

## **APLIKASI MODEL RASCH CAMPURAN DALAM MENGEVALUASI PENGUKURAN HARGA DIRI**

*Wahyu Widbiarso*

Universitas Gadjah Mada

Jl. Humaniora No. 1 Bulaksumur 55281

wahyu\_psy@ugm.ac.id

### **Abstrak**

Penelitian ini bertujuan untuk mengeksplorasi keberadaan kelompok responden yang menyebabkan estimasi parameter butir melalui pemodelan Rasch tidak invarian pada keseluruhan responden. Teknik analisis yang dipakai adalah model Rasch campuran yang merupakan penggabungan antara Model Rasch dan analisis kelas laten. Dengan menggunakan data hasil pengukuran harga diri didapatkan hasil analisis bahwa keseluruhan responden penelitian sebanyak 2.987 dapat dikategorikan menjadi tiga kelas berdasarkan pola respons mereka pada skala. Hasil estimasi parameter butir pada responden pada masing-masing kelas dengan menggunakan model kredit parsial menunjukkan bahwa ketiga kelas memiliki parameter butir yang berbeda. Dua kelas relatif sesuai dengan model, sedangkan satu kelas tidak sesuai karena responden pada kelas tersebut merespons skala dengan cara yang unik. Keberadaan responden dengan respons unik ini relatif kecil (12,5%) sehingga tidak mengganggu estimasi parameter pada keseluruhan butir.

**Kata kunci:** *model Rasch campuran, parameter butir, kelas responden*

## **RASCH MIXED MODEL APPLICATION IN EVALUATING THE MEASUREMENT OF SELF-ESTEEM**

*Wahyu Widhiarso*

Universitas Gadjah Mada

Jl. Humaniora No. 1 Bulaksumur 55281

wahyu\_psy@ugm.ac.id

### **Abstract**

This study aimed to explore the existence of groups of items that cause Estimation of item parameters using Rasch modeling was not invariant for all respondents. Mixed Rasch model which is the combination between Rasch Models and Latent Class Analysis was employed. By using data from measuring self-esteem found for overall respondents (N=2987) can be categorized into three classes based on their item respons patterns on entire scale. Results based on estimation of item parameters to the respondents in each class using the Partial Credit Model found that each class has different item parameters. Two classes supported the model while the other class did not; due to respondents on this class give a response on the scale in a unique way. The proportion of the respondents with a unique response is relatively small (12,5%) therefore they do not much interfere the estimation of item parameters on the overall items.

**Keywords:** *mixed rasch model, Item Estimation Parameter, Class*

## **Pendahuluan**

Proses mengevaluasi properti psikometris butir pengukuran dalam bidang ilmu sosial dengan menggunakan pemodelan Rasch kadang dianggap terlalu ketat. Hal tersebut dikarenakan proses pengukuran tidak lepas dari pengaruh-pengaruh luar, misalnya heterogenitas karakteristik individu dan interaksi individu dengan tes atau skala (Preinerstorfer & Formann, 2011). Heterogenitas karakteristik individu dapat berupa perbedaan cara individu dalam memahami butir di dalam skala. Misalnya, ketika merespons pernyataan di dalam skala psikologi, ada individu yang cenderung untuk menyetujui semua pernyataan yang diberikan (*yea saying*), yaitu dengan memilih kategori respons (e.g. “sesuai” atau “sangat sesuai”) dan ada juga individu yang cenderung memilih kategori respons tengah (e.g. “netral”). Interaksi individu dengan tes atau skala dapat ditunjukkan dengan kemungkinan adanya individu yang memilih kategori respons yang dinilai olehnya sebagai respons yang sesuai dengan kepatutan sosial (*social desirability*). Penelitian telah menunjukkan bahwa respons yang mengandung kepatutan sosial seringkali muncul ketika individu menghadapi pengukuran yang bersifat sensitif terhadap diri individu, misalnya pengukuran perilaku seks (van de Mortel, 2008) atau pengonsumsi alkohol (Cox, Swinson, Drenfeld, & Bourdeau, 1994).

Ketatnya persyaratan yang harus dipenuhi menyebabkan banyak ditemui peneliti yang mendapati butir-butir yang dianalisis tidak sesuai (*fit*) dengan model Rasch (Goldstein & Blinkhorn, 1982). Ketika peneliti mendapatkan datanya tidak cukup mendukung untuk didekati dengan model Rasch ada dua strategi yang biasa dipakai (Smit, Kelderman, & van der Flier, 1999). Pertama, menerapkan pendekatan teori respons butir (IRT) yang lebih kompleks. Kedua, mencoba untuk memisahkan sampel menjadi subkelompok yang dapat diskalakan dengan model Rasch. Dengan menggunakan pendekatan ini, heterogenitas dalam populasi dimodelkan dengan mengidentifikasi subkelompok yang lebih homogen agar sesuai dengan model pengukuran yang dipakai.

Persoalan lain yang perlu diperhatikan dalam pengukuran dalam bidang ilmu sosial adalah masalah konsistensi perilaku manusia. Bem dan

Allen (1974) menunjukkan bahwa pengukuran ciri kepribadian mungkin memiliki kekuatan prediksi yang tergantung pada bagaimana konsistensi individu dalam menilai perilaku mereka yang terkait dengan ciri kepribadian. Beberapa ahli mengatakan bahwa konsistensi perilaku manusia perlu dipertimbangkan agar pertanyaan apakah struktur ciri kepribadian yang diwujudkan dalam faktor kepribadian didapatkan dari inferensi analisis faktor berlaku pada semua individu pada populasi atau tidak (Rost, Carstensen, & von Davier, 1997).

Tulisan ini mendemonstrasikan penerapan Model Rasch Campuran (*Mixed Rasch Modeling*/MRM) yang dikembangkan oleh Rost (1988, 1990) untuk mengeksplorasi adanya subkelompok responden yang memiliki pola respons yang berbeda dalam pengukuran ciri kepribadian. Pola respons tersebut ditinjau berdasarkan model kredit parsial (PCM) dari Masters (1982) yang merupakan varian dari Model Rasch yang tepat untuk diaplikasikan pada data politomi. Ciri kepribadian yang diukur dalam penelitian ini adalah harga diri. Hal yang dieksplorasi adalah keberadaan kelompok responden yang menyebabkan estimasi parameter butir pada skala pengukuran kurang berlaku pada keseluruhan responden.

Secara umum, IRT mengasumsikan bahwa parameter butir untuk setiap butir adalah konstan untuk semua responden (Embretson & Reise, 2000). Ketika parameter butir tersebut tidak konstan untuk responden yang berbeda karakteristiknya, butir tersebut dinamakan butir yang terjangkit keberfungsian butir diferensial (DIF). Biasanya, prosedur investigasi butir yang terjangkit DIF dilakukan pada kelompok responden berdasarkan pendekatan apriori, bukan pendekatan posteriori. Dengan kata lain, kelompok responden tersebut ditentukan terlebih dahulu (e.g. gender atau suku) sebelum uji DIF dilakukan (Osterlind & Everson, 2009).

Perkembangan terbaru menunjukkan bahwa pengujian DIF berdasarkan kelompok posteriori memungkinkan untuk dilakukan (Eid & Zickar, 2007). Karena kelompok tersebut adalah kelompok yang belum ditentukan, kelompok tersebut adalah kelompok yang bersifat laten. Jika kelompok responden telah ditentukan sebelum uji DIF dilakukan maka pendekatan tersebut dinamakan dengan uji DIF konfirmatori, sedangkan ketika kelompok tidak ditentukan, pendekatan tersebut dinamakan uji DIF

eksploratori. Sama seperti analisis faktor eksploratori yang memunculkan faktor laten yang diturunkan dari seperangkat butir, pendekatan ini menghasilkan pembagian kelompok-kelompok responden. Jika analisis faktor memetakan butir yang memiliki kesamaan atribut ukur, pendekatan ini memetakan responden berdasarkan pola responsnya (*respons pattern*) terhadap butir di dalam skala.

Ada dua pendekatan yang mengakomodasi pendekatan posteriori dalam menguji konsistensi parameter butir pada sejumlah responden, yaitu Pemodelan Rasch Campuran (*Mixed Rasch Modeling*/MRM) dan Pemodelan Hibrida Rasch-Kelas Laten (*Hybrid Rasch-Latent Class Modeling*). MRM dikembangkan oleh Rost (1990) yang mengintegrasikan antara model Rasch dan analisis kelas laten dalam satu model. Model ini mengasumsikan bahwa model Rasch berlaku untuk setiap distribusi kelompok responden yang diturunkan dari campuran pemodelan Rasch dan analisis kelas laten. Model Hibrida dikembangkan oleh Yamamoto (1989) yang mengintegrasikan model kelas laten dan IRT yang memiliki perspektif yang berbeda dengan MRM. Bedanya, model ini mengasumsikan bahwa distribusi pemodelan Rasch berlaku pada satu kelompok responden sedangkan distribusi pemodelan analisis kelas laten berlaku pada distribusi lainnya. Model Rasch Campuran dijalankan dengan persamaan di bawah ini.

$$p(x_{vi} = 1|g) = \frac{\exp(\theta_{vg} - \rho_{ig})}{1 + \exp(\theta_{vg} - \rho_{ig})}$$

Keterangan

$\theta_{vg}$  = kemampuan subjek  $v$ , dengan asumsi subjek tersebut masuk dalam kelas  $g$

$\rho_{ig}$  = tingkat kesukaran butir untuk butir  $i$  di dalam kelas  $g$

Persamaan tersebut menunjukkan bahwa butir-butir dalam skala mengukur kemampuan yang unidimensional pada setiap kelas namun kelas-kelas tersebut berbeda dalam hal urutan peringkat dan tingkat kesulitan butir. Dengan demikian, dimensi-dimensi kemampuan yang bersifat laten tersebut diestimasi pada setiap kelas. Dari persamaan ini

dapat diketahui bahwa Model Rasch yang standar adalah identik dengan model hanya mengestimasi satu kelas saja.

MRM telah diaplikasikan dalam banyak penelitian. Dengan menerapkan MRM pada pengukuran kemampuan spasial, Rost (1997) menemukan bahwa penyelesaian tes kemampuan spasial hanya dapat dilakukan dengan strategi kognitif yang memanfaatkan rotasi mental stimulus. Analisis MRM menemukan adanya kelompok-kelompok individu yang menggunakan strategi berbeda dalam mengatasi soal yang diberikan. Individu yang menggunakan selain strategi kognitif yang memanfaatkan rotasi mental stimulus memiliki kemungkinan yang besar akan gagal dalam mengatasi butir di dalam tes. Dengan kata lain, perbedaan penggunaan strategi kognitif yang dimiliki individu mempengaruhi parameter tingkat kesulitan tes.

### **Metode Penelitian**

Data diambil dari penelitian yang dilakukan oleh Retnowati (2004) yang melibatkan remaja sekolah antara usia 15 tahun hingga 19 tahun. Rata-rata usia partisipan adalah 16 tahun dengan jumlah 2.987 orang yang terdiri dari 46% pria dan 54% wanita. Partisipan adalah siswa SLTP, SMU, SMK dan PSBR di Daerah Istimewa Yogyakarta dari empat wilayah antara lain Kabupaten Sleman, Bantul, Kulon Progo, Gunung Kidul dan Kodya Yogyakarta. Masing-masing Kabupaten dan Kodya Yogyakarta, diwakili oleh beberapa SMU dan SMK.

Alat ukur yang dipakai untuk mengukur harga diri dalam penelitian ini adalah skala harga diri yang diadaptasi dari skala harga diri yang dikembangkan oleh Rosenberg (1965). Skala ini memuat 10 butir yang menggunakan format Likert dengan 4 kategori yang diskor 1 hingga 4. Kategori respons yang diberikan adalah: sangat setuju; setuju; tidak setuju dan sangat tidak setuju. Pernyataannya antara lain: “Secara keseluruhan saya puas dengan diri saya”; “*Saya pikir saya sama sekali tidak baik*”. Hasil analisis butir skala harga diri menunjukkan dari 10 butir yang diujicobakan, korelasi butir total berkisar antara 0,2581–0,3917 dengan koefisien alpha sebesar 0,8689.

Tahap-tahap yang dilakukan untuk menganalisis data adalah sebagai berikut:

- (1) Mengidentifikasi jumlah kelas. Pada tahap ini berbagai model dengan jumlah kelas yang berbeda (1 hingga 4 kelas) dipakai untuk memodelkan data. Setiap model akan menghasilkan kriteria informasi yang menunjukkan ketepatan model dengan data. Penelitian ini menggunakan kriteria informasi Bayesian (*Bayesian Information Criterion/BIC*) sebagai dasar untuk menentukan model yang tepat karena kriteria ini mengakomodasi ukuran sampel dan menghindari parameterisasi yang berlebih.
- (2) Mengidentifikasi profil tiap kelas. Setelah jumlah kelas ditentukan, tahap selanjutnya adalah menjabarkan profil tiap kelas dan membandingkan kekhasan tiap kelas berdasarkan profil parameter butirnya.

Analisis data dilakukan dengan bantuan program WINMIRA (von Davier, 2001).

## Hasil Penelitian dan Pembahasan

### Statistik Deskripsi Skor

Tabel 1. Perbandingan Statistik Skor dari Skala dan Responden

Sumber	Skor Minimal	Skor Maksimal	Rerata	Deviasi. Std.
Skala (distribusi hipotetis)	10	40	25	5
Responden (distribusi empiris)	13,00	37,00	27,06	3,14

Tabel 1 menunjukkan perbandingan statistik skor yang didapatkan dari skala dan responden. Statistik skala yang merupakan distribusi hipotetis, sedangkan statistik responden didapatkan dari distribusi empiris

dari keseluruhan responden. Jumlah butir skala dalam penelitian ini adalah 10 butir dengan rentang skor butir antara 1 hingga 4. Dengan demikian skor minimal pada distribusi hipotetis adalah 10 ( $1 \times 10$ ) dan skor maksimalnya adalah 40 ( $4 \times 10$ ). Rerata skor distribusi hipotetis didapatkan dari luas interval skor dan deviasi standar didapatkan dari luas interval skor dibagi 6 (menyesuaikan standar deviasi distribusi normal).

Hasil perbandingan menunjukkan bahwa tidak ada responden yang mendapatkan skor minimal dan maksimal. Rerata distribusi skor empiris sedikit lebih tinggi dibanding dengan distribusi skor hipotetis namun masih dalam interval yang setara karena selisihnya tidak lebih dari  $+1SD$ . Di sisi lain, standar deviasi skor empirik masih di bawah distribusi skor hipotetis. Perbandingan ini menunjukkan bahwa distribusi skor pengukuran harga diri pada responden sesuai dengan distribusi hipotetisnya namun sedikit kurang bervariasi.

#### Perbandingan Antarmodel

Dari hasil perbandingan antara model dari model 1 kelas hingga 4 kelas diperoleh bahwa pembagian responden berdasarkan pola respons pada skala lebih cenderung mengarah pada model 3 kelas. Nilai BIC mengalami penurunan yang cukup besar ketika model berubah dari Model-1 ke Model-2 serta Model-2 ke Model-3 Kelas, namun tidak cukup ketika berubah dari Model-3 ke Model-4 (lihat Tabel 2). Pola ini juga terjadi pada indeks lainnya, yaitu indeks *Akaike Information Criterion* (AIC) dan CAIC.

Tabel 2. Perbandingan Indeks Ketepatan Model Berdasarkan Jumlah Kelas

Indeks Ketepatan Komparatif	Model 1 Kelas	Model 2 Kelas	Model 3 Kelas	Model 4 Kelas
Indeks BIC	54305,90	52237,68	51890,59	51950,09
Indeks AIC	54119,84	51859,55	51320,40	51187,84

### Profil Kelas pada Model 3 Kelas

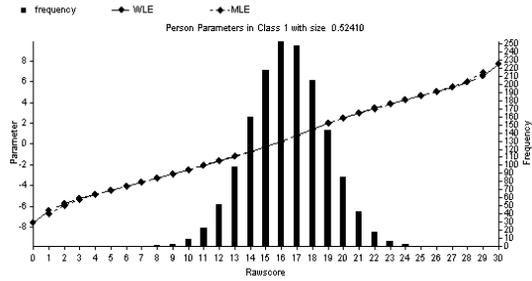
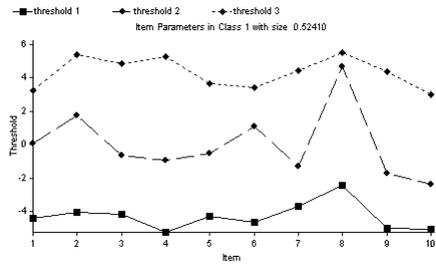
Setelah model 3 kelas disimpulkan yang hasilnya lebih mendukung data penelitian, langkah selanjutnya adalah mengidentifikasi profil model tersebut. Model 3 kelas mengkategorikan 53% responden dalam kelas 1 dengan probabilitas untuk menjadi anggota kelas yang cukup besar ( $p=0,93$ ). Hal ini menunjukkan bahwa dengan pola yang sama responden di dalam kelas ini memiliki kemungkinan kecil untuk masuk dalam kelas lainnya. Misalnya probabilitas untuk masuk menjadi anggota Kelas-2 adalah 0,05. Data selengkapnya dapat dilihat pada Tabel 3.

Tabel 3. Profil Model 3 Kelas

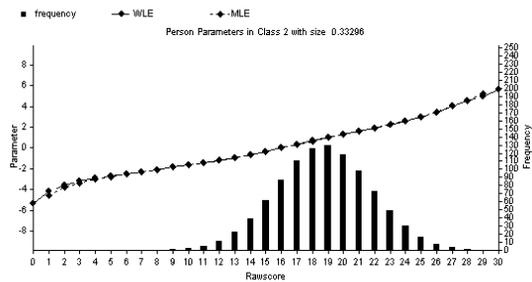
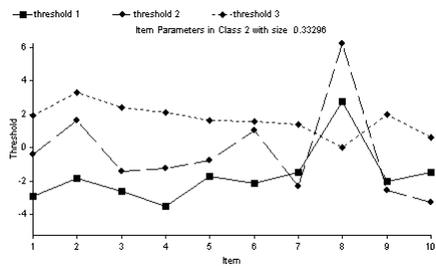
Kelas	Frekuensi Anggota Kelas (%)	Rerata Probabilitas	Probabilitas		
			Kelas 1	Kelas 2	Kelas 2
1	0,53	0,93	0,93	0,05	0,03
2	0,35	0,85	0,08	0,85	0,07
3	0,13	0,83	0,07	0,10	0,83

Profil tiap kelas dan distribusi skor responden yang masuk dalam kelas tersebut disajikan pada Gambar 1. Gambar di sisi kiri menunjukkan parameter ambang butir pada kesepuluh butir skala. Setiap butir memiliki 3 parameter ambang karena jumlah kategori respons skala berjumlah 4 kategori (politomi). Gambar di sisi kanan menunjukkan distribusi skor responden yang masuk dalam kelas tersebut. Dengan membandingkan profil ketiga kelas di dalam model dapat dipaparkan informasi beberapa temuan sebagai berikut.

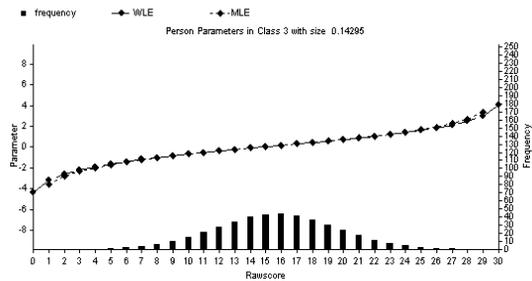
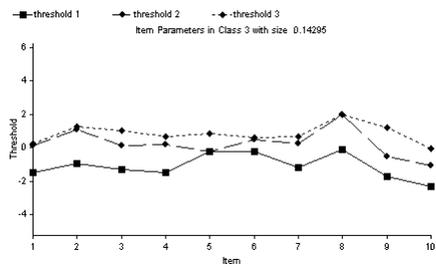
- (1) *Kelas-1*. Kelas-1 berisi responden yang merespon skala dengan cara sesuai dengan model karena nilai parameter-parameter ambang yang dihasilkan memiliki jarak interval antarparameter yang setara dan sesuai dengan urutannya. Parameter ambang butir pada butir 8 merupakan pengecualian karena nilai parameter ambang 2 hampir berhimpitan dengan parameter ambang 3.



(a) Kelas-1



(b) Kelas-2



(c) Kelas-3

Gambar 1. Profil Parameter Ambang Butir dan Distribusi Skor Tiap Kelas

(2) *Kelas-2*. Kelas-2 berisi responden yang mendukung model karena jarak interval antarparameter ambang konstan yang didukung dengan nilai ambang yang sesuai dengan urutannya. Pada butir tertentu (i.e butir 1,

2, 10) jarak antarintervalnya setara, akan tetapi pada butir lainnya jarak antarintervalnya tidak setara (i.e. butir 3, 5, 6). Di sisi lain ditemukan butir yang menghasilkan parameter ambang yang tidak sesuai dengan urutannya (yaitu butir 7, 8, 9).

- (3) *Kelas-3*. Kelas-3 berisi responden yang kurang mampu membedakan kategori respons pada skala sehingga menghasilkan parameter ambang butir yang tidak sesuai dengan model. Nilai parameter ambang pada dalam satu butir memiliki besaran yang hampir setara yang didukung dengan adanya beberapa butir yang memiliki nilai parameter ambang yang tidak sesuai dengan urutannya.

Melalui perbandingan antardistribusi skor responden tiap kelas dapat disimpulkan bahwa responden di dalam tiap kelas memiliki distribusi skor berbeda. Responden pada Kelas-1 memiliki skor skala yang bergerak antara 9 hingga 24, Kelas-2 bergerak antara 9 hingga 28 sedangkan Kelas-3 bergerak antara 5 hingga 26. Hasil analisis varians satu jalur menunjukkan ada perbedaan skor yang signifikan antarkelas ( $F=273,10$ ;  $p<0,001$ ). Hasil uji lanjutan (*post-hoc*) melalui Uji Tukey menunjukkan bahwa setiap pasangan perbandingan kelas berbeda antara satu dengan lainnya. Responden pada Kelas-2 memiliki skor yang lebih tinggi secara signifikan dibanding dengan dua kelas lainnya.

Tabel 4. Uji Perbandingan Skor Antarkelas

Kelas	N	Rerata	Deviasi Standar	Sesatan Standar	F
1	1571	26,381	2,446	0,062	273,10 ( $p<0,001$ )
2	1044	28,693	3,091	0,096	
3	372	25,379	3,872	0,201	

Perbandingan pola respons responden pada tiap kelas dapat dilihat pada Tabel 5. Responden pada Kelas-1 mendapatkan skor butir yang setara. Kelas-2 memiliki profil pola respons yang kurang lebih sama dengan Kelas-1, perbedaannya hanya pada butir tertentu responden tidak men-

dapatkan skor yang setara. Skor butir individu pada butir-butir skala relatif bervariasi. Pada butir tertentu mendapatkan skor maksimal (4), namun pada butir lainnya mendapatkan skor minimal (1). Responden pada Kelas-3 memiliki skor butir yang paling tidak konsisten dibanding dengan kelas lainnya yang terlihat dari variasi skor butir yang didapatkan responden sangat besar.

Tabel 5. Contoh Pola Respons Pada Tiap Kelas

Kelas-1		Kelas-2		Kelas-3	
Pola Respons	Skor	Pola Respons	Skor	Pola Respons	Skor
2212222132	19	2223113133	21	1122111211	13
3223323233	26	3223122133	22	3211223323	22
3332223233	26	2234223234	27	2222221244	23
2322233234	26	3234223243	28	3243222213	24
3233333233	28	4233422134	28	3333412232	26
3343444333	34	4444444144	37	4444414144	34

Penelitian ini bertujuan untuk mengeksplorasi keberadaan kelompok responden yang memiliki cara yang berbeda dalam merespon skala. Cara tersebut terlihat dari pola respons terhadap seperangkat butir di dalam skala. Individu yang memiliki pola respons yang sama akan masuk menjadi kelompok individu yang dinamakan dengan kelas. Hasil analisis menunjukkan bahwa pola respons individu pada skala terbagi menjadi tiga kelas. Dua kelas relatif konsisten dengan model (e.g. Kelas-1 dan Kelas-2) sedangkan satu kelas kurang konsisten (e.g. Kelas-3). Dari keanggotaan kelas dapat diketahui bahwa proporsi individu dalam kelas yang tidak konsisten sebesar 12 persen.

Hasil penelitian ini mengindikasikan bahwa terdapat responden yang memberikan respons dengan cara yang unik karena tidak konsisten dalam merespons butir dalam skala. Pada butir tertentu dia cenderung menyetujui pernyataan (mendapatkan skor butir yang tinggi) yang diajukan sedangkan pada butir lainnya cenderung menolak (mendapatkan skor butir yang rendah). Pola respons dari responden ini dikatakan unik atau tidak

konsisten karena butir-butir skala harga diri bersifat unidimensi dan memiliki presisi ukur yang setara.

Berdasarkan beberapa hasil penelitian, ketidakkonsistenan responden ini dapat diakibatkan oleh beberapa hal. Pertama, responden kurang mampu memahami pernyataan yang diberikan. Peacock, Ervin, & Daly (2009) melihat bahwa individu yang memiliki keterbatasan kemampuan penalaran akan kesulitan memahami butir pernyataan di dalam skala. Akibat kesalahan memahami pernyataan tersebut pola respons mereka kurang bisa dimodelkan. Kedua, responden memiliki strategi tertentu dalam merespons butir pernyataan, misalnya respons yang mengandung kepatutan sosial. Pada butir yang dinilai sensitif, mereka akan berusaha menunjukkan bahwa mereka adalah individu yang ideal dalam kaca mata sosial, namun pada butir yang tidak terlalu sensitif mereka memberikan respons yang jujur (Fisher, 1993). Ketiga, responden kurang memiliki motivasi dalam merespons skala. Rendahnya motivasi ini menyebabkan mereka memberikan respons asal-asalan. Dampaknya adalah respons mereka bervariasi dan tidak konsisten, tergantung pada *mood* mereka saat merespons butir.

Terkait dengan skala yang dipakai dalam penelitian, yaitu skala adaptasi dari Rosenberg *Self-Esteem Scale*, skala tersebut direspons dengan cara yang berbeda oleh ketiga kelas responden. Langkah selanjutnya adalah mengidentifikasi latar belakang masing-masing kelas, baik demografi maupun kondisi psikologisnya. Misal, butir 8 patut menghasilkan parameter butir yang ketepatan dengan model yang lemah. Pernyataan butir 8 dengan kalimat “Saya berharap saya dapat lebih dihargai” secara teoritik menunjukkan sebagai indikator individu yang memiliki harga diri yang rendah. Individu yang memiliki harga diri rendah diasumsikan memiliki harapan agar mereka dihargai dan mendapatkan tempat di mata masyarakat. Namun secara empiris, data di lapangan menunjukkan bahwa semua orang memiliki harapan untuk lebih dihargai. Baik mereka yang telah memiliki harga diri tinggi maupun rendah, mereka memiliki harapan yang sama. Ketidakkonsistenan ini menyebabkan parameter butir 8 tidak sesuai dengan model sehingga memerlukan modifikasi lebih lanjut.

## **Simpulan**

Penelitian ini mendemonstrasikan penggunaan Model Rasch Campuran untuk mengestimasi parameter butir. Dengan memadukan antara Model Rasch dengan Analisis Kelas Laten, parameter butir tersebut dikondisikan lebih spesifik dan berlaku pada kelas yang bersangkutan. Analisis lanjutan yang membandingkan parameter butir dari Model Rasch biasa (Model 1 Kelas) dan campuran menunjukkan bahwa parameter butir yang dihasilkan dari kedua model tersebut berbeda-beda. Parameter butir yang dihasilkan dari Model Rasch biasa hanya setara pada Kelas-1 pada Model Rasch Campuran 3 Kelas. Adanya responden-responden yang memberikan respons unik menyebabkan parameter butir dari Model Rasch biasa tidak bersifat invarian terhadap keseluruhan kelas. Penelitian dengan desain seperti ini dapat diaplikasikan pada berbagai jenis pengukuran, baik yang mengukur atribut abilitas maupun trait/ciri kepribadian.

## **Daftar Pustaka**

- Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 81(6), 506-552.
- Cox, B. J., Swinson, R. P., Drenfeld, D. M., & Bourdeau, D. (1994). Social desirability and self-reports of alcohol abuse in anxiety disorder patients. *Behaviour Research and Therapy*, 32(1), 175-178. doi: 10.1016/0005-7967(94)90100-7
- Eid, M., & Zickar, M. J. (2007). Detecting response styles and faking in personality and organizational assessments by mixed Rasch models. In M. Von Davier & C. Carstensen (Eds.), *Multivariate and Mixture Distribution Rasch Models*. New York: Springer.

- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists : Multivariate applications book series*. Mahwah (NJ): Lawrence Erlbaum Associates, Inc.
- Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, 20(2), 303-315.
- Goldstein, H., & Blinksorn, S. (1982). The Rasch Model Still Does Not Fit. *British Educational Research Journal*, 8(2), 167-170.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning*. Thousand Oaks: Sage Publications, Inc.
- Peacock, G. G., Ervin, R. A., & Daly, E. J. (2009). *Practical handbook of school psychology: Effective practices for the 21st*. New York, NY.: Guilford Press.
- Preinerstorfer, D., & Formann, A. K. (2011). Parameter recovery and model selection in mixed Rasch models. *British Journal of Mathematical and Statistical Psychology*, no-no. doi: 10.1111/j.2044-8317.2011.02020.x
- Retnowati, S. (2004). *Depresi Pada Remaja: Model Integrasi Penyebab Depresi Dan Pengobatan Depresi Pada Remaja*. Disertasi, Universitas Gadjah Mada, Yogyakarta.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rost, J. (1988). Rating scale analysis with latent class models. *Psychometrika*, 53(3), 327-348. doi: 10.1007/bf02294216
- Rost, J. (1990). Rasch Models in Latent Classes: An Integration of Two Approaches to Item Analysis. *Applied Psychological Measurement*, 14(3), 271-282. doi: 10.1177/014662169001400305

- Rost, J. (1997). Logistic mixture models. In W. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 449 - 463). New York: Springer-Verlag.
- Rost, J., Carstensen, C., & von Davier, M. (1997). Applying the mixed rasch model to personality questionnaires. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences*. Münster: Waxmann.
- Smit, A., Kelderman, H., & van der Flier, H. (1999). Collateral information and Mixed Rasch models. *Methods of Psychological Research Online*, 4(3).
- Van de Mortel, T. F. (2008). Faking it: social desirability response bias in self report research. *Australian Journal of Advanced Nursing*, 25(4), 40-48.
- von Davier, M. (2001). WINMIRA 2001. Kiel: Institute for Science Education.
- Yamamoto, K. (1989). *A Hybrid model of IRT and latent class models (ETS Research Report)*. Princeton, NJ: Educational Testing Service.