

KOMPARASI KETEPATAN ESTIMASI KOEFISIEN RELIABILITAS TEORI SKOR MURNI KLASIK

Wahyu Widhiarsop

Universitas Gadjah Mada

wahyu_psy@ugm.ac.id

Djemari Mardapi

Universitas Negeri Yogyakarta

djemarimardapi@gmail.com

Abstrak

Penelitian ini bertujuan untuk membandingkan ketepatan estimasi koefisien reliabilitas pada teori skor murni klasik. Peneliti membandingkan ketepatan estimasi koefisien-koefisien reliabilitas melalui data simulasi. Data simulasi dibangkitkan secara acak berdasarkan besarnya nilai reliabilitas murni, model pengukuran, ukuran sampel dan distribusi normal. Ukuran sampel yang dipakai terdapat empat jenis yaitu sebesar 50, 250, 1000 dan 5000. Reliabilitas murni yang dipakai terdiri dari lima kondisi yaitu 0,5; 0,6; 0,7; 0,8 dan 0,9. Hasil penelitian ini menunjukkan bahwa (a) koefisien reliabilitas dalam teori skor murni klasik yang dikembangkan oleh para ahli pengukuran memiliki ketepatan estimasi yang bervariasi. (b) Rata-rata koefisien reliabilitas memiliki ketepatan estimasi yang cukup tinggi yang dibuktikan dengan kecilnya rerata bias estimasi. (c) Koefisien reliabilitas komposit memiliki ketepatan yang tinggi pada model paralel dan kesetaraan tau, Koefisien Feldt pada model konjenerik, Koefisien Wang pada model korelasi antarsesatan dan Koefisien alpha berstrata pada model multidimensi.

Kata kunci: *ketepatan estimasi reliabilitas*

THE COMPARISON OF RELIABILITY COEFFICIENTS ESTIMATION AMONG CLASSICAL TEST THEORY

Wahyu Widhiarso

Psichology Dept Gadjah Mada University

wahyu_psy@ugm.ac.id

Djemari Mardapi

Graduate School of Yogyakarta State University

djemarimardapi@gmail.com

Abstract

This study aims was to compare the estimation precision among reliability coefficients classical test theory. Author compared the estimation precision reliability coefficients through simulation data. Simulation data was generated randomly based on the value of true reliability, measurement model, the sample size and normal distribution. Sample sizes used are in four types 50, 250, 1000 and 5000. True reliability was conducted in five conditions is 0.5; 0.6, 0.7, 0.8 and 0.9. The results of this study suggest that (a) the reliability coefficient in the classical test theory that developed by experts varied in estimation precision. (b) Almost all reliability coefficients have good estimation precision that proofed by the small average of bias estimation value. (c) The reliability of composites coefficient have high precision in parallel dan tau equivalence model, Feldt coefficient for congeneric model, Wang coefficient for error correlation model and stratified alpha coefficient for multidimensional models.

Key words: *precision estimation of reliability*

Pendahuluan

Wacana reliabilitas pengukuran dalam perspektif teori klasik masih menjadi kajian yang hangat di kalangan peneliti karena masih menyisakan permasalahan yang perlu diatasi. Permasalahan pertama adalah minimnya konsensus terhadap penerapan teori tes pada tataran yang lebih praktis (Slaney, 2006). Saat ini belum ada acuan yang tepat bagaimana memilih teknik yang tepat dalam mengevaluasi hasil pengukuran. Banyak formula dikembangkan untuk mengestimasi reliabilitas, namun konsensus belum juga tercapai. Bahkan, ada formula reliabilitas yang dikembangkan akan tetapi tujuan utamanya mulai bergeser kepada kemudahan dalam penghitungannya dibanding dengan keakuratannya sehingga Weis dan Davison (1981) mengatakan bahwa teori skor murni klasik sedikit kehilangan arahnya dalam mengkaji reliabilitas. Minimnya konsensus para ahli tersebut menunjukkan bahwa prosedur estimasi reliabilitas masih mengandung ketidakjelasan. Hal ini diperparah dengan masih banyaknya peneliti yang kurang tepat dalam melaporkan reliabilitas hasil pengukuran yang mereka lakukan (Thompson, 1994).

Permasalahan lain yang muncul adalah penggunaan koefisien reliabilitas oleh peneliti secara monoton tanpa mempertimbangkan asumsi yang mendasari koefisien tersebut sehingga diharapkan kajian mengenai reliabilitas tidak hanya terpaku pada satu koefisien saja melainkan juga melibatkan koefisien lain yang kemungkinan menggambarkan hasil yang lebih optimal (Ferketich, 1990). Socan (2000) mengatakan bahwa banyak di antara para peneliti yang hanya terpaku pada penggunaan koefisien Alpha Cronbach dalam mengestimasi reliabilitas tanpa memahami terlebih dahulu asumsi yang mendasari koefisien tersebut. Banyak di antara para peneliti yang tidak menyadari bahwa koefisien alpha menghendaki asumsi tertentu untuk dipenuhi. Jika asumsi ini tidak dipenuhi maka koefisien reliabilitas yang dihasilkan adalah nilai di batas estimasi terendah (Raykov, 1998).

Pemilihan formula estimasi reliabilitas secara monoton tersebut dapat diakibatkan oleh dua sebab. Pertama minimnya pemahaman peneliti mengenai koefisien reliabilitas lain yang dapat menjadi alternatif analisis. Kedua, masih minimnya keberadaan program komputasi yang dapat

mewadahi model pengukuran alternatif, misalnya model konjenerik. Feldt dan Ankenman (1999) mengatakan bahwa popularitas koefisien alpha lahir karena beberapa faktor, antara lain: (1) teknik komputasinya relatif mudah, karena hanya memerlukan informasi berupa varian butir dan varian skor total, (2) distribusi sampling sudah diketahui sehingga penentuan interval kepercayaan pada populasi sangat dimungkinkan. Koefisien alpha banyak dipakai dalam penelitian karena merupakan estimator yang moderat dalam mengestimasi reliabilitas.

Permasalahan yang telah dipaparkan di atas menunjukkan adanya kebutuhan bagi penyusun tes untuk mengetahui perbandingan ketepatan estimasi antarformula reliabilitas sehingga mereka dapat memilih formula mana yang paling akurat untuk pengukuran yang mereka lakukan. Kajian psikometri telah mengenal banyak formula reliabilitas yang masing-masing disusun berdasarkan asumsi, model dan pendekatan yang berbeda-beda. Adanya teknik analisis faktor konfirmatori telah memberikan sejumlah kemudahan bagi peneliti untuk mengidentifikasi model pengukuran yang sesuai dengan hasil pengukuran yang didapatkan. Analisis faktor konfirmatori juga telah menjadi dasar penyusunan berbagai model psikometris. Model konjenerik ataupun berbagai model alternatif, misalnya model pengukuran dengan korelasi antarsesatan (*correlated error*), model pengukuran dengan faktor berjenjang (*hierarchical model*) maupun model indikator formatif-reflektif (*formative-reflective indicator*) dapat dianalisis dan yang lebih penting adalah model pengukuran yang dihipotesiskan dapat diketahui.

Beberapa peneliti sudah melakukan studi perbandingan antarformula reliabilitas baik yang menggunakan data simulasi maupun data empirik. Minimnya program bantu kalkulasi, misalnya program komputer, membatasi peneliti untuk mengidentifikasi koefisien reliabilitas yang tidak tersedia pada program tersebut. Di sisi lain, adanya asumsi yang berbeda pada tiap koefisien reliabilitas memperlihatkan bahwa tiap koefisien reliabilitas tidak berada pada kontinum yang sama untuk dibandingkan. Namun eksplorasi tetap perlu dilakukan untuk menunjukkan bahwa penerapan koefisien reliabilitas yang bukan pada model pengukuran yang diasumsikan akan menghasilkan koefisien reliabilitas yang bias.

Dari studi literatur yang dilakukan penulis, perbandingan model pengukuran sudah dilakukan oleh banyak peneliti. Studi perbandingan model pengukuran diawali oleh Votaw yang menggunakan *English Composition Examination* pada 126 siswa yang dianalisis dengan menggunakan pendekatan SEM (Joreskog dan Sorbom, 1996). Dengan melakukan perbandingan model pengukuran tes *parallel, tau-equivalent* dan *konjenerik*. Fleishman dan Benson (1987) menggunakan teknik SEM untuk mengevaluasi model pengukuran dan reliabilitas pengukuran. Beberapa model teori skor murni klasik diaplikasikan termasuk model konjenerik. Bacon, Sauer, dan Young (1995) membandingkan koefisien omega dan koefisien theta dalam konteks model persamaan struktural. Ferketich (1990) membandingkan koefisien theta atau omega ketika mengestimasi reliabilitas konsistensi internal jika kondisi penelitian tidak memungkinkan untuk menggunakan koefisien alpha.

Raykov (1998) pernah menyusun pendekatan secara aljabar untuk menentukan seberapa rendah hasil estimasi koefisien alpha dari reliabilitas sesungguhnya, namun pendekatan tersebut memerlukan pendekatan model persamaan struktural (SEM). Socan (2000) pernah melakukan studi perbandingan tiga koefisien reliabilitas, yaitu Koefisien Alpha dari Cronbach, pendekatan *Analysis Congeneric Measure* (ACM) dari Joreskog, dan *Greatest Lower Bound Reliability* (GLB) dari Jackson dan Aguwamba. Yurdugül (2006) pernah membandingkan lima koefisien reliabilitas, yaitu antara Koefisien Alpha Cronbach (α), Koefisien Armor Theta (θ), Koefisien Omega (Ω) dari Heise dan Bohrnstedt, Koefisien Omega (ω) dari McDonald, dan Koefisien Beta (β) dari Revelle. Penelitian-penelitian yang dilakukan tersebut belum merangkum sebagian besar koefisien reliabilitas serta berbagai model pengukuran sehingga penelitian yang dapat merangkum semua koefisien reliabilitas dan model pengukuran sangat diperlukan.

Pada skenario pertama penelitian ini bertujuan untuk membandingkan ketepatan estimasi koefisien reliabilitas teori skor murni klasik yang didapatkan dari berbagai literatur (Tabel 1). Koefisien reliabilitas yang dikaji terdiri dari dua jenis yaitu koefisien reliabilitas model unidimensi

dan model multidimensi. Pada model unidimensi koefisien reliabilitas dispesifikkan menjadi tiga sub model yaitu model paralel, model kesetaraan nilai tau dan model konjenerik. Beberapa konsep yang diangkat dalam penelitian ini adalah ketepatan estimasi, ketepatan model, formula reliabilitas skor murni klasik dan dimensi pengukuran. Berikut ini akan dijelaskan pengertian konsep tersebut.

- a) *Ketepatan Estimasi*. Ketepatan estimasi yang dimaksud dalam penelitian ini adalah keakuratan sebuah formula reliabilitas terhadap reliabilitas murni. Besarnya ketepatan estimasi didapatkan melalui kesamaan nilai antara reliabilitas murni (ρ_{xx}) dan reliabilitas estimasi (r_{xx}). Semakin kecil selisih antara reliabilitas murni dan reliabilitas estimasi maka semakin tinggi ketepatan estimasi koefisien reliabilitas yang bersangkutan. Semakin besar selisih antara reliabilitas murni dan reliabilitas estimasi maka semakin rendah ketepatan estimasi koefisien reliabilitas yang bersangkutan. Apabila nilai reliabilitas estimasi lebih rendah dibanding dengan reliabilitas murninya ($\rho_{xx} > r_{xx}$), kondisi tersebut dinamakan dengan estimasi di atas batas (*overestimation*). Sebaliknya, jika nilai reliabilitas estimasi lebih tinggi dibanding dengan reliabilitas murninya ($\rho_{xx} < r_{xx}$) maka dinamakan dengan estimasi di bawah batas (*underestimation*).
- b) Formula reliabilitas teori skor murni klasik adalah formula yang dikembangkan oleh para ahli psikometri untuk mengestimasi reliabilitas dengan menggunakan dasar-dasar teori murni klasik.
- c) Model pengukuran adalah bentuk data hasil pengukuran yang memiliki asumsi-asumsi tertentu yang berbeda antara satu model dengan model lainnya. Model pengukuran yang dipakai dalam penelitian ini adalah model pengukuran dalam teori skor murni klasik yaitu model *parallel, essentially τ -equivalent, congeneric* dan *correlated error*.
- d) Dimensi pengukuran adalah jumlah dimensi yang ada di dalam data hasil pengukuran. Dimensi data hasil pengukuran dalam penelitian ini dibagi menjadi dua jenis yaitu unidimensi, yang menunjukkan adanya

dimensi tunggal di dalam data dan multidimensi yang menunjukkan adanya dimensi yang majemuk.

Metode Penelitian

Prosedur Penelitian

Peneliti melakukan perbandingan ketepatan estimasi koefisien reliabilitas terhadap reliabilitas murni (*true reliability*) pada data simulasi yang disusun oleh peneliti. Langkah-langkah yang dilakukan pada skenario ini antara lain sebagai berikut:

- a) *Penyusunan data simulasi.* Data simulasi didapatkan melalui pembangkitan angka acak dengan besaran reliabilitas, ukuran sampel dan model yang ditetapkan. Pada tahap ini peneliti menentukan skor murni (T) dan sesatan (E) terlebih dahulu sehingga reliabilitas murni dapat diketahui.
- b) *Mengestimasi reliabilitas.* Tiap koefisien reliabilitas dipakai untuk mengestimasi reliabilitas murni yang terdapat pada data simulasi. Pada tahap ini didapatkan ketepatan atau bias estimasi yang merupakan selisih antara reliabilitas murni dan reliabilitas hasil estimasi koefisien reliabilitas.
- c) *Menghitung ketepatan estimasi reliabilitas.* Ketepatan estimasi reliabilitas ditunjukkan dengan selisih antara nilai reliabilitas hasil estimasi koefisien reliabilitas dengan reliabilitas murni yang telah ditetapkan oleh peneliti.

Penyusunan Data Simulasi

Data simulasi disusun berdasarkan besarnya nilai reliabilitas murni, model pengukuran dan ukuran sampel. Ukuran sampel yang dipakai terdapat empat jenis yaitu sebesar 50, 250, 1000 dan 5000. Reliabilitas murni yang dipakai terdiri dari lima kondisi yaitu 0,5; 0,6; 0,7; 0,8 dan 0,9. Pemilihan kondisi ini berdasarkan pada nilai reliabilitas yang bergerak dari kategori sedang, baik dan istimewa (Crocker & Algina, 1986). Dengan adanya sampel dan nilai reliabilitas murni yang berbeda-beda maka terdapat 15 kombinasi (3 jenis ukuran sampel x 5 jenis reliabilitas murni = 15

kombinasi) data simulasi pada tiap-tiap model pengukuran (paralel, kesetaraan tau, konjenerik).

Untuk menyusun data simulasi, peneliti menggunakan program Microsoft Excel melalui menu *Random Number Generation* dan *Sampling*. Langkah pertama yang dilakukan adalah membuat enam variabel yang terdistribusi normal yang memuat angka acak dengan besar ukuran sampel sebesar 5000 ($N=5000$), rerata sebaran sebesar 0 ($\bar{x}=0$) dan deviasi standar sebesar 1 ($s=1,0$). Keenam variabel yang didapatkan dipakai untuk variabel skor murni (T) dan sisanya dipakai untuk variabel sesatan (E). Ukuran sampel sisanya, yaitu $N=50$, $N=250$ dan $N=1000$ didapatkan melalui menu *sampling* berdasar sebaran yang didapatkan dari ukuran sampel $N=50$. Data yang dihasilkan setelah diuji telah memenuhi asumsi teori murni klasik, antara lain tidak adanya korelasi antarsesatan ($r_{E1E2} = 0$) maupun korelasi antara sesatan dan skor murni ($r_{ET} = 0$).

Hasil Dan Pembahasan

Perbandingan Ketepatan Estimasi

Melalui data simulasi didapatkan besarnya ketepatan estimasi masing-masing koefisien. Besarnya perbandingan ketepatan estimasi pada tiap model yang dapat dilihat pada Tabel 1 sampai Tabel 3. Tabel tersebut menunjukkan statistik deskriptif ketepatan estimasi. Selengkapnya penjelasan hasil menurut model Tabel 1 sampai Tabel 3 pengukuran yang dipakai adalah sebagai berikut:

- a) *Model Paralel*. Pada model paralel didapatkan keterangan bahwa nilai estimasi terhadap reliabilitas murni oleh masing-masing formula mendekati nilai reliabilitas murni. Rata-rata bias estimasi 18 formula yang dipakai dalam penelitian ini adalah $-0,0373$. Hal ini menunjukkan bahwa sebagian besar koefisien reliabilitas mengestimasi pada batas bawah reliabilitas murni (*lower bound estimator*). Koefisien Guttman Lambda-1 memiliki bias yang terbesar yaitu $-0,1728$ sedangkan formula yang memiliki bias estimasi terendah adalah Koefisien Reliabilitas Komposit dengan rerata bias estimasi sebesar $-0,0102$.

- b) *Model Kesetaraan Tau*. Pada model kesetaraan tau didapatkan hasil bahwa nilai estimasi terhadap reliabilitas murni oleh masing-masing formula mendekati nilai reliabilitas murni. Rata-rata bias estimasi 18 formula yang dipakai dalam penelitian ini adalah sebesar $-0,0269$ yang mendekati rerata bias estimasi pada model paralel yang memiliki rerata sebesar $-0,0373$. Hasil estimasi ini menunjukkan bahwa sebagian besar koefisien reliabilitas mengestimasi pada batas bawah reliabilitas murni. Koefisien Reliabilitas Wang memiliki bias yang terbesar yaitu $-0,1983$ sedangkan formula yang memiliki bias estimasi terendah adalah Koefisien Reliabilitas Komposit dengan rerata bias estimasi sebesar $-0,0029$.
- c) *Model Konjenerik*. Pada model konjenerik didapatkan keterangan bahwa nilai estimasi terhadap reliabilitas murni oleh masing-masing formula mendekati nilai reliabilitas murni. Rerata bias estimasi semua formula adalah $0,1727$ yang menunjukkan bahwa koefisien reliabilitas mengestimasi pada batas atas reliabilitas murni. Koefisien Lambda-1 memiliki ketepatan estimasi tertinggi dengan rerata $0,0089$ sedangkan Koefisien Reliabilitas Maksimal memiliki ketepatan estimasi terendah dengan rerata bias estimasi sebesar $0,2208$.
- d) *Model Multidimensi*. Dari model multidimensi didapatkan keterangan bahwa nilai estimasi oleh koefisien reliabilitas bervariasi satu dengan lainnya. Rata-rata bias estimasi 10 koefisien yang dipakai dalam penelitian ini adalah $-0,0069$. Hal ini menunjukkan bahwa reliabilitas yang dilibatkan dalam penelitian ini mengestimasi pada batas atas reliabilitas murni (*upper bound estimator*). Koefisien Alpha Berstrata dan Koefisien Reliabilitas Komposit memiliki ketepatan estimasi paling tinggi yaitu dengan rerata $-0,0005$ dan $0,0026$ sedangkan Koefisien Wang dan Koefisien Omega memiliki ketepatan estimasi paling rendah yaitu dengan rerata bias estimasi sebesar $-0,1446$ dan $0,1879$.

Tabel 1. Perbandingan Ketepatan Estimasi (Model Paralel)

No	Koefisien Reliabilitas	Minimal	Maksimal	Rerata	Dev.Std
1.	Guttman Lambda-1	-0.2227	-0.0982	-0.1728	0.0364
2.	Guttman Lambda-4	-0.4007	-0.0345	-0.1467	0.1049
3.	Koefisien Winer	-0.2125	-0.0271	-0.1411	0.0463
4.	Guttman Lambda-6	-0.1360	-0.0229	-0.0613	0.0318
5.	Reliabilitas Konstrak	-0.2855	0.0139	-0.0440	0.0659
6.	Guttman Lambda-3	-0.1253	0.0081	-0.0387	0.0347
7.	Cronbach Alpha	-0.1254	0.0140	-0.0360	0.0365
8.	Revelle Beta	-0.1254	0.0140	-0.0360	0.0365
9.	Armor Theta	-0.1129	0.0105	-0.0335	0.0332
10.	Reliabilitas Maksimal 2	-0.1649	0.0448	-0.0273	0.0499
11.	Guttman Lambda-2	-0.0856	0.0153	-0.0250	0.0270
12.	Koefisien Feldt	-0.0862	0.0337	-0.0165	0.0263
13.	Koefisien Heise	-0.0751	0.0892	-0.0121	0.0408
14.	Guttman Lambda-5	-0.0725	0.0255	-0.0111	0.0255
15.	Reliabilitas Komposit	-0.0605	0.0252	-0.0102	0.0234
16.	McDonald Omega	-0.0115	0.2442	0.0384	0.0544
17.	Heise-Bohrnstedt Omega	-0.0149	0.1728	0.0501	0.0491
18.	Reliabilitas Maksimal 1	-0.0050	0.1351	0.0525	0.0438

Tabel 2. Perbandingan Ketepatan Estimasi (Model Kesetaraan Nilai Tau)

No	Koefisien Reliabilitas	Minimal	Maksimal	Rerata	Dev.Std
1.	Koefisien Winer	-0.3637	-0.0711	-0.1983	0.0802
2.	Guttman Lambda-1	-0.2144	-0.0838	-0.1717	0.0341
3.	Koefisien Heise	-0.2308	0.1897	-0.0499	0.1222
4.	Cronbach Alpha	-0.1605	0.0251	-0.0496	0.0458
5.	Revelle Beta	-0.1605	0.0251	-0.0496	0.0458
6.	Guttman Lambda-6	-0.1020	0.0538	-0.0407	0.0451
7.	Guttman Lambda-3	-0.0770	0.0251	-0.0372	0.0272
8.	Feldt	-0.0862	-0.0002	-0.0235	0.0243
9.	Armor Theta	-0.0642	0.0545	-0.0201	0.0347
10.	Reliabilitas Konstrak	-0.0916	0.0527	-0.0195	0.0371
11.	Guttman Lambda-2	-0.0378	0.0566	-0.0122	0.0247
12.	Guttman Lambda-4	-0.0525	0.1074	-0.0067	0.0425
13.	Reliabilitas Maksimal 2	-0.1225	0.1413	-0.0059	0.0651
14.	Reliabilitas Komposit	-0.0176	0.0332	0.0029	0.0125
15.	Guttman Lambda-5	-0.0192	0.0469	0.0042	0.0163
16.	Heise-Bohrnstedt Omega	0.0037	0.2684	0.0585	0.0766
17.	Reliabilitas Maksimal 1	0.0067	0.1829	0.0663	0.0503
18.	McDonald Omega	0.0125	0.3283	0.0681	0.0718

Tabel 3. Perbandingan Ketepatan Estimasi (Model Konjenerik)

No	Koefisien Reliabilitas	Minimum	Maximum	Mean	Dev.Std
1.	Guttman Lambda-1	-0.2231	-0.0921	-0.1722	0.0379
2.	Koefisien Winer	-0.2411	-0.0602	-0.1496	0.0498
3.	Guttman Lambda-4	-0.4060	-0.0309	-0.1453	0.1078
4.	Guttman Lambda-6	-0.1307	-0.0213	-0.0589	0.0306
5.	Koefisien Heise	-0.2289	0.1266	-0.0514	0.0964
6.	Guttman Lambda-3	-0.1219	0.0096	-0.0376	0.0345
7.	Armor Theta	-0.1002	0.0176	-0.0298	0.0323
8.	Cronbach Alpha	-0.1221	0.0202	-0.0297	0.0402
9.	Revelle Beta	-0.1221	0.0202	-0.0297	0.0402
10.	Reliabilitas Maksimal 2	-0.1698	0.0432	-0.0288	0.0523
11.	Reliabilitas Konstrak	-0.1112	0.0205	-0.0285	0.0353
12.	Guttman Lambda-2	-0.0845	0.0165	-0.0241	0.0271
13.	Reliabilitas Komposit	-0.0736	0.0277	-0.0109	0.0266
14.	Guttman Lambda-5	-0.0719	0.0264	-0.0102	0.0256
15.	Feldt	-0.0669	0.0365	-0.0083	0.0278
16.	McDonald Omega	-0.0096	0.1086	0.0385	0.0305
17.	Reliabilitas Maksimal 1	-0.0060	0.1424	0.0540	0.0452
18.	Heise-Bohrstedt Omega	-0.0160	0.2049	0.0578	0.0600

Secara umum dari hasil perbandingan ketepatan estimasi antar koefisien didapatkan beberapa temuan penelitian yaitu sebagai berikut.

- a) Pada model paralel, kesetaraan nilai tau dan tiap koefisien reliabilitas memiliki ketepatan yang cukup akurat karena rata-rata bias estimasi bergerak dari urutan desimal ke-dua dan cenderung di bawah batas estimasi (*underestimasi*), namun ketika diaplikasikan pada model korelasi antarsesatan, estimasi semua koefisien reliabilitas cenderung melambung (*overestimasi*).
- b) Koefisien reliabilitas yang dipakai dalam penelitian ini dapat dikategorikan menjadi dua kategori, yaitu koefisien yang berbasis pada varian-kovarian (Koefisien Alpha dan Koefisien Guttman) dan koefisien yang berbasis pada analisis faktor (Koefisien Armor Theta, Reliabilitas Konstrak dan Reliabilitas Komposit). Dari hasil

perbandingan ketepatan estimasi didapatkan informasi tidak adan perbedaan daya ketepatan antara koefisien berbasis pada varian-kovarian dan analisis faktor.

- c) Dari pengujian secara statistika melalui analisis varian mengenai ketepatan estimasi ditinjau dari koefisien reliabilitas dan ukuran sampel didapatkan keterangan bahwa terdapat interaksi yang signifikan antara koefisien reliabilitas dan ukuran sampel. Adanya interaksi tersebut menunjukkan bahwa terdapat perbedaan efek penambahan ukuran sampel terhadap ketepatan estimasi antar koefisien reliabilitas.
- d) Data simulasi yang dipakai dalam penelitian ini terdiri dari empat jenis ukuran sampel. Dari hasil perbandingan ketepatan estimasi ditinjau berdasarkan ukuran sampel didapatkan hasil yang bervariasi. Pada model paralel dan model *konjenerik* penambahan ukuran sampel meningkatkan daya estimasi koefisien reliabilitas sedangkan pada model kesetaraan tau dan korelasi antarsesatan, penambahan ukuran sampel tidak berperan terhadap peningkatan daya estimasi.

Tabel 4. Koefisien Reliabilitas yang Memiliki Daya Estimasi Tertinggi

No	Model Pengukuran	Koefisien Reliabilitas
1.	Model Paralel	Reliabilitas Komposit
		Koefisien Guttman Lambda-5
		Koefisien Heise
2.	Model Kesetaraan Nilai Tau	Reliabilitas Komposit
		Koefisien Guttman Lambda-5
		Reliabilitas Maksimal-2
3.	Model Konjenerik	Koefisien Feldt
		Koefisien Guttman Lambda-5
		Koefisien Reliabilitas Komposit
4.	Model Korelasi Antarsesatan	Koefisien Guttman Lambda-1
		Koefisien Winer
		Koefisien Guttman Lambda-4
5.	Model Multidimensional	Koefisien Alpha Berstrata
		Reliabilitas Komposit
		Koefisien Mosier

Pembahasan

Berdasarkan hasil perbandingan ketepatan estimasi terhadap reliabilitas murni didapatkan kesimpulan bahwa terdapat perbedaan ketepatan estimasi antar koefisien reliabilitas. Selisih bias estimasi antar koefisien reliabilitas adalah cukup rendah yang terlihat dari nilai selisih yang bergerak pada nilai desimal kedua ($\Delta \approx 0,0$). Koefisien Reliabilitas Komposit yang dikembangkan oleh Raykov (1998) untuk mengatasi beberapa kelemahan Koefisien Cronbach Alpha, menjadi estimator yang memiliki ketepatan yang tinggi untuk diaplikasikan pada model-model pengukuran teori klasik. Dengan menggunakan analisis faktor konfirmatori dalam pendekatan model persamaan struktural (SEM), koefisien ini mengasumsikan bahwa seperangkat butir merupakan representasi dari faktor umum yang tunggal. Penggunaan asumsi faktor tunggal ini merupakan salah satu kelebihan Koefisien Reliabilitas Komposit jika dibanding dengan Koefisien Cronbach Alpha. Karena terbatas pada varian butir dan varian skor Koefisien Alpha dalam hal ini hanya belum menjangkau representasi faktor umum yang tunggal. Hal ini sesuai dengan apa yang dijelaskan oleh McDonald (1981) yang mengatakan bahwa Koefisien Alpha belum mampu menjelaskan varian umum konstruk latent di dalam data.

Selain Reliabilitas Komposit, koefisien lain yang memiliki ketepatan cukup tinggi adalah Koefisien Guttman Lambda-5. Koefisien Guttman Lambda-5 adalah penyempurnaan Koefisien Guttman Lambda-1 yang merupakan *lower bound estimator*. Estimasi yang terlalu rendah dari Koefisien Guttman Lambda-1 kemudian diatasi oleh Guttman (1945) dengan menambahkan pembagian akar kuadrat pangkat dua rerata kovarian butir dengan varian skor total ($\sqrt{S_{ij}} / S_x^2$) sehingga koefisien Guttman Lambda-5 diformulasikan dalam persamaan $\lambda_5 = \lambda_1 + \sqrt{S_{ij}} / S_x^2$. Dalam penelitian ini estimasi reliabilitas yang terlalu rendah oleh Koefisien Guttman Lambda-1 teratasi oleh dengan adanya penambahan unsur baru tersebut sehingga Koefisien Guttman Lambda-5 menjadi salah satu estimator yang memiliki ketepatan estimasi reliabilitas murni yang cukup tinggi.

Koefisien Cronbach Alpha dalam penelitian ini memiliki nilai estimasi yang setara dengan Koefisien Revelle Beta dan Koefisien Lambda-3. Koefisien ini konsisten mengestimasi pada batas bawah dengan daya ketepatan yang moderat kecuali pada model korelasi antarsesatan. Meskipun memiliki daya estimasi yang setara, dibanding dengan Koefisien Revelle Beta dan Koefisien Lambda-3, Koefisien Alpha memiliki kelebihan dalam kesederhanaan prosedur komputasi yang menggunakan varian butir dan varian skor total. Hasil penelitian mengenai ketepatan estimasi Koefisien Alpha ini didukung oleh pernyataan Callender dan Osburn (1979) yang mengatakan bahwa Koefisien Alpha menghasilkan estimasi yang sama dengan Koefisien Guttman Lambda-3 dan cenderung mengestimasi pada batas bawah.

Pada model konjenerik, Koefisien Feldt merupakan estimator yang memiliki ketepatan yang paling tinggi di antara koefisien lain dalam penelitian ini. Koefisien Feldt yang merupakan penjabaran dari Koefisien Flanagan menekankan pada varian dan kovarian data. Hasil penelitian ini yang membuktikan bahwa dalam kasus model *konjenerik*, Koefisien Feldt memiliki ketepatan yang lebih tinggi dibanding dengan Koefisien Alpha juga dibuktikan oleh Sedere dan Feldt (1976) serta Feldt dan Charter (2003). Dalam penelitiannya dengan menggunakan data simulasi berupa 60 butir dengan ukuran sampel sebesar 100, 200 dan 400 yang dibelah dalam dua jenis belahan sama panjang dan tidak sama panjang, didapatkan kesimpulan bahwa Koefisien Feldt memiliki daya ketepatan estimasi yang lebih tinggi dibanding Koefisien Alpha.

Penelitian ini menemukan bahwa koefisien alpha dapat bernilai tinggi meskipun data bersifat heterogen yang ditunjukkan dengan dimensi pengukuran yang berbeda dalam skor tes. Hasil penelitian ini mendukung pernyataan Helms, Henze, Sass, dan Venus (2006) yang melihat bahwa koefisien alpha tidak menunjukkan seberapa jauh tes mengukur satu konstruk ukur. Cortina (1993) menambahkan bahwa koefisien alpha dapat bernilai tinggi pada data multidimensi apabila jumlah butir adalah minimal 14 butir.

Secara umum koefisien yang dilibatkan dalam penelitian ini menggunakan dua pendekatan yang berbeda. Kelompok koefisien pertama

menggunakan pendekatan varian-kovarian seperti Koefisien Alpha, Koefisien Guttman Lambda, Koefisien dan Reliabilitas Maksimal dan kelompok koefisien kedua yang menggunakan pendekatan analisis faktor baik eksploratori maupun konfirmatori adalah Koefisien Heisse-Bohrsted Omega, Reliabilitas Komposit dan Reliabilitas Konstrak. Hasil perbandingan estimasi menunjukkan bahwa tidak ditemukan perbedaan ketepatan antar kedua pendekatan tersebut.

Dalam penelitian ini koefisien yang mampu mengadopsi model pengukuran yang ditetapkan akan memiliki ketepatan etimasi yang tinggi. Hal ini sesuai dengan pernyataan Schmidt et.al (2003) yang mengatakan bahwa sebuah koefisien reliabilitas lebih tepat dipakai dibanding dengan koefisien yang lain dikarenakan koefisien tersebut memiliki model yang tepat dalam menjelaskan berbagai proses sesatan, misalnya sesatan acak, sesatan sementara dan faktor spesifik.

Penelitian ini memiliki beberapa keterbatasan yang menyangkut keluasan dan kedalaman tema penelitian maupun masalah metodologis. Koefisien reliabilitas yang dilibatkan dalam penelitian ini terbatas dan belum menjangkau koefisien reliabilitas yang banyak dikaji dalam bidang psikometri, misalnya Koefisien Spearman Brown (Spearman, 1904), Koefisien Kuder-Richardson (Kuder & Richardson, 1937), Koefisien Flanagan (Rulon, 1939), Koefisien Rulon (Rulon, 1939), Koefisien Horst (Horst, 1954), Koefisien Kristof (Kristof), Koefisien Raju (Raju, 1977), Koefisien Tarkonen Rho (Vehkalahti, Puntanen, & Tarkkonen, 2006). Peneliti tidak melibatkan koefisien tersebut dalam penelitian ini dikarenakan beberapa alasan seperti, a) model pembelahan yang memiliki kombinasi seperti Koefisien Spearman Brown dan Koefisien Rulon yang akan memiliki 10 kombinasi belahan jika dikenakan pada 5 butir; b) perbedaan jenis data, misalnya tidak digunakannya Koefisien KR-20 karena dalam penelitian ini lebih mengarah pada data kontinum; c) keterwakilan oleh koefisien lain seperti Koefisien Kristof dan Raju yang memiliki kesetaraan dengan Koefisien Feldt; d) tidak tersedianya program analisis yang mendukung, misalnya Koefisien Tarkonen yang belum ada perangkat lunak yang dapat mengakomodasi koefisien ini.

Simpulan

Penelitian ini menghasilkan beberapa kesimpulan mengenai ketepatan estimasi reliabilitas teori skor murni klasik. Kesimpulan tersebut adalah sebagai berikut:

1. Koefisien reliabilitas dalam teori skor murni klasik yang dikembangkan oleh para ahli pengukuran memiliki ketepatan estimasi yang bervariasi. Penelitian ini membuktikan adanya perbedaan ketepatan estimasi yang signifikan, baik ketika diterapkan pada model paralel, kesetaraan tau, konjenerik, dan korelasi antarsesatan.
2. Rata-rata koefisien reliabilitas memiliki ketepatan estimasi yang cukup tinggi yang dibuktikan dengan kecilnya rerata bias estimasi, yaitu sebesar $\Delta = -0,0337$. Sebagian besar koefisien reliabilitas tersebut mengestimasi pada batas bawah estimasi (*underestimated*). Koefisien-koefisien tersebut kurang dapat diaplikasikan pada model pengukuran korelasi antarsesatan karena memiliki ketepatan estimasi yang cukup rendah yaitu sebesar 0.1728 dan cenderung mengestimasi pada batas atas estimasi (*overestimated*).
3. Koefisien reliabilitas komposit memiliki ketepatan yang tinggi pada model paralel dan kesetaraan nilai tau, Koefisien Feldt memiliki ketepatan yang tinggi pada model konjenerik, Koefisien Wang memiliki ketepatan yang tinggi pada model korelasi antarsesatan dan Koefisien alpha berstrata memiliki ketepatan yang tinggi pada model multidimensi.

Daftar Pustaka

- Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite Reliability in Structural Equations Modeling. *Educational and Psychological Measurement*, 55(3), 394-406.
- Callender, J. C., & Osburn, H. G. (1979). An empirical comparison of Coefficient Alpha, Guttman's lambda - 2, and MSPLIT maximized split-half reliability estimates. *Journal of Educational Measurement*, 16(2), 89-99.

- Cortina, J. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.
- Feldt, L. S., & Ankenmann, R. D. (1999). Determining sample size for a test of the equality of alpha coefficients when the number of part-tests is small. *Psychological Methods*, 4(4), 366-377.
- Feldt, L. S., & Charter, R. A. (2003). Estimating the reliability of a test split into two parts of equal or unequal length. *Psychological Methods*, 8(1), 102-109.
- Ferketich, S. (1990). Focus on Psychometrics Internal Consistency Estimates of Reliability. *Researching Nursing & Health*, 13, 437-440
- Fleishman, J., & Benson, J. (1987). Using Lisrel to Evaluate Measurement Models and Scale Reliability. *Educational and Psychological Measurement*, 47(4), 925-939.
- Helms, J. E., Henze, T. K., Sass, T. L., & Venus, A. M. (2006). Treating Cronbach's Alpha Reliability Coefficients as data in counseling research. *The Counseling Psychologist*, 34(5), 630 - 660.
- Horst, P. (1954). The estimation of immediate retest reliability. *Educational and Psychological Measurement*, 14(3), 705-708.
- Kristof, W. The statistical theory of stepped-up reliability coefficients when a test has been divided into several equivalent parts. *Psychometrika*, 28(1), 221-238.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(8), 151-160.
- McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.
- Raju, N. S. (1977). A generalization of coefficient alpha. *Psychometrika*, 42, 549-565.

METODE *LINKING* DAN BUTIR *FALSE POSITIVE* PADA PENDETEKSIAN *DIF* BERDASARKAN TEORI RESPONSI BUTIR

Wardani Rahayu
PPs UNJ
wardani9164@yahoo.com

Abstract

Tujuan dari penelitian ini adalah untuk mengetahui pengaruh metode *linking* pada sejumlah butir *false positif* dalam pendeteksian *DIF* berdasarkan teori response butir. Metode *linking* yang digunakan adalah metode rerata dan sigma (RS), metode tegar rerata dan sigma (TRS), metode kurva karakteristik (KK). Pendeteksian *DIF* dilakukan dengan menggunakan metode Lord's Chi-Square Test. Hasil pendeteksian *DIF* melalui metode *Lord's Chi-Square* ini didasarkan pada banyaknya butir *false positive* (FP) yang mempengaruhi keakuratan metode *linking*. Jumlah butir false positif dihitung dari butir yang sebelumnya diperkirakan tidak *DIF* berdasarkan *Likelihood Ratio Test*, akan tetapi setelah pendeteksian *DIF* menggunakan Lord's Chi-Square Test, ternyata ditemukan bahwa item mengandung *DIF*. Hasil penelitian menunjukkan bahwa metode kurva karakteristik, metode rerata dan metode sigma adalah lebih akurat daripada metode tegar rerata dan metode sigma. Namun, metode kurva karakteristik memiliki akurasi yang sama dengan metode rerata dan metode sigma.

Kata kunci: *metode linking, butir false positive, DIF*