

## KOMPARASI METODE PENYETARAAN TES MENURUT TEORI RESPONS BUTIR

*Oleh:*

*Erna Miyatun dan Djemari Mardapi*

### **Abstrak**

Penelitian ini bertujuan menemukan metode penyetaraan tes dengan teori respons butir yang paling akurat. Sebelum dilakukan pengujian keakuratan metode penyetaraan tes, terlebih dahulu dilakukan: 1) Pemeriksaan setara atau tidaknya perangkat tes, 2) Penemuan persamaan konversi antarpaket soal yang tidak setara dan 3) Estimasi parameter butir. Penentuan metode penyetaraan tes yang akurat didasarkan pada besarnya kesalahan standar penyetaraan (SEE); yang paling akurat adalah metode penyetaraan tes yang mempunyai SEE paling kecil.

Dalam penelitian ini, terdapat dua macam objek penelitian yaitu paket soal dan jawaban peserta EBTANAS. Soal EBTANAS terdiri dari 3 paket soal utama, 1 paket soal susulan, dan 1 paket soal cadangan. Peserta EBTANAS SLTP di Jawa Tengah Tahun Ajaran 1998/1999 terdiri dari 370.187 siswa. Teknik *Purposive Sampling* digunakan untuk menentukan sampel paket soal, yaitu ketiga paket soal utama terpilih sebagai sampel penelitian. Teknik *Cluster Stratified Systematic Random Sampling* digunakan untuk menentukan sampel jawaban peserta EBTANAS. Jawaban peserta sebanyak 8.251, terpilih sebagai sampel penelitian. Paket soal EBTANAS SLTP bidang studi Matematika dan data jawaban peserta tes, diperoleh melalui pengambilan data dokumentasi. Keparalelan tes diuji dengan analisis varians, uji *Scheffe*, dan uji *Levene*. Penyetaraan tes dilakukan dengan metode: Rerata dan Sigma, Rerata dan Sigma Tegar, dan Kurva Karakteristik. Estimasi parameter butir dilakukan dengan analisis program *Bilog*.

Hasil analisis menunjukkan, pasangan paket soal utama-1 dan 2 serta pasangan paket soal utama-1 dan 3 tidak paralel, sehingga diperlukan proses penyetaraan. Pasangan paket soal utama-2 dan paket soal utama-3 bersifat paralel. Oleh karena itu tidak diperlukan proses penyetaraan. Penyetaraan paket soal utama-1 ke paket soal utama-2, dengan metode Rerata dan Sigma didapatkan persamaan konversi  $X_2 = 1,19495 X_1 + 0,25766$ ; dengan metode Rerata dan Sigma Tegar,  $X_2 = 0,71710 X_1 - 0,00081$ , sedangkan dengan metode Kurva Karakteristik,  $X_2 = 1,05199 X_1 + 0,03012$ . Penyetaraan paket soal utama-1 ke paket soal utama-3, dengan metode Rerata dan Sigma didapatkan persamaan konversi  $X_3 = 1,49581 X_1 + 0,42366$ ; dengan metode Rerata dan Sigma Tegar,  $X_3 = 0,60018 X_1 - 0,00157$ , sedangkan dengan metode Kurva Karakteristik diperoleh persamaan konversi  $X_3 = 1,05391 X_1 + 0,03232$ .

Penyetaraan paket soal utama-1 ke paket soal utama-2, dengan metode Rerata dan Sigma didapatkan nilai SEE = 0,030574, dengan metode Rerata dan Sigma Tegar, SEE = 0,04933, dan dengan metode Kurva Karakteristik, SEE = 0,032968. Penyetaraan paket soal utama-1 ke paket soal utama-3, dengan metode Rerata dan Sigma didapatkan nilai SEE = 0,045906, dengan metode Rerata dan Sigma Tegar SEE = 0,106431, dan dengan metode Kurva Karakteristik SEE = 0,05585. Hasil tersebut menunjukkan, metode Rerata dan Sigma merupakan metode paling akurat, diikuti metode Kurva Karakteristik, dan metode Rerata dan Sigma Tegar.

## **Pendahuluan**

Evaluasi Belajar Tahap Akhir (EBTANAS) memiliki fungsi penting dalam bidang pendidikan, antara lain: 1) untuk mengetahui sejauh mana tujuan kurikulum telah tercapai, 2) penentu standardisasi mutu pendidikan nasional, 3) acuan untuk kelulusan, dan 4) acuan seleksi penerimaan siswa baru ke jenjang pendidikan berikutnya (Depdikbud, 1998). Untuk tujuan keamanan soal EBTANAS tiap bidang studi disusun beberapa perangkat tes berdasarkan kisi-kisi sama. Paket soal EBTANAS terdiri dari: tiga paket soal utama, satu paket soal susulan, dan satu paket soal cadangan (Kanwil Depdikbud, 1998).

Informasi yang diperoleh melalui EBTANAS karena terdiri lebih dari satu paket soal harus benar-benar mencerminkan kemampuan peserta EBTANAS yang sebenarnya. Artinya, perbedaan skor yang diperoleh seorang peserta dengan peserta lainnya harus menunjukkan perbedaan kemampuan di antara mereka, bukan disebabkan karena faktor paket soal yang berbeda. Di Wilayah Propinsi Jawa Tengah ketiga paket soal utama diujikan, namun dalam pelaksanaannya dihindari menggunakan paket soal yang sama untuk daerah yang berdekatan yaitu dengan membentuk sistem rayonisasi.

Hambleton dan Swaminathan (1985) menegaskan bahwa sekalipun perangkat tes disusun berdasarkan kisi-kisi yang sama, namun jarang sekali atau hampir tidak pernah ditemukan perangkat tes yang benar-benar setara dalam sebaran serta tingkat kesukaran. Pendapat lain dikemukakan oleh Suryabrata (1997) serta Hayat dan Pranata (1995) bahwa dalam pelaksanaan penilaian penggunaan beberapa perangkat tes perlu dilakukan penyetaraan, karena dengan penyetaraan perangkat tes dapat dijamin keadilan bagi peserta tes.

Penyetaraan tes sangat dirasakan kegunaannya mengingat mutu pendidikan di Indonesia belum merata dan keadaan geografis Indonesia sebagai negara kepulauan yang cukup luas. Hal ini mengakibatkan pengukuran secara serentak dalam waktu yang sama tidak mudah untuk dilakukan. Adanya kebijaksanaan otonomi pendidikan pelaksanaan pendidikan sudah menjadi wewenang daerah setempat, namun pemerintah pusat tetap berkewajiban untuk mengendalikan kualitas pendidikan nasional. Misalnya, dengan penetapan kurikulum nasional dan kemampuan standar pada tingkat tertentu. Pengendalian akan mudah dilakukan jika perangkat tes di tiap-tiap daerah

diketahui tingkat kesetaraannya. Tingkat kesetaraan perangkat tes yang berbeda akan dapat diketahui melalui proses penyetaraan.

Keakuratan informasi kesetaraan antarpaket tidak terlepas dari keakuratan metode penyetaraan yang digunakan. Dalam teori respons butir terdapat empat metode penyetaraan tes, yaitu metode: Regresi, Rerata dan Sigma, Rerata dan Sigma Tegar, dan Kurva Karakteristik (Holland, 1982; Angoff, 1982; Potthoff, 1982; Lord, 1980, Holland dan Wightman, 1982; serta Naga, 1992). Keempat metode penyetaraan tersebut menggunakan prosedur yang berbeda-beda, sehingga ada kemungkinan konstanta konversi yang dihasilkan berbeda untuk penyetaraan paket soal yang sama. Penyetaraan paket soal yang sama walaupun dengan metode yang berbeda seharusnya didapatkan hasil yang sama pula. Hasil konstanta konversi yang berbeda ada kemungkinan terjadi karena perbedaan keakuratan dari metode penyetaraan yang digunakan.

Berdasarkan adanya kemungkinan konstanta konversi yang berbeda dari masing-masing metode penyetaraan dalam pengukuran paket soal yang sama tersebut, perlu diteliti metode penyetaraan mana yang paling akurat dalam proses penyetaraan tes.

### **Pembatasan Masalah**

Metode penyetaraan yang akan dikomparasikan adalah: Rerata dan Sigma, Rerata dan Sigma Tegar, dan Kurva Karakteristik. Metode Regresi tidak ikut dikomparasikan karena tidak bersifat timbal balik dalam penentuan konstanta konversi, sedangkan ketiga metode yang lain bersifat timbal balik.

Pendekatan teori respons butir dipilih karena teori bersifat *sample free* dan *respondent free*. Sifat semacam ini tidak terdapat pada teori pengukuran klasik.

### **Perumusan Masalah**

Berkaitan dengan kesetaraan paket soal EBTANAS dan konteks komparasi metode, maka masalah yang ada sebagai berikut:

1. Bagaimana konstanta konversi penyetaraan paket soal EBTANAS SLTP bidang studi Matematika Tahun Ajaran 1998/1999, menurut metode Rerata dan Sigma, metode Rerata dan Sigma Tegar, dan metode Kurva Karakteristik?

2. Metode penyetaraan mana yang paling akurat dalam penyetaraan paket soal?

### **Tujuan Penelitian**

Untuk memperoleh jawaban atas masalah yang ingin diungkap, maka tujuan penelitian ini, adalah untuk:

1. Memperoleh bukti empiris tentang tingkat kesetaraan paket soal EBTANAS SLTP bidang studi Matematika Tahun Ajaran 1998/1999.
2. Mengetahui konstanta konversi skor penyetaraan paket soal EBTANAS SLTP bidang studi Matematika Tahun Ajaran 1998/1999 dengan menggunakan metode Rerata dan Sigma, metode Rerata dan Sigma Tegar, dan metode Kurva Karakteristik.
3. Menentukan metode penyetaraan yang paling akurat dalam penyetaraan paket soal sesuai dengan asumsi-asumsi tes yang ada.

### **Manfaat Penelitian**

Diharapkan hasil penelitian dapat bermanfaat bagi:

1. Bidang pengukuran, yaitu mengetahui metode penyetaraan yang paling akurat dalam membuat penyetaraan perangkat tes.
2. Bidang pendidikan:
  - a. supaya konversi skor dari sejumlah perangkat tes yang digunakan mempunyai kesalahan yang sekecil mungkin sehingga keputusan yang diambil tidak keliru.
  - b. pengembangan bank soal yang terkalibrasi baik di tingkat wilayah maupun di tingkat nasional.
3. Puslitbang Sisjian sebagai bahan pertimbangan dalam melakukan kalibrasi paket soal EBTANAS baik tingkat SD, SLTP, dan SMU.
4. Bidang penelitian, sebagai titik tolak untuk penelitian lain mengenai penyetaraan skor dan penentuan metode penyetaraan yang tepat berdasarkan asumsi yang ada.

## **Kerangka Teori**

### **Metode Penyetaraan pada Teori Respons Butir**

Metode penyetaraan digunakan untuk penentuan konstanta konversi ( $\alpha$  dan  $\beta$ ). Langkah-langkah penentuan konstanta konversi dari masing-masing metode penyetaraan, adalah sebagai berikut:

#### **1) Metode Rerata dan Sigma**

Penentuan  $\alpha$  dan  $\beta$  dengan metode Rerata dan Sigma dengan memperhatikan nilai estimasi parameter butir tes yaitu  $b_x$  dan  $b_y$ . Menurut Hambleton dan Swaminathan (1985), hubungan antara estimasi parameter butir tes atau estimasi kemampuan peserta pada kedua perangkat tes yang akan disetarakan, memenuhi hubungan:

$$y = \alpha x + \beta$$

$$\bar{y} = \alpha \bar{x} + \beta$$

$$\alpha = \frac{s_y}{s_x}$$

$$\beta = \bar{y} - \alpha \bar{x}$$

keterangan:

$y$  : estimasi kemampuan atau estimasi parameter butir pada tes Y

$x$  : estimasi kemampuan atau estimasi parameter butir pada tes X

$\bar{y}, \bar{x}$  : rata-rata dari  $y$  dan  $x$

$s_x, s_y$  : standard deviasi dari  $x$  dan  $y$

Jika nilai  $\alpha$  dan  $\beta$  sudah ditetapkan, konversi parameter kemampuan  $\theta^*$  dan parameter item, ditempatkan dalam skala yang sama dengan menggunakan hubungan:

## *Komparasi Metode Penyetaraan Tes Menurut Teori Respons Butir*

$$\theta^*_Y = \alpha\theta_X + \beta$$

$$b^*_Y = \alpha b_X + \beta$$

$$a^*_Y = \frac{a_X}{\alpha}$$

Keterangan:

- $\theta^*_Y$  : kemampuan peserta tes pada tes Y sebagai kriteria
- $\theta_X$  : kemampuan peserta tes pada tes X sebagai prediktor
- $b^*_Y$  : tingkat kesukaran butir tes Y sebagai kriteria
- $b_X$  : tingkat kesukaran butir tes X sebagai prediktor
- $a^*_Y$  : diskriminasi butir tes Y sebagai kriteria
- $a_X$  : diskriminasi butir tes X sebagai prediktor

Metode Rerata dan Sigma bersifat timbal balik sehingga dengan cara yang sama hubungan dari y ke x dapat ditentukan.

### **2) Metode Rerata dan Sigma Tegar**

Prosedur penyetaraan metode Rerata dan Sigma Tegar dikembangkan oleh Linn, Levine, Hastings, dan Wardrop (Hambleton dan Swaminathan, 1991). Langkah-langkah dalam penentuan konstanta konversi dengan menggunakan metode Rerata dan Sigma Tegar adalah sebagai berikut:

a) Penentuan bobot parameter item ( $w_i$ ), pada setiap pasangan ( $b_{xi}, b_{yi}$ ), yaitu:

$$w_i = [\text{maks}\{v(x_i), v(y_i)\}]^{-1} \quad i = 1, 2, 3, \dots, k$$

$v(x_i)$  dan  $v(y_i)$  adalah varians estimasi parameter tingkat kesukitan tes X dan Y

b) Penentuan penskalaan bobot skala  $w'_i$  dengan menggunakan rumus berikut:

$$w'_i = \frac{w_i}{\sum_{j=1}^k w_j}$$

k : jumlah item anchor pada perangkat tes X dan Y

- c) Penghitungan estimasi berbobot tes X dan Y, dengan menggunakan rumus:

$$\begin{aligned} \bar{x} &= w_i x_i \\ \bar{y} &= w_i y_i \end{aligned}$$

- d) Penentuan rerata dan simpangan baku dari estimasi berbobot tes X dan Y, yaitu  $\bar{x}, \bar{y}, s_x, s_y$

- e) Penentuan konstanta konversi  $\alpha$  dan  $\beta$  dengan menggunakan rerata dan simpangan baku estimasi berbobot, dilakukan dengan mensubstitusikan rerata dan simpangan baku bobot estimasi pada persamaan penyamaan skala.

Menurut Stocking dan Lord (Hambleton, 1985) metode penyetaraan Rerata dan Sigma Tegar dapat diperbaiki dengan jalan memperhatikan skor kelompok ekstrim. Langkah-langkah penentuan konstanta konversi pada kelompok ekstrim pada dasarnya seperti langkah a) sampai e), dilanjutkan langkah-langkah berikut:

- f) dengan menggunakan nilai  $\alpha$  dan  $\beta$  yang sudah ditentukan jarak pasangan  $(x_i, y_i)$  terhadap garis penyetaraan, adalah:

$$d_i = \frac{|(y_i - \alpha x_i - \beta)|}{\sqrt{\alpha^2 + \beta^2}}$$

- g) Jika M adalah median dari  $d_i$ , kemudian dilakukan perhitungan bobot Tukey dengan fungsi:

$$T = \begin{cases} \left[ 1 - \left( \frac{d_i}{6M} \right)^2 \right]^2 & \text{untuk } d_i < 6M \\ 0 & \text{untuk } d_i \geq 6M \end{cases}$$

- h) Pembobotan tiap-tiap pasangan  $(x_i, y_i)$ , dengan rumus:

$$u_i = w_i T \left\{ \sum_{i=1}^n w_i T_i \right\}$$

- i) Ulangi langkah c) dengan menggunakan  $u_i$  sebagai pengganti  $w_i$ , kemudian ditentukan  $\alpha$  dan  $\beta$  seperti langkah e).

j) Ulangi langkah f) sampai i) sampai didapatkan hasil  $\alpha$  dan  $\beta$  lebih kecil dari  $\alpha$  dan  $\beta$  yang sudah ditentukan.

### 3) Metode Kurva Karakteristik

Haebara (1980), menyatakan bahwa Metode penyetaraan Kurva Karakteristik memperhatikan hubungan antara: parameter-parameter daya beda, tingkat kesukaran, dan *true score* peserta tes butir tes yang akan disetarakan. *True Score* ( $\tau_{xa}$ ) dari peserta tes dengan kemampuan  $\theta_a$  yang merespons k item dalam perangkat tes X dan tes Y adalah:

$$\tau_{xa} = \sum_{i=1}^k p(\theta_a, b_{xi}, a_{xi}, c_{xi})$$

$$\tau_{ya} = \sum_{i=1}^k p(\theta_a, b_{yi}, a_{yi}, c_{yi})$$

Setiap item pada perangkat tes X dan Y memenuhi persamaan:

$$b_{yi} = \alpha b_{xi} + \beta$$

$$a_{yi} = \frac{a_{xi}}{\alpha}$$

$$c_{yi} = c_{xi}$$

$$\text{atau} \quad \alpha = \frac{a_{xi}}{a_{yi}}$$

$$\beta = b_{yi} - \alpha b_{xi}$$

Konstanta  $\alpha$  dan  $\beta$  dipilih sedemikian sehingga fungsi F seperti tertera di bawah ini mencapai nilai minimal.

$$F = \frac{1}{N} \sum_{a=1}^N (\tau_{xa} - \tau_{ya})^2$$

Keterangan:

F : fungsi dari  $\alpha$  dan  $\beta$ , yang menunjukkan ketidaksesuaian antara  $\tau_{xa}$  &  $\tau_{ya}$

N : jumlah peserta tes

$\tau_{xa}$  : true score peserta tes pada kemampuan a pada perangkat tes X

$\tau_{ya}$  : true score peserta tes pada kemampuan a pada perangkat tes Y

Untuk menentukan nilai minimal fungsi F, digunakan pendekatan numerik "Golden Section" (Chapra dan Canale, 1996; Susila, 1994).

**Metode Penelitian**  
**Populasi dan Sampel**

Populasi dalam penelitian ini ialah: 1) lima paket soal EBTANAS SLTP bidang studi Matematika tahun ajaran 1998/1999 di Propinsi Jawa Tengah, dan 2) 370.187 respons jawaban siswa yang menempuh EBTANAS SLTP bidang studi Matematika tahun ajaran 1998/1999 di Propinsi Jawa Tengah.

Teknik *purposive sampling* digunakan untuk menentukan sampel paket soal EBTANAS. Paket soal yang dipilih sebagai sampel adalah paket soal utama. Sampel respons jawaban siswa peserta EBTANAS ditentukan dengan teknik *cluster stratified proportional sistematik sampling*. Klaster dalam penelitian ini adalah kelompok siswa SLTP yang mengerjakan soal EBTANAS paket utama-1, paket utama-2, dan paket utama-3. Strata kelompok yaitu status SLTP Negeri dan Swasta. Proporsional dalam penarikan sampel sekolah digunakan proporsi 50%, karena proporsi sebenarnya dari populasi belum diketahui. Acak Sistematik dilakukan untuk penarikan sampel respons siswa dari siswa sekolah terpilih. Menurut Isaac & Michael (1984), penentuan ukuran sampel minimal yang berkaitan dengan estimasi proporsi dapat ditentukan dengan menggunakan rumusan:

$$n \geq \frac{\chi^2 N.P.(1 - P)}{d^2(N - 1) + \chi^2.P.(1 - P)} \dots\dots\dots (1)$$

**Keterangan:**

- n : ukuran sampel minimal
- N : populasi
- $\chi^2$  : harga Chi kuadrat untuk d.b. = 1
- p : proporsi, p= 0,5
- d : tingkat ketelitian (error), d =5%

Dengan menggunakan rumusan (1), dihasilkan banyaknya sampel penelitian seperti tertera pada Tabel 1 berikut:

Tabel 1.  
Banyaknya sampel penelitian di tiap-tiap klaster pada strata  
SLTP negeri dan swasta

| Klaster       | Strata      |             |             |
|---------------|-------------|-------------|-------------|
|               | Negeri      | Swasta      | Jumlah      |
| I             | 1410        | 1324        | 2734        |
| II            | 1415        | 1341        | 2756        |
| III           | 1426        | 1335        | 2761        |
| <i>Jumlah</i> | <i>4251</i> | <i>4000</i> | <i>8251</i> |

### **Teknik Analisis Data**

#### **1. Pengujian kesetaraan paket soal**

Tingkat kesetaraan ditentukan dengan melakukan pengujian keparalelan tes, yaitu: a) uji beda rerata dengan analisis varians, dan uji *scheffe* untuk uji pasang, b) uji homogenitas varians dengan uji *Levene's* (Glass dan Hopkins, 1984; Sudjana, 1983). Dua paket soal yang tidak paralel dilakukan proses penyetaraan.

#### **2. Estimasi Parameter Butir dengan analisis program *Bilog***

Estimasi besarnya parameter dilakukan dengan menggunakan paket program *Bilog*.

#### **3. Proses Penyetaraan**

Proses penyetaraan untuk penentuan konstanta konversi dilakukan dengan ketiga metode penyetaraan yang sudah ditentukan. Program yang digunakan dalam proses penyetaraan terlebih dahulu dibuat oleh peneliti dengan menggunakan bahasa Pascal.

#### **4. Keakuratan Metode Penyetaraan**

Keakuratan metode penyetaraan diidentifikasi dengan nilai kesalahan baku penyetaraan yang kecil. Menurut Kolen dan Brennan (1995), kesalahan

baku penyetaraan dengan *Equivalent-group design*, ditentukan dengan persamaan berikut:

$$\text{var} \left[ \hat{l}_y(x_i) \right] \cong \frac{\sigma^2(Y)[1-\rho(X, Y)]}{N} \left\{ 2 + [1+\rho(X, Y)] \left[ \frac{x_i - \mu(X)}{\sigma(X)} \right]^2 \right\} \dots\dots (2)$$

keterangan:

- Var  $[\hat{l}_y(x_i)]$  : Varians skor penyetaraan dari skor tes X ke skor tes Y
- N : banyaknya nilai estimasi parameter butir tes yang disetarakan
- $\rho(X, Y)$  : korelasi antara skor X dan Y
- $\sigma^2(Y)$  : Varians skor tes Y
- $\sigma(X)$  : simpangan baku skor tes X
- $\mu(X)$  : rata-rata skor tes X

## **Hasil Penelitian**

### **A. Hasil Analisis Data**

#### **1. Kesetaraan Paket Soal EBTANAS**

Hasil pengujian beda rerata, terjadi perbedaan rerata yang signifikan pada pasangan  $X_1 - X_2$  dan pasangan  $X_1 - X_3$  pada taraf signifikansi 5%. Oleh karena itu pasangan  $X_1 - X_2$  dan pasangan  $X_1 - X_3$  bersifat tidak paralel sehingga perlu proses penyetaraan. Pasangan  $X_2 - X_3$  memiliki perbedaan rerata yang tidak signifikan, sehingga hanya perlu pengujian homogenitas varians pada pasangan ini.

Hasil uji homogenitas varians, didapatkan hasil nilai statistik  $F=41,455$  dengan signifikansi 0,000. Hasil ini menunjukkan bahwa varians skor populasi ketiga paket utama soal EBTANAS tidak homogen. Pengujian homogenitas varians antara paket utama-2 dan paket utama-3 diperoleh nilai  $F = 0,925$  pada signifikansi 0,336, sehingga tidak terjadi perbedaan varians yang signifikan antara skor siswa di paket utama-2 dan paket utama-3. Oleh karena itu pasangan paket utama-2 dan paket utama-3 bersifat paralel sehingga tidak perlu proses penyetaraan.

**2. Estimasi Parameter Butir**

Pada Fase pertama analisis *Bilog*, butir nomor 6 dan 31 di paket utama-1, butir nomor 1 dan 31 di paket utama-2, serta butir nomor 21 di paket utama-3 tidak diikuti dalam proses analisis berikutnya karena mempunyai nilai biserial yang kurang dari 0,3. Butir dengan nilai biserial kurang dari 0,3 jika diikuti dalam proses analisis *Bilog* akan mengganggu proses analisis, yaitu munculnya nilai negatif yang tidak signifikan sehingga akan menjadikan keputusan bias (Mislevy, 1990).

Fase kedua didapatkan nilai parameter butir berdasarkan model Logistik dua parameter, yaitu parameter daya beda ( $a_i$ ) dan parameter tingkat kesukaran ( $b_i$ ). Hasil estimasi  $a_i$  berada dalam interval  $[0, 2]$ , sehingga keseluruhan butir mempunyai cukup kemampuan untuk membedakan tingkat kemampuan peserta. Hasil estimasi  $b_i$  berada dalam interval  $[-2, +2]$ , jadi tiap butir soal dapat mengukur *latent trait* sesuai dengan kemampuan peserta.

**3. Proses Penyetaraan**

Proses penyetaraan dilakukan pada nomor butir yang ada di kedua perangkat tes yang disetarakan, karena pada analisis *Bilog* terdapat butir-butir soal yang terpaksa dihilangkan. Penyetaraan soal paket utama-1 ke soal paket utama-2 terdiri dari 37 butir, yaitu butir nomor: 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, dan 40.

Penyetaraan soal paket utama-1 ke soal paket utama-3 terdiri dari 37 butir, yaitu butir nomor: 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, dan 40.

Setelah dilakukan proses penyetaraan, hasil konstanta konversi yang didapatkan pada masing-masing metode seperti tertera dalam Tabel 2 berikut:

Tabel 2.  
Konstanta Konversi Masing-Masing Metode Penyetaraan

| Metode                 | Paket1-paket2 |          | Paket1- paket3 |         |
|------------------------|---------------|----------|----------------|---------|
|                        | $\alpha$      | $\beta$  | $\alpha$       | $\beta$ |
| Rerata dan Sigma       | 1.19495       | 0.25766  | 1.49581        | 0.42366 |
| Rerata dan Sigma Tegar | 0.71710       | -0.00081 | 0.60018        | 0.00157 |
| Kurva karakteristik    | 1.05199       | 0.03012  | 1.05391        | 0.03232 |

Konstanta konversi yang didapatkan dari proses penyetaraan dapat digunakan untuk menentukan persamaan konversi dari soal paket utama yang disetarakan. Persamaan konversi yang didapatkan untuk tiap-tiap metode penyetaraan, yaitu:

**a. Metode Rerata dan Sigma**

$$b_{x2} = 1.19495 b_{x1} + 0.25766$$

$$a_{x2} = a_{x1}/1.19495$$

$$b_{x3} = 1.49581 b_{x1} + 0.42366$$

$$a_{x3} = a_{x1}/1.49581$$

$$\theta_{x2} = 1.19495 \theta_{x1} + 0.25766$$

$$\theta_{x3} = 1.49581 \theta_{x1} + 0.42366$$

**b. Metode Rerata dan Sigma Tegar**

$$b_{x2} = 0.71710 b_{x1} - 0.00081$$

$$a_{x2} = a_{x1}/0.71710$$

$$b_{x3} = 0.60018 b_{x1} - 0.00157$$

$$a_{x3} = a_{x1}/0.60018$$

$$\theta_{x2} = 0.71710 \theta_{x1} - 0.00081$$

$$\theta_{x3} = 0.60018 \theta_{x1} - 0.00157$$

**c. Metode Kurva Karakteristik**

$$b_{x2} = 1.05199 b_{x1} + 0.03012$$

$$a_{x2} = a_{x1}/1.05199$$

$$b_{x3} = 1.05391 b_{x1} + 0.03232$$

$$a_{x3} = a_{x1}/1.05391$$

$$\theta_{x2} = 1.05199 \theta_{x1} + 0.03012$$

$$\theta_{x3} = 1.05391 \theta_{x1} + 0.03232$$

Berdasarkan persamaan konversi yang diperoleh, dapat dilakukan konversi parameter butir dan parameter kemampuan. Konversi dapat dilakukan dari paket utama-1 ke paket utama-2 dan paket utama-3 atau sebaliknya. Oleh karena itu parameter butir dari paket soal yang berbeda dan tidak paralel dapat diketahui tingkat kesetaraannya.

Hasil penyetaraan dapat dijadikan pedoman untuk menentukan urutan tingkat kesukaran soal EBTANAS. Persamaan konversi yang dihasilkan karena proses penyetaraan dengan metode Rerata dan Sigma dan metode Kurva Karakteristik, didapatkan urutan tingkat kesukaran soal EBTANAS berturut-turut dari yang mudah ke yang sulit adalah soal paket utama-1, soal

### *Komparasi Metode Penyetaraan Tes Menurut Teori Respons Butir*

paket utama-2, dan soal paket utama-3. Persamaan konversi yang dihasilkan karena proses penyetaraan dengan metode Rerata dan Sigma Tegar urutan tingkat kesukaran soal EBANAS berturut-turut, dari yang mudah ke yang sulit, adalah: soal paket utama-2, soal paket utama-3, dan soal paket utama-1. Terjadi perbedaan hasil proses penyetaraan dengan metode penyetaraan yang berbeda.

Perbedaan tersebut terjadi karena adanya perbedaan keakuratan metode yang ditunjukkan oleh kesalahan standard penyetaraan (SEE) dari masing-masing metode penyetaraan

#### **4. Kesalahan Standard Penyetaraan (SEE)**

Hasil analisis SEE dengan menggunakan persamaan (2), seperti tertera dalam Tabel 3 berikut:

**Tabel 3.**  
**Hasil Analisis SEE**

| <b>Metode Penyetaraan</b> | <b>Proses Konversi</b>                      |   |
|---------------------------|---|---|
|                           | <b><math>X_1</math> ke <math>X_2</math></b> | <b><math>X_1</math> ke <math>X_3</math></b> |
| Rerata dan Sigma          | 0.030574                                    | 0.045906                                    |
| Rerata dan Sigma Tegar    | 0.04933                                     | 0.106431                                    |
| Kurva Karakteristik       | 0.032968                                    | 0.05585                                     |

Dengan memperhatikan nilai SEE dari masing-masing metode penyetaraan, tingkatan keakuratan metode penyetaraan berturut-turut dari yang paling akurat adalah metode: Rerata dan Sigma, Kurva Karakteristik, dan Rerata dan Sigma Tegar.

Tingkat konsistensi keakuratan metode penyetaraan, ditunjukkan dari nilai korelasi SEE masing-masing metode. Nilai-nilai korelasi SEE dari metode penyetaraan yang berbeda, seperti tertera dalam Tabel 4 berikut:

**Tabel 4.**  
**Nilai-Nilai Korelasi SEE dari Metode Penyetaraan yang Berbeda**

| <i>Korelasi antara metode</i>    | <i>Penyetaraan<br/><math>X_1</math> ke <math>X_2</math></i> | <i>Penyetaraan<br/><math>X_1</math> ke <math>X_3</math></i> |
|----------------------------------|---|---|
| Rerata dan Sigma, T.Rerata       | 0,99819   | 0,981987  |
| T.Rerata, K.Karakteristik        | 0,934566  | 0,860837  |
| Rerata dan Sigma, K.Karakteistik | 0,941162  | 0,941536  |

### **Kesimpulan, Saran, dan Implikasi**

#### **Kesimpulan**

Berdasarkan temuan hasil penelitian, beberapa kesimpulan mengenai komparasi metode penyetaraan tes menurut Teori Respons Butir, sebagai berikut:

1. Secara empiris terbukti tiga perangkat soal EBTANAS bidang studi Matematika SLTP Tahun Ajaran 1998/1999, merupakan perangkat soal yang tidak paralel. Secara berpasangan soal EBTANAS paket utama-1 dan soal EBTANAS paket utama-2, soal EBTANAS paket utama-1 dan soal EBTANAS paket utama-3, bukan perangkat soal yang paralel. Konsekuensi perangkat tes yang tidak paralel ialah perlunya proses penyetaraan pada pasangan perangkat soal yang tidak paralel.
2. Tingkat kesetaraan dari perangkat tes yang secara empirik tidak paralel menurut proses penyetaraan pada masing-masing metode penyetaraan, yaitu:
  - a. Rerata dan Sigma, diperoleh tingkat kesetaraan parameter kesulitan butir dan parameter peserta di paket utama-1, paket utama-2, dan paket utama-3 berturut-turut sebesar 0; 0,25766; dan 0,42366.
  - b. Rerata dan Sigma Tegar, diperoleh tingkat kesetaraan parameter kesulitan butir dan parameter peserta di paket utama-1, paket utama-2, dan paket utama-3 berturut-turut sebesar 0; -0,00081; dan -0,00157.
  - c. Kurva Karakteristik, diperoleh tingkat kesetaraan parameter kesulitan butir dan parameter peserta di paket utama-1, paket utama-2, dan paket utama-3 berturut-turut sebesar 0; 0,03012; dan 0,03232.

### *Komparasi Metode Penyetaraan Tes Menurut Teori Respons Butir*

3. Nilai konversi penyetaraan paket utama-1 ke paket utama-2 dan paket utama-3, pada tiap-tiap metode penyetaraan dinyatakan dalam persamaan konversi berikut:
  - a. Rerata dan Sigma diperoleh persamaan konversi:  
$$X_2 = 1,19495 X_1 + 0,25766 \text{ dan } X_3 = 1,49581 X_1 + 0,42366$$
  - b. Rerata dan Sigma Tegar diperoleh persamaan konversi:  
$$X_2 = 0,71710 X_1 - 0,00081 \text{ dan } X_3 = 0,60018 X_1 - 0,00157$$
  - c. Kurva Karakteristik diperoleh persamaan konversi:  
$$X_2 = 1,05199 X_1 + 0,03012 \text{ dan } X_3 = 1,05391 X_1 + 0,03012$$
4. Penyetaraan dengan metode Rerata dan Sigma, dan Kurva Karakteristik, diperoleh urutan tingkat kesukaran soal dari yang paling mudah ke yang paling sukar adalah perangkat EBTANAS paket: utama-1, utama-2, dan utama-3.
5. Penyetaraan dengan metode Rerata dan Sigma Tegar, menyimpulkan urutan tingkat kesukaran soal dari yang paling mudah ke yang paling sukar adalah perangkat EBTANAS paket: utama-2, utama-1, dan utama-3.
6. Berdasarkan nilai SEE diperoleh urutan tingkat keakuratan metode penyetaraan dari yang paling akurat adalah: metode Rerata dan Sigma, metode Kurva Karakteristik, dan metode Rerata dan Sigma Tegar.

### **Saran**

Berdasarkan beberapa kesimpulan penelitian, dikemukakan saran-saran sebagai berikut:

1. Perlunya pengujian kesetaraan paket soal EBTANAS bidang studi Matematika SLTP Tahun Ajaran 1998/1999 dan perangkat tes yang lebih dari satu paket soal, tidak terbatas pada perangkat soal EBTANAS saja, misalnya UMPTN, dan tes Cawu.
2. Proses penyetaraan paket soal EBTANAS bidang studi Matematika SLTP Tahun Ajaran 1998/1999, hendaknya dilakukan dengan metode Rerata dan Sigma. Proses penyetaraan untuk data berbeda, sebaiknya disesuaikan dengan asumsi metode penyetaraan yang tepat sehingga kesalahan proses penyetaraan minimal.

3. Parameter butir soal EBTANAS bidang studi Matematika Tahun Ajaran 1998/1999 yang akan dimasukkan dalam pembuatan "Bank Soal" hendaknya melalui penyetaraan Rerata dan Sigma, dan untuk data yang berbeda sebaiknya diperhatikan asumsi metode dan karakteristik data. Pengadaan "Bank Soal" akan mempermudah dalam pengadaan soal yang terkendali kualitasnya sekaligus dapat menghemat biaya.
4. Perlunya penelitian lebih lanjut dengan menggunakan data berbeda untuk mengetahui konsistensi keakuratan metode penyetaraan.

### **Implikasi**

Berdasarkan hasil penelitian dan pengadaan program komputer untuk penyetaraan tes dapat diimplikasikan dalam bidang:

1. Program komputer yang dibuat untuk penyetaraan tes, akan membantu dan mempermudah dalam pengadaan bank soal baik di tingkat daerah maupun tingkat pusat.
2. Di bidang pengukuran, proses penyetaraan dapat dikendalikan keakuratannya. Metode penyetaraan yang tepat dan cocok dengan data akan memperkecil penyetaraan sehingga pengambilan keputusan menjadi tepat.

### **Daftar Pustaka**

- Angoff. (1982). *Summary and derivation of equating methods useds at ETS*, dalam Holland, Pul, W. dan Rubin, Donald. *Test equating*. New York: Academic press Inc.
- Chapra, S.C., dan Canale, R.P. (1996). *Numerical methods for engineers*, (2<sup>nd</sup> ed.) New York: McGraw-Hill Inc.
- Depdikbud. (1998). *Pedoman pelaksanaan kegiatan penulisan usul soal EBTANAS SD/MI, SLTP/MTS, SMU/MA, dan SMK tahun pelajaran 1998/1999*. Semarang: Kanwil Depdikbud Propinsi Jawa Tengah.

*Komparasi Metode Penyetaraan Tes Menurut Teori Respons Butir*

- Glass, G.V., & Hopkins, K.D. (1984). *Statistical methods in education and psychology*. New Jersey: Prentice-Hall.
- Hambleton, R.K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff Publishing.
- Hambleton, R.K., Swaminathan, H., dan Rogers, H.J. (1991). *Fundamental of item response theory*. London: Sage Publications Inc.
- Hayat, Bahrul dan Surya Paranata. (1995). *Analisis dan kalibrasi soal EBTANAS SMP tahun pelajaran 1993/1994*. Jakarta: Puslitbang Sisjian-Depdikbud.
- Holland, P.W., & Rubin, D.B. (1982). *Test equating*. New York: Academic Press Inc.
- Isaac, Stephen, & Michael, William B. (1984). *Handbook in research an evaluation* (2<sup>nd</sup> ed). San Diego: Edits Publisher.
- Kolen, M.J., & Brennan, R.L. (1995). *Test equating: methods and practices*. New York: Verlag.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates Company.
- Mislevy, R.J. (1990). *Bilog 3, item analysis and test scoring with binary logistic models* (2<sup>nd</sup> edition). Chicago: Mooresville.
- Naga, Dali. S. (1992). *Pengantar teori sekor pada pengukuran pendidikan*. Jakarta: Gunadarma.
- Susila, Nyoman. (1994). *Dasar-dasar metode numerik*. Jakarta: Depdikbud Dirjen Dikti.
- Suryabrata, Sumadi. (1987). *Pengembangan tes hasil belajar*. Jakarta: Rajawali Press.
- Sudjana. (1983). *Dasar-dasar statistika*. Bandung: Tarsito.