

JADWAL ACARA SEMINAR SEHARI
“BEBAN KERJA GURU DAN MUTU HASIL BELAJAR”
Hotel Borobudur Jakarta, Kamis, 10 Desember 2009

WAKTU	ACARA	PEMBICARA	PEMBAHAS	MODERATOR	NOTULIS
Kamis, 10 Desember 2009					
Pk. 08.00 – 08.30	Registrasi dan Coffee Break	PANITIA			
Pk. 08.30 – 08.45	Pembukaan	Kepala Balitbang			
Pk. 08.45 – 09.45	Presentasi I dan Tanya Jawab: - Standard Setting	Dr. Samsul Hadi (UNY)	Prof. Dr. Dali S. Naga (UNJ)	Arniati PH., M.Psi.	Mira Josy M.
Pk. 09.45 – 11.15	Presentasi II dan Tanya Jawab: - Time on Task/Teachers Workload Study to Assess Performance of Teacher and Develop Instruments for Quantitative Impact Study	Prof. Dr. Suhardjono (Unibraw)	Prof. Dr. Azmawi Zainul (UPI)	A. Hendriastuti, M.A.	Enny Hariani
Pk. 11.15 – 12.45	Presentasi III dan Tanya Jawab: Analysis of The Determinants of Learning Outcomes Using Data from The International PISA	Prof. Dr. Arif Tiro (UNM)	Prof. Dr. Furqon (UPI)	Linda D. Gonggalang	Rumondang
Pk. 12.45 – 13.45	ISHOMA				
Pk. 13.45 – 15.45	Presentasi II dan Tanya Jawab: - Analysis of The Determinants of Learning Outcomes Using Data from The International PIRLS & TIMSS (Content Analysis)	- Prof. Dr. Nuryani Rustaman (UPI) - Prof. Dr. Hendra Gunawan ITB - Dr. Suhendra Yusuf (UNINUS)	- Prof. Dr. Muslimin (UNESA) - Prof. Dr. Budiyono (UNS) - Prof. Dr. Pujiati Suyata (UNY)	Dr. Bastari	Rahmawati
Pk. 15.45 – 16.00	Coffee Break				
Pk. 16.00 – 17.30	Presentasi III dan Tanya Jawab: Analysis of The Determinants of Learning Outcomes Using Data from The International PIRLS & TIMSS	Dr. Jahja Umar (UIN Syarif Hidayatullah)	Prof. Dr. IGN Agung (UI)	Dr. Nonny Swediati	Erika Afiani
Pk. 17.30 – 17.45	Rumusan hasil seminar dan PENUTUPAN	Kapuspendik			



DEPARTEMEN PENDIDIKAN NASIONAL
BADAN PENELITIAN DAN PENGEMBANGAN
PUSAT PENILAIAN PENDIDIKAN

SERTIFIKAT

diberikan kepada:

Dr. Samsul Hadi

Atas partisipasinya sebagai PENYAJI dalam :

Seminar Sehari
“Beban Kerja Guru dan Mutu Hasil Belajar”

Yang dilaksanakan pada tanggal 10 Desember 2009
di Hotel Borobudur – Jakarta

Jakarta, 10 Desember 2009

Kepala Pusat Penilaian Pendidikan,



Dr. Nugan Yulia Wardhani S., M.Psi.

NIP. 19560724 198303 2 001

**MATERI SEMINAR SEHARI
BEBAN KERJA GURU DAN MUTU HASIL BELAJAR**

NO.	M A T E R I
1.	Standard Setting
2.	Time on Task/Teachers Workload Study to Assess Performance of Teacher and Develop Instruments for Quantitative Impact Study
3.	Analysis of The Determinants of Learning Outcomes Using Data from The International PISA
4.	Analysis of The Determinants of Learning Outcomes Using Data from The International PIRLS & TIMSS (Content Analysis)
5.	Analysis of The Determinants of Learning Outcomes Using Data from The International PIRLS & TIMSS



STANDARD SETTING

Oleh :
Samsul Hadi

**PUSAT PENILAIAN PENDIDIKAN
BADAN PENELITIAN DAN PENGEMBANGAN
DEPARTEMEN PENDIDIKAN NASIONAL
2009**

STANDARD SETTING*)

Samsul Hadi

Abstrak

Studi ini bertujuan untuk: menentukan kriteria kelulusan, yaitu yang berupa skor dan deskripsi kemampuan peserta didik pada mata pelajaran Matematika IPA, Matematika IPS, Bahasa Indonesia, Bahasa Inggris, Fisika, dan Ekonomi yang diujikan secara nasional.

Kegiatan ini dilakukan di Universitas Negeri Yogyakarta (UNY) dengan panelis dari 4 propinsi, yaitu Bengkulu, Daerah Istimewa Yogyakarta, Jawa Tengah, dan Nusa Tenggara Timur. Panelis tersebut merupakan guru SMA/MA program IPA/IPS yang mengajar Matematika, Bahasa Indonesia, Bahasa Inggris, Fisika, dan Ekonomi yang berpengalaman mengajar dalam bidang studi yang diteliti minimum 10 tahun dan mengajar kelas 12 minimal selama 5 tahun. Rata-rata jumlah guru setiap mata pelajaran adalah 11 orang dan tiap propinsi 2 orang, kecuali Yogyakarta 5 orang. Mereka mewakili mewakili berbagai kualitas sekolah. Dalam setiap kelompok mata pelajaran ada ketua kelompok yaitu dosen UNY yang minimal berijazah S2, pengalaman minimal 5 tahun, berasal dari program studi yang sama dengan program studi panelis yang dipimpinnya, dan telah mengikuti pelatihan *standard setting*.

Teknik analisis data yang digunakan dalam penelitian ini adalah prosedur Angoff dan *Bookmark*. Pada metode Angoff, teknik analisis data yang digunakan adalah menghitung rata-rata jumlah butir yang diasumsikan dapat oleh panelis dikerjakan oleh peserta tes. Selain itu juga menghitung persentase jumlah dari butir yang diperkirakan dapat dikerjakan oleh peserta tes. Pada metode *Bookmark*, teknik analisis yang digunakan adalah mengestimasi parameter butir menurut teori respons butir dengan rumus 1 parameter logistik (1PL), mengurutkan butir berdasarkan tingkat kesulitan, dan mengestimasi kemampuan peserta tes untuk dapat menjawab soal dengan probabilitas 2/3. Selain itu juga menghitung rata-rata estimasi kemampuan peserta yang menurut panelis merupakan batas kelulusan peserta tes (*cut score*).

Hasil penelitian menemukan bahwa: 1) Batas kelulusan dari 6 mata pelajaran berbeda-beda. Berdasarkan metode Angoff rerata KKM dari dua putaran untuk mata pelajaran Fisika, Ekonomi, Bahasa Inggris, Matematika IPA, Matematika IPS, dan Bahasa Indonesia masing-masing adalah: 57,10%; 61,60%; 65,4%; 58,44%; 53,43%; dan 74,25% jumlah soal benar. Berdasarkan metode *Bookmark* rerata KKM dari dua putaran untuk mata pelajaran Fisika, Ekonomi, Bahasa Inggris, Matematika IPA, Matematika IPS, dan Bahasa Indonesia masing-masing adalah: 29,566; 35,115; 29,260; 37,075; 41,220; dan 48,768. 2) Penelitian ini juga menemukan deskriptor kemampuan lulusan 6 mata pelajaran tersebut.

Kata Kunci: *standard setting*, metode Angoff, metode *Bookmark*.

I. PENDAHULUAN

A. Latar Belakang

Salah satu kegiatan yang penting dalam meningkatkan kualitas pendidikan adalah penilaian. Dalam banyak hal, hasil penilaian sering dipandang sebagai tolok ukur keberhasilan proses pembelajaran. Hampir pada setiap sistem ujian, hasil penilaian dikategorikan menjadi dua, yaitu lulus dan tidak lulus. Lulus berarti peserta didik telah memenuhi kriteria tertentu, sedang tidak lulus berarti belum memenuhi kriteria tertentu.

Kriteria kelulusan atau batas kelulusan memegang peran penting dalam meningkatkan kualitas pendidikan. Peserta didik yang dinyatakan lulus harus disertai dengan penjelasan kemampuan apa saja yang telah dimiliki peserta didik tersebut untuk semua mata pelajaran. Jadi kriteria kelulusan harus mencakup skor dan kemampuan yang dicapai. Selama ini prosedur yang digunakan adalah melalui *judgment* yang tidak berdasarkan data empirik dan prosedur yang telah teruji. Untuk itu perlu dicari cara yang dapat dipertanggungjawabkan dari teori pengukuran.

B. Tujuan Studi

1. Mengetahui batas kelulusan peserta didik untuk jenjang pendidikan SMA dan MA pada mata pelajaran Matematika IPA, Matematika IPS, Bahasa Indonesia, Bahasa Inggris, Fisika, dan Ekonomi.
2. Mengetahui kemampuan yang harus dimiliki peserta didik untuk dinyatakan lulus pada mata pelajaran Matematika IPA, Matematika IPS, Bahasa Indonesia, Bahasa Inggris, Fisika, dan Ekonomi.

II. METODE PENELITIAN

A. Tempat dan Waktu Penelitian

Kegiatan ini dilakukan di Universitas Negeri Yogyakarta (UNY) dengan sampel peserta dari 4 propinsi, yaitu Bengkulu, Daerah Istimewa Yogyakarta, Jawa Tengah, dan Nusa Tenggara Timur.

B. Populasi dan Sampel

Panelis kegiatan ini adalah guru SMA/MA program IPA/IPS yang mengajar Matematika, Bahasa Indonesia, Bahasa Inggris, Fisika, dan Ekonomi yang berpengalaman mengajar dalam bidang studi yang diteliti minimum 10 tahun dan mengajar kelas 12 pada mata pelajaran tersebut minimal selama 5 tahun. Jumlah guru setiap mata pelajaran adalah 11 orang yang diambil dari propinsi Bengkulu, Jawa Tengah, D.I. Yogyakarta, dan Nusa Tenggara Timur yang mewakili berbagai kualitas sekolah. Keempat propinsi tersebut dipilih karena pelaksanaan UN-nya dianggap bersih

dari kecurangan. Proporsi guru tiap mata pelajaran dilihat dari keempat propinsi tersebut ditunjukkan pada Tabel 1.

Dalam setiap kelompok mata pelajaran ada ketua kelompok yaitu dosen UNY yang minimal berijazah S2, pengalaman minimal 5 tahun, berasal dari program studi yang sama dengan program studi panelis yang dipimpinnya, dan telah mengikuti pelatihan *standard setting*.

Tabel 1. Proporsi Sampel Tiap Mata Pelajaran

Provinsi	Banyaknya Guru	Sekolah		Jumlah
		Prestasi	Asal	
Bengkulu	1	Sedang	Kota	2
	1	Rendah	Pinggiran/Desa	
NTT	1	Sedang	Pinggiran/Desa	2
	1	Rendah	Kota	
Jawa Tengah	1	Sedang	Kota	2
	1	Rendah	Pinggiran/Desa	
DI Yogyakarta	1	Sedang	Kota	5
	1	Rendah	Pinggiran/Desa	
	1	Sedang	Pinggiran/Desa	
	1	Sedang	Kota	
	1	Sedang	Pinggiran/Desa	
Total Guru Tiap Mata Pelajaran =				11

C. Metode Pengumpulan Data dan Instrumen

Metode yang digunakan untuk mengumpulkan data dalam penelitian ini adalah dokumentasi, wawancara, dan survei. Selaras dengan metode pengumpulan datanya, instrumen yang digunakan adalah daftar dokumentasi, pedoman wawancara, dan kuesioner.

D. Teknik Analisis Data

Teknik analisis data yang digunakan dalam penelitian ini adalah prosedur Angoff dan *Bookmark* pada mata pelajaran Bahasa Indonesia, Bahasa Inggris, Matematika, dan Fisika untuk program Ilmu Pengetahuan Alam (IPA) dan Bahasa Indonesia, Bahasa Inggris, Matematika, dan Ekonomi untuk program Ilmu Pengetahuan Sosial.

Pada metode Angoff, teknik analisis data yang digunakan adalah menghitung rata-rata jumlah butir yang diasumsikan dapat oleh panelis dikerjakan oleh peserta tes. Selain itu juga menghitung persentase jumlah dari butir yang diperkirakan dapat dikerjakan oleh peserta tes.

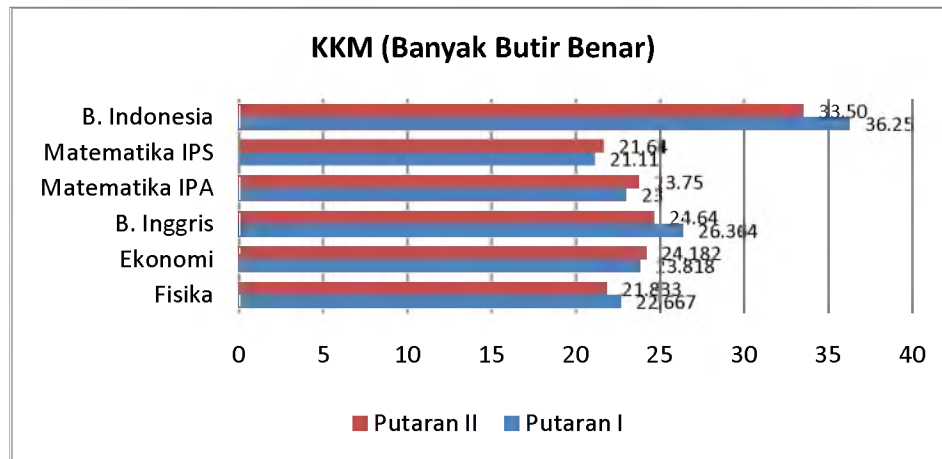
Pada metode *Bookmark*, teknik analisis yang digunakan adalah mengestimasi parameter butir, mengurutkan butir berdasarkan tingkat kesulitan, dan mengestimasi kemampuan peserta tes untuk dapat menjawab soal dengan probabilitas 2/3. Selain itu juga menghitung rata-rata estimasi kemampuan peserta yang menurut panelis merupakan batas kelulusan peserta tes (*cut score*).

III. HASIL PENELITIAN DAN PEMBAHASAN

A. Batas Kelulusan

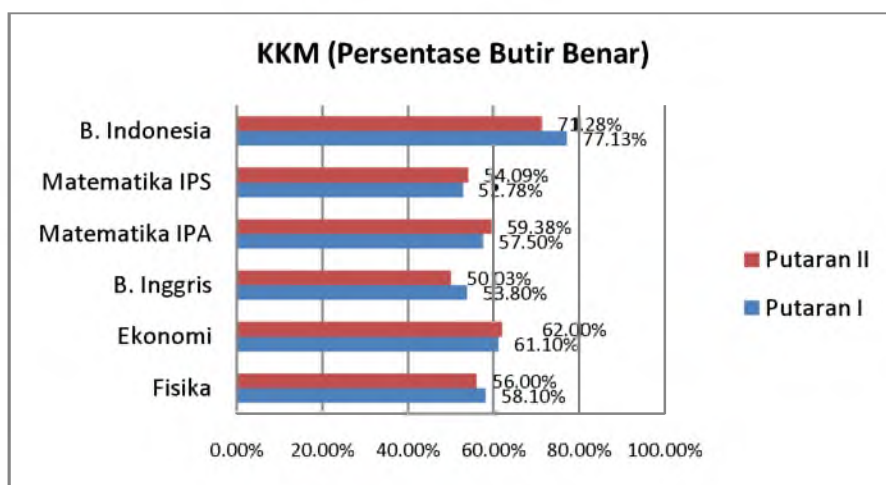
1. Metode Angoff

Rangkuman Kriteria Kelulusan Minimal (KKM) berdasarkan banyaknya butir benar hasil standard setting dengan metode Angoff dari 6 mata pelajaran ditunjukkan pada Gambar 1. Berdasarkan gambar tersebut, pada putaran pertama dan kedua Bahasa Indonesia memiliki KKM yang paling tinggi disusul dengan Bahasa Inggris, Ekonomi, Matematika IPA, Fisika, dan Matematika IPS. Tetapi urutan tinggi KKM dari keenam mata pelajaran tersebut sebetulnya tidak dapat langsung diperbandingkan. Hal ini disebabkan adanya jumlah butir soal UN yang berbeda dari tiap mata pelajaran.



Gambar 1. KKM (Banyak Butir Benar) Hasil Metode Angoff

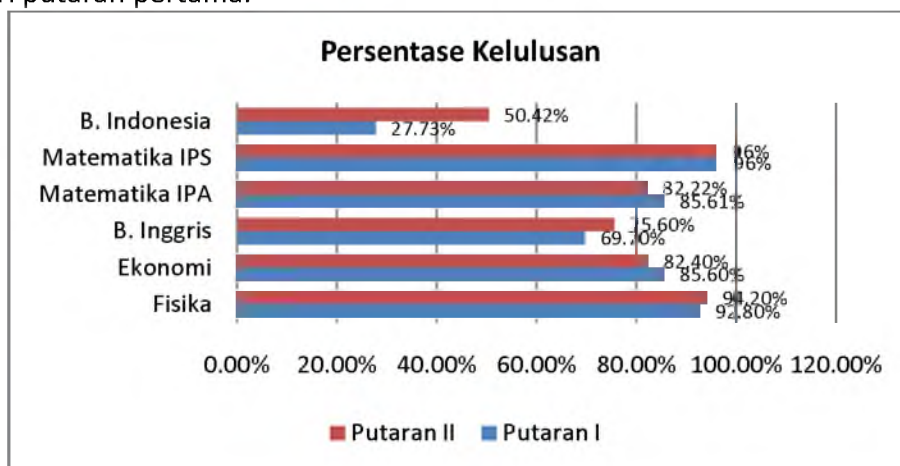
Agar KKM dari keenam mata pelajaran tersebut dapat dibandingkan, pada penelitian ini dihitung KKM yang didasarkan pada persentase butir benar. KKM dari keenam mata pelajaran yang dihitung berdasarkan persen butir benar ditunjukkan pada Gambar 2. Berdasarkan cara ini pada putaran pertama ternyata KKM Bahasa Indonesia tetap yang paling tinggi disusul Ekonomi, Fisika, Matematika IPA, Matematika IPS, dan terakhir Bahasa Inggris.



Gambar 2. KKM (Persentase Butir Benar) Hasil Metode Angoff

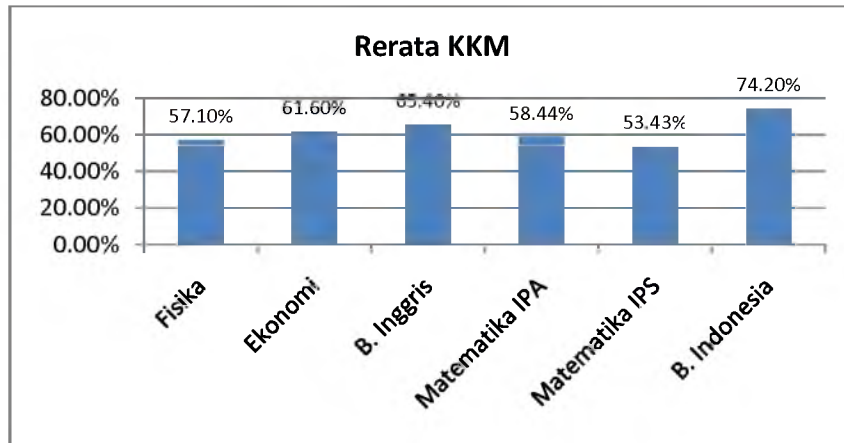
Pada putaran kedua KKM tertinggi adalah Bahasa Indonesia, disusul Ekonomi, Matematika IPA, Fisika, Matematika IPS, dan terakhir Bahasa Inggris. Perubahan urutan ini mungkin disebabkan para panelis telah mengetahui dampak penilaian butir soal yang telah ia lakukan terhadap persentase kelulusan peserta ujian seperti yang ditunjukkan pada Gambar 3. Panelis yang menginginkan banyak siswa yang lulus akan mengurangi jumlah butir soal yang harus dapat dikerjakan oleh peserta ujian.

Namun tentunya tidak semua panelis melakukan hal tersebut, karena pengurangan butir soal yang benar berarti juga pengurangan deskriptor. Selain itu, proses penilaian soal dilakukan secara individual, sehingga dampak tindakan seorang panelis hanya dapat diketahui setelah diskusi kelompok yang memaparkan KKM berdasar soal yang benar, KKM berdasar persentase soal benar, dan persentase kelulusan. Bukti tentang hal ini ditunjukkan pada Gambar 4, yaitu pada putaran kedua ternyata 3 dari 6 mata pelajaran pada putaran kedua persentase soal benarnya lebih tinggi dari putaran pertama.



Gambar 3. Persentase Kelulusan Hasil Metode Angoff

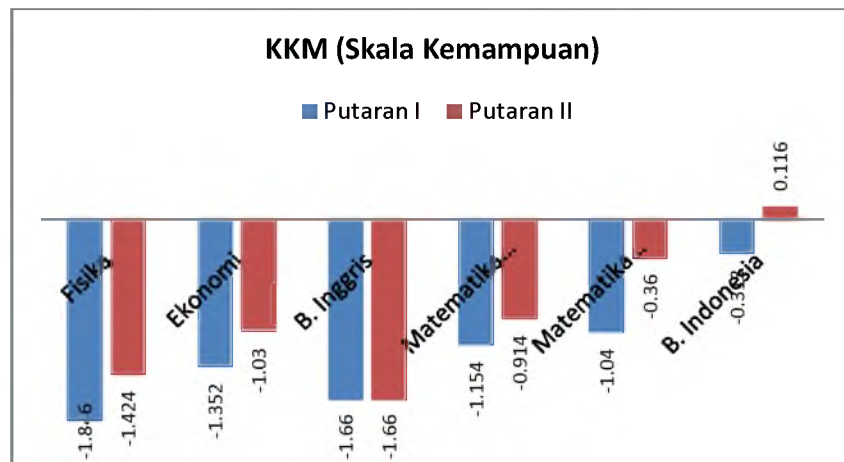
Rerata KKM standart setting 2 putaran metode Angoff ditunjukkan pada Gambar 4. Terlihat pada gambar tersebut rerata KKM-nya semua di atas 50% soal benar. Rerata KKM 6 mata pelajaran tersebut semua berada di atas batas kelulusan 42,5 skala 100 atau 4,25 skala 10 dari Keputusan BSNP No 1512/BSNP/XII/2008.



Gambar 4. Rerata KKM Hasil Metode Angoff Dua Putaran

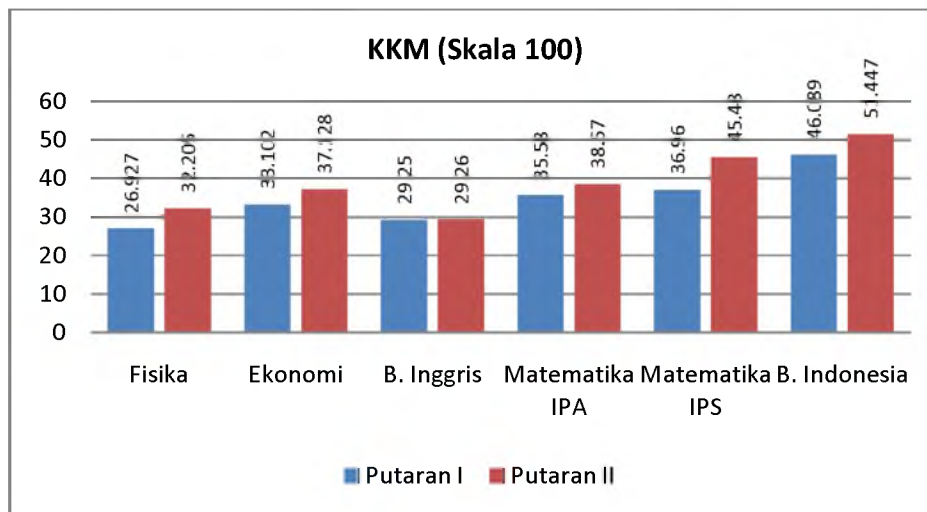
2. Metode Bookmark

Rangkuman KKM berdasarkan kemampuan peserta ujian hasil standard setting dengan metode Bookmark dari 6 mata pelajaran dalam skor baku ditunjukkan pada Gambar 5. Gambar tersebut menunjukkan bahwa pada putaran pertama Bahasa Indonesia memiliki KKM yang paling tinggi, disusul dengan Matematika IPS, Matematika IPA, Ekonomi, Bahasa Inggris, dan Fisika. Pada putaran kedua Bahasa Indonesia tetap memiliki KKM yang paling tinggi, disusul dengan Matematika IPS, Matematika IPA, Ekonomi, Fisika, dan Bahasa Inggris.

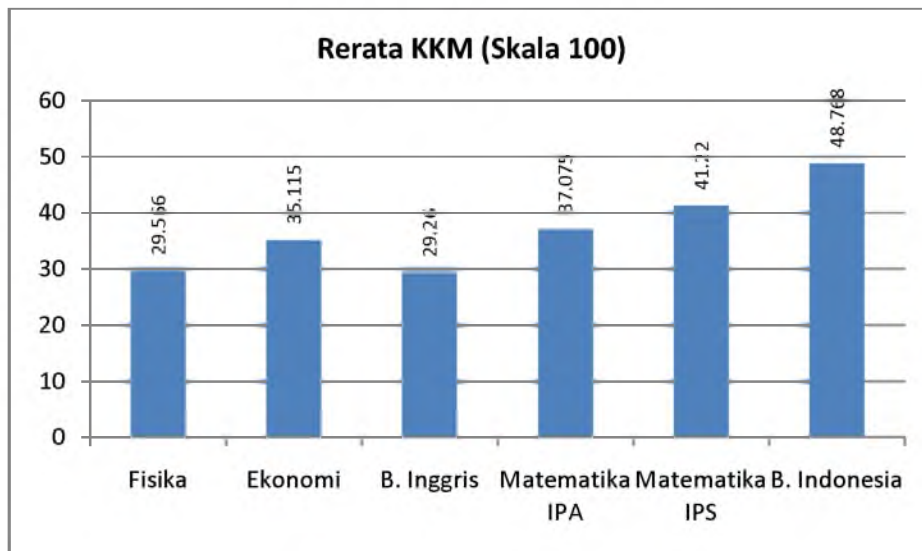


Gambar 5. KKM (Skala Kemampuan) Hasil Metode Bookmark

Rangkuman KKM berdasarkan kemampuan peserta ujian hasil standard setting dengan metode Bookmark dari 6 mata pelajaran dalam skala 100 ditunjukkan pada Gambar 6. Identik dengan Gambar 5, Gambar 6 menunjukkan bahwa pada putaran pertama Bahasa Indonesia memiliki KKM yang paling tinggi, disusul dengan Matematika IPS, Matematika IPA, Ekonomi, Bahasa Inggris, dan Fisika. Pada putaran kedua Bahasa Indonesia tetap memiliki KKM yang paling tinggi, disusul dengan Matematika IPS, Matematika IPA, Ekonomi, Fisika, dan Bahasa Inggris. Rerata KKM dalam skala 100 dari keenam mata pelajaran tersebut dari dua putaran metode Bookmark ditunjukkan pada Gambar 7.

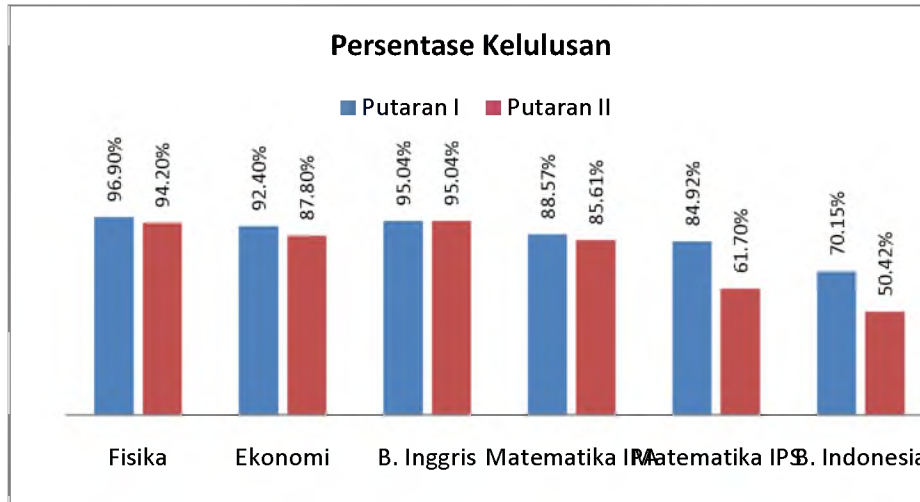


Gambar 6. KKM (Skala 100) Hasil Metode Bookmark



Gambar 7. Rerata KKM (Skala 100) Hasil Metode Bookmark Dua Putaran

Semakin tinggi KKM menyebabkan semakin rendahnya tingkat kelulusan. Hal ini dapat dilihat pada Gambar 8. Pada gambar tersebut terlihat bahwa pada putaran pertama persentase kelulusan Fisika paling tinggi, disusul Bahasa Inggris, Ekonomi, Matematika IPA, Matematika IPS, dan Bahasa Indonesia. Pada putaran kedua Bahasa Inggris paling tinggi, disusul Fisika, Ekonomi, Matematika IPA, Matematika IPS, dan Bahasa Indonesia.



Gambar 8. Persentase Kelulusan Hasil Metode Bookmark

Rendahnya KKM yang ditentukan dengan *bookmark* disebabkan karena beberapa hal, yang merupakan kelemahan dari metode *bookmark*. Penentuan metode *bookmark* dilakukan terlebih dahulu dengan mengestimasi parameter butir, urutan tingkat kesulitan hasil estimasi berbeda dibandingkan dengan urutan tingkat kesulitan berdasarkan pengalaman panelis. Hal ini menyebabkan panelis berhenti dan membuat *bookmark* pada butir yang siswa tidak bisa mengerjakan, padahal butir selanjutnya ada yang bisa dikerjakan siswa. Dengan kata lain, tingkat kesulitan butir antar daerah berbeda-beda, sehingga hasil kalibrasi parameter kesulitan menggunakan keseluruhan siswa berbeda dengan tingkat kesulitan menurut panelis.

Rendahnya *cut score* dengan metode *bookmark* dapat disebabkan pula oleh rendahnya tingkat kesulitan butir. Berdasarkan hasil analisis dengan model 1 parameter, diperoleh bahwa rerata tingkat kesulitan mata pelajaran Bahasa Inggris sebesar -1,850, Fisika sebesar -1,807, Matematika IPA sebesar -1,371, dan Matematika IPS sebesar -2.039. Tingkat kesulitan ini mempengaruhi kemampuan yang diestimasi pada saat peluang menjawab benar siswa sebesar 0,67. Semakin rendah tingkat kesulitan, semakin rendah pula kemampuan pada saat proporsi kemampuan sebesar 0,67.

C. Kemampuan Lulusan

Deskriptor kemampuan lulusan mata pelajaran Fisika yang paling banyak disebut panelis adalah a) membaca pengukuran alat ukur panjang (jangka sorong) dan menentukan ketelitiannya, b) menentukan frekuensi bunyi bila ada sumber bunyi yang bergerak dan diam,

dan c) menyebutkan faktor yang mempengaruhi energi foton berdasarkan teori kuantum Planck. Pada pelajaran Ekonomi, deskriptor tersebut adalah: a) menyebutkan sumber-sumber pendapatan asli daerah, b) mencatat transaksi-transaksi perusahaan jasa ke dalam jurnal umum, dan c) menyebutkan usaha-usaha yang dilakukan wirausaha dalam memperoleh keuntungan.

Pada pelajaran Bahasa Inggris, deskriptor kemampuan lulusan yang paling banyak disebut panelis adalah: a) menentukan informasi rinci tersurat *announcement*, b) menentukan informasi rinci tersurat tentang *advertisement*, dan c) menentukan rujukan kata dalam teks *message*. Deskriptor kemampuan lulusan Matematika IPA yang paling banyak disebut panelis adalah: a) menentukan negasi pernyataan yang diperoleh dari penarikan kesimpulan, b) menggunakan aturan pangkat, akar, dan logaritma, c) menentukan persamaan kuadrat baru, d) menentukan ukuran pemusatan data, e) menghitung peluang satu kejadian, f) menentukan komposisi dua fungsi, g) menghitung limit fungsi aljabar, h) menyelesaikan masalah sistem persamaan linier, dan i) menyelesaikan masalah program linier.

Deskriptor kemampuan lulusan mata pelajaran Matematika IPS yang banyak disebut panelis adalah: a) menentukan penyelesaian SPL 2 variabel, b) menyelesaikan masalah matriks yang berkaitan dengan determinan, c) menentukan nilai kebenaran suatu pernyataan majemuk, d) menentukan kesimpulan dari beberapa premis, e) menyelesaikan masalah yang berkaitan dengan pencacahan, f) menyelesaikan masalah yang berkaitan dengan permutasi, g) menentukan nilai peluang dan frekuensi harapan suatu kejadian, h) menentukan unsur-unsur diagram lingkaran, i) menghitung nilai ukuran pemusatan dari data dalam bentuk tabel atau diagram, dan menentukan ukuran penyebaran. Pada Bahasa Indonesia deskriptor tersebut adalah: a) menentukan isi & bagian paragraf, b) menentukan unsur paragraf, c) menentukan unsur paragraf, d) menentukan kata penghubung dalam paragraf, dan e) melengkapi paragraf sebab akibat.

Mencermati deskriptor masing-masing mata pelajaran, baik dengan menggunakan metode Angoff maupu Bookmark, dapat diketahui kesulitan-kesulitan siswa dalam mengerjakan butir-butir soal masing-masing pelajaran. Kesulitan-kesulitan ini mengindikasikan berbagai kendala selama pembelajaran. Kesulitan pada masing-masing mata pelajaran disajikan berikut.

Kesulitan siswa dalam mengerjakan fisika terkait dengan memahami permasalahan dari butir soal dan dituangkan ke rumus fisika. Penerapan perhitungan matematis untuk menyelesaikan masalah juga menghambat siswa untuk menjawab benar butir soal. Membaca grafik, tabel, diagram, dan juga memaknai soal terkait dengan fisika juga menjadi kendala siswa dalam mengerjakan butir-butir fisika UN.

Pada mata pelajaran ekonomi kesulitan yang dialami siswa ada beberapa hal. Karena terbatasnya kemampuan dalam konsep matematika siswa kesulitan dalam mengaplikasikan matematika ekonomi dalam pemecahan masalah. Penerapan dan implementasi konsep dan teori ekonomi dalam kehidupan juga merupakan hal yang sulit. Juga terkait dengan soal yang memerlukan berbagai hubungan logika untuk menyelesaikan masalah, dan panelis pun juga mengalami kesulitan. Siswa juga sering terjebak dengan alternatif jawaban yang mirip, meskipun jawaban pada butir hanya ada satu. Ketelitian dalam persoalan akuntansi, misal penjumlahan dalam membuat neraca rugi-laba, juga dialami oleh siswa.

Pada pelajaran bahasa Inggris, permasalahan besar yang dialami siswa yakni mengetahui makna yang tersirat. Makna yang tersirat yang sulit diketahui siswa ini terjadi tidak hanya pada *reading*, namun juga pada *listening* dan memahami iklan dan

lain-lain. Kesulitan siswa juga pada butir-butir *listening*. Hal ini bisa dipahami, karena panelis sendiri mengakui bahwa media pembelajaran untuk *listening* terbatas.

Kesulitan siswa pada pembelajaran matematika, baik IPA ataupun IPS lebih banyak disebabkan karena kurangnya pemahaman konsep. Kurangnya pemahaman konsep ini menyebabkan siswa akan bingung jika pertanyaan tidak lazim atau berbeda dari yang biasanya. Interpretasi soal yang berbentuk cerita, misalnya pada program linear bagi siswa juga merupakan hal yang sulit. Langkah dalam penyelesaian soal yang terlalu banyak menyebabkan siswa bingung terhadap apa yang akan dilakukan terlebih dahulu.

Pada pelajaran bahasa Indonesia, kesulitan siswa ketika mengerjakan soal bahasa Indonesia terkait dengan tiga hal. Siswa perlu waktu yang lama untuk membaca tabel, grafik, diagram dan sejenisnya. Dalam memahami grafik pada soal, ketelatenan siswa teruji. Membaca dan memaknai puisi juga merupakan hal yang sulit, karena interpretasi puisi yang berbeda-beda. Membaca paragraf panjang, siswa juga enggan membaca, sehingga perlu waktu lama untuk memahami bacaan.

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

1. Batas kelulusan dari 6 mata pelajaran berbeda-beda. Berdasarkan metode Angoff rerata KKM dari dua putaran untuk mata pelajaran Fisika, Ekonomi, Bahasa Inggris, Matematika IPA, Matematika IPS, dan Bahasa Indonesia masing-masing adalah: 57,10%; 61,60%; 65,4%; 58,44%; 53,43%; dan 74,25% jumlah soal benar. Berdasarkan metode Bookmark rerata KKM dari dua putaran untuk mata pelajaran Fisika, Ekonomi, Bahasa Inggris, Matematika IPA, Matematika IPS, dan Bahasa Indonesia masing-masing adalah: 29,566; 35,115; 29,260; 37,075; 41,220; dan 48,768.
2. Kemampuan lulusan mata pelajaran Fisika yang banyak disebut panelis adalah: a) membaca pengukuran alat ukur panjang (jangka sorong) dan menentukan ketelitiannya, b) menentukan frekuensi bunyi bila ada sumber bunyi yang bergerak dan diam, dan c) menyebutkan faktor yang mempengaruhi energi foton berdasarkan teori kuantum Planck. Pada pelajaran Ekonomi, kemampuan tersebut adalah: a) menyebutkan sumber-sumber pendapatan asli daerah, b) mencatat transaksi-transaksi perusahaan jasa ke dalam jurnal umum, dan c) menyebutkan usaha-usaha yang dilakukan wirausaha dalam memperoleh keuntungan.
Pada pelajaran Bahasa Inggris, kemampuan tersebut adalah: a) menentukan informasi rinci tersurat *announcement*, b) menentukan informasi rinci tersurat tentang *advertisement*, dan c) menentukan rujukan kata dalam teks *message*. Kemampuan lulusan pelajaran Matematika IPA yang paling banyak disebut panelis adalah: a) menentukan negasi pernyataan yang diperoleh dari penarikan kesimpulan, b) menggunakan aturan pangkat, akar, dan logaritma, c) menentukan persamaan kuadrat baru, d) menentukan ukuran pemusatan data, e) menghitung peluang satu kejadian, f) menentukan komposisi dua

fungsi, g) menghitung limit fungsi aljabar, h) menyelesaikan masalah sistem persamaan linier, dan i) menyelesaikan masalah program linier.

Kemampuan lulusan Matematika IPS adalah: a) menentukan penyelesaian SPL 2 variabel, b) menyelesaikan masalah matriks yang berkaitan dengan determinan, c) menentukan nilai kebenaran suatu pernyataan majemuk, d) menentukan kesimpulan dari beberapa premis, e) menyelesaikan masalah yang berkaitan dengan pencacahan, f) menyelesaikan masalah yang berkaitan dengan permutasi, g) menentukan nilai peluang dan frekuensi harapan suatu kejadian, h) menentukan unsur-unsur diagram lingkaran, i) menghitung nilai ukuran pemusatan dari data dalam bentuk tabel atau diagram, dan menentukan ukuran penyebaran. Kemampuan lulusan Bahasa Indonesia kemampuan tersebut adalah: a) menentukan isi & bagian paragraf, b) menentukan unsur paragraf, c) menentukan unsur paragraf, d) menentukan kata penghubung dalam paragraf, dan e) melengkapi paragraf sebab akibat.

B. Saran

1. Perlu menggunakan batas kelulusan yang berbeda untuk mata pelajaran yang berbeda dalam UN.
2. Perlu mengembangkan berbagai upaya untuk mencapai kemampuan lulusan seperti yang tertuang dalam SKL. Upaya tersebut antara lain: menambah variasi metode pembelajaran, meningkatkan rasa suka membaca, menggunakan media listening dalam pelajaran Bahasa Inggris, dan meningkatkan kemampuan listening bagi gurunya.
3. Soal-soal yang ada di UN perlu disusun dengan tingkat kesulitan yang memadai bagi siswa.

C. Keterbatasan

1. Jumlah panelis terbatas dengan sebaran hanya 3 atau 4 provinsi pada setiap pelajaran dan panelis lebih banyak berasal dari guru sekolah yang prestasinya sedang atau rendah.
2. Perlunya penelitian lebih lanjut tentang penentuan standard setting, dan dalam skala yang lebih luas.
3. Mengingat ada beberapa tingkatan kemampuan siswa di sekolah dan ada berbagai level rerata kemampuan sekolah, perlu adanya penelitian standard setting dengan pendekatan multilevel.

DAFTAR PUSTAKA

Angoff, W. H. (1971). Scale, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508-600). Washington, DC: American Council on Education.

- Anonim. 2009. *Standard Setting*. Diambil dari <http://bute-induction.st-andrews.ac.uk/> pada tanggal 16 September 2009.
- Depdikbud. (1998). *Laporan Hasil Evaluasi Pelaksanaan Program Direktorat Jenderal Pendidikan Dasar dan Menengah tahun 1994/1995 s.d. 1997/1998*. Jakarta: Depdikbud.
- Djemari Mardapi dkk. (1999). *Survei Kegiatan Guru dalam Melakukan Penilaian di Kelas. Laporan Penelitian*. Yogyakarta: Pusbangsisjian Lemlit IKIP Yogyakarta dan Pusiasjian balitbang Depdikbud.
- Djemari Mardapi, dkk,. (2001). *Sistem ujian akhir dalam otonomi daerah. Laporan penelitian*. Yogyakarta: Universitas Negeri Yogyakarta.
- Ebel, Robert L. (1972). *Essentials of educational measurement*. Englewood Cliffs: Prentice-Hall.
- Glas, C.A.W. et al.(1997).A Step Model to Analyze Partial Credit. In Hambleton, K.R & vander Linden W.J. (eds, 1997), *Handbook of Modern Item Response Theory*. New York: Springer.
- Goodwin, L. D. (1999). Relations between observed item difficulty levels and Angoff minimum passing levels for a group of borderline examinees. *Applied Measurement in Education*, 12, 13-28.
- Hambleton, K.R. et al. (1991). *Fundamentals of Item Response Theory*. New York: Sage Publications.
- Hattie, J.A., & Brown, G. T. L. (2003, August). *Standard setting for asTTle reading: A comparison of methods*. asTTle Technical Report #21, University of Auckland/Ministry of Education.
- Impara James C. & Plake Barbara S. (2000). A Comparison of Cut Scores using Multiple Standard Setting Methods. *Paper presented at the Large Scale Assessment Conference*. Snowbird, UT June, 2000. at [http://www.unl.edu/BIACO/coop/lsac/aeramillardsimpfinal .pdf](http://www.unl.edu/BIACO/coop/lsac/aeramillardsimpfinal.pdf) diambil tanggal 27 November 2005
- Impara, J.C., & Plake, B.S. (1997). Standard-setting: An alternative approach. *Journal of Educational Measurement*, 34, 353–366.
- Impara, J.C., Barbara S Plake. 1998. *Teachers' Ability to Estimate Item Difficulty: A Test of the Assumptions in the Angoff Standard Setting Method*. University of Nebraska-Lincoln. *Journal of Educational Measurement* Spring, 1998. Vol.35, No. 1
- Impara, J.C., Barbara S Plake. 2000. *A Comparison of Cut Scores using Multiple Standard Setting Methods*, Universitas Nebraska- Lincoln, Paper presented at the Large Scale Assessment Conference. Snowbird, UT, June, 2000.
- Jaeger, R. M. (1989). Certification of student competence. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 485–514). New York: American Council on Education/Macmillan.

- Jaeger, R. M. (1991). *Selection of judges for standard-setting*. *Educational Measurement: Issues and Practice*, 10(2), 3-6, 10.
- Lin, J. (tth). *The Bookmark Standard Setting Procedure: Strengths and Weaknesses*. Alberta: The Centre for Research in Applied Measurement and Evaluation, The University of Alberta. Diambil pada tanggal 30 Oktober 2009 dari http://www.education.ualberta.ca/educ/psych/crame/files/standard_setting.pdf
- Lin, Jie. *The Book Standard Setting Procedure: Strengths and Weaknesses*. http://www.education.ualberta.ca/educ/psych/crame/files/standard_setting.pdf diambil 25 Agustus 2005.
- Mardapi (1999). Estimasi Kesalahan Pengukuran dalam Bidang Pendidikan dan Implikasinya pada Ujian Nasional. *Pidato Pengukuhan Guru Besar* di UNY Yogyakarta tanggal 11 September 1999.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark Procedure: Psychological Perspectives. In G.J. Cizek (Ed.), *Setting Performance Standards*. Mahwah, NJ.
- Nedelsky, L. (1954). Absolute grading standards for objective test. *Educational and Psychological Measurement*, 14, 3-19.
- Plake B. S., Impara, J. C., & Irwin, P. (2000). Consistency of Angoff-based predictions of item performance: Evidence of technical quality of results from the Angoff standard setting method. *Journal of Educational Measurement*, 37(4), 347–355.
- Plake, B. S., Melican, G. J., & Mills, C. N. (1991). Factors influencing intrajudge consistency during standard-setting. *Educational measurement: Issues and Practice*, 10(2), 15-16, 22, 25.
- Ricker, K. L. 2009. Setting Cut Scores: Critical Review of Angoff and Modified-Angoff Methods. Edmonton (Alberta, Canada): Centre for Research in Applied Measurement and Evaluation University of Alberta.
- Stahl, J. A. 2008. *Standard Setting Methodologies: Strengths and Weaknesses*. Illinois: Pearson VUE. Diambil pada tanggal 28 September 2009 dari www.iaea2008.cambridgeassessment.org.uk/ca/digitalAssets/180502_Stahl.pdf
- Undang-undang No.20 tahun 2003 tentang Sistem Pendidikan Nasional.