

**PENENTUAN BATAS LULUS
PADA KEMAMPUAN MINIMAL UJIAN KOMPETENSI KEJURUAN
DENGAN METODE STANDARD SETTING**

¹Soffan Nurhaji, ²Haryanto
PEP UNY, Universitas Negeri Yogyakarta
¹soffansatu@gmail.com, ²haryanto.ftuny@gmail.com

Abstrak

Penelitian ini bertujuan untuk mengetahui batas skor kelulusan (*cutscore*) yang tepat bagi peserta Uji Kompetensi Kejuruan (UKK) SMK Kabupaten Brebes. Penelitian dilaksanakan pada bulan Maret sampai dengan April 2015. Metode yang digunakan adalah *Angoff*, *Direct Consensus Method* and *Contrasting Group Methods*. akurasi dari hasil perbandingan metode berdasarkan nilai standar deviasi dan reliabilitas dari dua putaran masing-masing metode. Sumber data penelitian ini adalah peserta tes 879 dan panelis (*Judgement*). *Judgement* digunakan untuk menentukan *cutscore*. Keakuratan metode dalam menentukan *cutscore* ditunjukkan dengan kecilnya nilai standar deviasi dan reliabilitas yang diperoleh dari masing-masing metode. Hasil penelitian menunjukkan hasil analisis karakteristik butir soal berdasarkan teori tes klasik dan teori respon butir pada perangkat soal UKK baik. *Cutscore* yang diperoleh dari metode *Angoff* sebesar 53,95, *Direct Consensus Method* sebesar 60,19 dan *Contrasting Group Method* sebesar 21 52,50. *Direct Consensus Method* merupakan metode yang paling akurat, Hal ini terlihat dari nilai standar deviasi dan reliabilitas relatif paling kecil, yaitu 0,991.

Kata kunci: *batas lulus (cutscore), standard setting, metode Angoff, Direct Consensus Method, Contrasting Group Method*

**DETERMINING THE CUTSCORE OF VOCATIONAL COMPETENCE TEST SUBJECT USING
STANDARD SETTING METHOD**

¹Soffan Nurhaji, ²Haryanto
PEP UNY, Universitas Negeri Yogyakarta
¹soffansatu@gmail.com, ²haryanto.ftuny@gmail.com

Abstract

The purpose of this research is to find out the cutscore in Teory-Productive Subjects at SMK of Brebes regency. The research was conducted in March to April 2015. Based on Standard Setting by Angoff Direct Consensus and Contrasting Group Methods. the accuracy of the results of the comparison method is based on the standard deviation value and reliability of two rounds each method. The data source of this research is 879 test participants and panelists (Judgement). Judgement is used to determine cutscore obtained from each method. The results showed the results of the analysis of the characteristics of items based on classical test theory and item response theory in in Teory-Productive is good. Cutscore obtained from Angoff method amounted to 53.95, Direct Method Consensus Group amounted to 60.19 and Contrasting Method at 21 52.50. Consensus Direct Method is the most accurate method, It is seen from the standard deviation and the relative reliability obtained that is the smallest (0.991), of the three methods.

Keywords: *cutscore, standard setting, Angoff method, direct consensus method, contrasting group method*

Pendahuluan

Penilaian akan menyediakan informasi secara komprehensif tentang keberhasilan dalam proses pembelajaran. Tinggi rendahnya hasil belajar siswa yang dihasilkan melalui penilaian, selanjutnya akan dievaluasi sebagai sebuah *judgment* terhadap pembelajaran yang dilaksanakan dalam kurun waktu tertentu. Proses penilaian terhadap keberhasilan pembelajaran tersebut tentunya tidak terlepas dari proses pengukuran yang dilaksanakan dengan instrumen tertentu. Hasil pengukuran yang dilakukan dengan instrumen tertentu itulah yang selanjutnya dinilai dan dievaluasi.

Kuantifikasi dari pembelajaran dibutuhkan sebuah instrumen yang baik. Instrumen yang baik adalah instrumen yang memiliki validitas dan reliabilitas tinggi serta memiliki error yang sekecil mungkin dalam menjangking informasi tentang keberhasilan proses pembelajaran. Informasi yang didapatkan melalui pengukuran merupakan hasil dari proses analisis data dengan teknik tertentu. Selanjutnya barulah guru melakukan penilaian secara intensif untuk melihat keberhasilan dari proses pembelajaran, apakah angka-angka yang didapatkan melalui pengukuran sudah memenuhi standar keberhasilan atau sebaliknya.

Salah satu upaya yang bisa dilakukan agar butir soal memiliki validitas dan reliabilitas tinggi adalah dengan cara menganalisis butir soal pengukuran secara kontiniu. Hasil analisis tersebut akan memberikan informasi kepada guru tentang kelayakan atau validitas butir soal yang digunakan dalam proses pengukuran. Selain validitas, hasil analisis juga memberikan informasi kepada guru tentang reliabilitas soal yang dirancang guru. Informasi tentang reliabilitas soal juga penting diketahui oleh guru. Hal tersebut dikarenakan butir soal dikatakan baik apabila telah memenuhi kedua syarat tersebut. Meskipun pada dasarnya butir soal yang memiliki validitas tinggi pada umumnya reliabel, akan tetapi pengujian reliabilitas perlu dilakukan untuk melihat kejelasan atau kehandalan soal tersebut dalam menjangking informasi keberhasilan siswa dalam proses pembelajaran.

Kaitannya dengan pelaksanaan ujian kompetensi keahlian bagi siswa SMK seringkali ditemukan permasalahan, baik tentang standar isi, perangkat uji yang digunakan maupun pola penyelenggaraan. Naskah soal teori dan kompetensi keahlian siswa yang diujikan terkadang tidak sesuai. Masalah lain yang ditemukan

adalah kualitas soal yang rendah, seperti grafik, gambar, dan lain-lain tidak jelas (Laporan monev uji kompetensi, 2013).

Kelulusan peserta didik dari program atau satuan pendidikan salah satunya ditentukan oleh hasil ujian. Peserta ujian dinyatakan lulus ujian apabila nilai rata-rata yang didapatkan memenuhi skor batas kelulusan (*Cutscore*) yang sudah ditetapkan oleh pemerintah sebelumnya. Penentuan kelulusan yang selama ini digunakan adalah berdasarkan *judgment* saja. Cara ini mengandung banyak kelemahan, diantaranya tidak memperhatikan kondisi siswa. Daerah yang sangat luas dan kondisi sekolah yang bervariasi menuntut agar batas kelulusan harus didasarkan pada kemampuan peserta didik. Oleh karena itu penetapan standar minimum yang harus dicapai peserta agar lulus harus menggunakan cara yang lebih baik. Oleh karena itu perlu dilakukan evaluasi terhadap skor batas yang telah ditentukan oleh pemerintah apakah benar-benar sudah mengukur kompetensi siswa.

Batas minimal merupakan batas kemampuan minimal yang harus dimiliki oleh peserta didik untuk dapat dinyatakan lulus dari suatu jenjang pendidikan tertentu. Batas minimal atau disebut juga batas lulus ini digunakan untuk melihat ketercapaian kompetensi peserta didik. Lulus mengartikan bahwa peserta didik tersebut memiliki kompetensi yang telah dipersyaratkan, sedangkan tidak lulus berarti peserta didik belum memiliki kompetensi tersebut. Penentuan batas minimal kelulusan peserta didik ini sangatlah penting, karena skor batas kelulusan minimal ini akan menggambarkan kompetensi dari peserta didik yang dilihat dari lulus dan tidak lulus. Oleh karena itu perlu dilakukan peninjauan kembali terhadap skor batas yang telah ditentukan sebelumnya. Peninjauan ini dilakukan untuk melihat apakah skor batas yang telah ditentukan tersebut benar-benar menggambarkan kompetensi siswa. Hal ini perlu dilakukan secara berkelanjutan dikarenakan tingkat kelulusan dimasing-masing daerah setiap tahunnya berbeda-beda.

Permasalahan dari sisi tujuan uji kompetensi, hasil uji kompetensi memiliki dua sisi kepentingan yaitu sebagai persyaratan kelulusan dari satuan pendidikan dan untuk kepemilikan ijazah. Penilaian dengan uji kompetensi ini tidak mengukur kinerja sesuai standar kompetensi yang dipersyaratkan industri, karena masih lebih mementingkan aspek kognitif, sedangkan aspek afektif dan psikomotorik masih ditelantarkan. Uji kompetensi seharusnya

mengukur pencapaian kompetensi siswa atau peserta didik dan dibandingkan dengan standar atau kebutuhan kerja yang telah ditetapkan.

Sebagai bagian dari sistem pendidikan nasional, pendidikan menengah kejuruan mengutamakan pengembangan kemampuan peserta didik untuk dapat bekerja dalam bidang tertentu, kemampuan beradaptasi di lingkungan kerja, melihat peluang kerja dan mengembangkan diri di kemudian hari (Undang-Undang Nomor 20, 2003). Pendidikan kejuruan akan efisien jika: (1) lingkungan di mana siswa dilatih merupakan replika lingkungan di mana nanti ia akan bekerja, (2) pendidikan kejuruan yang efektif untuk setiap profesi, jabatan atau pekerjaan hanya dapat diberikan kepada seseorang yang memerlukannya, yang menginginkannya, dan yang dapat untung darinya, (3) pendidikan kejuruan yang efektif hanya dapat diberikan dimana tugas-tugas latihan dilakukan dengan cara, alat dan mesin yang sama seperti yang ditetapkan di tempat kerja, (4) pendidikan kejuruan akan efektif jika pengalaman latihan untuk membentuk kebiasaan kerja dan kebiasaan berfikir yang benar diulangkan sehingga sesuai seperti yang diperlukan dalam pekerjaan nantinya.

Analisis hasil Ujian Nasional untuk memperoleh informasi mengenai karakteristik butir soal perlu dilakukan karena beberapa hal tersebut saling berkaitan untuk menentukan skor batas kelulusan minimal (*Cut of score*) yang mampu membedakan siswa yang layak untuk lulus maupun tidak, diharapkan berlaku secara umum bagi semua siswa dan dapat mengukur kemampuan siswa sebenarnya pada suatu satuan pendidikan di wilayah manapun. Kegiatan penentuan skor batas kelulusan minimal tersebut dinamakan *Standard Setting*. (Cizek dan Bunch, 2007).

Analisis tes yang dapat digunakan untuk keperluan *Standard Setting* diantaranya adalah model Teori Tes Klasik. Pada model ini tidak menjamin akan dapat diperoleh keputusan terbaik bagi validitas pengambilan keputusan dari suatu analisis tes, karena domain kemampuan siswa tidak ada hubungannya dengan tingkat kesulitan butir soal. Selain itu, tingkat kesulitan butir soal tidak tetap bergantung kepada kemampuan siswa yang mengerjakan soal tes. Oleh karena itu dibutuhkan suatu model analisis tes yang lebih baik sehingga hasilnya pun akan menjadi lebih baik. Model analisis tes tersebut yaitu model Teori Responsi Butir atau *Item Response Theory* (IRT).

Model IRT sangat membantu dalam keperluan *Standard Setting*. Dengan model ini skor batas lulus dapat disajikan dalam rentang interval 0 sampai 1 pada kemampuan atau *ability*, selain itu tingkat kesulitan dan kemampuan siswa berada pada skala yang sama. Tidak seperti pada Teori Tes Klasik, IRT ini tidak bergantung pada kelompok sampel, artinya estimasi kemampuan yang diperoleh dari suatu perangkat tes yang berbeda akan selalu sama dan parameter butir yang diperoleh dari kelompok subjek yang berbeda juga akan selalu sama. Pada model ini kemudian muncul beberapa metode analisis yang dapat digunakan untuk memudahkan analisis tes, dalam hal ini memudahkan keperluan *Standard Setting*. Berkaitan dengan hal ini, maka perlu ditentukan batas kelulusan yang tidak hanya berdasarkan pada *judgement* pemerintah semata, namun juga berdasarkan data empirik di lapangan, atau berdasarkan kemampuan peserta didik. Hal yang harus dicermati bahwa respon peserta didik dalam ujian dapat digunakan untuk mengestimasi abilitasnya.

Berdasarkan hal tersebut perlu diadakan sebuah penelitian untuk melihat karakteristik butir tes UKK dalam peninjauan secara empiris untuk menentukan skor batas lulus dengan metode *standard setting*. Selain faktor tersebut, terdapat faktor lain yang melatar belakangi peneliti ingin meneliti penentuan skor batas lulus dengan metode *standard setting*, yaitu dikarenakan belum pernah dilakukan penelitian ini Kabupaten Brebes. Guru tidak mengetahui bagaimana penentuan batas lulus dengan bukti empiris dan dapat dipertanggungjawabkan dengan teori pengukuran. Sehingga dengan adanya penelitian ini diharapkan dapat memberi kontribusi besar bagi perkembangan dan kemajuan pendidikan SMK di Kabupaten Brebes.

Berbagai penelitian tentang *Standard Setting* telah banyak dilakukan oleh pemerhati di dunia pendidikan maupun bidang lain untuk menentukan *Cut Score* ujian nasional untuk mata pelajaran Matematika. Ternyata secara empirik *Cut Score* ujian nasional mata pelajaran Matematika adalah sebesar 71,6 (dengan metode Grup Kontras) dan sebesar 72,25 (dengan metode Ebel) semuanya pada skala 100. Sementara *Cut Score* yang ditentukan pemerintah adalah sebesar 4,50 (untuk tahun 2008/2009). Artinya terjadi perbedaan batas lulus (*Cut Score*) menurut *judgement* pemerintah dan berdasarkan kemampuan riil peserta didik (Natalina, 2010, p.7).

Giraud & Impare (2005) melakukan studi kualitatif yang melibatkan guru, administrator dan konsultan secara bersama-sama dalam menyusun *cutscore*. Tujuan studi ini menyusun *cutscore* yang nantinya akan direkomendasikan pada Dewan Direksi Sekolah Willtown. Studi ini difokuskan pada penentuan *cutscore* untuk membaca dan menulis pada level 5. Hasil studi ini menyarankan untuk melanjutkan pelatihan dan memberikan keyakinan terhadap guru dalam proses penyusunan *cutscore*.

Carlson, Tomkowiak & Stilp (2009) melakukan penelitian tentang penggunaan metode *Angoff* dalam menentukan *cutscore* pada standar kinerja pendidikan asisten dokter atau *Phydicion Assisten* (PA). Penelitian ini bertujuan untuk mengeksplorasi reliabilitas dan kredibilitas standar kinerja berdasarkan ujian berbasis kinerja pendidikan asisten dokter dengan menggunakan metode *Angoff*. Panel yang terlibat yaitu anggota dari fakultas PA dengan menggunakan metode *Angoff* dalam menentukan *cutscore*. Reliabilitas item tes dinilai menggunakan *Cronbach's Alpha*, sedangkan persetujuan dari panel dievaluasi dengan menggunakan statistik *Kappa*. Kredibilitas panel dinilai menggunakan *Pearson Correlation* dengan membandingkan antara prestasi siswa pada item tes dan tingkat kesukaran butir yang didefinisikan oleh metode *Angoff*. Hasil dari penelitian ini menunjukkan bahwa Metode *Angoff* terbukti menjadi metode yang dapat diandalkan dan dipercaya untuk menetapkan *cutscore* lewat ujian dengan metode *Angoff* adalah 62% (skala 100) dengan tingkat kelulusan 100%, 70% (presentase kelulusan), dan *cutscore* yang ditetapkan adalah 72% (presentase kelulusan 88%). Nilai *Cronbach's Alpha* untuk tes item adalah 0,75; koefisien *Kappa* adalah 0,71 dan *Pearson Correlation* antara kesulitan butir yang sebenarnya dan kesulitan seperti yang diperkirakan oleh metode *Angoff* menunjukkan korelasi positif yang signifikan (+0.44, $p < 0,05$).

Banyak metode yang berkembang yang bisa digunakan untuk menganalisis karakteristik butir dalam pengukuran dan pengujian baik pada Teori Tes Klasik maupun Teori Respons Butir. Perbedaan asumsi dan teknik yang terdapat pada masing-masing metode tersebut menyebabkan hasil pengukuran yang menghasilkan akurasi yang berbeda. Begitu juga dengan metode *standard setting*, banyak metode *standard setting* yang ada dan dimodifikasi dalam penggunaannya. Salah satunya menggunakan

metode *standard setting* yang berpusat pada tes dan berpusat pada peserta tes.

Istilah *standard* banyak digunakan dalam program pengujian tes, salah satunya yaitu *standard setting*. Menurut Cizek & Bunch (2007, p.13) *standard setting* merupakan proses mengidentifikasi satu atau lebih skor batas (*cutscore*) pada suatu tes.

Bejar (2008, p.1) mengungkapkan bahwa *standard setting* adalah suatu metodologi yang digunakan untuk menentukan tingkat pencapaian kompetensi yang ditunjukkan dengan *cutscores*. Berdasarkan kedua pendapat ahli tersebut jelas bahwasanya *standard setting* merupakan suatu metode yang digunakan dalam menentukan satu atau lebih skor batas (*cutscore*) untuk mengetahui pencapaian kompetensi peserta tes.

Tujuan utama penggunaan metode *standard setting* yaitu untuk membantu para pengambil kebijakan dalam menentukan skor batas kelulusan secara komprehensif. Brandon (2002, p.1) dalam artikelnya mengungkapkan bahwa tujuan dari *standard setting* yaitu menetapkan penilaian *judgment* yang disebut dengan *cutscore*. *Cutscore* merupakan suatu titik pada skala skor dimana skor yang berada atau di atas titik tersebut dalam kategori yang berbeda. Horn (2000, p.3) mengungkapkan bahwa *cutscore* merupakan titik yang terdapat pada skala skor tertentu yang memisahkan antara level kinerja (*performance*) yang satu dengan yang lainnya.

Penentuan *cutscore* tidak terlepas dari apa yang namanya *judgements*. Hal ini diperkuat oleh pendapat yang dikemukakan oleh Zieky, Perie & Livingston (2008, p.10): *every method of setting cutscore depend of judgment at some point in the process*. Pertimbangan yang perlu dilakukan ketika menentukan *cutscore* adalah panelis (ahli) yang akan ikut serta dalam penentuan *cutscore*. Panelis yang memenuhi kualifikasi akan memberikan informasi yang tepat dalam pengambilan keputusan, sehingga *cutscore* yang diperoleh akan menjadi lebih bermakna dan realistis.

Secara umum metode dalam *standard setting* itu dibedakan menjadi lima. Namun, dari ke-5 metode tersebut, metode berpusat pada tes (*test-cetered*) dan pertimbangan peserta tes (*test-takers*) yang akan digunakan dalam penelitian ini. Penggunaan kedua pendekatan metode ini selain mudah digunakan, kedua metode ini juga lebih efisien.

Pendekatan metode *standard setting* yang berpusat pada tes dibatasi pada metode Angoff dan *Direct Consensus Method*. Metode yang berpusat pada peserta tes dibatasi pada metode grup kontras (*Contrasting Groups Method*).

Metode Angoff pertama kali diperkenalkan pada tahun 1971 oleh Williem Angoff. Metode ini mengasumsikan bahwa panelis mampu menentukan peluang menjawab benar dari peserta tes pada *borderline* dengan meminta pertimbangan ahli untuk melakukan *judgement* terhadap peluang peserta tes yang mampu menjawab butir dengan benar. Cizek & Bunch (2007, p.82) mengungkapkan bahwa metode Angoff meminta panelis untuk menelaah masing-masing butir yang ada dalam perangkat tes yang digunakan dan mengestimasi peluang peserta tes yang mampu menjawab butir dengan benar. *Cutscore* yang diperoleh pada metode Angoff didasarkan pada peluang minimal peserta tes yang mampu menjawab butir dengan benar. Kegiatan ini dilakukan untuk setiap butir yang ada dalam perangkat tes yang digunakan. Selanjutnya hasil estimasi dari masing-masing butir tes dihitung besarnya rerata. Hasil rerata dari estimasi masing-masing butir tersebut yang kemudian disebut dengan *cutscore*.

Kelebihan metode Angoff adalah mudah digunakan karena perhitungan *cutscore*nya relatif lebih sederhana. Selain itu metode ini dapat digunakan pada perangkat tes yang berbentuk pilihan ganda dan bentuk uraian sehingga penggunaannya lebih fleksibel. Kelemahan metode ini adalah pada penentuan *borderline*. Metode ini hanya berdasar pada *test-centered* tidak berdasarkan data performansi peserta tes sesungguhnya. *Direct Consensus Method* merupakan salah satu perkembangan terbaru dalam metode *standard setting* yang berpusat pada tes. Metode ini pertama kali diperkenalkan oleh Sireci, Hambleton & Pitoniak pada tahun 2004. Sireci et. al juga mengungkapkan bahwa metode ini digunakan untuk mengatasi kelemahan dalam metode Angoff dan memberikan panelis kontrol langsung dalam merekomendasikan batas kelulusan yang akan ditetapkan (Cizek & Bunch, 2007, p.97).

Metode ini mengasumsikan bahwa panelis sudah mengenal tujuan dari tes tersebut dan familiar terhadap standar isi tes tersebut (Cizek & Bunch, 2007, p.98). Sedikit berbeda dengan metode Angoff, *judgement* yang diberikan pada metode ini tidak item ke item pada

perangkat tes, melainkan bentuk tes diorganisir ke dalam subtes. Subtes yang sudah tersusun tersebut ditinjau kembali dan dinilai oleh panelis. Penilaian *judgement* dilakukan dengan menentukan banyaknya butir yang dapat dijawab dengan benar oleh peserta tes secara individu pada masing-masing subtes dengan benar.

Kelebihan penggunaan metode *Direct Consensus Method* yaitu dari segi waktu, metode ini lebih efisien karena proses penilaian terhadap item tes tidak dilihat dari item ke item, tetapi penilaian dilakukan berdasarkan subtes yang sudah terbentuk. Lebih mudah digunakan karena hanya didasarkan pada keputusan benar dan salah. Selain itu panelis dapat memberikan kontrol langsung terhadap *cutscore* yang direkomendasikan. Kelemahan metode ini yaitu metode ini hanya dapat digunakan pada tes pilihan ganda dan tes yang dinilai terdiri dari beberapa subtes, sehingga penggunaannya kurang fleksibel.

The Contrasting Groups Method diperkenalkan oleh Berk pada tahun 1976. Metode ini merupakan metode yang berpusat pada peserta tes (*test-taker*). Metode ini membagi sebuah kelompok tes kedalam dua kelompok yaitu kelompok yang menguasai (*master*) dan yang tidak menguasai (*nonmaster*). Cizek & Bunch (2007, p.106) mengungkapkan bahwa *The Contrasting Groups Method* merupakan prosedur kelompok yang digunakan dalam menentukan perbedaan *cutscore* antara peserta tes yang terlatih dan tidak terlatih atau peserta tes yang telah menguasai materi dan yang belum menguasai materi. Metode ini melibatkan pelaksanaan ujian dua kelompok siswa dimana mereka menerima pembelajaran secara efektif mencakup materi tes yang akan diujikan. Distribusi dari prestasi peserta tes dari kedua kelompok ini selanjutnya dipotongkan untuk memperoleh *cutscore*.

Kelebihan menggunakan metode ini selain mudah, juga dapat memberikan hasil lebih akurat dalam penentuan skor pemisah antara kelompok yang menguasai dan yang tidak menguasai. Hal ini dikarenakan *cutscore* yang diperoleh berdasarkan kondisi nyata dari peserta tes. Kelemahan metode ini adalah sulit mendapatkan evaluasi peserta yang dapat diperbandingkan untuk wilayah yang luas seperti misalnya ujian tingkat nasional.

Metode Penelitian

Jenis penelitian ini termasuk dalam penelitian deskriptif eksploratif. Penelitian ini

dimaksudkan untuk mendeskripsikan hasil bukti empirik mengenai batas lulus Mata Pelajaran Produktif Teori dalam UKK SMK TKR di Kabupaten Brebes.

Penelitian ini dilakukan di SMK TKR Kabupaten Brebes dengan menghadirkan Panelis. Penelitian ini dilaksanakan Maret sampai dengan April 2015. Sumber data yang digunakan dalam penelitian ini terdiri dari dua sumber yaitu data set respon peserta *try out* UKK SMK TKR di Kabupaten Brebes, Panelis dari Musyawarah Guru Mata Pelajaran Produktif (MGMP) Kabupaten Brebes sebagai *judgement*. Panelis sebagai *judgement* dengan pengalaman mengajar rata-rata 10 tahun, pendidikan terakhir minimal S-1 dan mampu mengkaji atau memahami secara mendalam mengenai level konseptual dalam bidangnya.

Subjek dalam penelitian ini adalah peserta *try out* UKK SMK TKR dan MGMP di Kabupaten Brebes Tahun Pelajaran 2014/2015. Objek dalam penelitian ini adalah respon *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2014/2015. Tabel 4 menunjukkan daftar sekolah dan jumlah peserta tes yang ikut dalam *try out* UKK SMK Otomotif Teknik Kendaraan Ringan di Kabupaten Brebes.

Sebelum dilakukan penentuan *cutscore*, langkah pertama yaitu analisis karakteristik perangkat soal *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2013/2014, Karakteristik butir akan dikaji secara kuantitatif melalui analisis empirik. Pada penelitian ini dilakukan analisis butir soal berdasarkan teori tes klasik dengan menggunakan program Itean versi 3.0 dan teori respon butir dengan model dua parameter (2-PL).

Penentuan *cutscore* dilakukan dalam tiga tahap. Tahap pertama dilakukan untuk memberikan pelatihan kepada panel tentang pengertian, tujuan dan penggunaan metode *standard setting* dalam menentukan *cutscore*. Tahap kedua pelaksanaan panelis yaitu dengan praktek menentukan *cutscore* dengan menggunakan perangkat soal *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2013/2014 pada putaran I dan tahap ketiga pelaksanaan panelis untuk putaran II. Untuk menentukan *cutscore* dalam penelitian ini digunakan tiga metode sebagai berikut ini.

Prosedur penentuan *cutscore* dengan menggunakan metode *Angoff* yaitu; Panelis menelaah butir tes secara individual, Panelis menentukan skor butir dengan mengestimasi peluang minimal peserta tes yang mampu

menjawab butir dengan benar (dinyatakan dengan persen), Menentukan skor total berdasarkan rata-rata persentase skor butir, Menentukan *cutscore* berdasarkan rata-rata skor total masing-masing panelis, Panelis berdiskusi mengenai *cutscore* yang diperoleh kemudian mendeskripsikan kompetensi minimal yang diperlukan siswa untuk dapat dinyatakan lulus.

Prosedur penentuan *cutscore* dengan metode *Direct Consensus Method* yaitu; Panelis menentukan subtes dari Perangkat soal *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2013/2014, subtes dikelompokkan berdasarkan 6 dimensi sesuai dengan SKL yang telah ditentukan oleh pemerintah yaitu: peralatan, perakitan, pengelasan, perbaikan, perawatan, dan control process berbasis peralatan elektronik pada kendaraan, Panelis menentukan banyaknya item yang dapat dijawab dengan benar oleh peserta tes untuk masing-masing subtes, Panelis secara langsung menjumlahkan item yang dapat dijawab benar dari masing-masing subtes, Panelis menentukan *cutscore* berdasarkan rata-rata skor total yang diperoleh oleh masing-masing panelis, Panelis berdiskusi mengenai *cutscore* yang diperoleh kemudian mendeskripsikan kompetensi minimal yang diperlukan siswa untuk dapat dinyatakan lulus.

Prosedur penentuan *cutscore* dengan menggunakan *Contrasting Groups Method* yaitu; Panelis mendiskusikan indikator master dan nonmaster, Selanjutnya Panelis membagi grup peserta tes menjadi dua yaitu grup master dan grup nonmaster, Menghitung hasil skor tes, Menyiapkan distribusi frekuensi, Membuat grafik distribusi frekuensi pada setiap grup, Titik potong dari kedua kelompok tersebut merupakan *cutscore*.

Hasil Penelitian dan Pembahasan

Karakteristik Perangkat Tes

Karakteristik perangkat tes Perangkat soal *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2013/2014. Karakteristik perangkat tes dilihat berdasarkan teori tes klasik dan teori respon butir. Data set yang digunakan analisis karakteristik perangkat tes yaitu berupa respon siswa SMK TKR sebanyak 879 respons yang berasal dari 7 SMK yang ada di Kabupaten Brebes. Data ini merupakan data dokumentasi yang diperoleh peneliti dari MKKS Kabupaten Brebes. Perangkat Tes Perangkat soal *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2013/2014

terdiri atas 40 butir soal pilihan ganda dengan 5 alternatif jawaban.

Hasil analisis karakteristik butir soal berdasarkan teori tes klasik menunjukkan bahwa 35 butir (87,5%), memiliki tingkat kesulitan butir sedang, 3 butir termasuk mudah (7,5%) dan 2 butir termasuk sulit (5%). Daya beda butir yang dapat diterima (baik) yaitu 32 butir (80%), sedangkan sisanya kurang baik 8 butir (20%) yaitu butir 3, 5, 6, 17, 18, 32, 38 dan 39. Jika dilihat dari keberfungsian pengecoh, terdapat 38 butir (95%) yang berfungsi dengan baik, sedangkan 2 butir (5%) yaitu butir 1 dan 2 kurang baik.

Hasil analisis karakteristik butir soal berdasarkan teori respon butir menunjukkan bahwa yang digunakan 39 butir (97,5%) memiliki tingkat kesukaran baik ($-2 \leq b_t \leq 2$) dan satu butir termasuk sulit yaitu butir 17 ($b_t = 3,971$). Jika dilihat berdasarkan daya beda butir, diketahui 40 butir (100%) memiliki daya beda yang baik ($0 \leq a_t \leq 2$). Hal ini berarti bahwa perangkat soal UKK yang di ujikan pada *try out* Tahun Pelajaran 2013/2014, hasil MGMP Kabupaten Brebes mampu membedakan siswa dengan kemampuan yang rendah. Kesesuaian butir dengan model menunjukkan bahwa 38 butir (95%) cocok dengan model, sedangkan kedua butir yang tidak sesuai dengan model yaitu butir 1 dan butir 29.

Penentuan *Cutscore*

Proses penentuan *cutscore* dengan menggunakan metode *standard setting* dilakukan selama dua hari. Pada hari pertama dilakukan kegiatan seminar. Pada kegiatan ini pertama memberikan penjelasan mengenai maksud dan tujuan dari pelaksanaan kegiatan, pemaparan mengenai *standard setting* dan metode yang digunakan untuk menentukan *cutscore* dalam *standard setting* (Angoff, *Direct Consensus Method* dan *Contrasting Groups Method*). Selanjutnya dilakukan sesi tanya jawab terhadap apa yang sudah dipaparkan dan terakhir yaitu latihan (simulasi) penentuan *cutscore* dengan metode *standard setting* pada masing-masing metode.

Pada hari kedua yaitu praktek penentuan batas kelulusan UKK. Perangkat soal *try out* UKK SMK TKR di Kabupaten Brebes Tahun Pelajaran 2013/2014 dengan jumlah butir soal 40 butir. Panel yang hadir (7) dibagi menjadi dua kelompok. Masing- masing kelompok menentukan *cutscore* dengan metode yang berbeda

yaitu metode *Angoff* dan *Direct Consensus Method*. Pada metode *Contrasting Groups Method* keseluruhan panel bergabung menjadi satu untuk menentukan kelompok master (menguasai) dan kelompok nonmaster (belum menguasai).

Cutscore Berdasarkan Metode Angoff

Pelaksanaan metode ini diawali dengan pemaparan mengenai proses pelaksanaan. Panel yang terlibat dalam metode ini yaitu sebanyak 7 panel. Selanjutnya masing-masing panel diberikan perangkat tes dengan jumlah butir soal 40 butir. Panel secara individu mengestimasi peluang minimal siswa yang mampu menjawab butir dengan benar (dinyatakan dalam persen). Kegiatan ini dilakukan sebanyak dua kali putaran. Hasil rerata estimasi *cutscore* masing-masing panel pada masing-masing putaran disajikan pada Tabel 1.

Tabel 1. *Cutscore* Berdasarkan Metode Angoff

Panel	Putaran		Rerata
	I	II	
Panel 1	58,25	50,95	54,60
Panel 2	54,00	49,75	51,88
Panel 3	56,63	50,33	53,48
Panel 4	57,75	52,25	55,00
Panel 5	56,03	50,15	53,09
Panel 6	57,95	50,43	54,19
Panel 7	58,75	52,20	55,48
Estimasi Kemampuan Siswa	57,08	50,89	53,95
Standard Deviasi	1,60	1,00	1,24
Varians	2,60	0,98	1,50

Pada tabel 1 *cutscore* yang diperoleh masing-masing panel pada putaran I dan II tampak berbeda. Putaran pertama rerata *cutscore* dari 7 panel yaitu 57,08. Hasil ini jauh lebih tinggi dari pada rerata *cutscore* yang diperoleh pada putaran II yaitu 50,89. Hal ini disebabkan karena hasil diskusi setelah putaran I, panel berpendapat bahwa *cutscore* 57,08 masih banyak siswa yang dapat melampaui *cutscore* tersebut.

Hasil rerata putaran I dan II diketahui bahwa rerata *cutscore* dari kedua putaran yaitu 53,95 (skala 100) dengan standar deviasi 1,24. Hal ini menunjukkan bahwa batas lulus pelajaran produktif untuk teori UKK SMK TKR berdasarkan soal Tahun Pelajaran 2013/2014 yang di uji *try out* di Kabupaten Brebes adalah 53,95. Artinya bahwa, siswa yang dinyatakan lulus

jika siswa mampu menjawab butir soal minimal 53,95% dari keseluruhan butir tes yang ada.

Deskripsi Level Kompetensi Siswa

Tahap selanjutnya yaitu mendiskusikan tentang deskripsi level kompetensi minimal siswa berdasarkan hasil *cutscore* yang diperoleh dengan metode *Angoff*. Deskriptor yang digunakan merupakan kompetensi yang diukur berdasarkan tiap butir soal hasil diskusi dan disesuaikan dengan rerata panelis yang dalam mengestimasi peluang menjawab butir soal dengan benar. Deskripsi kemampuan siswa dapat dilihat pada Tabel 2.

Tabel 2. Deskripsi Level Kompetensi Siswa Berdasarkan metode Angoff

No	Deskripsi
1	Menjelaskan proses-proses mesin konversi energi
2	Melaksanakan prosedur pengelasan, pematrian, pemotongan dengan panas dan pemanasan
3	Memelihara/service sistem pendingin dan komponen-komponennya
4	Memelihara/servis sistem bahan bakar bensin
5	Memperbaiki sistem injeksi bahan bakar diesel
6	Memelihara/servis <i>engine</i> dan komponen
7	Memperbaiki unit kopling dan komponen-komponen sistem pengoperasian
8	Memelihara transmisi
9	Memelihara unit <i>final drive</i> /gardan
10	Memperbaiki poros penggerak roda
11	Memperbaiki roda dan ban
12	Memperbaiki sistem rem
13	Memperbaiki sistem kemudi
14	Memperbaiki sistem suspensi
15	Memelihara baterai
16	Memperbaiki kerusakan ringan pada rangkaian/ sistem kelistrikan, pengaman, dan kelengkapan tambahan
17	Memperbaiki sistem pengapian
18	Memperbaiki sistem <i>starter</i> dan pengisian
19	Memelihara/servis sistem AC (<i>Air Conditioner</i>)

Cutscore berdasarkan Direct Consensus Method

Proses pelaksanaan penentuan batas lulus dengan *Direct Consensus Method* diawali dengan pemaparan mengenai tata cara proses pelaksanaan. Selanjutnya masing-masing panel diberikan perangkat UKK SMK TKR Tahun

Pelajaran 2013/2014 yang di uji *try out* sub-area/subtes yang meliputi peralatan, perakitan, pengelasan, perbaikan, perawatan, dan control process berbasis peralatan elektronik pada kendaraan, sesuai dengan SKL yang sudah ditetapkan. Selanjutnya masing-masing panel mengestimasi banyaknya butir yang dapat dijawab dengan benar pada setiap subtes kegiatan ini dilakukan sebanyak dua kali hasil rerata rerata estimasi *cutscore* masing-masing subtes dari ke-7 panel pada masing-masing putaran adalah sebagai berikut.

Tabel 3. *Cutscore* Berdasarkan *Direct Consensus Method*

Subtes (Σ butir)	Putaran (%)		Rerata
	I	II	
Peralatan (2)	80,95	66,67	73,81
Perakitan (18)	59,86	51,70	55,78
Pengelasan (5)	64,29	61,90	63,1
Perbaikan (4)	71,43	60,71	66,07
Perawatan/ <i>Troubleshooting</i> (8)	69,64	55,36	62,5
<i>control process</i> berbasis peralatan elektronik pada kendaraan (3)	78,57	61,90	70,24
Estimasi kemampuan	60,68	59,71	60,19
Standar Deviasi	0,55	0,39	0,47
Varians	0,01	0,03	0,05

Hasil rerata putaran I dan II batas lulus perangkat tes dapat dilihat pada Tabel 3. Rerata *cutscore* yang diperoleh pada putaran I dan II tampak berbeda. Pada putaran I diperoleh rerata *cutscore* yaitu $30,27 \cong 30$ (skala 40) atau 60,68%. Sedangkan pada putaran II lebih rendah yaitu $25,79 \cong 26$ (skala 40) atau 59,71%. Perbedaan ini terjadi karena panel dalam mengestimasi butir soal dalam masing-masing subtes dilakukan secara individu sesuai persepsi panel dan hasilnya diketahui setelah dikumpulkan menjadi satu pada ketua kelompok. Selain itu, beberapa panel masih terbawa pada kemampuan minimal pada sekolah masing-masing. Sehingga menyebabkan standar deviasi pada putaran I lebih tinggi dibandingkan Putaran II.

Hasil rerata dari ke-2 putaran diperoleh batas kelulusan UKK sebesar 28,00 (skala 40) atau 60,19 (skala 100). Hasil rerata ke-2 putaran, estimasi tertinggi pada subtes pada subtes peralatan yaitu 73,81%. Hal ini berarti bahwa pada kompetensi ini termasuk mudah sehingga akan banyak peserta tes yang mampu menjawab butir-butir soal dengan benar pada subtes ini.

Estimasi terendah terjadi pada subtes perakitan yaitu 55,78%. Hal ini menunjukkan bahwa dari 21 butir soal yang terdapat dalam subtes perakitan, peserta tes hanya mampu menjawab butir dengan benar 55,78%.

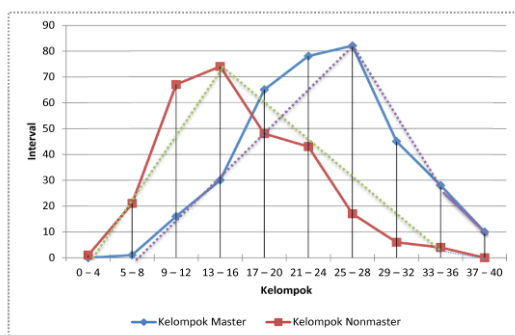
Deskripsi level kompetensi siswa disesuaikan dengan hasil estimasi penguasaan kompetensi siswa pada masing-masing subtes. Deskripsi level kompetensi siswa ini hanya memuat subtes-subtes yang minimal harus dikuasai oleh siswa.

Tabel 4. Deskriptor Kemampuan Siswa Berdasarkan *Direct Consensus Method*

No	Subtes	Presentase
1	Peralatan	73,81%
6	peralatan elektronik pada kendaraan berbasis <i>control process</i>	70,24%
4	Perbaikan	66,07%

Cutscore Berdasarkan Contrasting Groups Method

Langkah pertama dalam menentukan *cutscore* dengan *Contrasting Group Method* yaitu seluruh panel mendiskusikan mengenai kelompok “master” dan “nonmaster”. Indikatornya berdasarkan *raw score* dari masing-masing sekolah *judgement* panel. *Judgement* diberikan berdasarkan pengalaman panel selama menjadi guru produktif di sekolah serta mempertimbangkan input dan output dari masing-masing sekolah. Daftar sekolah yang digunakan dalam penentuan master dan nonmaster merupakan beberapa sekolah SMK Otomotif Teknik Kendaraan Ringan sebanyak 7 sekolah. Berdasarkan hasil diskusi panel, panel setuju dari 7 sekolah SMK negeri dan swasta, terdapat 4 sekolah yang termasuk kelompok *master* dan 3 sekolah yang termasuk kelompok *nonmaster*.



Gambar 1. Grafik Distribusi Frekuensi Kelompok Master dan Nonmaster

Gambar 1 menunjukkan perpotongan grafik antara kedua kelompok belum tampak jelas. Titik potong kedua kelompok menunjukkan bahwa *cutscore* terletak antara 21-24. Oleh sebab itu agar diperoleh *cutscore* yang akurat maka diperlukan metode penghalusan dengan menggunakan persamaan regresi logistik.

Persamaan regresi logistik ini merupakan cara yang paling umum digunakan dalam menentukan titik potong antara dua distribusi yang terbentuk, dalam penelitian ini yaitu distribusi skor kelompok master dan nonmaster. Regresi logistik digunakan untuk memprediksi probabilitas kejadian suatu peristiwa dengan mencocokkan data pada fungsi logit kurva logistik (y^*). Pada kasus ini, regresi logistik digunakan untuk memperoleh nilai x yang menggambarkan skor yang terletak tepat di tengah-tengah antara dua klasifikasi yang ada dengan nilai probabilitas dari kedua kelompok 0,50.

Hasil dari analisis regresi logistik, diperoleh *cutscore* 21,36; mendekati 21 pada skala 40 atau 52,50% pada skala 100. Skor benar ini merupakan skor minimal yang direkomendasikan oleh *Contrasting Group Method*, artinya siswa yang memperoleh skor benar ≥ 21 dapat dinyatakan lulus.

Deskripsi level kompetensi siswa pada metode ini dilihat berdasarkan nilai proporsi menjawab benar (*Prop. Correct*) pada output analisis butir soal dengan menggunakan Microcat Iteman. Deskripsi level kompetensi siswa dengan menggunakan *Contrasting Group Method*.

Keakuratan metode dilihat berdasarkan nilai standar deviasi dan reliabilitas yang diperoleh dari masing-masing metode. Berdasarkan ke-3 metode yang digunakan metode *Direct Consensus Method* merupakan metode yang paling akurat. Hal ini ditunjukkan dengan nilai SEM yang diperoleh relatif paling kecil dibandingkan dengan ke-2 metode lainnya (0,991). Keakuratan metode ini dalam menentukan *cutscore* disebabkan karena panel dapat mengontrol secara langsung *cutscore* yang diperoleh. Ini terlihat dari *cutscore* yang diperoleh berdasarkan penilaian panelis pada masing-masing subtes.

Cutscore yang diperoleh berdasarkan perangkat perangkat UKK SMK TKR Tahun Pelajaran 2013/2014 yang di uji *try out* dari masing-masing metode tidak terlalu jauh berbeda. *Cutscore* dengan metode Angoff sebesar 53,95 (skala 100), *Direct Consensus Method* 60,19 (skala 100) dan *Contrasting Groups Method* 52,50 (skala 100). Jika dicermati *cut-*

score yang diperoleh dengan *Contrasting Groups Method* lebih rendah dibandingkan dengan kedua metode lainnya. Pada metode ini, *cutscore* yang diperoleh dipengaruhi oleh konsep belajar teori produktif siswa dalam menguasai materi yang diujikan. Siswa yang memiliki pemahaman konsep yang baik dalam penguasaan materi dan pemahaman konsep yang baik tentunya akan dengan mudah menjawab butir-butir soal yang terdapat dalam perangkat tes sehingga skor yang diperoleh oleh peserta tes akan tinggi dan sebaliknya. Hal ini yang nantinya akan berpengaruh terhadap distribusi skor yang diperoleh siswa.

Tabel 5. Deskripsi level kemampuan siswa berdasarkan *Contrasting Groups Method*

No	Prop. Correct	Deskripsi
1	0,880	Menjelaskan proses-proses mesin konversi energi
2	0,826	Melaksanakan prosedur pengelasan, pematrian, pemotongan dengan panas dan pemanasan
3	0,804	Memelihara/service sistem pendingin dan komponen-komponennya
4	0,778	Memelihara/servis sistem bahan bakar bensin
5	0,727	Memperbaiki sistem injeksi bahan bakar diesel
6	0,690	Memeliharaan/servis <i>engine</i> dan komponen
7	0,665	Memperbaiki unit kopling dan komponen-komponen sistem pengoperasian
8	0,652	Memelihara transmisi
9	0,632	Memelihara unit <i>final drive</i> /gardan
10	0,619	Memperbaiki poros penggerak roda
11	0,529	Memperbaiki roda dan ban
12	0,586	Memperbaiki sistem rem
13	0,585	Memperbaiki sistem kemudi
14	0,569	Memperbaiki sistem suspensi
15	0,560	Memelihara baterai
16	0,545	Memperbaiki kerusakan ringan pada rangkaian/ sistem kelistrikan, pengaman, dan kelengkapan tambahan
17	0,540	Memperbaiki sistem pengapian
18	0,521	Memperbaiki sistem <i>starter</i> dan pengisian
19	0,513	Memelihara/servis sistem AC (<i>Air Conditioner</i>)

Direct Consenses Method merupakan metode yang paling akurat dalam menentukan *cutscore*. Hal ini terlihat dari hasil nilai SEM yang diperoleh relatif paling kecil jika diban-

dingkan dengan kedua metode lainnya. Paling akurat dalam menentukan batas kelulusan dengan menggunakan metode *standard setting*. Pada penelitian ini, penentuan *cutscore* dilaksanakan selama dua putaran. Masing-masing panel menentukan *cutscore* dengan memberikan penilaian pada perangkat perangkat UKK SMK TKR Tahun Pelajaran 2013/2014 yang di uji *try out* sebanyak dua kali untuk masing-masing metode. Stabilitas penilaian yang diberikan panel ditunjukkan dengan *cutscore* yang yang diperoleh dari ke-dua putaran relatif sama. Untuk memperoleh *cutscore* yang relatif sama, maka harus sebisa mungkin menghindari kesalahan pengukuran yang mungkin terjadi. Oleh karena itu keandalan dari sebuah alat ukur dimana dalam penelitian ini adalah metode yang digunakan dalam penentuan *cutscore* haruslah memiliki tingkat kesalahan pengukuran sekecil mungkin.

Simpulan dan Saran

Simpulan

Berdasarkan hasil penelitian dan pembahasan di atas, maka dapat diambil kesimpulan sebagai berikut; Analisis berdasarkan karakteristik butir teori tes klasik sudah baik, terlihat dari 35 butir (87,5%) memiliki tingkat kesukaran baik, 32 butir (80%) butir memiliki daya beda yang baik dan 38 butir (95%) memiliki distraktor yang dapat berfungsi dengan baik. Berdasarkan teori respon butir sudah baik, terlihat dari 97,5% butir memiliki tingkat kesukaran baik, 40 butir (100%) memiliki daya beda butir yang baik dan 95% butir cocok dengan model.

Cutscore yang diperoleh dengan metode Angoff sebesar 53,95 (skala 100), yang berarti bahwa kemampuan minimal yang dimiliki peserta tes untuk menjawab butir soal dengan benar sebesar 53,95% agar peserta tes tersebut dapat dinyatakan lulus Ujian Kompetensi Kejuruan teori. *Cutscore* yang diperoleh dengan metode *Direct Consensus Method* sebesar 60,19 (skala 100), yang berarti bahwa peserta tes harus dapat mengerjakan minimal 60,19% butir soal dari keseluruhan butir soal yang terdapat pada masing-masing subtes untuk dinyatakan lulus Ujian Kompetensi Kejuruan teori. *Cutscore* yang diperoleh dengan metode *Contrasting Group Method* yaitu 21 (skala 40) atau 52,50 (skala 100), yang berarti bahwa peserta tes harus dapat mengerjakan minimal 21 butir soal dari 40 butir soal yang ada untuk

dinyatakan lulus Ujian Kompetensi Kejuruan teori.

Metode yang paling akurat dalam menentukan Cutscore berdasarkan perangkat tes Uji Kompetensi Kejuruan SMK Otomotif Teknik Kendaraan Ringan hasil MGMP di Kabupaten Brebes Tahun 2013/2014 adalah *Direct Consensus Method*, dengan tingkat berdasarkan nilai standar deviasi dan reliabilitas yang diperoleh yaitu 0,991.

Saran

Berdasarkan hasil penelitian dan kesimpulan di atas, saran yang peneliti dapat berikan adalah. Bagi peserta didik, hasil penelitian ini diharapkan dapat dijadikan sebagai motivasi bagi peserta didik dalam untuk belajar lebih giat dalam meningkatkan kemampuannya agar menguasai standar kompetensi yang telah direkomendasikan. Bagi Guru, Hasil penelitian ini dapat dijadikan sebagai gambaran mengenai deskripsi level kompetensi yang sudah dikuasai oleh siswa dan belum dikuasai oleh siswa, sehingga pada proses pembelajaran guru akan memberikan strategi ataupun metode pembelajaran yang tepat sesuai akan kebutuhan siswa. Perlu adanya berbagai pelatihan mengenai penentuan skor batas kelulusan (*cutscore*) bagi kelompok guru mata pelajaran sehingga guru memiliki tambahan pengetahuan yang lebih dan akhirnya mampu menerapkan metode-metode tersebut dalam menentukan *cutscore*. Bagi penelitian selanjutnya penelitian ini hanya sebatas pada lingkup pembelajaran Produktif pada SMK Otomotif Teknik Kendaraan Ringan sehingga perlu dilakukan penelitian sejenis untuk ruang lingkup yang lebih luas, seperti pada Mata Pelajaran SMK Kejuruan bidang lain dan Mata Pelajaran Umum. Penentuan skor bass kelulusan (*cutscore*) dapat mengkombinasikan berbagai metode seperti menggunakan metode berpusat pada tes lainnya dengan pendekatan modern seperti Bookmark dan Item Mapping. Penelitian ini hanya dilakukan dalam dua kali putaran, bagi penelitian selanjutnya dapat menambah banyaknya putaran sehingga guru lebih paham dalam mengestimasi kemampuan seluruh peserta tes dan tentunya hasil yang diperoleh akan lebih optimal.

Daftar Pustaka

Bejar, I. I. (2008). Standard setting: what is it? why is it important?. A Primer on Setting Cut Scores on Tests of Edu-

cational Achievement. Diakses pada 12 Nopember 2014, dari https://wvw.ets.org/Media/Research/pdf/RD_Connections7.pdf

Brandon, P. L. (2002). Using test standard-setting methods in educational program evaluation: Addressing the issue of how good is good Enough [Versi elektronik]. *Journal of Multi Displinary Evaluation*, 3,1-28.

Brataningrum, N. (2013). Komparasi Standard Setting Metode Group Contrast dan Bookmark pada Mata Pelajaran Akuntansi. *Jurnal Penelitian dan Evaluasi Pendidikan*, 14(2). Retrieved from <http://journal.uny.ac.id/index.php/jpep/article/view/1080>

Cizek, G. J. & Bunch, M.B. (2007). *Standard setting: A guide to establishing and evaluating performance standard on test*. New Delhi: Saga Publitions.

Mardapi, D. (2012). *Pengukuran, penilaian dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

Gardner, J. (2012). *Assesment and learning* (2th ed). London: SAGE.

Gueskey, T. R & Bailey, J. M. (2010). *Developing standard-based report card*. Thousand Oaks, California: Corwin.

Gueskey, T. R & Jung, L. A. (2013). *Answer to essential questions ababout standards, assessments, grading, & reporting*. Thousand Oaks, California: Corwin.

Horn, C, et al. (2000). Cut scores: results may vary. *The national board on educational testing and public policy*. Vol. 1, No. 1, 16-24. Bonston Collage: NBETPP.

Nasstrom, G. & Nystrom, P. (2008). Practical assesment, research & evaluation: A comparison of two different methods for setting performance standards for a test with constructed-response items [Electronic Journal]. *Pratical Assessment Research & Evaluation*. Vol. 13, No. 9, 1-12. Diakses pada tangga 24 September 2014 dari <http://pareonline.net/pdf/v13n9/pdf>.

Sirecie, S. G., Hambleton, R. K., & Pitoniak, M. J. (2004). *Advances in Standard Setting for*

Professional Licensure Examinations.
Diambil pada tanggal 10 Januari 2014,
dari
[http://www.performancetest.org/uploads/r
esources/Advances_in_Standard_Setting_
for_Professiona_Licensure_Examinations.
pdf](http://www.performancetest.org/uploads/resources/Advances_in_Standard_Setting_for_Professiona_Licensure_Examinations.pdf)

Zieky, M. J., Perie, M., & Livingstone, S. A.
(2008). *Cut score: a manual for setting
standard of performance on educatio-
nal and occupational test*. New Delhi:
SAGE Publications.