

**PENYETARAAN HORIZONTAL PERANGKAT TES UJICоба
UJIAN NASIONAL MATEMATIKA SMA PROGRAM IPA
DI SMAN KOTA YOGYAKARTA
TAHUN PELAJARAN 2009/2010**

Haryani
SMA Negeri 1 Yogyakarta
haryanimpd@gmail.com

Abstrak

Penelitian ini bertujuan menentukan kesetaraan tes ujicoba Ujian Nasional matematika SMA program IPA di SMAN Kota Yogyakarta tahun ajaran 2009/2010. Sumber data berasal dari lembar jawaban siswa kelas XII IPA di SMAN kota Yogyakarta yang mengikuti ujicoba Ujian Nasional Matematika IPA tahun pelajaran 2009/2010 putaran 1, 2, dan 3 dengan jumlah sampel 1396. Karakteristik tes dilakukan Bilog MG, Analisis untuk kesetaraan meliputi: analisis varians untuk menguji kesamaan rata-rata, uji Tukey untuk uji pasangan, dan uji Levene untuk menguji homogenitas varians. Proses penyetaraan menggunakan kurva karakteristik dari Haebra dengan bantuan program Excel 2007. Hasil penelitian menunjukkan bahwa secara bersama-sama tes putaran 1, 2, dan 3 tidak paralel. Tes putaran 1 dan putaran 2 tidak paralel dengan persamaan kesetaraan tes putaran satu (X) ke putaran dua (Y) adalah $\theta_y = 1,5824\theta_x - 1,8653$, demikian juga dengan tes putaran 1 dan putaran 3 juga tidak paralel dengan persamaan kesetaraan soal putaran satu (X) ke putaran tiga (Z) adalah $\theta_z = 0,99046\theta_x - 1,2212$.

Kata Kunci: *Penyetaraan horizontal, test matematika*

1. PENDAHULUAN

Penilaian pendidikan pada jenjang pendidikan dasar dan menengah terdiri atas penilaian hasil belajar oleh pendidik, satuan pendidikan, dan penilaian hasil belajar oleh pemerintah. Penilaian pembelajaran oleh pemerintah disebut sebagai Ujian Nasional (UN). UN sebagai salah satu proses pengukuran hasil belajar tingkat Nasional, memiliki tujuan dan kegunaan yang penting dalam bidang pendidikan salah satunya adalah untuk penentuan kelulusan siswa dari satuan pendidikan.

Perangkat soal yang baik, sebelum perangkat soal digunakan untuk mengukur

kemampuan siswa, seharusnya dilakukan tahap telaah kualitatif dan empiris. Analisis butir soal secara kualitatif, dilakukan untuk menilai butir soal ditinjau dari aspek materi, konstruksi dan bahasa.

Analisis soal secara empiris dengan pendekatan teori tes klasik dan Teori Respons Butir. Pengujian secara klasik akan mendapatkan informasi yang berkaitan dengan daya beda, reliabilitas, tingkat kesukaran, efektifitas distraktor, *Standar Error Measurement (SEM)*, dan lain sebagainya. Pengujian dengan Teori Respons Butir akan mendapatkan informasi tentang parameter butir (daya beda, tingkat

kesukaran, tebakan semu) dan parameter kemampuan peserta. Butir soal yang memenuhi syarat dapat disimpan untuk pengembangan bank soal di daerah.

Dalam Peraturan Pemerintah Tahun 2007 Nomor 20 disebutkan bahwa, instrumen penilaian UN harus memenuhi persyaratan substansi dari aspek materi, konstruksi, bahasa, dan memenuhi bukti validitas empiris serta menghasilkan skor yang dapat diperbandingkan antar sekolah, antar daerah dan antar tahun. Mengacu pada ketentuan ini, uji coba Ujian Nasional yang merupakan simulasi dari UN, sebisa mungkin harus memenuhi persyaratan yang sesuai dengan instrumen penilaian UN.

Berdasarkan kenyataan yang ada selama ini, untuk membuat soal uji coba UN Matematika di SMA kota Yogyakarta, dilakukan oleh guru-guru bidang studi Matematika melalui forum MGMP atau penunjukan langsung tanpa melalui telaah soal dan ujicoba terbatas. Kondisi ini memungkinkan tidak diketahuinya karakteristik soal yang digunakan, baik yang bersifat kualitatif maupun empiris. Selain itu, soal uji coba UN yang digunakan belum melalui proses penyeteraan, sehingga hasil pada setiap putaran uji coba tidak dapat dibandingkan.

Data empirik hasil penelitian, membuktikan bahwa beberapa paket tes, yang dikembangkan dengan kisi-kisi yang

sama terbukti tidak setara, dengan *SEM* yang cenderung besar. Kenyataan seperti ini menjadi sulit untuk membuat keputusan yang berkaitan dengan kemampuan peserta didik. Hal tersebut dikarenakan nilai enam pada salah satu paket, belum tentu sama dengan nilai enam pada paket yang lain.

Masalah lain dalam pemberian paket yang berbeda meskipun dari kisi-kisi yang sama adalah, tidak ada jaminan bahwa tes yang dikembangkan dari kisi-kisi yang sama menghasilkan tingkat kesukaran yang sama. Perbedaan paket soal memungkinkan terjadinya perbedaan karakteristik, sifat, dan tingkat kesukaran yang di ukur.

Berbagai kelemahan di atas, akan mempengaruhi hasil evaluasi yang diperoleh, sehingga kesimpulan yang diambil tidak akurat. Informasi yang akurat hanya dapat diperoleh jika kualitas alat yang digunakan mempunyai memenuhi syarat. Persyaratan tersebut di samping sah dan handal, hasil pengukuran dapat dibandingkan. Untuk mencapai hal tersebut, perlu dikembangkan sejumlah perangkat tes yang berkualitas sehingga hasilnya memberikan informasi tentang perkembangan mutu pembelajaran. Teknik penyamaan skor (*equiting*) baik horisontal maupun vertikal dapat digunakan untuk keperluan tersebut.

Tujuan penelitian ini adalah mengetahui kesetaraan soal uji coba Ujian

Nasional Matematika program IPA putaran satu sampai tiga tahun pelajaran 2009/2010 di SMAN kota Yogyakarta menggunakan Teori Respon Butir. Temuan penelitian ini dapat menjadi bahan masukan bagi MGMP Matematika kota Yogyakarta SMA IPA. agar konversi skor dari sejumlah perangkat tes yang digunakan mempunyai kesalahan yang sekecil mungkin sehingga keputusan yang diambil tidak keliru.

2. KAJIAN TEORI

a. Teori Respon Butir (TRB)

Teori respons butir memuat dua parameter, yaitu parameter butir dan parameter peserta. Parameter cirri peserta θ menyatakan cirri peserta dengan kemampuan θ . Menurut Kolen & Brennan (2004: 158) Parameter kemampuan (θ) atau *ability*, terletak dalam interval $-\infty \leq \theta \leq \infty$, dan diskalakan mendekati distribusi normal dengan rata-rata 0 dan simpangan baku 1, tetapi pada praktiknya, kemampuan seseorang (θ) terletak antara $-3 \leq \theta \leq 3$.

Pengujian dengan Teori Respons Butir akan mendapatkan informasi tentang parameter butir (daya beda, tingkat kesukaran, tebakan semu) dan parameter kemampuan peserta. Secara teoritis daya beda butir terletak pada

skala $-\infty \leq a \leq \infty$. Parameter a_i adalah ciri butir yang berkaitan dengan daya pembedaan yaitu kemampuan butir untuk mempertegas perbedaan di antara peserta yang dapat menjawab dengan benar dan menjawab dengan salah. Daya beda negatif disebabkan sesuatu yang salah pada butir, jika peluang menjawab dengan benar menurun sementara kemampuan peserta meningkat. Menurut Hambleton, Swaminathan (1985: 36) parameter daya beda terletak pada $[0,2]$.

Parameter b_i adalah ciri butir yang berkenaan dengan tingkat kesukaran. yaitu sukar atau kurang sukarnya butir untuk dijawab oleh peserta. Tingkat kesukaran butir merupakan fungsi dari kemampuan seseorang (Djemari Mardapi, 1991: 11). Besarnya tingkat kesulitan butir (b_i) terletak pada $-\infty \leq a \leq \infty$. Tetapi dalam prakteknya, range b_i terletak antara -2,0 sampai +2,0 (Hambleton, Swaminathan & Rogers, 1991: 13). Nilai b_i mendekati -2,0 mengindikasikan bahwa butir tersebut mudah, sedang nilai b_i mendekati + 2,0 mengindikasikan bahwa butir tersebut sukar.

Parameter c_i adalah *lower asymptote* dari kurva karakteristik butir dan menunjukkan peluang peserta dengan kemampuan rendah menjawab

butir dengan benar (Hambleton & Swaminathan, 1985: 38). Selain itu, parameter c_i menyatakan besarnya peluang menjawab benar bagi peserta dengan kemampuan rendah Djemari Mardapi (1991: 9).

b. Penyetaraan

Peterson, Kolen & Hoover (Linn, 1989: 242) menyatakan, penyetaraan adalah suatu prosedur empiris yang diperlukan untuk mentransformasi skor dari tes yang satu ke skor tes yang lain. Dengan demikian, penyetaraan adalah suatu cara untuk mengubah skor-skor dari perangkat tes yang berbedadalam rerata tingkat kesukaran dan rerata daya beda butir diantara dua atau lebih perangkat tes, agar menjadi sama setelah dilakukannya penyetaraan. Sementara itu, Hambleton, Swaminathan, & Rogers, 1991: 94) menyatakan bahwa penyetaraan merupakan proses mentransformasikan skor tes X ke matriks skor tes Y, atau sebaliknya, sehingga dari hasil penyetaraan ini kedua skor dapat dibandingkan.

Syarat penyetaraan menurut Petersen et al. (Linn, 1986: 242) skor-skor dari perangkat tes X dan Y dapat disetarakan jika terpenuhi empat syarat, yaitu 1) kedua tes harus mengukur hal yang sama, 2)

kemampuan peserta tes harus sama, 3) populasinya harus *invariant*, 4) transformasinya harus simetris. Proses penyetaraan dari beberapa perangkat tes (*equating*) dapat dilakukan dengan dua cara, yaitu penyetaraan secara horizontal dan penyetaraan secara vertikal. Proses penyetaraan yang diperoleh dari dua perangkat tes yang berbeda tetapi mengukur hal yang sama dinamakan penyetaraan horizontal.

Rancangan penyetaraan terdiri dari: rancangan *Single-Group-Design* menggunakan satu kelompok peserta yang merespon dua perangkat tes (X dan Y), rancangan Kelompok Ekuivalen (*Equivalent-Group-Design*), menggunakan dua kelompok peserta ekuivalen (K_1 dan K_2) dan dua perangkat tes (X dan Y), kelompok peserta K_1 mengerjakan perangkat tes X dan kelompok peserta K_2 mengerjakan perangkat tes Y dan rancangan Tes Anchor (*Anchor Test Design*) menggunakan dua perangkat tes (X dan Y) dan dua kelompok peserta (K_1 dan K_2), yang masing-masing perangkat tes ditambahkan item – item tes *anchor* Z, sehingga kedua perangkat tes menjadi (X + Z) dan (Y+Z).

Metode penyetaraan terdiri dari dua pendekatan, yaitu de Penyetaraan Teori Tes Klasik dan Teori Respon

Butir (TRB). Penyetaraan tes dengan menggunakan Pendekatan teori klasik dapat dilakukan dengan cara linier dan ekipersentil, sedang pendekatan TRB menggunakan metode regresi, metode, reratasigma, Tegar rerata dan Kurva karakteristik (Hambleton dan Swaminathan, 1985: 206).

Menurut Kolen & Brenan (2004: 45), hubungan antara parameter yang disetarakan adalah:

$$\theta_y = \alpha\theta_x + \beta \dots\dots\dots 1$$

$$\text{dengan } \alpha = \frac{\sigma(b_y)}{\sigma(b_x)} = \frac{\mu(a_x)}{\mu(a_y)} \dots\dots\dots 2$$

$$\beta = \mu(b_y) - \alpha\mu(b_x) \dots\dots 3$$

- θ_x : estimasi parameter kemampuan tes X
- θ_y : estimasi parameter kemampuan tes Y
- $\mu(b_y)$ dan $\mu(b_x)$: rata-rata kemampuan pada tes Y dan tes X
- $\sigma(b_y)$ dan $\sigma(b_x)$: simpangan baku tingkat kesulitan tes Y dan tes X
- α dan β : konstanta penyetaraan

3. METODE PENELITIAN

Sumber data berasal dari lembar jawaban siswa kelas XII IPA di SMAN kota Yogyakarta yang mengikuti ujicoba Ujian Nasional Matematika IPA tahun pelajaran 2009/2010 putaran 1, 2, dan 3 dengan jumlah sampel 1396. Karakteristik tes dilakukan Bilog MG, Analisis untuk

kesetaraan meliputi: analisis varians untuk menguji kesamaan rata-rata, uji Tukey untuk uji pasangan, dan uji Levene untuk menguji homogenitas varians. Proses penyetaraan dilakukan dengan bantuan program Microsoft Excell 2007. Proses penyetaraan dilakukan pada soal dengan nomor soal yang mengukur kemampuan sama di kedua paket soal yang disetarakan.

Untuk mengetahui karakteristik soal secara empiris dengan pendekatan Teori Respons Butir menggunakan program BILOG 3.0 pada masing masing paket soal. Analisis empiris dengan BILOG untuk mengestimasi parameter butir (item). Analisis dengan program Bilog terdiri tiga fase. Pada fase pertama. diperoleh info tentang banyaknya peserta tes yang menjawab benar, proporsi peluang menjawab benar dibagi peluang menjawab salah serta koefisien korelasi biserial. Apabila suatu item memiliki nilai koefisien biserial kurang dari 0,3 maka tidak diikutkan dalam analisis berikutnya.

Fase kedua, estimasi parameter Teori Respon Butir. Pada fase ini diperoleh tentang informasi parameter butir menurut Teori Respon Butir yang banyaknya sesuai dengan model yang digunakan. Berdasarkan kecocokan model penelitian ini menggunakan model logistik 2 parameter.

Item yang baik, menurut Hambleton&Swaminathan (1985: 36) jika mempunyai nilai daya beda dalam interval [0,2] dan nilai tingkat kesukaran dalam interval [-2,+2]. Selain parameter butir, pada fase kedua juga dihasilkan statistik kecocokan suatu butir dengan model atau *goodness of fit statistic*.

4. HASIL DAN PEMBAHASAN

a. Hasil Penelitian

Pada uji beda rerata didapatkan F= 1,017 dengan taraf signifikansi 0,000. Hal ini menunjukkan bahwa ketiga soal berbeda secara signifikan. Kondisi ini memerlukan uji berikutnya, yaitu uji pasang untuk mengetahui pasangan manakah yang berbeda. Hasil pengujian secara empiris dengan uji Tukey, pada tes putaran 1, 2, dan 3 dengan jumlah masing-masing tes sebanyak 40 item, pada skala 0 sampai 10, dapat dilihat dalam Tabel 1

Tabel 1. Hasil Uji Beda Rata-Rata Paket Soal

Pasangan	Beda rerata	Signifikan si
Putaran 1-putaran 2	-2.284	0.00
Putaran 1 -putaran 3	-2,372	0.00
Putaran 2 -putaran 3	-0.087	0.429

Berdasarkan Tabel 2, pasangan putaran 1- putaran 2 dan pasangan putaran 1-putaran 3 terjadi perbedaan rerata yang signifikan pada taraf signifikansi 5%, sedang putaran 2-putaran 3 perbedaan rerata yang ada tidak signifikan pada taraf 5%.

Hasil uji homogenitas varians dengan uji *Levene* didapatkan nilai statistik dengan taraf signifikansi 0,000. Hal ini menunjukkan varians tidak homogen. Selanjutnya dilakukan uji pasang untuk menguji homogenitas varians untuk Putaran 2 - putaran 3. Hal ini dilakukan karena pada ujibeda rerata pasangan tersebut tidak signifikan sehingga diperlukan pengujian lebih lanjut. Hasil pengujian ,diperoleh F= 0,161 sehingga tidak terjadi perbedaan varians yang signifikan antara skor siswa di paket putaran 2-putaran 3. Hasil pengujian secara empiris, berdasarkan hasil ujibeda rerata dan uji homogenitas, kesetaraan paket soal dapat dilihat dalam Tabel 2 berikut.

Tabel 2. Hasil Uji Kesetaraan Tes Putaran 1, 2 dan 3

Nomor	Aspek Uji	Put 1- Put 2	Put 1 – Put 3	Put 2 – Put 3	Put 1-Put 2- Put 3
1.	rerata	berbeda	berbeda	Tidak berbeda	berbeda
2.	Varians	tidak sama	Tidak sama	Sama	Tidak sama
kesimpulan		tidak paralel	Tidak paralel	Paralel	Tidak paralel

Mencermati Tabel 2, terbukti bahwa secara berpasangan Putaran 1- putaran 2, dan putaran 1- putaran 3 bersifat tidak paralel, sedangkan pasangan putaran 2-

putaran 3 bersifat paralel. Oleh karena itu, secara bersama-sama pasangan Putaran 1-Putaran 2-Putaran 3 tidak paralel. Bukti ini sekaligus mendukung pandangan teoritis tentang kesetaraan perangkat tes, yaitu pada dasarnya tidak pernah ada perangkat tes yang paralel, kendatipun kesemuanya telah menggunakan kisi-kisi yang sama. Sehubungan dengan hasil pembuktian kesetaraan perangkat tes yang diteliti terbukti tidak setara, maka perlu dilakukan proses penyetaraan antar paket soal yang tidak paralel.

Penyetaraan dilakukan pada nomor item yang mempunyai kesamaan kemampuan yang diuji berdasarkan Standar Kelulusan Ujian Nasional SMA IPA tahun pelajaran 2009/2010 dalam Peraturan Menteri Nomor 75 tahun 2009. Penyetaraan tes putaran 1, 2, dan 3 Uji coba Ujian Nasional dengan metode kurva karakteristik Haebera, melibatkan parameter tingkat kesukaran dan daya beda. Untuk memperoleh konstanta α dan β , penghitungannya menggunakan rerata tingkat kesukaran dan rerata simpangan baku dari parameter tingkat kesukaran tes.

Pada proses penyetaraan tes putaran 1 ke putaran 2 yang berdasarkan kesamaan isi materi yang diujikan, penyetaraan dilakukan pada 12 item di putaran 1 yaitu nomor: 2, 3, 4, 6, 7, 9, 17, 21, 28, 30, 38, 40 disetarakan dengan 12

item di putaran 2 yaitu nomor: 3, 5, 7, 8, 9, 18, 22, 29, 30, 38, dan 40. Proses penyetaraan menggunakan kurva karakteristik menghasilkan persamaan konstanta penyetaraan $\alpha=1,5824$ dan konstanta $\beta=-1,8653$ sehingga diperoleh persamaan kesetaraan putaran 1 (X) ke putaran 2 (Y) yaitu:

$$\theta_y = 1,5824\theta_x - 1,8653$$

$$b_y = 1,5824b_x - 1,8653$$

$$a_y = \frac{a_x}{1,5824}$$

Konversi hasil penyetaraan tingkat kesukaran tes putaran 1 ke 2 disajikan pada Tabel 3

Tabel 3. Konversi Tingkat Kesukaran Tes Putaran 1 ke Putaran 2

Putaran 1		Putaran 2	
No item	Tingkat Kesukaran mula - mula	No item	Tingkat Kesukaran Setelah disetarakan
2	0.326	2	-1.3494
3	0.379	3	-1.26553
4	0.528	5	-1.02974
6	0.603	7	-0.91106
7	0.549	8	-0.99651
9	0.286	9	-1.4127
17	1.391	18	0.335904
21	-0.525	22	-2.69606
28	0.693	29	-0.76864
30	1.531	30	0.557446
38	0.68	38	-0.78921
40	0.81	40	-0.5835
μ	0.6043		-0.9091
σ	0.5263		0.8329

Selanjutnya dapat diketahui posisi tes putaran 1 dibandingkan dengan putaran 2. Rangkuman parameter butir setelah proses penyetaraan berdasar kesamaan isi/indikator dengan kurva karakteristik dari Haebera disajikan dalam Tabel 4

Tabel 4. Rangkuman Parameter Butir Setelah Proses Penyetaraan Berdasar Isi Tes Putaran 1 ke 2

Statistik	Putaran 2		Putaran 1	
	a	b	a	b
μ	05490	-0,9091	0,30375	-0,9091
σ	0,134	0,8329	0,07387	0,8329

Berdasarkan Tabel 4, dapat diketahui bahwa rerata tingkat kesukaran tes putaran 1 dan putaran 2 mempunyai rerata tingkat kesukaran yang sama. Sedangkan berdasarkan daya beda, tes putaran 2 mempunyai rerata daya beda yang lebih besar dari tes putaran 1.

Proses penyetaraan putaran 1 ke putaran 3 berdasar kesamaan isi materi yang diujikan, hanya dilakukan pada 10 item tes putaran 1 yaitu nomor: 4, 6, 7, 9, 11, 15, 17, 28, 30, dan 34 disetarakan dengan 10 item pada putaran 3 yaitu nomor: 5, 7, 8, 10, 12, 17, 18, 29, 30, dan 34. Proses penyetaraan menggunakan kurva karakteristik menghasilkan persamaan konstanta penyetaraan $\alpha= 0,9905$ dan konstanta $\beta= -1,2212$, sehingga diperoleh persamaan kesetaraan putaran 1 (X) ke putaran 3 (Z) sebagai berikut:

$$\theta_z = 0,9905\theta_x - 1,2212$$

$$b_z = 0,9905b_x - 1,2212$$

$$a_z = \frac{a_x}{0,9905}$$

Berdasarkan konstanta α dan β yang telah diperoleh, dilakukan transformasi parameter butir, sehingga diperoleh parameter butir hasil penyetaraan. Konversi tingkat kesukaran tes putaran 1 ke putaran 3 berdasar kesamaan materi disajikan pada Tabel 5

Tabel 5. Konversi Tingkat Kesukaran Tes Putaran 1 ke Putaran 3

Putaran 1		Putaran 3	
No item	Tingkat Kesukaran mula - mula	No item	Tingkat Kesukaran Setelah disetarakan
4	0.528	5	-0.69819
6	0.603	7	-0.62391
7	0.549	8	-0.67739
9	0.286	10	-0.93788
11	-1.098	12	-2.30868
15	0.094	17	-1.12805
17	1.391	18	0.156576
28	0.693	29	-0.53476
30	1.531	30	0.295241
34	0.294	34	-0.92996
μ	0,3173		-0,6095
σ	0,7838		0,8451

Selanjutnya, rangkuman parameter butir setelah proses penyetaraan berdasar kesamaan isi, tes putaran 1 ke putaran 3 disajikan dalam Tabel 6.

Tabel 6. Rangkuman Parameter Butir Setelah Proses Penyetaraan Berdasarkan Isi Tes Putaran 1 ke 3

Statistik	Putaran 3		Putaran 1	
	a	b	a	b
μ	0.6391	-0.7387	0.5001	-0.7387
σ	0.1496	0.7167	0.1107	0.7167

Berdasarkan Tabel 6, dapat diketahui bahwa setelah disetarakan berdasarkan kesamaan isi/indikator, rerata tingkat kesukaran tes putaran 1 dan putaran 3 mempunyai rerata tingkat kesukaran yang sama, sedangkan berdasarkan daya beda tes putaran 3 mempunyai rerata daya beda yang lebih besar dari tes putaran 1.

b. Pembahasan

Hasil penyetaraan menunjukkan bahwa secara bersama-sama, ketiga soal ujicoba Ujian Nasional Matematika IPA tahun ajaran 2009/2010 terbukti tidak paralel karena mempunyai rerata skor dan varians yang tidak sama. Hasil penelitian ini mendukung penelitian yang dilakukan Nauli (2009) yang meneliti tentang kesetaraan perangkat UAS SMP kelas 7 bidang studi Bahasa Inggris di Provinsi DIY dengan menggunakan metode rerata dan sigma, metode rerata dan rerata, dan kurva karakteristik dari Lord, yang menyatakan bahwa dengan menggunakan ketiga metode tersebut perangkat UAS SMP kelas 7 bidang studi Bahasa Inggris di Provinsi DIY tidak setara. Hal ini memperkuat pernyataan yang

dikemukakan oleh Hambleton & Swaminathan (1985: 2) yang menyatakan bahwa tes yang paralel, sulit dipenuhi, walaupun disusun dengan kisi-kisi yang sama.

Proses penyetaraan menghasilkan kesimpulan bahwa tes putaran 1 setelah disetarakan dengan tes putaran 2 menghasilkan tingkat kesukaran yang sama, demikian juga tes putaran 1 yang disetarakan dengan tes putaran 3 menghasilkan tingkat kesukaran sama. Hasil penyetaraan ini memberikan penafsiran, adanya kenaikan rata-rata pada putaran 2 dan 3 dibandingkan dengan putaran 1, sebenarnya bukan dipengaruhi oleh tingkat kesukaran soal, tetapi disebabkan faktor – faktor lain. Faktor tersebut antara lain, adanya proses pembelajaran dan latihan soal. Menurut Thorndike (Bower & Hilgard, 1981: 26) ada dua prinsip belajar yang mempengaruhi peningkatan stimulus respon yaitu : 1) hukum Latihan, dan 2) kesiapan. Hukum latihan menyatakan bahwa suatu hubungan atau rangasangan dan perilaku akan makin kukuh apabila sering dilakukan latihan. Hukum kesiapan menyatakan bahwa hubungan antara rangsangan dengan perilaku akan menjadi lebih kukuh apabila disertai dengan kesiapan individu. Atas dasar hukum ini, maka pembelajaran akan lebih efektif apabila memberikan hasil yang memuaskan, disertai dengan banyak latihan dan memiliki

kesiapan untuk melakukan aktivitas pembelajaran.

5. KESIMPULAN

Secara bersama-sama tes putaran 1, 2, dan 3 tidak paralel. Tes putaran 1 dan putaran 2 tidak paralel dengan persamaan kesetaraan tes putaran satu (X) ke putaran dua (Y) adalah $\theta_y = 1,5824\theta_x - 1,8653$, demikian juga dengan tes putaran 1 dan putaran 3 juga tidak paralel dengan persamaan kesetaraan soal putaran satu (X) ke putaran tiga (Z) adalah $\theta_z = 0,99046\theta_x - 1,2212$.

Saran

Perlunya diperhatikan kesetaraan soal pada paket ujicoba, sehingga hasilnya dapat diperbandingkan dan perbaikan pembelajaran lebih efektif

6. REFERENSI

- Bower, G.H., & Hilgard, E.R. (1981). *Theories of learning in* . USA: Prentice-Hall, Inc.
- Brennan, R.L. (2006). *Educational measurement*. USA: Praeger Publisher.
- Crocher, L. & Algina, J. (1981). *Introduction to classical and modern test theory*. New York: CBS College Publishing.
- Djemari Mardapi .(2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendikia.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamental of item response Theory*. Newburg Park, LA: Sage Publication ICC.
- Kolen, M.J., & Brennan, R.L. (2004). *Test equating, scaling, and linking*. USA: Springer.
- Linn, R.L. (1989). *Educational measurement*. New York: Macmillan Publishing Company.
- Nauli. (2009). *Penyetaraan perangkat UAS SMP kelas X bidang studi Bahasa Inggris di Provinsi DIY*: Tesis magister, tidak diterbitkan, Universitas Negeri Yogyakarta, Yogyakarta.
- Reckase, M.D. (2009). *Multidimensional item response theory*. USA: Springer.
- Rustam. (2000). *Penyetaraan perangkat tes matematika program D2 PGSD Universitas Terbuka*. Laporan penelitian. Jakarta: Lembaga Penelitian Universitas Terbuka.