

## Aplikasi Pemodelan Persamaan Struktural dalam Pengujian Model Pengukuran Psikologi<sup>1</sup>

Wahyu Widhiarso<sup>2</sup>  
Universitas Gadjah Mada

### Abstrak

Penelitian ini bertujuan untuk menguji model-model pengukuran psikologi melalui pemodelan persamaan struktural (SEM). Dalam teori skor murni klasik psikometri dikenal empat model pengukuran yaitu paralel, kesetaraan nilai tau, konjenerik dan korelasi antar sesatan. Model paralel dan kesetaraan nilai tau mengasumsikan setiap butir memiliki kapasitas, bobot atau unit pengukuran yang setara dalam mengungkap target ukur, sedangkan model konjenerik mengasumsikan kapasitas ukur yang berbeda. Sebagian besar pengujian properti psikometris pengukuran psikologi menggunakan model paralel dan kesetaraan nilai tau yang dibuktikan dengan penggunaan koefisien alpha Cronbach dalam melaporkan properti psikometris instrumen yang dilibatkan. Asumsi paralel dan kesetaraan nilai tau sulit untuk dipenuhi karena perilaku manusia merupakan konstruk yang kompleks dan dinamis, sehingga butir-butir pengukuran yang merupakan sampel indikator perilaku ukur akan lebih cenderung memiliki kapasitas ukur yang berbeda daripada memiliki kapasitas yang sama. Berdasarkan penjelasan tersebut penulis mengajukan hipotesis yang menyatakan bahwa model konjenerik dan korelasi antar sesatan lebih menggambarkan pengukuran psikologi dibanding model paralel dan kesetaraan nilai tau. Instrumen yang dipakai sebagai sampel adalah *Big Five Inventori* (BFI-44) yang mewakili pengukuran kepribadian dan *Beck Depression Inventory* (BDI) yang mewakili pengukuran klinis, Skala Sikap terhadap Perubahan Kebijakan Organisasi yang mewakili pengukuran sikap. BFI-44 dan BDI diberikan kepada mahasiswa UGM Yogyakarta (N=185 dan 184) sedangkan skala sikap diberikan kepada karyawan PT TELKOM Bandung. Analisis dengan menggunakan teknik analisis faktor konfirmatori melalui program LISREL 8.30 menghasilkan nilai kai-kuadrat model korelasi antar sesatan dan konjenerik lebih tinggi dibanding model paralel dan kesetaraan nilai tau. Dengan kata lain, model korelasi antar sesatan dan model konjenerik lebih menggambarkan model pengukuran psikologi dibanding dengan model paralel dan kesetaraan nilai tau. Temuan ini menyarankan kepada peneliti yang mengembangkan pengukuran psikologi untuk (a) mengidentifikasi potensi perbedaan kapasitas ukur setiap butir instrumen melalui analisis faktor dan (b) menggunakan formula psikometris yang menggunakan asumsi yang sesuai dengan model pengukuran instrumen yang dipakai agar menghasilkan informasi properti psikometris yang lebih akurat.

**Kata Kunci.** *Pengukuran Psikologi, Model Pengukuran, Analisis Faktor Konfirmatori*

Banyak ditemui peneliti yang mengevaluasi properti psikometris pengukuran dengan menggunakan formula-formula psikometris tanpa memverifikasi asumsi yang mendasari formula tersebut terpenuhi atukah tidak. Fenomena ini juga dirasakan oleh Socan (2000) dan Graham (2006) yang mengatakan bahwa kebanyakan peneliti yang hanya terpaku pada

---

<sup>1</sup> Makalah disampaikan pada Temu Ilmiah Nasional dan Kongres XI Himpsi di Surakarta tanggal 18 – 20 Maret 2010.

<sup>2</sup> Dosen Fakultas Psikologi Universitas Gadjah Mada

penggunaan koefisien alpha Cronbach dalam mengestimasi reliabilitas dan tidak menyadari bahwa koefisien alpha menghendaki asumsi tertentu yang harus dipenuhi. Jika asumsi koefisien alpha tidak dipenuhi maka koefisien reliabilitas yang dihasilkan adalah nilai di bawah batas estimasi (*underestimate*). Para ahli psikometri telah banyak mengingatkan bahwa dibalik formula psikometri terdapat asumsi-asumsi yang perlu dipenuhi agar formula tersebut menghasilkan informasi yang akurat mengenai pengukuran yang dilakukan.

Baik teori skor murni klasik maupun teori respons butir (IRT), masing-masing memiliki perspektif tertentu dalam meninjau sebuah data hasil pengukuran yang lebih dikenal dengan model pengukuran. Model tersebut menjelaskan hubungan berbagai unsur pengukuran, seperti hubungan antara konstruk ukur dan indikatornya dalam teori skor murni klasik, dan hubungan antara konstruk ukur dengan probabilitas mendapatkan skor maksimal. Model pengukuran dibangun berdasarkan asumsi-asumsi tertentu mengenai proses dan keluaran hasil pengukuran (data hasil pengukuran) yang dilakukan. Model pengukuran menjadi dasar munculnya berbagai formula untuk mengevaluasi properti psikometris pengukuran yang dilakukan. Paparan ini menunjukkan bahwa formula-formula psikometri muncul dari proses yang bertahap, yang diawali dengan pengembangan asumsi terhadap proses dan hasil pengukuran, pengembangan model pengukuran yang kemudian menghasilkan formula psikometri.

Secara garis besar teori skor murni klasik psikometri menjelaskan bahwa properti psikometris lebih mudah digali jika pengukuran dilakukan secara majemuk (*multiple measure*) atau melibatkan indikator yang majemuk (*multiple indicators*). Pengukuran majemuk, misalnya berimplikasi pada pengukuran berulang-ulang sedangkan indikator majemuk berimplikasi jumlah butir yang majemuk. Pengukuran dikatakan memiliki keandalan ukur yang memuaskan jika pengulangan pengukuran yang dilakukan dan butir-butir yang dilibatkan memiliki korelasi tinggi antar pengukuran atau antar indikator ketika dipakai mengukur konstruk ukur sama. Model pengukuran menjelaskan asumsi mengenai hubungan antar pengukuran yang dilakukan dan antar indikator yang dilibatkan. Contohnya, model pengukuran paralel mengasumsikan bahwa pengukuran-pengukuran yang dilakukan memiliki keandalan ukur yang setara dan skor hasil pengukuran berada pada skala yang sama (McPherson & Rotolo, 1995). Pada tulisan ini, penulis menyetarakan penggunaan istilah pengukuran dan butir. Keduanya memiliki makna yang sama dalam pembahasan mengenai model pengukuran karena dalam model pengukuran seperangkat indikator manifes dapat ditafsirkan sebagai pengukuran atau butir majemuk.

#### **A. Model Pengukuran**

Diskusi mengenai model pengukuran (*measurement model*) teori skor murni klasik terkadang dijelaskan dengan menggunakan istilah-istilah lain seperti model tes (*test model*) contohnya model tes paralel (Meredith, Frederiksen, & McLaughlin, 1974) dan teori (*theory*) contohnya teori tes konjenerik (Lucke, 2005). Meski berbeda, kesemuanya mengacu pada maksud yang sama. Selain istilah model dan teori, penggunaan istilah tes dan pengukuran juga terkadang saling tumpang tindih. Tulisan ini menggunakan istilah model pengukuran karena istilah sesuai dengan teknik analisis yang dipakai yaitu pemodelan persamaan struktural (SEM). Model pengukuran merupakan salah satu bentuk model di dalam SEM. Dalam teori skor murni klasik dikenal lima model pengukuran, yaitu model pengukuran paralel (*parallel*), nilai tau-setara (*tau-equivalent*), nilai tau-setara esensial (*essentially tau-equivalent*), konjenerik (*congeneric*) dan korelasi antar sesatan (*correlated error*). Dari kelima model

tersebut model paralel, nilai tau-setara dan konjenerik lebih banyak dikenal dan dipakai dalam penelitian pengembangan alat ukur psikologi (Millsap & Everson, 1991).

*Model Pengukuran Paralel.* Traub (1994) mengatakan bahwa model pengukuran paralel mengasumsikan bahwa skor harapan, varians dan kovarians antar pengukuran/butir besarnya setara. Menurut DeVellis (1991), model ini mengasumsikan: nilai semua varians butir dan varians sesatan butir sama, koefisien jalur terstandarisasi dari skor murni menuju semua butir adalah sama, nilai korelasi antar butir adalah setara, nilai korelasi antar butir setara dengan nilai koefisien jalur yang dikuadratkan dari skor murni menuju tiap butir, setiap butir memiliki rerata dan varian yang setara. Melihat asumsi-asumsi di dalam model pengukuran paralel, maka dapat disimpulkan bahwa tiap pengukuran dilakukan atau butir yang dilibatkan dalam pengukuran memiliki kesamaan target ukur (*unidimensionality*), kesetaraan kapasitas ukur (*amount of measure*) dan kesamaan skala ukur (Raykov, 1997). Kapasitas ukur yang setara tersebut ditunjukkan oleh derajat keandalan dan sesatan ukur yang setara. Traub (1994) mengatakan bahwa keandalan ukur yang sama yang dapat ditunjukkan dengan nilai reliabilitas dan muatan faktor yang sama. Model pengukuran paralel adalah contoh khusus dari model pengukuran konjenerik ketika setiap skor mengandung unsur yang identik (Lord & Novick, 1968). Masing-masing pengukuran yang bersifat paralel dapat digunakan secara bergantian jika memiliki reliabilitas dan validitas yang sama dalam memprediksi kriteria tertentu.

*Model Pengukuran Nilai Tau-Setara.* Traub (1994) mengatakan bahwa model pengukuran nilai tau-setara menghendaki asumsi yang lebih mudah dipenuhi dibanding dengan model paralel, karena model ini tidak menghendaki besarnya varians sesatan yang sama. DeVellis (1991) juga mengatakan bahwa model pengukuran nilai tau-setara mengasumsikan varians sesatan setiap butir dapat berbeda namun varians skor murni tetap harus sama. Model pengukuran nilai tau-setara pada dasarnya muncul dari gagasan untuk mengembangkan pengukuran majemuk terhadap konstruk ukur yang menghasilkan skor dalam unit skala yang sama namun masih memungkinkan kadar sesatan yang berbeda dalam pengukuran yang dilakukan. McPherson dan Rotolo (1995) mengatakan bahwa kesamaan unit skala yang sama tersebut terlihat dari nilai muatan faktor tidak terstandar dalam analisis faktor. Pernyataan ini juga sejalan dengan McDonald (1999) yang mengatakan bahwa secara operasional, model tau-setara menunjukkan kesetaraan keandalan ukur tiap indikator serta unidimensionalitas pengukuran. Keandalan dan unidimensionalitas pengukuran ini ditunjukkan dengan nilai muatan faktor (*factor loading*) tiap indikator yang tinggi.

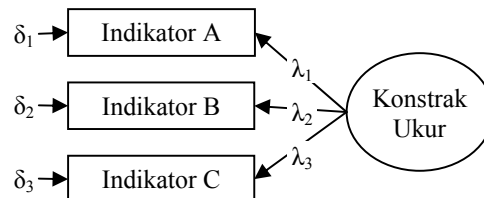
*Model Pengukuran Konjenerik.* Ciri model pengukuran konjenerik yang membedakan dengan model pengukuran lainnya salah satunya adalah hubungan skor murni sesuai dengan hubungan linier (Grujter dan Kamp, 2005). Model pengukuran konjenerik mengasumsikan bahwa hubungan antara satu komponen pengukuran dengan komponen lainnya mengikuti pola hubungan linier (Socan, 2000). Pendapat ini sesuai dengan pernyataan Steyer (1989) yang mengatakan bahwa dua pengukuran A dan pengukuran B dapat dikatakan konjenerik jika kedua skor murni yang dihasilkan memiliki hubungan sesuai dengan fungsi linier serta kedua sesatan pengukuran tersebut tidak berkorelasi.

Tabel 1 merangkum perbedaan statistik masing-masing model. Model pengukuran paralel adalah model yang cukup membatasi (*restricted*) karena menghendaki nilai harapan, varians skor, reliabilitas pengukuran dan korelasi antar pengukuran/butir yang sama. Sebaliknya model yang memiliki asumsi paling longgar adalah model konjenerik, karena statistik pengukuran/butir diperbolehkan memiliki besaran yang berbeda.

Tabel 1.  
Perbandingan Karakteristik Model Pengukuran dalam Pendekatan SEM

Model Pengukuran	Nilai Harapan ( $\mu$ )	Varians ( $\sigma$ )	Reliabilitas ( $r_{ii}$ )	Korelasi ( $r_{ij}$ )
Paralel	$\mu_1 = \mu_2 = \mu_3 = \mu$	$\sigma_1 = \sigma_2 = \sigma_3 = \sigma_i$	$r_{11} = r_{22} = r_{33} = r_{ii}$	$r_{12} = r_{13} = r_{23} = r_{ij}$
Nilai Tau-Setara	$\mu_1 = \mu_2 = \mu_3 = \mu$	$\sigma_1 = \sigma_2 = \sigma_3 = \sigma_i$	$r_{11} \neq r_{22} \neq r_{33} \neq r_{ii}$	$r_{12} = r_{13} = r_{23} = r_{ij}$
Konjenerik	$\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu$	$\sigma_1 \neq \sigma_2 \neq \sigma_3 \neq \sigma_i$	$r_{11} \neq r_{22} \neq r_{33} \neq r_{ii}$	$r_{12} \neq r_{13} \neq r_{23} \neq r_{ij}$

Melalui pendekatan SEM, Joreskog (1997) membedakan ketiga model pengukuran tersebut berdasarkan nilai muatan faktor dan varians sesatan pada analisis faktor konfirmatori (CFA). Gambar 1 menunjukkan model pengukuran dalam skema analisis faktor konfirmatori dengan tiga indikator, yaitu A, B dan C yang mengukur satu konstruk ukur. Tiap indikator memiliki muatan faktor ( $\lambda_i$ ) dan varians sesatan ( $\delta_i$ ). Muatan faktor menunjukkan seberapa besar porsi konstruk ukur di dalam indikator sedangkan varians sesatan ukur menunjukkan seberapa besar porsi sesatan dalam indikator tersebut. Indikator yang memiliki keandalan ukur yang baik adalah indikator yang memiliki porsi konstruk ukur yang tinggi dan sesatan yang sedikit. Ini adalah konsep dasar teori psikometri yang mengatakan bahwa di dalam varians skor tampak (X), dalam hal ini adalah indikator, terdapat varians skor murni (T), dalam hal ini adalah konstruk ukur, dan varians sesatan (E).



Gambar 1. Model Pengukuran dalam Skema Analisis Faktor Konfirmatori

Tabel 2. Perbandingan Karakteristik Model Pengukuran dalam Pendekatan SEM

Model Pengukuran	Muatan Faktor ( $\lambda_i$ )	Varians Sesatan ( $\delta_i$ )
Paralel	$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_i$	$\delta_1 = \delta_2 = \delta_3 = \delta_i$
Nilai Tau-Setara	$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_i$	$\delta_1 \neq \delta_2 \neq \delta_3 \neq \delta_i$
Konjenerik	$\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_i$	$\delta_1 \neq \delta_2 \neq \delta_3 \neq \delta_i$

Menurut Joreskog (1997), perbedaan model pengukuran ditandai dengan perbedaan besarnya muatan faktor dan varians sesatan di dalam model yang dapat dilihat pada Tabel 2. Model pengukuran paralel ditandai dengan nilai muatan faktor dan varians sesatan yang sama pada semua indikator, model nilai tau-setara ditunjukkan dengan nilai muatan faktor yang sama

namun memiliki nilai varian sesatan yang berbeda, sedangkan model pengukuran konjenerik memiliki nilai muatan faktor dan varian sesatan yang berbeda-beda.

#### B. Tinjauan Kesesuaian Model Pengukuran dengan Pengukuran Psikologi

Konstrak-konstrak psikososial seperti kepribadian, tingkat depresi atau sikap merupakan satu kesatuan dari beberapa indikator perilaku. Indikator-indikator tersebut tidak dapat ditentukan secara apriori akan tetapi didapatkan dari teori yang menjelaskan. Jika sebuah teori menjelaskan bahwa sebuah konstrak memuat beberapa indikator maka pengukuran terhadap konstrak tersebut harus mengukur indikator-indikator tersebut (Lucke, 2005). Pengukuran dalam ilmu sosial memiliki banyak keterbatasan dibandingkan ilmu eksakta. Terkait dengan indikator dari konstrak ukur, keterbatasan tersebut terletak pada banyaknya indikator yang dapat dimanifestasikan oleh konstrak ukur sehingga pengukuran dilakukan terbatas pada sampel indikator saja. Operasionalisasi indikator menjadi butir-butir ukur tidak memiliki standar baku sehingga satu indikator dapat diturunkan menjadi butir-butir ukur yang berbeda. Misalnya dalam penyusunan skala psikologi, satu indikator dapat dijabarkan menjadi berbagai butir pernyataan yang bervariasi.

Keterbatasan ini berpotensi menyebabkan butir-butir skala memiliki kesamaan target ukur namun memiliki keandalan ukur dan skala ukur yang berbeda-beda. Sama-sama mengukur harga diri, antara butir pernyataan *"Saya memiliki banyak kelebihan yang membanggakan"* dan *"Saya banyak memiliki kelebihan"*, keduanya memiliki kesamaan domain ukur akan tetapi berpotensi memiliki keandalan ukur yang berbeda. Pernyataan pertama lebih kuat dibanding dengan pernyataan pertama sehingga arah dan bentuk distribusi data pengukuran dari kedua butir tersebut berbeda. Perbedaan arah distribusi data yang ditunjukkan oleh rerata skor menyebabkan perbedaan keandalan ukur, sedangkan perbedaan bentuk distribusi yang ditunjukkan oleh nilai varians akan mempengaruhi perbedaan skala ukur antar kedua butir di atas (Graham, 2006).

Ilustrasi di atas menunjukkan bahwa upaya untuk mencapai kesetaraan keandalan dan skala ukur merupakan hal sulit untuk dilakukan. Dengan kata lain, asumsi-asumsi dalam model pengukuran paralel dan nilai tau setara sulit untuk diaplikasikan pada tataran operasional. Bahkan Kamata, Turhan, dan Darandari (2003) mengatakan bahwa hampir tidak mungkin model nilai tau setara dapat dicapai dalam hal kesetaraan kekuatan diskriminasi dan unidimensionalitas antar pengukuran-pengukuran yang dilakukan. Penelitian ini bertujuan untuk memverifikasi berbagai model pengukuran yang tepat dengan pengukuran di psikologi. Pertanyaan yang diajukan adalah diantara model pengukuran yang ada, model pengukuran mana yang lebih mewakili pengukuran psikologi. Penelitian ini dilakukan karena masih banyak ditemui peneliti yang secara monoton menggunakan formula properti psikometris tertentu tanpa mempertimbangkan asumsi dan model pengukuran yang mendasari formula tersebut. Dari ketiga model pengukuran di atas, penulis menghipotesiskan bahwa model konjenerik lebih mewakili pengukuran psikologi dibanding dengan model lainnya.

## METODE

### *Prosedur*

Penelitian ini dilakukan dengan menggunakan data sekunder yang bersumber dari tiga penelitian. Data pengukuran depresi diambil dari penelitian Retnowati (2004) yang mengembangkan model integrasi depresi pada remaja, Wicaksono (2003) yang meneliti tentang hubungan antara etos kerja dan sikap terhadap perubahan organisasi dan Widhiarso (2009) yang meneliti respons tipuan pada pengukuran psikologi.

#### *Responden*

Responden dalam penelitian ini berbeda-beda yang disesuaikan dengan instrumen pengukuran yang dikenakan. Pengukuran kepribadian dilakukan pada responden Mahasiswa Fakultas Psikologi UGM (N=90) yang dipilih dengan menggunakan teknik pengambilan sampel purposif. Responden yang berjenis kelamin laki-laki sebanyak 39 orang dan perempuan sebanyak 51 orang dengan rentang usia antara 19-21. Pengukuran depresi dilakukan pada remaja di Yogyakarta (N=2586) dengan rentang usia antara 12-19 tahun yang sedang bersekolah di SMA/SMK dengan jenis kelamin pria sebanyak 46 persen dan wanita sebanyak 54 persen. Pengukuran sikap dilakukan terhadap karyawan PT.Telkom di lingkungan DIVLAT Bandung (N=100) yang berusia 26 hingga 55 tahun dengan dengan jenis kelamin pria sebanyak 73 persen dan wanita sebanyak 14 persen.

#### *Pengukuran.*

*Kepribadian.* Pengukuran kepribadian dilakukan dengan menggunakan *Big Five Inventory (BFI)* yang telah diterjemahkan dalam Bahasa Indonesia oleh penulis. BFI dikembangkan oleh John, Donahue, & Kentle (1991), menggunakan model skala Likert dengan lima alternatif respons dan pengukurannya dengan menggunakan pelaporan mandiri (*self report*) yaitu dengan meminta responden merespon butir pernyataan. Terdapat lima alternatif respons dalam BFI yaitu dari sangat setuju hingga sangat tidak setuju dengan penyekoran butir bergerak dari 1 hingga 5. Skala ini mengukur lima faktor kepribadian antara lain ekstrasversi (*extroversion*), keramahan (*agreeableness*), keuletan (*conscientiousness*), neurotisisme (*neuroticism*) dan keterbukaan (*openness*). Contoh butir pada masing-masing faktor antara lain sebagai berikut. "Menjadi penggerak kelompok" (faktor ekstrasversi), "Memiliki sedikit kepedulian terhadap orang lain" (keramahan), "Selalu mempersiapkan diri sebelum melakukan sesuatu" (keuletan), "Mudah meredakan perasaan tertekan" (emosi stabil) dan "Memiliki banyak cadangan kosa kata" (ekstrasversi). Penyekoran butir pada faktor neurotisisme dalam penelitian ini dilakukan secara terbalik sehingga semakin tinggi skor faktor ini menunjukkan semakin tinggi kestabilan emosi subjek. Prosedur ini sesuai dengan pernyataan John, Donahue, & Kentle (1991) yang mengatakan bahwa penyekoran BFI pada faktor neurotisisme dapat dilakukan secara terbalik. BFI versi Bahasa Indonesia telah diujicobakan oleh peneliti pada sampel mahasiswa (N=185) yang menghasilkan nilai reliabilitas ( $\alpha$ ) sebagai berikut ekstrasversi (0.839), keramahan (0.789), keuletan (0.924), kestabilan emosi (0.848) dan keterbukaan (0.807). Hasil ini mirip hasil penelitian John dan Srivastava (1999) yang melaporkan bahwa BFI memiliki reliabilitas ( $\alpha$ ) antara 0.75 hingga 0.80 dan reliabilities tes-tes ulang antara 0.80 hingga 0.90. Validitas BFI pada versi asli yang dikorelasikan dengan NEO-FFI dan TDA menghasilkan rata-rata korelasi sebesar 0.83 hingga 0.91 (John & Srivastava, 1999).

*Depresi.* Instrumen pengukuran depresi yang dipakai adalah *Children's Depression Inventory (CDI)* yang telah diterjemahkan dalam Bahasa Indonesia. CDI berbentuk pilihan ganda dengan yang terdiri atas tiga pilihan jawaban yang diskor dengan nilai 0, 1, 2. Semakin tinggi skor yang didapatkan pada sebuah item menunjukkan tingkat depresi yang semakin tinggi. Respon terhadap item di dalam CDI didasarkan atas pilihan yang paling

menggambarkan keadaan diri subjek dalam waktu seminggu terakhir. CDI mencakup aspek kognitif, afektif, psikomotor dan vegetatif dari gangguan depresi. Item-item yang dimuat dalam CDI meliputi kelompok simtom suasana hati negatif, problem interpersonal, ketidakmampuan, anhedonia, dan harga diri negatif. Reliabilitas CDI versi Bahasa Indonesia ini diestimasi dengan teknik konsistensi internal dan mendapatkan nilai reliabilitas keseluruhan item yang cukup tinggi yaitu sebesar  $\alpha=0,7135$  (N=252). Reliabilitas CDI dengan berdasarkan data yang dipakai dalam tulisan ini adalah sebesar 0,746 (N=2.987). Validitas CDI dibuktikan oleh Retnowati (2003) melalui pengujian validitas kriteria dan validitas konstruk. Validitas kriteria dibuktikan dengan nilai korelasi yang cukup tinggi antara CDI dengan BDI (*Beck Depression Inventory*) dengan  $r_{xy} = 0,561$  (N=109) sedangkan validitas konstruk dianalisis dengan analisis faktor konfirmatori yang menghasilkan indeks ketepatan model yang memuaskan menurut kriteria model fit yang dipaparkan Hair, Anderson, Tatham, & Black, (1995), misalnya indeks GFI=0,995 (>0,90); AGFI=0,985 (>0,90); CFI=0,987 (>0,95); dan RMSEA=0,047 (<0,08). Hasil ini menunjukkan bahwa masing-masing butir pada kelima faktor di dalam CDI dapat menjelaskan konstruk depresi.

*Sikap.* Pengukuran sikap diwakili oleh pengukuran sikap terhadap perubahan organisasi. Skala ini dikembangkan oleh Wicaksono (2003) yang berisi 44 butir yang menggunakan model Likert dengan empat alternatif respon dari sangat setuju hingga sangat tidak setuju. Skala ini mengukur tiga aspek sikap yaitu pola pikir, keahlian dan kondisi lingkungan kerja. Validasi skala ini dilakukan dengan menggunakan validitas isi dan reliabilitas pengukuran diestimasi dengan menggunakan teknik konsistensi internal. Skala ini memiliki properti psikometris yang memuaskan. Validasi secara internal dibuktikan oleh persetujuan pakar yang memahami bidang psikologi industri dan organisasi terhadap butir pernyataan yang dilibatkan dalam skala sedangkan reliabilitas yang diestimasi dengan menggunakan teknik alpha menghasilkan nilai 0.969. Koefisien korelasi butir total bergerak antara 0,3249 hingga 0,8867 menunjukkan bahwa pernyataan dalam skala dapat membedakan responden yang memiliki sikap yang positif dan negatif terhadap perubahan organisasi. Butir pernyataan yang dilibatkan dalam skala ini misalnya, “*Saya akan merasa bangga bila dapat terlibat dalam program-program baru*” (butir 2), “*Kebijakan yang baru hanya akan merusak ritme kerja yang telah ada*” (butir 18), dan “*Perubahan kebijakan akan menciptakan pengalaman-pengalaman baru dalam melakukan pekerjaan*” (butir 25).

#### *Analisis Data*

Data dianalisis dengan menggunakan teknik analisis faktor konfirmatori melalui metode estimasi *maximum likelihood* dengan menggunakan program LISREL 8.0. Setiap data pengukuran akan diuji berdasarkan tiga model pengukuran, yaitu model paralel, nilai tau setara dan konjenerik. Model pengukuran ditentukan berdasarkan parameter mana yang ditetapkan (*constrained*) untuk memiliki besaran yang sama. Model paralel memiliki muatan faktor dan sesatan pengukuran yang sama sedangkan model nilai tau setara kesamaan nilai parameter hanya pada nilai muatan faktor, sedangkan model konjenerik tidak menetapkan parameter yang nilainya sama. Ketepatan masing-masing model pengukuran dengan data pada ketiga pengukuran konstruk psikologi dibandingkan. Nilai yang dipakai untuk membandingkan adalah kai-kuadrat, GFI dan RMSEA.

Pengujian ketepatan model dilakukan pada masing-masing faktor variabel yang dilibatkan dalam penelitian ini karena pengujian dilakukan pada pengukuran yang bersifat unidimensi. Perbandingan antar ketepatan model dilakukan dengan menggunakan uji kai kuadrat. Kai

kuadrat hitung didapatkan dari selisih nilai kai kuadrat antar dua model sedangkan derajat kebebasan didapatkan dari selisih derajat kebebasan kedua model. Nilai kai kuadrat yang signifikan ( $p < 0.05$ ) menunjukkan tidak ada perbedaan ketepatan model yang signifikan.

## HASIL

Hasil pengujian ketepatan data menunjukkan bahwa ketepatan model bervariasi pada antar pengukuran. Secara umum pengukuran depresi memiliki ketepatan model yang lebih besar dibanding dengan pengukuran kepribadian dan sikap. Hasil selengkapnya ketepatan model pengukuran yang dihitung dengan menggunakan kai kuadrat, GFI dan RMSEA pada masing-masing pengukuran dapat dilihat pada Tabel 3.

Hasil pengujian perbandingan ketepatan antar model menunjukkan bahwa model konjenerik memiliki ketepatan model yang lebih kuat dibanding dengan model paralel dan nilai tau setara. Hal ini ditunjukkan dengan nilai selisih kai kuadrat yang menunjukkan adanya perbedaan yang signifikan ( $p > 0.05$ ). Hanya pada faktor 1 dan faktor 2 pengukuran kepribadian, model pengukuran konjenerik tidak berbeda signifikan dengan model paralel dan nilai tau setara ( $p < 0.05$ ). Di sisi lain, model nilai tau setara secara umum memiliki nilai ketepatan model yang lebih kuat dibanding dengan model paralel. Dari tiga belas uji perbandingan antara model paralel dan nilai tau setara yang dilakukan, ada empat perbandingan yang tidak menunjukkan perbedaan signifikan ( $p < 0.05$ ), yaitu pada pengukuran kepribadian faktor 1 dan 2, pengukuran depresi pada faktor 2 dan pengukuran sikap pada faktor 1. Model paralel adalah model yang memiliki nilai ketepatan model yang paling rendah dibanding dengan model lainnya. Nilai kai-kuadrat model ini jauh di atas nilai kai kuadrat model lainnya sehingga uji statistik yang dilakukan banyak menunjukkan model paralel lebih lemah dibanding dengan kedua model lainnya.

Tabel 3. Hasil Identifikasi Ketepatan Model Pengukuran Melalui Analisis Faktor Konfirmatori

Variabel	Paralel			Nilai Tau Setara			Konjenerik		
	$\chi^2$ (df)	GFI	RMSEA	$\chi^2$	GFI	RMSEA	$\chi^2$	GFI	RMSEA
<i>Kepribadian</i>									
Faktor 1	138.61 (34)	0.72	0.22	127.12 (27)	0.73	0.21	123.40 (20)	0.74	0.28
Faktor 2	78.82 (43)	0.80	0.84	75.39 (35)	0.84	0.11	65.24 (27)	0.86	0.17
Faktor 3	168.25 (43)	0.74	0.16	145.65 (35)	0.78	0.16	102.65 (27)	0.80	0.18
Faktor 4	262.18 (34)	0.58	0.27	131.85 (27)	0.73	0.21	38.21 (20)	0.90	0.10
Faktor 5	278.82 (53)	0.61	0.22	215.95 (44)	0.67	0.21	140.00 (35)	0.76	0.18
<i>Depresi</i>									
Aspek 1	539.91 (19)	0.94	0.10	552.26 (14)	0.94	0.11	122.35 (9)	0.99	0.07
Aspek 2	115.67 (8)	0.98	0.07	113.34 (5)	0.98	0.08	89.35 (2)	0.99	0.12
Aspek 3	150.16 (8)	0.98	0.07	138.82 (5)	0.09	0.98	10.26 (2)	0.99	0.03
Aspek 4	960.32 (34)	0.93	0.10	885.52 (27)	0.93	0.10	396.11 (20)	0.97	0.08
Aspek 5	244.44 (13)	0.97	0.08	196.95 (9)	0.97	0.08	64.38 (5)	0.99	0.06
<i>Sikap</i>									
Aspek 1	316.40 (118)	0.61	0.16	293.13 (104)	0.62	0.19	255.55 (90)	0.67	0.16
Aspek 2	291.12 (118)	0.66	0.13	208.79 (104)	0.72	0.14	169.87 (90)	0.74	0.14
Aspek 3	253.15 (103)	0.64	0.15	206.33 (90)	0.68	0.17	175.14 (77)	0.72	0.14



Tabel 4. Hasil Uji Statistik Perbandingan Ketepatan antar Model Pengukuran

Variabel	Paralel vs Nilai Tau Setara	Paralel vs Konjenerik	Nilai Tau Setara vs Konjenerik
<b>Pribadi</b>			
Faktor 1	11.49 (7)	15.21 (14)	3.71 (7)
Faktor 2	3.42 (8)	13.57 (16)	10.15 (8)
Faktor 3	22.5 (8)**	65.6 (16)**	43 (8)**
Faktor 4	130.33 (7)**	223.97 (14)**	93.63 (7)**
Faktor 5	62.87 (9)**	138.82 (18)**	75.94 (9)**
<b>Depresi</b>			
Faktor 1	12.35 (5)**	417.55 (10)**	429.9 (5)**
Faktor 2	2.32 (3)	26.32 (6)**	23.99 (3)**
Faktor 3	11.34 (3)**	139.9 (6)**	128.56 (3)**
Faktor 4	74.80 (7)**	564.21 (14)**	489.40 (7)**
Faktor 5	47.49 (4)**	180.06 (6)**	132.57 (4)**
<b>Sikap</b>			
Faktor 1	23.26 (14)	60.8 (28)**	37.57 (14)**
Faktor 2	82.33 (14)**	121.25 (28)**	38.91 (14)**
Faktor 3	46.81 (13)**	78.01 (26)**	31.19 (13)**

Keterangan : \*\*= $p < 0.01$ ; \*= $p < 0.05$

## DISKUSI

Penelitian ini bertujuan untuk membandingkan ketepatan model pengukuran teori skor murni klasik yaitu antara model pengukuran paralel, nilai tau setara dan konjenerik. Penelitian ini tidak memusatkan pada pengujian ketepatan model pengukuran, akan tetapi lebih pada pengujian perbandingan ketepatan model pengukuran. Oleh karena itu meski mendapatkan beberapa nilai ketepatan model pengukuran yang rendah, misalnya pada pengukuran kepribadian dan sikap, pengujian perbandingan ketepatan model tetap dilakukan. Hal ini dikarenakan penulis ingin mendapatkan informasi mengenai sejauh mana model pengukuran yang diuji berbeda antara satu dengan lainnya. Dari hasil pengujian model didapatkan bahwa model pengukuran konjenerik memiliki ketepatan model yang lebih baik dibanding dengan model paralel dan nilai tau setara. Temuan ini menunjukkan bahwa pengukuran psikologi merupakan sesuatu yang kompleks sehingga memerlukan asumsi yang lebih longgar untuk mampu menjelaskan variasi data yang didapatkan dari pengukuran.

Hasil penelitian ini mendukung pernyataan Widhiarso (2009) yang mengatakan bahwa pengukuran dalam bidang psikologi memiliki karakteristik yang unik dari sisi karakteristik konstruk psikologi, aspek-aspek yang dilibatkan dalam penyusunan alat ukur, jumlah item di dalam instrumen dan kemungkinan adanya satuan pengukuran yang berbeda sehingga pengukuran dalam bidang psikologi lebih menggambarkan model konjenerik dibanding dengan pengukuran lainnya. Model konjenerik dibuktikan menggambarkan fenomena pengukuran dalam bidang psikologi dengan baik karena mewedahi adanya potensi satuan ukur yang berbeda antara satu aitem satu dengan aitem lainnya dalam sebuah instrumen ukur yang

sama. Hal ini didukung dengan antara item satu dengan item lainnya memiliki kapabilitas yang berbeda sebagai indikator konstruk ukur. Teori Psikometri berfokus terutama pada tes yang homogen, hanya mengukur satu atribut dari entitas psikososial. Namun, kompleksitas perilaku psikososial sering memerlukan tes yang heterogen, berukuran lebih dari satu atribut. Dalam presentasi ini, keandalan dan konsistensi internal diperpanjang menjadi tes heterogen di bawah rubrik teori uji konjenerik.

#### DAFTAR PUSTAKA

- DeVellis, R. (1991). *Scale Development*. Thousand Oaks: Sage Publications.
- Graham, J. M. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability: What They Are and How to Use Them. *Educational and Psychological Measurement*, 66(6), 930-944.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate Data Analysis*. New Jersey: Prentice Hall.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory - Versions 4a and 54*. California: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research (2nd ed)*. New York: Guilford Press.
- Joreskog, K. G. (1997). *Lisrel 8: User's Reference Guide*: Scientific Software.
- Kamata, A., Turhan, A., & Darandari, E. (2003). *Estimating Reliability for Multidimensional Composite Scale Scores*. Paper presented at the Annual meeting of American Educational Research Association.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley Publishing Company.
- Lucke, J. F. (2005). The  $\alpha$  and the  $\omega$  of congeneric test theory: An extension of reliability and internal consistency to heterogeneous tests. *Applied Psychological Measurement*, 29(1), 65-81.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- McPherson, J. M., & Rotolo, T. (1995). Measuring the Composition of Voluntary Groups: A Multitrait-Multimethod Analysis. *Social Forces*, 73(3), 1097-1115.
- Meredith, W. M., Frederiksen, C. H., & McLaughlin, D. H. (1974). Statistics and Data Analysis. *Annual Review of Psychology*, 25(1), 453-505.
- Millsap, R. E., & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research*, 26(3), 479-598.
- Raykov, T. (1997). Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement*, 21(2), 173-184.
- Socan, G. (2000). Assessment of Reliability when Test Items are not Essentially t-Equivalent. In A. Ferligoj & A. Mrvar (Eds.), *Developments in Survey Methodology Editors*. Ljubljana: FDV.
- Traub, R. E. (1994). *Reliability for the social sciences: theory and applications*. Thousand Oaks: Sage Publications.
- Widhiarso, W. (2009). Koefisien reliabilitas pada pengukuran kepribadian yang bersifat multidimensi. *Psikobuana*, 1, 39-48.

