



ANALISIS FAKTOR KONFIRMATORI GUNA MENGESTIMASI RELIABILITAS MULTIDIMENSI

Gaguk Margono

Universitas Negeri Jakarta, Jl. Rawamangun Muka, Rawamangun, Jakarta 13220
Surel: g_margono@yahoo.com

ABSTRACT

The purpose of this paper was to draw on the application of Confirmatory Factor Analysis (CFA) in Structural Equation Modeling (SEM) to estimate the reliability of students satisfaction as an internal customer instrument. Student satisfaction is the feeling of student performance resulting from comparing a product's perceived (outcome) in relation to the expectations of students. Reliability refers to accuracy and precision of a measurement instrument. Multidimensional reliability is rarely used in the realm of research. Measurement and computation is described in this article uses an instrument measuring student satisfaction as an internal customers. Survey methods used in research and sampling used simple random sampling. This instrument has been tested 75 students. It can be concluded that the instrument gauges student satisfaction as internal customers using multidimensional reliability coefficient has a high accuracy when compared with a unidimensional reliability coefficient. Expected in advanced research used another formula multidimensional reliability, included if used SEM.

Keywords: CFA, multidimensional reliability coefficient, satisfaction as an internal customers instrument

PENDAHULUAN

Untuk mengukur kepuasan mahasiswa digunakan suatu ukuran subyektif atau *soft measures* sebagai indikator mutu atau kualitas. Ukuran ini disebut lunak (*soft*), sebab ukuran-ukuran ini berfokus pada persepsi dan sikap daripada hal-hal yang konkret yang disebut kriteria obyektif. Oleh karena berfokus pada persepsi dan sikap maka alat pengukur yang digunakan dapat berupa kuesioner kepuasan mahasiswa yang dapat diukur melalui mutu atau kualitas pelayanan dari institusi pendidikan tinggi tersebut.

Mutu atau kualitas (*quality*) merupakan suatu istilah yang dinamis yang terus bergerak; jika bergerak maju dikatakan mutunya bertambah baik, sebaliknya jika bergerak mundur dikatakan mutunya merosot. Mutu berarti dapat berarti *superiority* atau *excellence* yaitu melebihi standar umum yang berlaku. Sesuatu dikatakan bermutu jika terdapat kecocokan antara syarat-syarat yang dimiliki oleh benda atau jasa yang dikehendaki dengan maksud dari orang yang menghendaknya. Menurut Idrus *et al.*

(2000) “...*the fitness purpose as perceived by the custome.*” Misalnya, mutu proses belajar cocok dengan apa yang diharapkan oleh mahasiswa; makin jauh melampaui apa yang diharapkan makin bermutu, jika terjadi sebaliknya, makin tidak bermutu.

Langkah pertama mengukur kualitas pelayanan adalah mengidentifikasi karakteristik kualitas pelayanan. Daftar karakteristik ini dapat digeneralisasi dalam berbagai cara dengan menggunakan berbagai sumber informasi. Salah satu cara adalah mencari literatur seperti jurnal yang mungkin memuat dimensi mutu jasa. Peneliti-peneliti seperti Parasuraman, Zeithaml, dan Berry (1985) telah menyimpulkan bahwa mutu jasa dapat diuraikan dengan dasar 10 dimensi. Mereka mencoba untuk mengukur sepuluh dimensi, ternyata pelanggan hanya dapat membedakan 5 dimensi yang disebut *ServQual* (Parasuraman, Zeithaml, dan Berry, 1988) memberi kesan bahwa dimensi 10 yang asli saling tumpang-tindih satu sama lain. Lima dimensi mutu pelayanan adalah sesuatu yang terwujud (*tangible*), kehandalan (*reliability*), daya tanggap (*responsiveness*), jaminan (*assurance*), dan empati (*empathy*). Lebih lanjut tentang dimensi ini dapat dibaca dari publikasi pada kualitas pelayanan jasa oleh Zeithaml, Parasuraman dan Berry (1990).

Dimensi pertama dari kualitas pelayanan menurut konsep *ServQual* ini adalah *tangible* karena suatu jasa tidak dapat dicium dan tidak dapat diraba, maka *tangible* menjadi penting sebagai ukuran terhadap pelayanan. *Tangible* merupakan kemampuan untuk memberi fasilitas fisik kampus dan perlengkapan perkuliahan yang memadai menyangkut penampilan karyawan/dosen dan pejabat serta sarana umum. Misalnya: ketersediaan ruang menyangkut kelengkapan dan ketersediaan peralatan, kenyamanan dan kecanggihan kampus, fasilitas komputer dan internet, perpustakaan, ruang kuliah, ruang seminar, ruang dosen, media perkuliahan, laboratorium, unit produksi, kantin, pusat bimbingan karir, layanan kesehatan, tempat ibadah, tempat istirahat dan tempat parkir, serta sarana transportasi. Mahasiswa akan menggunakan indra penglihatan untuk menilai suatu kualitas pelayanan dari segala sarana dan fasilitas yang ada.

Kedua, dimensi *reliability* yaitu dimensi yang mengukur kehandalan dari pendidikan tinggi dalam memberikan pelayanan kepada mahasiswa. Ada dua aspek dari dimensi ini yakni: (1) kemampuan perguruan tinggi untuk memberikan pelayanan seperti yang dijanjikan, dan (2) seberapa jauh perguruan tinggi memberikan pelayanan yang akurat atau tidak *error*. Dengan kata lain *reliability* merupakan kemampuan

pejabat, karyawan/dosen dalam memberikan pelayanan sesuai dengan yang dijanjikan (tepat waktu), dengan segera, relevan, dan akurat sehingga memuaskan mahasiswa. Contoh: pengembangan administrasi, kurikulum dan penawaran mata kuliah sesuai tuntutan keterampilan, profesi dan dunia kerja, perkuliahan berlangsung lancar sesuai jadwal, penilaian hasil studi obyektif, fair, dan tepat waktu.

Ketiga, *responsiveness* adalah dimensi kualitas pelayanan yang dinamis. Harapan mahasiswa terhadap kecepatan pelayanan hampir dapat dipastikan akan berubah dengan kecenderungan naik dari waktu ke waktu. *Responsiveness* merupakan kesediaan para pejabat, dosen/karyawan untuk membantu dan memberikan pelayanan sesuai kebutuhan mahasiswa. Contoh: pejabat mudah ditemui untuk diminta bantuan, dosen mudah ditemui untuk keperluan konsultasi, proses belajar mengajar berlangsung interaktif dan variatif serta memungkinkan para mahasiswa mengembangkan kapasitas dan kreativitas, pengelola memberi fasilitas yang memadai sesuai dengan kebutuhan mahasiswa dan dunia kerja.

Dimensi keempat dari 5 dimensi kualitas pelayanan yang menentukan kepuasan pelanggan adalah *assurance*, yaitu dimensi jaminan kualitas yang berhubungan dengan kemampuan institusi dan perilaku *front-line* staf dalam menanamkan rasa percaya dan keyakinan kepada para mahasiswa. *Assurance* mencakup kompetensi, pengetahuan, keterampilan, kesopanan, hormat terhadap setiap orang, dan sifat percaya yang dimiliki para staf. Contoh: para dosen menyampaikan kuliah sesuai dengan bidang keahlian/pengalamannya, dosen selalu berusaha menambah wawasan dengan membaca, menghadiri seminar, mengikuti pelatihan, studi lanjut, melakukan penelitian, memiliki sikap dan perilaku baik, serta seluruh jajaran, dan organisasi mencerminkan profesionalisme sesuai dengan yang diatur dalam standar. Berdasarkan banyak riset, ada 4 aspek dari dimensi ini, yaitu keramahan, kompetensi, kredibilitas, dan keamanan.

Kelima, empati adalah kemampuan pejabat, karyawan/dosen sehingga memberi pelayanan sepenuh hati, antara lain kemudahan dalam berkomunikasi, perhatian secara pribadi dan pemahaman akan kebutuhan spesifik individual mahasiswa. Contoh: dosen berusaha mengenal nama mahasiswanya, dosen penasehat akademis sungguh-sungguh berperan sebagai konselor, dan sebagai supervisor bukan sekedar editor bahasa, dan pejabat mudah dihubungi baik di ruang kerja, via telepon, email dan sebagainya. Empati ini berkaitan dengan teori perkembangan kebutuhan manusia dari Maslow. Pada tingkat

kebutuhan semakin tinggi, kebutuhan manusia tidak lagi dengan hal-hal yang primer seperti kebutuhan fisik, keamanan dan sosial terpenuhi, maka dua kebutuhan lagi akan dikejar oleh manusia yaitu kebutuhan akan ego dan aktualisasi diri. Dua kebutuhan terakhir inilah yang banyak berhubungan dengan dimensi empati.

Di bidang pendidikan, ekonomi, bisnis maupun manajemen, penilaian yang baik memerlukan pengukuran yang dapat diandalkan atau dipercaya. Demikian juga pada bidang pendidikan dan psikologi. Menurut Naga (1992) untuk pengukuran pendidikan dan psikologi mencakup beberapa hal. Pertama, mengukur ciri terpendam yang tak terlihat yang ada pada responden. Kedua, untuk mengukur ciri terpendam tersebut responden diberi stimulus berupa kuesioner atau alat ukur yang tepat. Ketiga, stimulus direspons oleh responden dengan harapan respons mencerminkan dengan benar ciri terpendam yang ingin diukur. Keempat, respons diskor dan dapat ditafsirkan secara memadai. Kemudian, perlu dipertanyakan sejauh manakah skor yang diperoleh dapat mencerminkan secara tepat ciri terpendam yang hendak diukur? Apakah instrumen yang dipakai sebagai stimulus itu mampu mengungkap secara benar ciri terpendam yang tak tampak itu? Kedua pertanyaan tersebut berkenaan dengan validitas. Sedang yang berkaitan dengan reliabilitas, apakah tanggapan yang diberikan oleh para peserta sudah dapat dipercaya untuk digunakan sebagai bahan penskoran bagi atribut psikologis itu?

Menurut Wiersma (1986), reliabilitas ialah konsistensi suatu instrumen mengukur sesuatu yang hendak diukur. Reliabilitas menunjukkan sejauh mana hasil pengukuran dengan alat tersebut dapat dipercaya. Oleh karena itu reliabilitas merupakan indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya atau dapat diandalkan. Bila suatu instrumen dipakai berulang-ulang untuk mengukur gejala yang sama dan hasil yang diperoleh relatif stabil atau konsisten, maka instrumen tersebut terpercaya. Dengan kata lain hasil pengukuran itu diharapkan sama apabila pengukuran diulang.

Dengan pendekatan varians, Kerlinger (2000) menyusun dua definisi tentang reliabilitas: (1) adalah proporsi varians “yang sebenarnya” terhadap varians total yang diperoleh untuk data yang didapatkan dengan suatu instrumen pengukur dan dapat dituliskan dalam persamaan $r_{tt} = v_{\infty} / v_t$ dengan catatan v_{∞} adalah varians murni dan v_t adalah varians total, dan (2) adalah proporsi varians keliru yang dihasilkan dengan suatu instrumen pengukur yang dikurangkan pada 1.00, dengan indeks 1.00 menunjukkan

koefisien reliabilitas sempurna, dan dapat ditulis dengan persamaan $r_{tt} = 1 - v_e/v_t$, di mana v_e adalah varians keliru dan v_t adalah varians total. Oleh karena itu reliabilitas merupakan indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya atau dapat diandalkan.

Secara garis besar ada tiga kategori besar dalam pengukuran reliabilitas: (1) tipe stabilitas (misalnya: tes ulang, bentuk paralel, dan bentuk alternatif), (2) tipe homogenitas atau internal konsistensi (misalnya: belah dua, Kuder-Richardson, *alpha* Cronbach, *theta* dan *omega*), dan (3) tipe ekuivalen (misalnya: butir-butir paralel pada bentuk alternatif dan reliabilitas antar penilai (*inter-rater reliability*)). Instrumen diberikan kepada sekelompok subjek satu kali lalu dengan cara tertentu dihitung estimasi reliabilitasnya. Pendekatan pengukuran satu kali ini menghasilkan informasi mengenai konsistensi internal instrumen. Konsistensi internal merupakan pernyataan-pernyataan tersebut mengukur aspek yang sama atau merefleksikan homogenitas butir-butir pernyataan.

Makin tinggi koefisien reliabilitas, makin dekat nilai skor amatan dengan skor yang sesungguhnya, sehingga nilai skor amatan dapat digunakan sebagai pengganti komponen skor yang sesungguhnya. Ukuran tinggi atau rendahnya koefisien reliabilitas tidak hanya ditentukan oleh nilai koefisien. Tafsiran tinggi rendahnya nilai koefisien diperoleh melalui perhitungan, ditentukan juga oleh standar pada cabang ilmu yang terlibat di dalam pengukuran itu. Makin tinggi koefisien reliabilitas suatu instrumen, maka kemungkinan kesalahan yang terjadi akan makin kecil kalau orang membuat keputusan berdasarkan skor yang diperoleh dalam instrumen tersebut.

Pada umumnya pengukuran karakteristik afektif memberikan koefisien reliabilitas yang lebih rendah daripada pengukuran ranah kognitif, karena karakteristik kognitif cenderung lebih stabil daripada karakteristik afektif. Menurut Gable (1986) koefisien reliabilitas instrumen ranah kognitif biasanya kira-kira 0,90 atau lebih, sedangkan koefisien reliabilitas instrumen ranah afektif kurang dari 0,70. Koefisien reliabilitas pada taraf 0,70 atau lebih biasanya dapat diterima sebagai reliabilitas yang baik (Litwin, 1995).^[10] Sedangkan menurut Naga (1992) koefisien reliabilitas yang memadai sebaiknya terletak di atas 0,75.

Pada setiap penelitian yang menggunakan pengukuran psikologis selalu menerapkan pengujian validitas dan reliabilitas. Namun dalam perjalanannya di bidang

psikometri, para ahli belum ada kesepakatan tentang koefisien reliabilitas atau rumus yang mana untuk reliabilitas antar peneliti. Pertama, masih banyak peneliti yang dinilai cukup kompeten masih banyak yang kurang tepat dalam melaporkan reliabilitas hasil pengukuran mereka (Thompson, 1994).

Kedua, masalah yang muncul adalah penggunaan koefisien reliabilitas oleh para peneliti secara monoton tanpa mempertimbangan asumsi yang mendasari koefisien tersebut. Para peneliti tanpa sadar menggunakan koefisien alpha yang juga dengan tanpa sadar bahwa untuk koefisien ini memerlukan asumsi yang sulit dipenuhi. Jika asumsi tidak dipenuhi maka koefisien alpha yang dihasilkan adalah nilai di batas estimasi terendah. Banyak peneliti hanya terpaku pada penggunaan koefisien alpha dalam mengestimasi reliabilitas. Popularitas koefisien alpha Cronbach ini lahir karena faktor: 1) teknik komputasi relatif mudah, karena hanya memerlukan informasi berupa varians skor total, dan 2) distribusi sampling sudah diketahui sehingga penentuan interval kepercayaan pada populasi sangat dimungkinkan (Feld dan kawan-kawan, 1987).

Ketiga, permasalahan yang berhubungan dengan asumsi yang menjadi syarat dalam mengestimasi reliabilitas. Pada ranah empiris selain persyaratan adanya sifat paralel, persyaratan *tau-equivalent* merupakan tantangan yang cukup berat bagi peneliti dalam menyusun instrumen pengukuran. Hal ini didukung oleh Kamata dan kawan-kawan (2003) yang menemukan bahwa asumsi kesetaraan, daya diskriminasi antar komponen tes dan unidimensionalitas pengukuran merupakan hal relatif sulit dicapai. Jika asumsi *essentially tau-equivalent* tidak dapat dipenuhi maka koefisien alpha menghasilkan nilai reliabilitas yang sangat kecil, sehingga koefisien tersebut di bawah estimasi.

Keempat, wacana pengukuran adalah masalah unidimensionalitas pengukuran. Unidimensionalitas adalah aspek penting dalam mengestimasi reliabilitas. Hasil pengukuran psikologis yang bersifat unidimensi sangat sulit dicapai, terutama dalam konteks domain kepribadian yang kebanyakan memuat area varians-varians *traits* yang luas. Socan (2000) menulis bahwa analisis faktor yang dilakukan dari beberapa penelitian banyak kasus multidimensi dibanding dengan unidimensi.

Masalah asumsi bukan menjadi masalah utama dalam menyusun model konsistensi internal, namun masalah ini menjadi bahan kajian banyak peneliti dalam pengkajian reliabilitas. Seperti penelitian Vehkahlati (2000) yang menyimpulkan bahwa

asumsi yang tidak cukup realistis pada teori skor murni klasik adalah asumsi unidimensionalitas skor murni yang secara praktis sulit dibuktikan. Jadi kajian multidimensionalitas pengukuran muncul ke permukaan karena banyak kasus ditemui bahwa juga adanya korelasi antar butir di dalam dimensi tersebut kadang-kadang lebih tinggi dibanding dengan korelasi antar butir dalam tes.

Pada pengembangan instrumen pengukuran dalam bidang pendidikan banyak mengasumsikan penggunaan pengukuran yang bersifat unidimensi yang secara konseptual dirumuskan bahwa hanya ada satu jenis faktor kemampuan, kepribadian, sifat, maupun sikap yang diukur oleh satu instrumen pengukuran. Tetapi, banyak penelitian menunjukkan bahwa asumsi unidimensi tersebut sulit dipenuhi dengan ditemukannya beberapa faktor baru yang ikut diukur dalam satu instrumen. Dengan kata lain, instrumen yang bersifat psikologis yang sering dipakai peneliti cenderung bersifat multidimensi.

Beberapa alasan pentingnya pengukuran reliabilitas yang bersifat multidimensi seperti dikemukakan oleh Widhiarso dan Mardapi (2010) dengan uraian sebagai berikut: pertama, karakteristik konstruk psikologis yang umumnya bersifat multidimensi. Kedua, adanya pelibatan aspek-aspek dalam penyusunan instrumen psikologis biasanya diawali dengan penurunan butir-butir dari beberapa aspek teoretis dan kecenderungannya bersifat multidimensi. Ketiga, jumlah butir di dalam instrumen. Jumlah butir yang terlalu banyak dapat menambah potensi penambahan varians *error* dalam butir sehingga memunculkan dimensi baru dari dimensi yang ditetapkan semula. Jumlah butir dan juga bentuk skala mempengaruhi sikap responden terhadap butir yang kemudian mempengaruhi tanggapan mereka terhadap instrumen. Keempat, teknik penulisan butir. Spector dan kawan-kawan (1997) menemukan bahwa teknik penulisan butir yang memiliki arah yang terbalik antara positif (*favorable*) dan negatif (*unfavorable*) dapat membentuk dimensi ukur baru padahal dalam pengambilan data banyak skala psikologi menggunakan teknik penulisan butir yang berbeda arah. Kelima, satuan pengukuran yang berbeda. Pengukuran bidang psikologis cenderung memiliki satuan ukur yang berbeda antara butir satu dengan butir lainnya memiliki kapabilitas yang berbeda sebagai indikator konstruk ukur. Kondisi ini akan menyebabkan hasil pengukuran cenderung bersifat multidimensi.

Dengan demikian dapat disimpulkan bahwa pengukuran psikologis, baik mengukur konstruk kognitif maupun nonkognitif sangat rentan terhadap kemajemukan atribut yang diukur (multidimensi). Selanjutnya dengan memahami kecenderungan pengukuran psikologis lebih pada model pengukuran multidimensi dibandingkan dengan model unidimensi, maka diharapkan proses pengukuran psikometris juga melibatkan teknik analisis yang menggunakan model multidimensi.

Selanjutnya menurut Latan (2012), model persamaan struktural atau *Structural Equation Modeling* (SEM) merupakan suatu teknik analisis *multivariate* generasi kedua yang menggabungkan antara analisis faktor dan analisis jalur sehingga memungkinkan peneliti untuk menguji dan mengestimasi secara simultan hubungan antara *multiple exogenous* dan *endogenous* variabel dengan banyak indikator. Hasil penelitian Joreskog pada tahun 1970an membawa teori statistika pada analisis struktural linear yang lebih dikenal dengan sebutan model persamaan struktural atau SEM. Sumber penting yang digunakan dalam menganalisis adalah struktur kovarian sehingga terkadang pendekatan ini dinamakan dengan *covariant structure model* (CSM). Model yang disusun memuat variabel tak terukur yang dinamakan dengan konstruk laten yang dibangun oleh serangkaian variabel terukur yang dinamakan dengan konstruk terukur. *Error* pengukuran yang merefleksikan reliabilitas skor pengukuran dilihat sebagai konstruk unik dan menjadi bagian yang penting dalam analisis SEM, *error* pengukuran yang dilibatkan dalam analisis SEM inilah yang kemudian menjadi kelebihan SEM dibanding dengan teknik analisis lainnya (Capraro, *et al.*, 2001). SEM dapat mengestimasi varians *error* skor hasil pengukuran secara aktual mengestimasi reliabilitas. Menurut Gefen *et al.*, (2001), SEM sebagai teknik statistik multivariat yang mengkombinasikan antara regresi berganda yang mengidentifikasi hubungan antara konstruk dan analisis faktor yang mengidentifikasi konsep tak terukur melalui beberapa indikator manifest yang keduanya dipakai secara simultan.

SEM memiliki beberapa kelebihan dibanding dengan teknik analisis lainnya. Dalam menguji hubungan antara variabel, SEM secara otomatis mereduksi efek *error* pengukuran. Capraro *et al.*, (2001) mengatakan bahwa pengaruh variabel independen terhadap variabel dependen dipengaruhi oleh efek atenuasi. Nilai efek ini tidak dapat melebihi batas koefisien reliabilitas skor tes yang digunakan. Pendekatan pertama adalah koreksi korelasi atenuasi yang disebabkan oleh *error* pengukuran dan

pendekatan kedua adalah model persamaan struktural dalam konteks analisis faktor konfirmatori. Lee dan Song (2001) mengatakan bahwa SEM adalah salah satu pendekatan untuk menegaskan model pengukuran. Pada model pengukuran SEM menghubungkan antara konstruk laten dengan dengan konstruk empirik. Konstruk empirik dinyatakan oleh kombinasi konstruk laten. Di samping dapat dan mampu menangani *generalizability theory* dan *item response theory*, SEM mampu membandingkan model pengukuran dan memfasilitasi investigasi ketepatan model.

SEM mempunyai dua komponen dasar. Pertama, model pengukuran didefinisikan sebagai hubungan antara variabel laten dan sekelompok variabel penjelas yang dapat diukur langsung. Kedua, model struktural didefinisikan sebagai hubungan antara variabel laten yang tidak dapat diukur secara langsung. Variabel-variabel tersebut juga dibedakan sebagai variabel bebas dan variabel tidak bebas.

Semua variabel tersebut dikelompokkan ke dalam 4 bagian, yaitu q variabel penjelas bebas, p variabel penjelas tidak bebas, n variabel laten bebas, dan m variabel laten tidak bebas. Variabel laten tak bebas dan variabel laten bebas mempunyai hubungan linier struktural sebagai berikut:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (1) \text{ (Wiyanto, 2008)}$$

Keterangan:

B = matriks koefisien variabel laten tidak bebas berukuran $m \times m$

Γ = matriks koefisien variabel laten bebas berukuran $m \times m$

η = vektor variabel laten tak bebas berukuran $1 \times m$

ξ = vektor variabel laten bebas berukuran $1 \times m$

ζ = vektor sisaan acak berukuran $1 \times m$

Hal ini berimplikasi $E(\eta) = E(\zeta) = 0$, $E(\xi) = 0$ dan diasumsikan ζ tidak berkorelasi dengan ξ serta (I-B) tak singular.

Ada dua persamaan matriks yang digunakan untuk menjelaskan model pengukuran.

Persamaan pertama untuk variabel penjelas tidak bebas yaitu :

$$y = \Lambda_y \eta + \varepsilon \quad (2) \text{ (Wiyanto, 2008)}$$

Keterangan:

y = vektor variabel penjelas tidak bebas yang berukuran $p \times 1$

Λ_y = matriks koefisien yang mengindikasikan pengaruh variabel laten tak bebas terhadap variabel penjelas tak bebas yang berukuran $p \times m$

η = vektor variabel laten tak bebas yang berukuran $m \times 1$

ε = vektor kesalahan pengukuran variabel penjelas tidak bebas yang berukuran $p \times 1$

Persamaan kedua untuk variabel penjelas bebas yaitu:

$$x = \Lambda_x \xi + \delta \quad (3) \text{ (Wiyanto, 2008)}$$

Keterangan:

x = vektor variabel penjelas tidak bebas yang berukuran $q \times 1$

Λ_x = matriks koefisien yang mengindikasikan pengaruh variabel laten tak bebas terhadap variabel penjelas tak bebas yang berukuran $q \times n$

ξ = vektor variabel laten bebas yang berukuran $n \times 1$

δ = vektor kesalahan pengukuran variabel penjelas bebas yang berukuran $q \times 1$

Kesalahan pengukuran ε dan δ dianggap tidak berkorelasi satu sama lain, serta tidak berkorelasi dengan variabel-variabel laten. Gefen *et al.* (2001) mengatakan bahwa model pengukuran adalah sub model di dalam SEM yang mengidentifikasi konstruk laten dengan indikator-indikatornya yang dapat dipakai untuk mengetahui reliabilitas setiap konstruk yang dilibatkan dalam model.

SEM juga dapat mengidentifikasi reliabilitas konstruk yang terlihat melalui nilai butir *loading* yang dihasilkan. Berdasarkan perspektif SEM reliabilitas konstruk dapat dihitung melalui persamaan sebagai berikut:

$$CR = \frac{\left(\sum_{i=1}^i \lambda_i \right)^2}{\left(\sum_{i=1}^i \lambda_i \right)^2 + \left(\sum_{i=1}^i \delta \right)}$$

Keterangan:

CR = reliabilitas konstruk

λ_i = *factor loading* terstandarisasi indikator ke- i

δ = *error* standar pengukuran

McDonald (1981) merumuskan sebuah koefisien reliabilitas yang kemudian diberi nama koefisien reliabilitas skor komposit McDonald yang juga dinamakan

koefisien omega (ω). Koefisien reliabilitas ini berbasis pada analisis faktor konfirmatori yang merupakan bagian dari menu pemodelan SEM. Reliabilitas skor komposit McDonald ini menjelaskan besarnya proporsi indikator dalam menjelaskan konstruk ukur. Adapun formula untuk mendapatkan koefisien reliabilitas konstruk adalah sebagai berikut:

$$\omega = \frac{\left(\sum_{i=1}^i \lambda_i \right)^2}{\left(\sum_{i=1}^i \lambda_i \right)^2 + \left(\sum_{i=1}^i 1 - \lambda_i^2 \right)}$$

Keterangan:

λ_i = *factor loading* terstandarisasi indikator ke- i

Bila dibandingkan antara reliabilitas konstruk dengan reliabilitas skor komposit McDonald akan memberikan hasil yang sama karena $\delta = 1 - \lambda^2$.

Koefisien reliabilitas multidimensi berikut merupakan koefisien reliabilitas konstruk yang dikembangkan oleh Hancock dan Mueller (2000), yang menunjukkan seberapa jauh indikator instrumen mampu merefleksikan konstruk yang hendak diukur. Koefisien ini merupakan modifikasi dari koefisien reliabilitas konstruk McDonald yang tidak mampu mengakomodasi bobot yang berbeda antardimensi. Hasil modifikasi dinamakan koefisien reliabilitas konstruk berbobot sebagai berikut:

$$\Omega_w = \frac{\sum_{i=1}^p \frac{l_i^2}{(1-l_i^2)}}{1 + \sum_{i=1}^p \frac{l_i^2}{(1-l_i^2)}}$$

Keterangan:

l_i = koefisien dimensi ke- i terstandar

Koefisien reliabilitas ini dapat diartikan sebagai korelasi kuadrat antara dimensi dengan skor komposit linier optimal, sehingga beberapa ahli menamakannya dengan reliabilitas maksimal (*maximal reliability*).

Pada penelitian Widhiarso dan Mardapi (2010) model multidimensi untuk koefisien reliabilitas memiliki ketepatan pengukuran yang tinggi bila dibandingkan dengan reliabilitas unidimensi. Oleh karena itu dalam penelitian ini, peneliti hanya

difokuskan pada koefisien konsistensi internal seperti α untuk reliabilitas unidimensi dan ω , CR dan Ω_w .

Berdasarkan uraian di atas maka untuk organisasi pendidikan seperti Universitas Negeri Jakarta (UNJ) dimunculkan berbagai pertanyaan seperti: Bagaimanakah reliabilitas internal konsistensi multidimensi dari instrumen pengukur kepuasan mahasiswa sebagai pelanggan internal? Bagaimanakah komparasi antara reliabilitas multidimensi dan unidimensi? Manakah yang lebih akurat sebagai pengukur reliabilitas?

Penelitian ini bertujuan untuk menghasilkan suatu konstruksi tingkat kepuasan mahasiswa sebagai pelanggan internal. Konstruksi tersebut berupa kuesioner yang objektif, valid dan reliabel. Selanjutnya pengukuran ini akan bermanfaat bagi FT UNJ sebagai berikut:

- (1) memiliki rasa berhasil dan berprestasi, yang kemudian diterjemahkannya menjadi pelayanan yang prima kepada mahasiswa,
- (2) dapat dijadikan dasar menentukan standar kinerja dan standar prestasi yang harus dicapai, yang akan mengarahkan FT UNJ menuju kualitas yang semakin baik dan kepuasan mahasiswa yang meningkat,
- (3) memberikan umpan balik segera kepada pelaksana atau institusi,
- (4) memberitahu kepada institusi apa yang harus dilakukan untuk memperbaiki kualitas dan kepuasan mahasiswa serta bagaimana harus melakukannya, dan
- (5) memotivasi institusi untuk melakukan dan mencapai tingkat produktivitas yang lebih tinggi.

METODE

Metode yang dipergunakan dalam penelitian ini adalah metode survei. Survei digunakan di dalam pengumpulan data dan tidak dibuat perlakuan (*treatment*) atau pengkondisian terhadap variabel yang diteliti, namun hanya mengungkap fakta berdasarkan gejala yang ada pada mahasiswa ataupun responden yang lainnya. Survei sampel di dalam penelitian ini merupakan survei sampel terhadap hal-hal yang tidak nyata (*intangible*) yakni bila survei menyangkut pengukuran konstruk psikologis atau sosiologis dan membandingkan anggota-anggota populasi yang besar dimana variabelnya tidak dapat langsung diamati. Oleh karena penelitian ini mengukur konstruk

psikologis secara tidak langsung dari sampel populasi, maka jelas penelitian ini disebut survei sampel terhadap hal-hal yang tidak nyata (*sample survey of intangibles*).

Instrumen dalam penelitian skala dibuat dua kolom dengan rincian, untuk kolom pertama ini merupakan kenyataan (realitas) atau fakta yang ada dan dipersepsi oleh mahasiswa terhadap kualitas pelayanan yang memuaskannya dengan lima alternatif jawaban mulai dari sangat tidak puas (STPs) nilai 1, tidak puas (TPs) nilai 2, netral (N) nilai 3, puas (Ps) nilai 4, dan sangat puas (SPs) nilai 5. Penskalaan ini untuk instrumen kinerja. Untuk kolom kedua, harapan mahasiswa terhadap institusi dengan skala lima alternatif berdasarkan tingkat kepentingan mahasiswa dengan jawaban mulai dari sangat tidak penting (STPt) nilai 1, tidak penting (TPt) nilai 2, Biasa-biasa (Bb) nilai 3, penting (Pt) nilai 4, dan sangat penting (SPt) nilai 5. Sedangkan ini untuk penskalaan instrumen harapan. Penelitian ini akan dilaksanakan di FT UNJ, terdiri dari 75 mahasiswa Jurusan Teknik Elektro. Jadi bisa dikatakan data penelitian ini data *ex post facto*.

HASIL DAN PEMBAHASAN

Reliabilitas Multidimensi Untuk Instrumen Kinerja

Instrumen kinerja ini terdiri atas 30 butir pernyataan kuesioner kepuasan mahasiswa sebagai pelanggan internal. Ketiga puluh butir instrumen ini merupakan hasil penelitian yang telah divalidasi oleh peneliti sendiri yang semula 33 butir dan gugur 1 butir. Instrumen yang terdiri dari 30 butir ini dapat dirinci sebagai berikut: 6 butir untuk dimensi *tangible*, 7 butir *reliability*, 5 butir *responsiveness*, 7 butir *assurance*, dan 5 butir *empathy*. Untuk reliabilitas konsistensi internal alpha Cronbach diperoleh langsung menggunakan program SPSS sebesar 0,927.

Untuk reliabilitas skor komposit McDonald, dengan menggunakan program

LISREL 8.7 dan Excel didapat: $\sum_{i=1}^i \lambda_i = 17,640$ dan $\sum_{i=1}^i 1 - \lambda_i^2 = 19,057$; jadi

$$\omega = \frac{(17,640)^2}{(17,640)^2 + (19,057)} = 0,942.$$

Reliabilitas konstruk diperoleh hasil yang sama sebagai berikut: $\sum_{i=1}^i \lambda_i = 17,640$

dan $\sum_{i=1}^i \delta = 19,180$; jadi $CR = \frac{(17,640)^2}{(17,640)^2 + (19,180)} = 0,942$.

Berikut untuk reliabilitas berbobot, dengan menggunakan dengan menggunakan program LISREL 8.7 dan Excel diperoleh: $\sum_{i=1}^p \frac{l_i^2}{(1-l_i^2)} = 13,946$, sehingga dapat

dihitung sabagai berikut: $\Omega_w = \frac{13,946}{1+13,946} = 0,933$.

Reliabilitas Multidimensi Untuk Instrumen Harapan

Instrumen harapan ini terdiri atas 30 butir pernyataan kuesioner tingkat kepuasan mahasiswa sebagai pelanggan internal. Untuk reliabilitas konsistensi internal alpha Cronbah diperoleh langsung menggunakan program SPSS sebesar 0,934.

Untuk reliabilitas skor komposit McDonald, dengan menggunakan dengan menggunakan program LISREL 8.7 dan Excel diperoleh: $\sum_{i=1}^i \lambda_i = 18,480$ dan

$\sum_{i=1}^i 1-\lambda_i^2 = 18,193$; jadi $\omega = \frac{(18,480)^2}{(18,480)^2 + (18,193)} = 0,949$.

Reliabilitas konstruk diperoleh hasil yang sama sebagai berikut: $\sum_{i=1}^i \lambda_i = 18,480$

dan $\sum_{i=1}^i \delta = 18,310$; jadi $CR = \frac{(18,480)^2}{(18,480)^2 + (18,310)} = 0,949$.

Berikut untuk reliabilitas berbobot, dengan menggunakan dengan menggunakan program LISREL 8.7 dan Excel diperoleh: $\sum_{i=1}^p \frac{l_i^2}{(1-l_i^2)} = 15,371$, dan sehingga dapat

dihitung sabagai berikut: $\Omega_w = \frac{15,371}{1+15,371} = 0,939$.

Dari hasil uraian di atas dapat dirangkum dalam bentuk tabel sebagai berikut:

Tabel 1. Rangkuman Hasil

Reliabilitas	α	$CR = \omega$	Ω_w
Kinerja	0,927	0,942	0,933
Harpan	0,925	0,949	0,939

Perhitungan untuk dua buah instrumen di atas diperoleh koefisien reliabilitas multidimensi relatif lebih tinggi atau lebih tepat dibandingkan dengan koefisien reliabilitas unidimensi. Hal ini belum ada kesepakatan antar ahli psikometri. Namun di kalangan peneliti di Indonesia sebaiknya setelah mengetahui alat yang paling tepat sebaiknya mulai memakai/menggunakan alat tersebut secara benar dan memadai.

Memang sebagian besar peneliti di kalangan dosen maupun mahasiswa S2 maupun S3 belum mengetahui formula untuk menghitung koefisien reliabilitas konstruk, omega ataupun reliabilitas maksimal tersebut. Jadi kali ini saatnya untuk mengenalkan dan juga menggunakan formula tersebut. Dengan alasan sudah tahu rumusnya dan kebanyakan konstruk psikologis, kepribadian, pendidikan, dan sosial adalah multidimensi, sehingga seluruh peneliti baik mahasiswa maupun dosen berkembang dan makin berkembang untuk menggali lebih dalam lagi tentang koefisien reliabilitas yang lainnya.

Interpretasi koefisien reliabilitas merupakan evaluasi kecermatan skor tes, bukan sekedar keajegannya saja. Juga dalam menginterpretasikan tingginya koefisien reliabilitas, paling tidak ada dua hal yang perlu dipahami, yakni: (1) reliabilitas yang diestimasi dengan menggunakan suatu kelompok subjek dalam situasi tertentu akan menghasilkan koefisien yang tidak sama dengan estimasi tes tersebut pada kelompok subjek lain, dan (2) koefisien reliabilitas hanyalah mengindikasikan besarnya inkonsistensi skor hasil pengukuran, bukan menyatakan langsung sebab-sebab inkonsistensi itu.

Pengukuran bidang pendidikan merupakan sesuatu yang cukup rumit. Berbagai tulisan di dalam jurnal pengukuran pendidikan berkisar pada cara pengukuran yang diharapkan memberikan hasil yang valid, reliabel, dan akurat. Usaha para pakar tidaklah mudah karena para pakar tersebut makin lama membawa pengukuran pendidikan itu jauh ke dalam kawasan matematika. Tanpa menguasai dengan baik matematika yang tinggi dan rumit, kita tidak dapat memahami berbagai jurnal pengukuran pendidikan.

Sejauh ini, kita sangat tertinggal di bidang pengukuran pendidikan. Sangat sedikit pakar ilmu pendidikan yang mampu memahami isi jurnal pengukuran pendidikan yang bertaburkan matematika tingkat tinggi. Oleh karena itu perlu diusahakan peningkatan para pakar ilmu pendidikan di bidang pengukuran pendidikan.

Usaha itu dapat dimulai dengan mengubah persepsi kita selama ini yang sejak lama, para pendidik dikalangan kita memiliki anggapan bahwa ilmu pendidikan dan psikologi tidak memerlukan matematika. Matematika adalah garapan MIPA dan Teknik dan bukan garapan ilmu pendidikan. Kini, berhadapan dengan pengukuran pendidikan para pendidik dikalangan kita perlu mengubah persepsi mereka terhadap matematika. Para pendidik perlu menyadari bahwa ada bagian ilmu pendidikan yang hampir tidak menggunakan matematika, tetapi ada juga bagian ilmu pendidikan yang sangat memerlukan matematika, seperti contoh di atas statistika multivariat yang memerlukan kemampuan matematika tinggi.

KESIMPULAN

Berdasarkan hasil ujicoba di dalam penelitian dapat disimpulkan beberapa hal sebagai berikut: pertama, kepuasan mahasiswa sebagai pelanggan internal memiliki 5 dimensi yaitu: (a) sesuatu yang terwujud (*tangible*), (b) kehandalan (*reliability*), (c) daya tanggap (*responsiveness*), (d) jaminan (*assurance*), dan (e) empati (*empathy*). Oleh karena memiliki 5 dimensi tersebut penelitian mencoba menghitung koefisien reliabilitas konstruk dan omega serta reliabilitas maksimal yang memang sudah seharusnya bila koefisien reliabilitas multidimensi digunakan. Dengan kata lain pengukuran yang lebih tepat menggunakan koefisien reliabilitas konstruk, omega atau reliabilitas maksimal.

Saran yang dapat diajukan adalah sebagai berikut: pertama, estimasi instrumen ini perlu diuji lebih lanjut dengan menggunakan rumus lainnya yang tidak berbasis SEM. Kedua, oleh karena penelitian ini menggunakan skala lima maka bila perlu dilanjutkan menggunakan berbagai skala lain, misalnya skala diferensial semantik, skala dikotomi, skala Thurstone, dan sebagainya.

Ketiga, instrumen ini perlu diuji dengan menggunakan sampel yang lebih besar dengan populasi dan *setting* yang lebih luas serta melibatkan beberapa propinsi



sekaligus, juga dengan jenjang sekolah dan jenis universitas atau perguruan tinggi yang berbeda.

DAFTAR PUSTAKA

- Capraro, M. M., Capraro, R. M., and Herson, R. K. 2001. Measurement Error of Score on The Mathematics Anxiety Rating Scale Across Studies. *Educational and Psychological Measurement*, 61: 373–386.
- Feld, I. S., Woodruff, D. J., and Salih, F. A. 1987. Statistical Inference for Coefficient Alpha. *Applied Psychological Measurement*, Vol. II: 93 – 103.
- Gable, R. K. 1986. *Instrument Development in The Affective Domain*. Amsterdam: Kluwer Nijhoff Publishing.
- Geffen, D., Straub, D. W. & Boudreau, M. D. 2001. Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of AIS*, Volume 4: Article 7.
- Hancock, G. R. & Mueller, R. O. 2000. Rethinking Construct Reliability Within Latent Variable Systems. Di dalam *Structural Equation Modeling: Present and Future*, R. Cudek, S. H. C. duToit, dan D. F. Sorbom (Eds.), Chicago: Scientific Software International.
- Idrus, N., *et al.*, 2000. *Quality Assurance*. Jakarta: Directorate General of Higher Education.
- Kamata, A., Turhan, A. & Darandari, E. April 2003. *Estimating Reliability for Multidimensional Composite Scales Scores*. Paper presented in Annual Meeting of American Educational Research Association at Chicago.
- Kerlinger, F. N. 2000. *Asas-asas Penelitian Behavioral*, terjemahan Landung Simatupang. Yogyakarta: Gadjah Mada University Press.
- Latan, Hengky. 2012. *Structural Equation Modeling Konsep dan Aplikasi Menggunakan Program Lisrel 8.80*. Bandung: Alfabeta.
- Lee, S. Y. & Song, X. Y. January 2001. Hypothesis Testing and Model Comparison in Two-Level Structural Equation Model. *Multivariate Behavioral Research*, Volume 36 (4): 639–655.
- Lewis, R. G. & Smith, D. H. 1994. *Total Quality in Higher Education*. Florida: St. Lucie Press.
- Litwin, M. S. 1995. *How to Measure Survey Reliability and Validity*. London: Sage Publications.



- McDonald, R. P. 1981. The Dimensionality of Test and Items. *British Journal of Mathematical and Statistical Psychology*, 34: 100 – 117.
- Naga, D. S. 1992. *Teori Sekor*. Jakarta: Gunadarma Press.
- Parasuraman, A. 1988. Servqual: A Multi-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, Vol. 64(1): 12 – 37.
- Parasuraman, A., Zeinhaml, V. A. & Berry, L. L. 1985. A Conceptual Model of Service Quality and Its Implications For Future Research. *Journal of Marketing*, Vol. 49: 41 – 50.
- Socan, G. 2000. Assessment of Reliability When Test Items Are Not Essentially T-Equivalent. Di dalam *Development in Survey Methodology*, Anuska Feligoj and Andrej Mrvar (Eds.), Ljubljana: FDV.
- Spector, P., Brannick, P. & Chen, P. 1997. When Two Factors Don't Reflect Two Constructs: How Item Characteristics Can Produce Artifictual Factors. *Journal of Management*, Vol. 23 (5): 659 – 668.
- Thompson, B. 1994. Guidelines for Author. *Educational and Psychological Measurement*, Vol. 54: 837 – 847.
- Vehkalahti, K. 2000. *Reliability of measurement scales tarkkonnen's general method supersedes cronbach's alpha*. Academic Dissertation, University of Helsinki.
- Widhiarso, W. & Mardapi, D. 2010. Komparasi Ketepatan Estimasi Koefisien Reliabilitas Teori Skor Murni Klasik. *Jurnal Penelitian dan Evaluasi Pendidikan*, Vol. 14 (1): 1 – 19.
- Widhiarso, Wahyu. 2009. Koefisien Reliabilitas Pada Pengukuran Kepribadian Yang Bersifat Multidimensi. *Psikobuana*, Vol. 1 (1): 39 – 48.
- Wiersma, W. 1986. *Research Methods in Education: An Introduction*. London: Allyn and Bacon, Inc.
- Wijanto, Setyo Hari. 2008. *Structural Equation Modeling Dengan LISREL 8.8*. Yogyakarta: Graha Ilmu.
- Zeinhaml, V. A., Parasuraman, A., and Berry, L. L. 1990. *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: The Free Press.