasan jumlah soal. Jumlah soal pilihan ganda hanya terbatas maksimal 50 soal dan soal esai maksimal 10 soal. Hal ini tentu membatasi pengguna jika soal yang dimiliki melebihi batas yang ditentukan. *Keempat*, tidak tersedia soal dengan kode yang berbeda (soal paralel) sehingga sulit diaplikasikan untuk keperluan soal yang memiliki lebih dari satu jenis soal paralel dengan seri yang berbeda.

SIMPULAN

Hasil penelitian menunjukkan bahwa *software* AnBuso yang dihasilkan terbukti sangat layak dilihat dari aspek kepraktisan dan kemudahan, aspek kebermanfaatan, aspek substansi isi dan aspek tampilan. *Software* ini juga sangat bermanfaat bagi guru karena dapat dengan mudah digunakan

dan dipraktikkan. *Output* yang dihasilkan dari *software* ini juga sesuai dengan kebutuhan dan keinginan guru.

Beberapa kendala yang ditemukan dalam menggunakan *software* ini di antaranya adalah penguasaan guru terhadap program *Microsoft Excel* masih kurang. Guru juga kurang terbiasa melakukan analisis butir soal karena pemahaman terhadap konsep analisis butir soal yang masih terbatas.

DAFTAR PUSTAKA

- Alagumalai, S., & Curtis, D.D. 2005. Classical Test Theory. In Alagumalai, S., Curtis, D.D., and Hungi, N. (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.
- Bers, T.H. 2008. "The Role of Institutional Assessment in Assessing Student Learning Outcomes". *New Directions for Higher Education*, 141, 31-39.
- Chemuturi, M. 2011. *Mastering Software Quality Assurance: Best Practices, Tools and Techniques for Software Developers*. Florida: J. Ross Publishing.
- 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.

- Harlen, W. 2010. "What is Quality Teacher Assessment?". In Gardner, J., Harlen, W., Hayward, L., & Stobart, G. (with Montgomery, M.) (Eds.) *Developing Teacher Assessment*, pp. 30-31. New York: Open University Press.
- Kaplan, R.M., & Saccuzzo, D.P. 2005 *Psychological Testing: Principles, Applications, and Issues*. (6th ed.). Belmont, CA: Thomson Wadsworth.
- Kim, H.Y, Jerath, K., & Sheldon, F.T. 2003. "Assessment of High Integrity Components for Completeness, Consistency, Fault-Tolerance and Reliability". In Cechich, A., Piattini, M., & Vallecillo, A. (Eds.). *Component-Based Software Quality: Methods and Techniques*, pp. 259-286. New York: Springer.
- Kubiszyn, T., & Borich, G.D. 2013. *Educational Testing and Measurement: Classroom Application and Practice*. (10th ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Mardapi, D., Haryanto, & Hadi, S. 2012. "Pengujian Hasil Belajar dan Penilaian Pendidikan Berbantuan Komputer". *Jurnal Kependidikan*, 42(2), 130-143.
- Miller, M.D., Linn, R.L., & Gronlund, N.E. 2009. *Measurement and Assessment in Teaching*. (10th ed.). 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. LPPM UNY.
- Nunnally, J.C., & Bernstein, I.H. 1994. *Psychometric Theory*. (3rd edition). New York: McGraw-Hill, Inc.
- 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.
- Russell, M.K., & Airasian, P.W. 2012 *Classroom Assessment: Concepts and Applications*. (7th ed.). New York: McGraw-Hill.
- Stiggins, R.J., & Chappuis, J. 2012. *An Introduction to Student Involved Assessment for Learning*. (6th ed.). Boston: Pearson Assessment Training Institute.