

PENETAPAN STANDARD KOMPETENSI PERFORMANCE STANDARD SETTING ✓

Oleh:
SRI HARTINI *)

ABSTRACT

Standard is meant the answer of "how good is enough". In the field of education (teaching and learning), passing point or cut score is a common terminology/concept. It is used as a standard to determine one's achievement (success). Standard is divided into academic content standard and performance standard. The former means what is supposed to know and to do by students. Where as the latter concerns how far students master knowledge and skills. There are three methods of setting performance standard (1) Nedelsky, (2) Angoff, (3) Ebel. The three methods are based on the fact that there is a group of students that belongs to border line. Based on the above explanation, it is concluded as follows : (1) passing point (cut score) is needed to determine student's competence for the sake of certification ; (2) The standard is based on border-line student ; (3) There is no best method. Method is determined by appropriateness due to certain/specific condition required.

Keyword : Performance and Standard Setting

**) Dosen FKIP UNISRI Surakarta*

Pendahuluan

Kita telah mengenal standar untuk berbagai macam pekerjaan dan produk pembangunan gedung, jembatan, jalan raya, dan irigasi harus mengikuti standar tertentu karena dampaknya jika tidak mengikuti akan mengancam keselamatan jiwa manusia. Demikian juga dengan praktek pengobatan dan produk obat-obatan, makanan, minuman, bahkan mobil. Standar diperlukan sebagai upaya preventif untuk *safety* dan *reliability*. Baru-baru ini muncul berbagai lembaga dan program akreditasi untuk standarisasi mutu manajemen perusahaan seperti sertifikasi ISO, atau akreditasi kursus-kursus pelatihan dan program

pendidikan tinggi. Konsep standar dengan demikian, telah meluas bukan hanya mencakup aspek *safety* dan *reliability* tetapi sebagai jaminan untuk mencapai mutu yang diharapkan.

Standar adalah jawaban dari pertanyaan "seberapaakah dapat dianggap cukup". Dengan mempergunakan "Nilai Batas Lulus" atau NBL". Itu berarti juga telah menetapkan standar untuk menentukan tingkat keberhasilan seseorang. Di dalam dunia penilaian pada umumnya "batas" itu tidaklah tegas, sehingga sulit untuk menentukan berapa (nilai) sesungguhnya yang dianggap cukup. Bagaimanakah cara menetapkan standar siswa yang mampu dan yang

tidak mampu dan siswa yang berhasil dan yang tidak berhasil. Untuk menjawab pertanyaan tersebut diperlukan *judgement* (penilaian), yang dilakukan oleh orang-orang yang mempunyai ketrampilan menilai sesuai dengan tujuan tes itu sendiri.

Pembahasan

Pengertian dan fungsi standar kompetensi

Dalam standar pendidikan dikenal dua jenis standar yaitu standar akademis (*academic content standart*) dan standar kompetensi (*performance standart*). Standar akademis mendefinisikan apa yang seharusnya dikuasai (diketahui dan dikerjakan) oleh siswa. Standar akademis merefleksikan pengetahuan dan ketrampilan esensial setiap disiplin ilmu yang harus dipelajari oleh seluruh siswa. Sedangkan standar kompetensi menjelaskan sejauh mana siswa seharusnya menguasai permasalahan atau keterampilan (*"how good is good enough?"*). Standar kompetensi ditunjukkan dalam bentuk tugas atau kegiatan yang didemonstrasikan sebagai penerapan dari pengetahuan dan ketrampilan yang telah dipelajarinya. Standar kompetensi harus dicapai oleh setiap siswa pada suatu kurikulum, untuk siswa di Palangkaraya mungkin tidak perlu sama dengan standar kompetensi siswa di Bandar Lampung. Demikian juga, standar akademis dapat sama untuk seluruh siswa tetapi standar kompetensi siswa biasa dengan siswa luar biasa dapat berbeda. Standar akademis berfokus pada tujuan pokok-pokok

bahasan yang esensial dari suatu disiplin ilmu dengan mengacu pada pernyataan isi pengetahuan dan ketrampilan.

Standar yang belum (baca : baru akan) diterapkan dalam sistem pendidikan nasional adalah standar kompetensi. Sebagai contoh, salah satu tujuan pokok bahasan dalam kurikulum mata pelajaran IPA kelas V Sekolah Dasar, menyatakan "Siswa dapat mengembangkan kemampuan merancang dan membuat suatu benda yang dapat bekerja dengan menerapkan pengetahuannya tentang gaya dan energi". Pertanyaannya adalah, sejauh mana siswa harus dapat mengembangkan kemampuan merancang dan membuat benda tersebut? Benda yang bekerja dengan prinsip gaya dan energi, banyak sekali dari yang sederhana sampai dengan yang kompleks. Batasan kompetensi seperti ini yang belum dijelaskan secara jelas dan terbuka dalam system pendidikan kita.

Kaitan antara standar kompetensi dan tes kompetensi

Standar kompetensi merupakan standar yang menyediakan jaminan kualifikasi yang memadai tentang siswa dalam melakukan sesuatu. Standar kompetensi mencoba mendefinisikan hasil belajar minimum siswa dalam ketrampilan dan kecakapan berkenaan dengan satu set pengetahuan atau ketrampilan yang terdefinisi secara jelas. Untuk menentukan apakah siswa dapat mendemonstrasikan sejumlah pengetahuan atau ketrampilan minimal

yang dianggap dapat menerangkan kompetensinya, maka digunakan tes kompetensi.

Tes kompetensi merupakan usaha yang dilakukan untuk meyakinkan, bahwa siswa dapat menunjukkan penguasaannya tentang kemampuan tertentu yang dibutuhkan untuk melaksanakan tugas-tugas yang akan dihadapinya di masa depan. Adapun untuk tujuan mengukur standar kompetensi tertentu, dikembangkan tes kompetensi minimum. Melalui tes ini, dapat dijelaskan apakah masing-masing siswa sekurang-kurangnya telah memiliki "kompetensi minimal" pada saat lulus sekolah? Melalui tes ini berarti sudah berlangsung suatu mekanisme sertifikasi terhadap siswa dengan cara mengujinya yang menunjukkan bahwa ia menguasai kemampuan atau ketrampilan minimal tertentu untuk menerika ijazah/sertifikat. Sertifikasi menghendaki skor-skor tes yang didapat, dievaluasi terhadap standar kemampuan (*passing point* atau *cut score*) yang telah ditetapkan sebelumnya agar kompeten tidaknya siswa yang dites dapat diputuskan.

Penetapan Standar Kompetensi

Ada tiga metode atau cara penetapan standar kompetensi yang dibahas dalam makalah ini, yaitu metode Nedelsky, Angoff, dan Ebel. Ketiga cara tersebut didasarkan pada kosep bahwa ada kelompok siswa yang tergolong "border line" atau "pas-pasan" untuk lulus. Kelompok siswa ini adalah mereka

yang pengetahuan dan keterampilannya berada pada batas kelompok ssiwa yang mampu dan kelompok siswa yang kurang mampu. Sesuai dengan pikiran logis maka kelompok siswa yang pandai akan memperoleh skor yang lebih tinggi dan kelompok siswa yang kurang akan memperoleh skor yang lebih rendah dari pada kelompok siwa yang termasuk "pas-pasan untuk lulus". Nilai Batas Lulus (NBL), dengan demikian berada pada nilai yang diperoleh diantara kelompok siswa pas-pasan ini. Penilaian untuk NBL, pada metode ini dicari secara khusus dari pertanyaan-pertanyaan pada ujian itu, seberapa seharusnya kemampuan siswa mahasiswa "pas-pasan" itu menjawab pertanyaan-pertanyaan sesuai dengan perkiraan dari para penguji.

Metode ini sesungguhnya cukup mudah dapat dilaksanakan baik sebelum maupun sesudah pelaksanaan ujian. Lagi pula data-data yang diperlukan berupa penilaian dari pertanyaan ujian itu sendiri selalu dapat kita peroleh. Kesulitannya ialah untuk membuat perkiraan seberapa seharusnya kemampuan kelompok siswa "pas-pasan" itu untuk menjawab pertanyaan tersebut.

Langkah-langkah yang diambil untuk ketiga metode ini ada beberapa tahap, sebagai berikut :

1. Memilih para penilai atau juri
2. Membuat definisi tentang apa yang disebut pengetahuan dan keterampilan yang "pas-pasan".
3. Melatih penilai tentang metode ini
4. Mengumpulkan hasil penilaian

5. Menggabungkan hasil-hasil penilaian sehingga diperoleh NBL.

Juri atau para penilai yang dipilih hendaknya mereka yang mempunyai kualifikasi untuk menentukan tingkat pengetahuan dan ketrampilan yang diperlukan seperti yang dikehendaki oleh suatu ujian. Contoh kasus dikemukakan sebagai berikut : Bila ujian dimaksudkan untuk menguji asisten praktikum fisika, maka para penilai harus mengetahui dan dapat menentukan pengetahuan dan keterampilan yang diperlukan untuk dapat melakukan pekerjaan sebagai asisten praktikum, misalnya mengoperasikan alat-alat ukur mekanik dan alat ukur listrik, menerangkan cara kerja alat, mengetahui cara mengkalibrasi alat, mengetahui bahaya akibat alat ukur tersebut, dan cara-cara melindungi diri, dan lain-lainnya. Berkaitan dengan penentuan jumlah penilai atau juri, makin banyak sesungguhnya makin naik tetapi 5 orang sudah dianggap cukup. Kemudian masing-masing juri diminta untuk memberi definisi atau melukiskan kemampuan siswa yang "pas-pasan". Pekerjaan ini memang sulit, namun untuk memudahkan, misalnya ini dapat dimulai dengan pertanyaan :Apakah seorang asisten praktikum fisika yang "pas-pasan" harus mampu mengoperasikan alat ukur mekanik dan alat ukur listrik serta menerangkan cara kerjanya? Ataupun sudah cukup kalau dapat mengoperasikan alat saja dan seterusnya. Berikanlah waktu yang cukup untuk membuat definisi dan menentukan kemampuan mahasiswa yang "pas-

pasan" ini. Bila tidak dapat tercapai kesepakatan dan terdapat perbedaan yang tajam, proses dapat dilanjutkan saja. Bila para penilai mencapai kesepakatan mengenai definisi itu, tulislah dan lengkapi dengan beberapa contoh, sehingga terdapat pernyataan tertulis tentang standar yang harus dicapai.

Metode Nedelsky

Metode ini disarankan penggunaannya oleh Leo Nedelsky pada tahun 1954 dan hanya dilaksanakan pada ujian soal-soal pilihan ganda karena memerlukan penilaian tentang "Kemungkinan jawaban yang salah". Tugas para penilai ialah melihat pertanyaan dan menemukan jawaban yang benar atau jawaban-jawaban yang salah, yang oleh siswa "pas-pasan" seharusnya dapat dikenal sebagai jawaban yang salah atau bukan merupakan jawaban yang paling benar. Sebagai contoh perhatikanlah pertanyaan tentang gerak dan gaya, dimana para peserta tes (siswa) harus menjawab dengan jawaban yang paling benar.

Dibawah ini, merupakan contoh gerak benda yang setiap selang waktu tertentu kecepatannya berubah secara teratur dari lambat menjadi cepat adalah

- a. anak panah yang lepas dari busur
- b. buah durian jatuh
- c. mobil yang sedang direm
- d. roket yang diluncurkan ke angkasa
- e. jarum jam berputa

Dalam hal demikian maka penilai memutuskan bahwa ketiga jawaban c, d, dan e adalah jawaban yang salah dan siswa yang "pas-pasan" harus dapat menyisihkan ketiga jawaban itu. A, dan b mungkin dianggapnya sulit oleh penilai bagi siswa "pas-pasan" itu untuk memilih mana yang benar. Dalam hal ini maka penilai memutuskan bahwa ketiga jawaban c, d dan e adalah jawaban yang salah yang harus diketahui oleh siswa yang "pas-pasan".

Nilai batas lulus diambil dari peluang mereka untuk jawaban yang benar, yaitu dari dua jawaban (*option*) yang meragukan. Dalam hal ini kesepakatan mereka itu ialah satu dari dua option yang meragukan ($1/2$).

Demikian seterusnya pada soal pilihan berganda dengan 5 pilihan : Jika satu jawaban harus diketahui salah atau ditolak oleh mahasiswa yang "pas-pasan", maka batas lulus ($1/4$).

Jika dua jawaban harus ditolak, batas lulusnya $1/3$

Jika tiga jawaban harus ditolak, batas lulusnya $1/2$

Jika empat jawaban harus ditolak, batas lulusnya $1/1 = 1$

Para juri (penilai) mungkin berbeda pendapat tentang jawaban-jawaban mana yang harus ditolak oleh mahasiswa "pas-pasan" itu. Untuk diusahakan tercapainya kesepakatan. Akan tetapi kalau tidak tercapai, maka proses pun dapat jalan terus dengan tidak berakibat buruk.

Tabel 1. Perhitungan nilai batas lulus cara Nedelsky

Pertanyaan	Jawaban	Jumlah jawaban yang tidak ditolak	Nilai batas lulus butir
1	A, B, C, D, E	2	$\frac{1}{2} = 0,50$
2	A, B, C, D, E	1	$1/1 = 1,00$
3	A, B, C, D, E	2	$\frac{1}{2} = 0,50$
4	A, B, C, D, E	3	$1/3 = 0,33$
5	A, B, C, D, E	1	$1/1 = 1,00$
6	A, B, C, D, E	5	$1/5 = 0,20$
7	A, B, C, D, E	4	$1/4 = 0,25$
8	A, B, C, D, E	4	$1/4 = 0,25$
9	A, B, C, D, E	5	$1/5 = 0,20$
10	A, B, C, D, E	5	$1/5 = 0,20$
		Nilai batas lulus ujian = 4,43	Jumlah = 4,43

Untuk menentukan nilai batas lulus dari ujian tinggal menjumlahkan nilai batas lulus masing-masing soal (lihat tabel 1). Setiap penilai melakukan hal yang sama, sehingga kemudian akan memperoleh nilai batas lulus dari ujian

yang ditetapkan terakhir sebagai rata-rata jumlah nilai batas lulus yang diperoleh semua penilai. Angka rata-rata ini dapat digunakan nilai mean, median atau "trimmed mean" seperti terlihat dalam tabel 2.

Tabel 2. Menghitung nilai batas lulus rata-rata.

Penilai 1 (tertinggi)	92,50	Penilai 2 : 77,25
Penilai 2	77,25	Penilai 3 : 67,00
Penilai 3	67,00	Penilai 4 : <u>66,67</u>
Penilai 4	66,67	Jumlah : 210,92
Penilai 1 (terendah)	<u>65,33</u>	
Jumlah	368,75	

Mean = $368,75/5 = 73,75$

Media = 67,00

Trimmed mean = $210,92/3 = 70,3$

Untuk penghitungan "trimmed mean", penilaian dari penilai 1-2 orang teratas dan terbawah tidak diikutsertakan dan angka rata-rata dihitung dari penilai yang lain-lainnya.

Metode Angoff

Metode Angoff dikemukakan oleh William Angoff (Buckendahl, 2000). Pada metode ini nilai batas lulus untuk masing-masing pertanyaan diperkirakan dari kemungkinan jawaban salah masing-masing pertanyaan. Pada penilaian memperkirakan seberapa kemungkinan

kelompok siswa yang "pas-pasan" dapat menjawab soal dengan benar, atau kalau ada 100 siswa yang "pas-pasan" berapa orang kiranya yang dapat menjawab benar. Persentase yang menjawab benar adalah NBL dari soal itu dan ini berkisar antara 0,00 dan 1,00. NBL ujian adalah penjumlahan dari NBL butir-butir soal itu. Masing-masing penilai akan menghitung NB, ujian itu lalu dihitung rata-ratanya untuk memperoleh NBL yang akhir dari ujian tersebut, seperti terlihat dalam table 3 dan 4.

Tabel 3. Contoh format salah satu penilai akan menghitung NBL.

No. Butir Soal	Kemungkinan menjawab benar (%)
1	95
2	80
3	90
4	60
5	75
6	40
7	50
8	25
9	25
10	40
Jumlah	580

$$\text{Nilai batas lulus (NBL)} = 580/10 = 58$$

Tabel 4. Menghitung nilai batas metode Angoff

Penilai	NBL
1	58
2	75
3	78
4	72
5	77
Jumlah	360

$$\text{Rata-rata : } 360/5 = 72 \text{ (Nilai batas lulus ujian)}$$

Metode Ebel

Agak berbeda dengan metode terdahulu, metode Ebel dilaksanakan dalam 2 tahap. Soal-soal dikelompokkan ke dalam kelompok-kelompok sesuai dengan derajat kesukarannya, kemudian masing-masing dikelompokkan lagi menurut relevansinya atau penting tidaknya. Kesukaran soal dikelompokkan ke dalam 3 kategori, yaitu kelompok soal-

soal yang mudah, sedang, dan sukar. Berdasarkan relevansinya butir soal dikategorikan menjadi 4 kelompok : amat penting, penting, dapat diterima, dan meragukan. Sesudah dikelompokkan seperti itu, maka masing-masing penilai mengatakan penilaian terhadap butir-butir soal tersebut, seperti contoh pada tabel di bawah ini.

Tabel 5. Contoh pengelompokan butir soal pada metode Ebel

Relevansinya	Tk Kesukaran Mudah	Sedang	Sukar
Amat penting	1,4,7,8,13 Penilaian 95% benar	11, 15 dan 22 Penilaian 85% benar	21 Penilaian 80% benar
Penting	2,6 dan 9 Penilaian 90% benar	10, 14 dan 20 Penilaian 75% benar	16 dan 25 Penilaian 60% benar
Dapat diterima	5 Penilaian 80% benar	12 dan 18 Penilaian 55% benar	19 dan 23 Penilaian 35% benar
Meragukan	3 Penilaian 50% benar	Tidak ada	17 dan 24 Penilaian 95% benar

Penilaian terhadap butir soal tersebut berdasarkan perkiraan, yaitu bilamana seseorang mahasiswa yang

“pas-pasan” menjawab sejumlah butir soal yang demikian, berapa persen kira-kira ia benar.

Tabel 6. Menghitung nilai batas lulus metode Ebel

Kategori	Persentase benar	Jumlah soal	Jumlah skor nilai
Amat Penting			
- Mudah	95	5	$95 \times 5 = 475$
- Sedang	85	3	$85 \times 3 = 255$
- Sukar	80	1	$80 \times 1 = 80$
Penting			
- Mudah	90	3	$90 \times 3 = 270$
- Sedang	75	3	$75 \times 3 = 225$
- Sukar	60	2	$60 \times 2 = 120$
Dapat diterima			
- Mudah	80	1	$80 \times 1 = 80$
- Sedang	55	2	$55 \times 2 = 110$
- Sukar	35	2	$35 \times 2 = 70$
Meragukan			
- Mudah	50	1	$50 \times 1 = 50$
- Sedang	-	-	-
- Sukar	20	2	$20 \times 2 = 40$
Jumlah			

Selanjutnya untuk menentukan batas lulus terakhir, dihitung rata-rata dari penilaian masing-masing penilai (rater) seperti pada metode Nedelsky dan Angoff.

Kesimpulan

Berdasarkan pembahasan di atas dapat ditarik kesimpulan sebagai berikut :

1. Untuk keperluan sertifikasi, seharusnya skor-skor tes yang telah diperoleh siswa dievaluasi terhadap standar kemampuan (*passing point* atau *cut score*) yang telah ditetapkan sebelumnya agar kompeten tidaknya siswa yang dites dapat diputuskan. Dengan demikian, penetapan standar kemampuan penting dilakukan melalui pertimbangan atau penilaian para ahli bidang masing-masing.
2. Penentuan standar selalu didasarkan pada siswa yang berkemampuan "pas-pasan". Alasannya karena kemampuan mereka ini berada pada garis batas yang mencapai standar kompetensi dan yang tidak. Penentuan ini didasarkan pada teori tes klasik dimana sebaran skornya diasumsikan normal.
3. Tidak ada metode penetapan standar yang paling baik untuk semua keadaan. Pilihan metode tertentu harus didasarkan kepada penilaian jenis mana yang mungkin dikerjakan sesuai fasilitas yang ada dan keyakinan akan manfaatnya.

DAFTAR PUSTAKA

Buckendhal. 2000. "A comparison of the Angoff and Bookmark standardsetting methods". Paper presented at annual meeting of the Mid-Western educational Research Association in Chicago, IL. October, 2000.

Duarsa, N. Wirya. 1986. "Cara Pemberian Skor dan Penentuan Standar Keberhasilan Belajar" dalam Penilaian Keberhasilan Belajar dalam Bidang Kesehatan. Penyunting E. Sukardi dan W.F. Maramis. Surabaya : Airlangga University Press.

Kiplinger, Vonda L. 1997. *Standard Setting Procedures for the Specification of Performance Levels on a Standards-based Assessment*. <http://www.cde.state,2002>.

Linn, Robert L. 1989. *Educational Measurement*. 3rd edition. New York : MacMilan Publishing Company.

Anonim, 2000. *Comparison of the advantages and concerns regarding the modified Angoff method and the bookmark procedure*. <http://www.cde.state,2002>