

PENENTUAN BATAS LULUS MATA PELAJARAN MATEMATIKA SMA/MA PROGRAM IPA MENGGUNAKAN METODE *STANDARD SETTING*

Anista Ika Surachman, Djemari Mardapi
Prodi Penelitian dan Evaluasi Pendidikan PPs UNY, Universitas Negeri Yogyakarta
anist_ta@yahoo.co.id, djemarimardapi@gmail.com

Abstrak

Penelitian ini bertujuan: (1) menentukan skor batas lulus Mata Pelajaran Matematika SMA/MA Program IPA di Kabupaten Purworejo menggunakan metode Angoff, *Direct Consensus Method* dan *Contrasting Group Method*; dan (2) mengetahui metode paling akurat dalam menentukan skor batas kelulusan Mata Pelajaran Matematika SMA/MA Program IPA di Kabupaten Purworejo. Penelitian ini termasuk deskriptif eksploratif. Sumber data berupa data set respons peserta Uji Coba Ke-3 UN Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo 2012/2013 dan panelis (ahli). *Judgement* panelis digunakan untuk menentukan *cutscore* menggunakan metode Angoff, *Direct Consensus Method* dan *Contrasting Group Method*. Keakuratan metode ditentukan dengan kecilnya nilai kesalahan pengukuran (SEM). Hasil penelitian menunjukkan (1) skor batas lulus pada Mata Pelajaran Matematika SMA/MA Program IPA dengan metode Angoff 53,95 (skala 100), *Direct Consensus Method* 54,46 (skala 100) dan *Contrasting Group Method* 52,50 (skala 100); (2) *Direct Consensus Method* merupakan metode yang paling akurat dengan SEM yang diperoleh paling kecil (0,999).

Kata kunci: *standard setting, metode angoff, direct consensus method, contrasting group method*

CUTSCORE OF MATHEMATIC SUBJECT USING *STANDARD SETTING METHOD* AT SMA/MA SCIENCE PROGRAM

Anista Ika Surachman, Djemari Mardapi
Prodi Penelitian dan Evaluasi Pendidikan PPs UNY, Universitas Negeri Yogyakarta
anist_ta@yahoo.co.id, djemarimardapi@gmail.com

Abstract

This study aims to: (1) determine the cutscore of mathematic subject at SMA/MA science program in Purworejo Regency using the Angoff Method, Direct Consensus Method, and Contrasting Group Method; and (2) find out which method is the most accurate to determine the cutscore of mathematic subject at SMA/MA science program in Purworejo Regency. This study used descriptive explorative technique. The source of data was sets of responses of the third participants of Try Out of National Examination in Mathematic subject of science program SMA/MA of MKKS Purworejo Regency 2012/2013 and panelists. The judgement of the panelists are used to determine the cutscore with Anggoff Method, Direct Consensus Method and Contrasting Group Method. The accuration of this method is shown by the smallest value of the Standard Error of Measurement (SEM). The result of the study are as follows: (1) the cutscore of Mathematic Subject at SMA/MA Science Program obtained from the Angoff method is 53.95 (scale 100), Direct Consensus Method for 54.46 (scale of 100) and Contrasting Group Method of 54.50 (scale of 100); and (2) the Direct Consensus Method is the most accurate method which seen from the SEM values obtained that are relatively the smallest 0.999.

Keywords: *standard setting, Angoff method, Direct Consensus Method, Contrasting Group Method*

Pendahuluan

Salah satu upaya pemerintah dalam meningkatkan kualitas pendidikan nasional yaitu dengan mengeluarkan Peraturan Pemerintah No. 19 tahun 2005 tentang Standar Nasional Pendidikan (SNP). Standar Nasional Pendidikan adalah kriteria minimal tentang sistem pendidikan di seluruh wilayah hukum Negara Kesatuan Republik Indonesia (PP No. 19/2005). SNP berfungsi sebagai dasar dalam perencanaan, pelaksanaan, dan pengawasan pendidikan dalam rangka mewujudkan pendidikan nasional yang berkualitas. SNP bertujuan untuk menjamin mutu pendidikan nasional dalam rangka mencerdaskan kehidupan bangsa dan membentuk watak serta peradaban bangsa yang bermartabat. Melalui SNP diharapkan kualitas pendidikan di Indonesia akan menjadi lebih baik.

SNP terdiri atas delapan standar yang harus dicapai oleh satuan pendidikan. Kedelapan standar tersebut terdiri dari standar kompetensi lulusan, isi, proses, tenaga kependidikan, sarana dan prasarana, pengelolaan, pembiayaan dan penilaian. Kedelapan standar tersebut dijadikan sebagai pedoman bagi pendidik dan tenaga kependidikan dalam mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat.

Salah satu dari delapan standar tersebut, yang penting yaitu Standar Kompetensi Lulusan (SKL). SKL digunakan sebagai acuan utama pengembangan standar isi, standar proses, standar penilaian pendidikan, standar pendidik dan tenaga kependidikan, standar sarana dan prasarana, standar pengelolaan, dan standar pembiayaan. SKL merupakan kriteria kompetensi lulusan minimal yang berlaku di Indonesia. SKL tersebut meliputi standar kompetensi lulusan minimal satuan pendidikan dasar dan menengah, standar kompetensi lulusan minimal kelompok mata pelajaran, dan standar kompetensi lulusan minimal mata pelajaran. SKL untuk satuan pendidikan dasar dan menengah digunakan sebagai pedoman penilaian dalam menentukan kelulusan peserta didik.

Penilaian merupakan komponen penting dalam penyelenggaraan pendidikan. Melalui sistem penilaian yang baik, mampu mendorong pendidik untuk menentukan strategi pembelajaran yang tepat serta dapat memotivasi peserta didik agar dapat belajar lebih baik lagi. Gueskey & Jung (2013, p.17) bagi guru, kegiatan penilaian ini sangatlah penting dalam membuat keputusan terutama dalam menentukan

efektifitas dari perbedaan strategi pembelajaran, atifitas kelas, dan materi pembelajaran. Penilaian merupakan suatu alat yang dapat digunakan dalam memperoleh informasi tentang sejauh mana hasil belajar peserta didik atau ketercapaian kompetensi peserta didik. Gardner (2012, p.87) menyebutkan bahwa penilaian dalam konteks pendidikan termasuk dalam menentukan, mengumpulkan dan membuat keputusan tentang bukti yang berkaitan dengan tujuan pembelajaran yang akan dinilai. Penilaian terhadap prestasi belajar peserta didik disebutkan dalam peraturan pemerintah nomor 19 tahun 2005 pasal 63 ayat (1) bahwa penilaian hasil belajar jenjang pendidikan dasar dan menengah terdiri dari (a) penilaian hasil belajar oleh pendidik, (b) penilaian hasil belajar oleh satuan pendidikan dan (c) penilaian hasil belajar oleh pemerintah.

Seiring dengan perkembangan dalam dunia pendidikan, sistem penilaian yang digunakan mengacu pada sistem penilaian dengan acuan kriteria. Mardapi (2012, p.186) menyebutkan tujuan penilaian acuan kriteria adalah untuk mengetahui kemampuan seseorang menurut kriteria tertentu. Pendekatan penilaian acuan kriteria berarti membandingkan skor-skor hasil tes peserta didik dengan kriteria mutlak telah ditetapkan oleh guru. Hasil penilaian dengan acuan kriteria berupa lulus dan tidak lulus atau yang menguasai dan belum menguasai. Ketentuan lulus jika nilai yang diperoleh siswa lebih dari sama dengan dari batas minimal yang telah ditentukan dan tidak lulus jika nilai yang diperoleh siswa lebih kecil dari batas minimal yang telah ditentukan.

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. 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 sudah ditentukan tersebut sudah benar-benar menggambarkan kompetensi siswa. Hal perlu dilakukan secara berkelanjutan dikarenakan tingkat lulusan dimasing-masing daerah untuk setiap tahun-

nya berbeda-beda. Misalkan saja untuk Kabupaten Purworejo pada tahun 2010 persentase kelulusan SMA/MA sebesar 84,4%, tahun 2011 sebesar 98,93% dan tahun 2012 yaitu sebesar 99,65% (Dindikpora Kabupaten Purworejo).

Kaitannya dengan peninjauan terhadap penentuan skor batas kelulusan, maka perlu dilakukan penentuan skor batas kelulusan dengan metode yang tepat sesuai dengan karakteristik peserta didik pada masing-masing daerah. Salah satu metode yang dapat digunakan yaitu metode *standard setting*. Metode *standard setting* merujuk pada prosedur kebijakan dalam menentukan skor batas (*cutscore*). Menurut Cizek & Bunch (2007, p.13) *standard setting* merupakan proses mengidentifikasi satu atau lebih skor batas (*cut score*) pada suatu tes. Bagian penting dari *standard setting* yaitu menetapkan satu atau lebih skor batas, namun untuk membuat kategori lulus dan tidak lulus maka digunakan satu skor batas.

Penentuan skor batas dengan menggunakan metode *standard setting* ini dapat diterapkan pada berbagai mata pelajaran, salah satunya yaitu Mata Pelajaran Matematika. Menurut siswa Mata Pelajaran Matematika merupakan salah satu mata pelajaran tergolong sulit, apalagi untuk tingkat SMA/MA Program IPA. Hal ini terlihat dari banyak siswa yang mengutarakan bahwa Mata Pelajaran Matematika lebih sulit jika dibandingkan mata pelajaran lainnya. Kompleksitas perhitungan untuk mata pelajaran ini membuat rata-rata nilai mata pelajaran matematika cenderung lebih rendah.

Penentuan skor batas kelulusan dengan menggunakan metode *standard setting* ada beberapa macam, namun yang sering digunakan dalam praktisi pendidikan yaitu metode yang berdasarkan *test centered* (berpusat pada tes) dan *test-takers* (berpusat pada peserta tes). Kedua metode tersebut dirasa lebih sederhana dalam pelaksanaannya sehingga sering digunakan oleh praktisi pendidikan. Selain itu, pemilihan model Angoff juga didukung dari hasil penelitian Ricker (2006) tentang "*Setting Cut-Scores: A Critical Review of the Angoff and Modified Angoff Methods*". Hasil penelitian secara keseluruhan metode Angoff adalah metode yang baik dalam penentuan *cutscore*, terutama ketika waktu dan sumber daya yang tersedia untuk melatih panel sudah benar. Modifikasi prosedur *Originally* Angoff menghasilkan hasil yang beragam. Secara umum, modifikasi tidak selalu cocok untuk setiap situasi, tetapi dapat

meningkatkan prosedur ketika digunakan pada konteks yang sesuai.

Metode yang berpusat pada tes merupakan metode *standard setting* yang berdasarkan pada pertanyaan tes (soal). Metode ini mengasumsikan bahwa panelis (ahli) mampu menggambarkan kemampuan peserta tes terhadap item-item yang terdapat dalam perangkat tes yang digunakan. Metode berpusat pada peserta tes (*examinee*) merupakan metode yang didasarkan pada peserta tes/produk. *Judgement* mengenai peserta tes pada metode ini didasarkan pada pengetahuan dan keterampilan yang menjadi sasaran tes yang bersangkutan.

Selama ini penentuan skor batas kelulusan mata pelajaran masih berdasarkan pada keputusan kebijakan pemerintah. Penentuan yang seperti ini cenderung memiliki kelemahan yaitu tidak adanya bukti empirik dalam penentuan skor batas kelulusan siswa. Oleh karena itu perlu dilakukan evaluasi terhadap skor batas yang telah ditentukan oleh pemerintah apakah benar-benar sudah mengukur kompetensi siswa.

Berdasarkan pada latar belakang di atas, maka peneliti tertarik untuk melakukan penelitian kaitannya dengan penentuan skor batas kelulusan Mata Pelajaran Matematika SMA/MA Program IPA secara empirik. Bukti empirik dilakukan dengan menggunakan metode *standard setting* yang berpusat pada tes dan berpusat pada peserta tes

Berbagai penelitian tentang *standard setting* telah banyak dilakukan praktisi pendidikan ataupun bidang lainnya. Nostrom & Nystrom (2008) melakukan penelitian yang bertujuan untuk membandingkan validitas kedua metode yang berbeda dalam menentukan *cutscore* pada Ujian Nasional Matematika di Swedia. Metode yaitu metode Angoff (berpusat pada tes) dan metode *Borderline Group Method* (berpusat pada peserta tes). Penelitian ini didasarkan pada bukti prosedural (langkah-langkah penentuan *standard setting*), eksternal (perbandingan sumber eksternal, seperti hasil dari metode lain dalam *standar setting*) dan internal (konsistensi hasil). Hasil penelitian ini menunjukkan bahwa kedua metode memberikan pendekatan yang masuk akal dan dapat dipercaya dalam *standard setting*. Brataningrum (2010) melakukan penelitian dengan tentang komparasi *standard setting* metode *group contrast* dan bookmark pada Mata Pelajaran Akuntansi. Subjek dalam penelitian ini yaitu respon jawaban tes UN Akutansi jenjang SMK akutansi 2008/2009 di Propinsi DIY.

Hasil dari penelitian ini yaitu bahwa *cutscore* dengan metode bookmark diperoleh hasil 68,96 pada Respon Probability (RP) 0,67 dan sebesar 67,36 (skala 100) pada RP 0,5 sementara pada metode *group contrast* menghasilkan *cutscore* sebesar 63,09 (skala 100). Metode Bookmark lebih memberikan hasil *cut score* yang relatif lebih akurat karena dalam prosedurnya mempertimbangkan setidaknya dua aspek, yakni parameter tingkat kesukaran butir dan estimasi *expert* dalam mengestimasi tiap butir dengan respon probabilitas tertentu.

Standard diartikan sebagai ukuran tertentu yang digunakan sebagai patokan yang telah disepakati. Merriam-webster *dictionary* dalam Guesky & Jung (2013, p.2) menyebutkan bahwa *standard* adalah sesuatu yang mengatur dan ditetapkan oleh otoritas tertentu sebagai aturan untuk mengukur kuantitas, bobot, tingkat, nilai atau kualitas. Guesky & Baily (2010, p.13) menjelaskan bahwa spesifikasi dari *standard* adalah pengetahuan, ketrampilan, kemampuan dan disposisi yang harapannya siswa dapat memperolehnya dari interaksi dengan guru ataupun sesama siswa lainnya di lingkungan sekolah. *Standard* dalam bidang akademik terdiri dari dua komponen yaitu menggambarkan spesifikasi tertentu dari sebuah isi (*content*) yang menunjukkan apa yang dipelajari oleh siswa selama mengikuti proses pembelajaran dan menggambarkan level *performance* yang menunjukkan apa yang diperoleh siswa dapat lakukan setelah mengikuti proses pembelajaran.

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-centered*) dan pertimbangan peserta tes (*test-takers*) yang akan digunakan dalam penelitian ini. Metode yang berpusat pada tes merupakan metode *standard setting* yang berdasarkan pada pertanyaan tes (soal). Metode ini mengasumsikan bahwa panelis (ahli) mampu menggambarkan dengan teliti mengenai kemampuan peserta tes terhadap item-item yang terdapat dalam perangkat tes yang digunakan. Metode berpusat pada peserta tes (*examinee*) merupakan metode yang didasarkan pada peserta tes/produk. *Judgement* mengenai peserta tes pada metode ini didasarkan pada pengetahuan dan keterampilan yang menjadi sasaran tes yang bersangkutan. 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 menela-

ah 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

Penelitian ini termasuk dalam penelitian deskriptif eksploratif. Penelitian ini dimaksudkan untuk mendeskripsikan hasil bukti empirik mengenai batas lulus Mata Pelajaran Matematika SMA Program IPA menggunakan metode *standard setting* dengan pendekatan *test-centered* (metode Angoff dan *Direct Consensus Method*) dan *test-taker* (*Contrasting Groups method*).

Tempat dan Waktu Penelitian

Penelitian ini dilakukan di Kabupaten Purworejo karena di selama ini penentuan batas lulus atau Kriteria Ketuntasan Minimal (KKM)

mata pelajaran di kabupaten masih didasarkan pada keputusan *judgement* sehingga membutuhkan informasi mengenai cara penentuan batas lulus (*cutscore*) secara empiris dengan menggunakan metode *standard setting* Penelitian ini dilaksanakan pada Maret sampai dengan April 2014 dengan menghadirkan panelis.

Sumber Data

Sumber data yang digunakan berupa data set respon peserta Uji Coba Ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA hasil MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 dan panelis (penilai ahli). Panelis yang ditunjuk sebagai sumber yaitu guru Mata Pelajaran Matematika SMA/MA dengan pengalaman mengajar rata-rata 10 tahun dengan kualifikasi pendidikan terakhir minimal S-1 Matematika/ Pendidikan Matematika, dan mampu mengkaji atau memahami secara mendalam mengenai level konseptual dalam bidangnya.

Subjek dan Objek Penelitian

Subjek dalam penelitian ini adalah peserta Uji Coba Ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013. Objek dalam penelitian ini adalah respons dan lembar soal Uji Coba Ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013.

Teknik dan Instrumen Pengumpulan Data

Teknik pengumpulan data yang digunakan yaitu dokumentasi dan FGD dari panelis yang bertujuan untuk memperoleh konsensus mengenai *cutscore* yang akan direkomendasikan. Instrumen pengumpulan data yang digunakan yaitu perangkat soal Uji Coba Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2011/2012, lembar isian untuk masing-masing metode dan lembar isian untuk dokumentasi hasil.

Teknik Analisis Data

Sebelum dilakukan penentuan *cutscore*, langkah pertama yaitu analisis karakteristik perangkat tes Uji Coba Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran

2012/2013, 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* atau skor batas kelulusan 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 tes Uji Coba Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 pada putaran I dan tahap ketiga pelaksanaan panelis untuk II. Untuk menentukan *cutscore* dalam penelitian ini digunakan tiga metode sebagai berikut ini.

Metode Angoff

Prosedur penentuan *cutscore* dengan menggunakan metode *Angoff* yaitu sebagai berikut: (1) Panelis menelaah butir tes secara individual; (2) Panelis menentukan skor butir dengan mengestimasi peluang minimal peserta tes yang mampu menjawab butir dengan benar (dinyatakan dengan persen); (3) menentukan skor total berdasarkan rata-rata persentase skor butir; (4) menentukan *cutscore* berdasarkan rata-rata skor total masing-masing panelis; (5) Panelis berdiskusi mengenai *cutscore* yang diperoleh kemudian mendiskripsikan kompetensi minimal yang diperlukan siswa untuk dapat dinyatakan lulus.

Direct Consensus Method

Prosedur penentuan *cutscore* dengan metode *Direct Consensus Method* yaitu sebagai berikut: (1) Panelis menentukan subtes dari perangkat tes yang digunakan. Perangkat Tes Uji Coba Ke-3 Mata Pelajaran Matematika SMA/MA Program IPA berdasarkan hasil MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013, subtes dikelompokkan berdasarkan 6 dimensi sesuai dengan SKL yang telah ditentukan oleh pemerintah yaitu: Logika, Aljabar, Geometri, Trigonometri, Kalkulus dan Statistika dan Peluang; (2) Panelis menentukan banyaknya item yang dapat dijawab dengan benar oleh peserta tes untuk masing-masing subtes; (3) Panelis secara langsung menjumlah-

kan item yang dapat dijawab benar dari masing-masing subtes; (4) Panelis menentukan *cutscore* berdasarkan rata-rata skor total yang diperoleh oleh masing-masing panelis; (5) Panelis berdiskusi mengenai *cutscore* yang diperoleh kemudian mendiskripsikan kompetensi minimal yang diperlukan siswa untuk dapat dinyatakan lulus.

Contrasting Groups Method

Prosedur penentuan *cutscore* dengan menggunakan *Contrasting Groups Method* yaitu sebagai berikut: (1) Panelis mendiskusikan indikator master dan nonmaster; (2) selanjutnya Panelis membagi grup peserta tes menjadi dua yaitu grup master dan grup nonmaster; (3) menghitung hasil skor tes; (4) menyiapkan distribusi frekuensi; (5) membuat grafik distribusi frekuensi pada setiap grup; (6) titik potong dari kedua kelompok tersebut merupakan *cutscore*.

Hasil Penelitian dan Pembahasan

Karakteristik Perangkat Tes

Karakteristik perangkat tes Uji Coba Ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013. Karakteristik perangkat tes dilihat berdasarkan teori tes klasik dan teori respon butir. Data set yang digunakan analisis karakteristik perangkat tes yaitu berupa respon siswa SMA/MA Program IPA sebanyak 623 respons yang berasal dari 21 SMA/MA yang ada di Kabupaten Purworejo. Data ini merupakan data dokumentasi yang diperoleh peneliti dari MKKS Kabupaten Purworejo. Perangkat Tes Uji Coba Ke-3 Mata Pelajaran Matematika SMA/MA Program IPA berdasarkan hasil MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 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 kesukaran sedang (baik), 3 butir termasuk mudah (7,5%) dan 2 butir termasuk sulit (5%). Ditinjau dari daya beda butir, menunjukkan 32 butir (80%) butir memiliki daya beda yang baik dan 8 butir (20%) kurang baik. Ditinjau dari keberfungsian distraktor, 38 butir (95%) memiliki distraktor yang dapat berfungsi dengan baik dan 2 butir kurang baik (5%).

Hasil analisis karakteristik butir soal berdasarkan teori respon butir menunjukkan bahwa 39 butir (97,5%) memiliki tingkat kesukaran baik dan satu butir termasuk sulit yaitu butir 17 ($b_i = 3,971$). Ditinjau dari daya beda butir, 100% butir memiliki daya beda yang baik. Jika dilihat dari kesesuaian butir dengan model menunjukkan bahwa 38 butir (95%) cocok dengan model, sedangkan 5% butir tidak sesuai dengan model.

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, pemamparan 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 Mata Pelajaran Matematika SMA/MA Program IPA. Perangkat tes yang digunakan yaitu tes Uji Coba ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 dengan jumlah butir soal 40 butir. Panel yang hadir (17) 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 Uji Coba ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 dengan jumlah butir soal 40 butir yang telah disusun oleh peneliti. Panel

secara individu mengestimasi peluang minimal siswa yang mampu menjawab butir dengan benar (dinyatakan dalam persen). Kegiatan ini dilakukan sebanyak dua kali putaran.

Pada Tabel 1 terlihat bahwa batas lulus Mata Pelajaran Matematika SMA/MA Program IPA berdasarkan perangkat tes Uji Coba ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 pada putaran I diperoleh 57,05 dan Putaran II 50,86%. *Cutscore* final yang diperoleh yaitu 53,95% yang berarti bahwa untuk dapat dinyatakan lulus Mata Pelajaran Matematika SMA/MA Program IPA, maka setiap siswa harus memiliki kemampuan menjawab butir soal dengan benar minimal 53,95% dari total butir yang terdapat dalam perangkat tes yang digunakan. Persentase kelulusan yang diperoleh yaitu 43,95%.

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	57,05	50,86	53,95
Standard Deviasi	1,64	1,00	1,24
Varians	2,69	0,99	1,53

Hasil estimasi setiap panel tampak jelas berbeda untuk setiap putarannya. Pada putaran I hasil estimasi *cutscore* jauh lebih tinggi dari pada Putaran II. Perbedaan hasil pada putaran I

dan II disebabkan pada putaran I panel mengestimasi kemampuan siswa secara individu dan panel mengetahui hasil final *cutscore* setelah dikumpulkan menjadi satu pada ketua kelompok. Selain itu, pada putaran pertama, beberapa panel mengalami kesulitan dalam mengestimasi kemampuan minimal seluruh peserta tes. Panel masih terpengaruhi oleh level kemampuan minimal dari peserta tes yang ada di sekolahnya masing-masing, sehingga pada putaran I diperoleh rerata *cutscore* dari 7 panel lebih tinggi dari putaran II yaitu 57,05%. Hasil ini masih dirasa cukup tinggi oleh panel mengingat banyaknya status sekolah negeri dan swasta hampir sama, akibatnya pada putaran II panel menurunkan hasil estimasi kemampuan peserta tes dalam menjawab butir soal dengan rerata *cutscore* yang diperoleh sebesar 50,86%.

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 tes Uji Coba ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 yang sudah dikelompokkan menjadi 6 subarea/subtes yang meliputi logika, aljabar, geometri, trigonometri, kalkulus dan peluang statistika sesuai dengan SKL yang sudah ditetapkan. Selanjutnya masing-masing panel mengestimasi banyaknya butir yang dapat dijawab dengan benar untuk setiap subtes. Kegiatan ini dilakukan sebanyak dua kali putaran. Hasil rerata estimasi *cutscore* masing-masing subtes dari ke-7 panel pada masing-masing putaran disajikan pada Tabel 2.

Tabel 2. Berdasarkan Direct Consensus Method

Subtes/ \sum Butir	Putaran (%)				Rerata <i>Cutscore</i>
	I		II		
	Rerata Jumlah Butir	%	Rerata Jumlah Butir	%	
Logika / 2	1,57	78,57	1,00	50,00	64,29
Aljabar / 21	12,14	59,86	11,14	51,02	55,44
Geometri / 2	1,14	57,82	1,00	53,06	53,57
Trigonometri / 4	2,43	60,71	1,86	46,43	53,57
Kalkulus / 8	3,86	48,21	3,43	42,86	45,54
Peluang dan Statistika / 3	2,14	71,43	1,86	61,90	66,67
Estimasi Kemampuan Siswa	23,29	58,21	20,28	50,70	53,95

Pada putaran I hasil estimasi *cutscore* dari keseluruhan subtes adalah 58,21% sedangkan putaran ke-II 50,71%. Pada putaran I *cutscore* yang diperoleh masih cukup tinggi, sehingga pada putaran berikutnya panel sepakat untuk menurunkan *cutscore* yang diperoleh pada putaran sebelumnya dengan harapan bahwa *cutscore* yang diperoleh pada putaran kedua akan sesuai dengan kemampuan minimal keseluruhan peserta tes. Rerata dari ke-2 putaran diperoleh batas kelulusan Mata Pelajaran Matematika SMA/MA Program IPA sebesar 54,46 (skala 100) atau $21,79 \cong 22$ (skala 40)

Tabel 2 menunjukkan bahwa rata-rata tertinggi terdapat pada subtes peluang dan statistika (66,67%). Hal ini berarti bahwa 66,67% dari ke-3 butir soal yang terdapat pada subtes ini harus dapat dikerjakan oleh siswa dengan benar. Tingginya estimasi yang diberikan panel pada subtes ini mengindikasikan bahwa peserta tes sebagian besar dapat mengerjakan butir-butir yang terdapat dalam subtes ini. Estimasi terendah yang diberikan panel yaitu pada subtes kalkulus yaitu dengan rata-rata kedua putaran 45,54%. Materi yang terdapat dalam subtes kalkulus ini meliputi limit fungsi (aljabar dan trigonometri), kalkulus diferensial dan kalkulus integral. Dari 8 butir soal yang terdapat dalam subtes ini panel menilai hanya 45,54% yang dapat dijawab oleh peserta tes.

Cutscore Berdasarkan Contrasting Groups Method

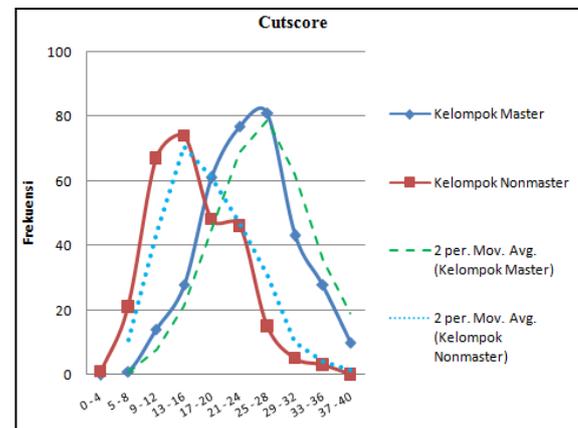
Langkah pertama dalam menentukan *cutscore* dengan *contrasting groups method* yaitu seluruh panel mendiskusikan mengenai kelompok “master” dan “nonmaster”. Indikator master dan nonmaster berdasarkan *judgement* panel. *Judgement* diberikan berdasarkan pengalaman panel selama menjadi guru matematika di Kabupaten Purworejo serta mempertimbangkan input dan *output* dari masing-masing sekolah.

Berdasarkan hasil diskusi dari 17 panel, bahwa dari 21 sekolah SMA/MA negeri dan swasta di Kabupaten Purworejo, terdapat 6 sekolah yang termasuk kelompok master (343 siswa) dan 15 sekolah termasuk kelompok nonmaster (280 siswa). Selanjutnya membuat distribusi frekuensi untuk masing-masing kelompok. Distribusi yang digunakan yaitu distribusi kelompok dari jumlah jawaban benar

siswa (*raw score*) selanjutnya membuat grafik distribusi frekuensi dari kedua kelompok.

Tabel 3. Daftar Distribusi Frekuensi Masing-masing Kelompok

No	Interval <i>Raw score</i>	Kelompok	
		Master	Nonmaster
1	0 - 4	0	1
2	5 - 8	1	21
3	9 - 12	14	67
4	13 - 16	28	74
5	17 - 20	61	48
6	21 - 24	77	46
7	25 - 28	81	15
8	29 - 32	43	5
9	33 - 36	28	3
10	37 - 40	10	0
Jumlah		343	280



Gambar 1. Grafik Distribusi Frekuensi Kelompok Master dan Nonmaster

Gambar 1 menunjukkan perpotongan grafik antara kedua kelompok belum tampak begitu jelas. *Series point* 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 yang dihasilkan dari regresi logistik berdasarkan output sebagai berikut: $y^* = -3,644 - 0,194(x)$; dengan $y^* = 0,5$ maka didapat x sebesar 21,36; mendekati 21 pada skala 40 atau 52,50 pada skala 100. Nilai ini merupakan *cutscore* yang direkomendasikan menurut *Contrasting Groups Method* dengan persentase kelulusan 50,06%.

Keakuratan metode dilihat berdasarkan nilai SEM yang diperoleh. 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 (0,999). 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 tes Uji Coba ke-3 Ujian Nasional Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 dari masing-masing metode tidak terlalu jauh berbeda. *Cutscore* dengan metode Angoff sebesar 53,95 (skala 100), *Direct Consensus Method* 54,46 (skala 100) dan *Contrasting Groups Method* 52,50 (skala 100).

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

Direct Consensus Method merupakan metode yang paling akurat dalam menentukan *cutscore*. Hal ini terlihat dari hasil nilai SEM yang diperoleh relatif paling kecil jika dibandingkan 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 Tes Uji Coba Ke-3 Mata Pelajaran Matematika SMA/MA Program IPA sebanyak dua kali untuk masing-masing metode. Stabilitas penilaian yang diberikan panel ditunjukkan dengan *cutscore* 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

Berdasarkan hasil penelitian dan pembahasan di atas, maka dapat diperoleh kesimpulan sebagai berikut: (1) *Cutscore* yang diperoleh dengan metode Angoff sebesar 53,95 (skala 100), *Direct Consensus Method* sebesar 54,46 (skala 100) dan *Contrasting Group Method* yaitu 21 (skala 40) atau 52,50 (skala 100); dan (2) metode yang paling akurat dalam menentukan *cutscore* berdasarkan perangkat tes Uji Coba ke-3 UN Mata Pelajaran Matematika SMA/MA Program IPA MKKS Kabupaten Purworejo Tahun Pelajaran 2012/2013 adalah *Direct Consensus Method*, karena metode ini memiliki tingkat kesalahan pengukuran yang terkecil yaitu 0,999.

Berdasarkan hasil penelitian dan kesimpulan di atas, saran yang peneliti dapat berikan: (1) dapat dijadikan sebagai motivasi bagi peserta didik untuk belajar lebih giat dalam meningkatkan kemampuannya agar menguasai standar kompetensi yang telah direkomendasikan; (2) 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*; (3) perlu dilakukan berbagai upaya dalam meningkatkan level kompetensi siswa sesuai dengan apa yang sudah ditetapkan dalam SKL. Upaya-upaya tersebut diantaranya perbaikan proses pembelajaran dengan memberikan strategi ataupun metode pembelajaran yang tepat sesuai akan kebutuhan siswa; dan (4) perlunya penelitian lebih lanjut mengenai penentuan batas kelulusan menggunakan metode *standard setting* dalam cakupan yang lebih luas.

Daftar Pustaka

- Bejar, I. I. (2008). *Standard setting: what is it? why is it important?*. ETS. Diakses pada 12 Nopember 2013, dari https://www.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 Disiplinary Evaluation, 3, 1-28.
http://journals.sfu.ca/jmde/index.php/jmde_1/article/viewFile/99/114.
- Cizek, G. J. & Bunch, M.B. (2007). *Standard setting: A guide to establishing and evaluating performance standard on test*. New Delhi: Saga Publitions.
- Djemari Mardapi. (2012). *Pengukuran, penilaian dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.
- Gardner, J. (2012). *Assesment and learning (2th ed)*. London: SAGE.
- Guskey, T. R & Bailey, J. M. (2010). *Developing standard-based report card*. Thousand Oaks, California: Corwin.
- Guskey, 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. Bonstton Collage: NBETPP.
- Brataningrum, N. P. (2010). Komparasi standard setting metode group contrast dan bookmark pada mata pelajaran akuntansi. *Jurnal HEPI*, Tahun 14 (2), 233-234.
- 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 Sepember 2013 dari <http://pareonline.net/pdf/v13n9/pdf>.
- Zieky, M. J., Perie, M., & Livingstone, S. A. (2008). *Cut score: a manual for setting standard of performance on educational and occupational test*. New Delhi: SAGE Publications.