

PENGEMBANGAN DAN VALIDASI SKALA PENGUKURAN KUALITAS LAYANAN TRIDARMA PERGURUAN TINGGI

Hisyam Ihsan & Ruslan

Jurusan Matematika Fakultas MIPA UNM
hisyamihsan@gmail.com & ruslan@gmail.com

Abstrak

Paper ini mengemukakan hasil penelitian yang mengembangkan dan memvalidasi skala pengukuran kualitas layanan pendidikan tinggi, yang merupakan penjabaran dari tridarma perguruan tinggi. Berdasarkan adaptasi model *Service Quality* 5-dimensi sebagaimana dalam Parasuraman, Zeithaml dan Barry (1993), instrumen 16-butir kualitas layanan pengajaran (*teachqual*), 14-butir kualitas layanan penelitian (*resqual*), dan 11-butir kualitas layanan pengabdian masyarakat (*pubqual*) dibuat untuk menilai kualitas layanan pendidikan. Masing-masing instrumen menggambarkan sikap dan perilaku mahasiswa mengenai kinerja layanan perguruan tinggi. Instrumen tersebut menilai kualitas secara fungsional (*process*) maupun teknikal (*output-outcome*). Tiga ratus sebelas mahasiswa program sarjana tahun terakhir sebagai responden instrumen. Unidimensional, reliabilitas, validitas isi, validitas terkait-kriterium, validitas konvergen, validitas diskriminan, dan validitas nomologikal diperiksa dengan menggunakan *Structural Equation Modeling*. Instrumen KLTPT telah dikembangkan dan berpotensi aplikasi untuk eksplorasi bagi para praktisi dan peneliti.

Kata kunci: *pengukuran kualitas layanan tridharma perguruan tinggi*

DEVELOPMENT AND VALIDATION A MEASUREMENT SCALE OF HIGHER EDUCATION SERVICE QUALITY

Hisyam Ihsan & Ruslan
Matematics Dept. Makassar State University
hisyamihsan@gmail.com & ruslan@gmail.com

Abstract

This paper is the research results that develops and validates a measurement scale of higher education service quality. It is reflection of Higher Education Tridahrma. Based on adaptation of 5-dimension Service Quality model of Parasuraman, Zeithaml, and Barry (1993), a 16-item instrument of teaching service quality (teachqual), a 14-item instrument of research service quality (resqual), and an 11-item instrument of public service quality (pubqual) were created to assess educational service quality. Each of them describes student's attitude and behavior of higher education service performance. The instruments assessed the quality both functionally (perceptions of educational process) and technically (perceptions of output-outcome). Three hundred and eleven fourth year undergraduate students evaluated three block of the instrument as respondents. Unidimensionality, content validity, criterion-related validity, convergent validity, discriminant validity, and nomological validity are checked out by using Structural Equation Modeling. The norms of the instrument (KLTPT) were then developed, and the potential applications for practitioners and researchers explored.

Kata kunci: *measurement of higher education service quality*

mengemukakan produk sampingannya (*by-product*) berupa *personal self discovery, career choice and placement, dan direct satisfactions and enjoyments*.

Dill (2003) mengemukakan bahwa perguruan tinggi dapat dipandang sebagai bagian dari industrial ekonomi yang menyediakan derajat akademik, penelitian, dan pengabdian. Hal ini sejalan dengan Tri Dharma Perguruan tinggi berupa pengajaran, penelitian dan pengabdian masyarakat. Pengajaran dalam Tri Dharma Perguruan tinggi tersebut tercermin dalam kurikulum (kurikuler dan ekstrakurikuler) yang akan membekali mahasiswa untuk terjun ke masyarakat, penelitian memberi bekal mencari dan menghasilkan ilmu pengetahuan agar dapat meningkatkan, menyebarkan, menginterpretasikan, dan mengaplikasikan ilmu pengetahuan, serta membudayakan kreativitas ilmiah. Sementara itu, pengabdian masyarakat dapat berupa pemeliharaan lingkungan, kebersihan, konsultan di berbagai bidang dan sebagainya.

Kemampuan dan efektivitas lembaga pendidikan khususnya dalam pendidikan tinggi untuk meningkatkan tingkat kemampuan dan kecakapan tenaga kerja dalam penyesuaian diri dan tumbuh secara kompetitif akan memiliki pengaruh besar terhadap kapasitas ekonomi. Berlanjutnya keyakinan seperti ini pada sektor manufaktur dan industri jasa untuk mencapai perkembangan ekonomi, niscaya akan meningkatkan permintaan di sektor pendidikan.

Dalam EFQM (1995: 5) dikemukakan bahwa produk pendidikan tinggi berkaitan dengan nilai tambah pada pengetahuan, kecakapan, dan pengembangan personal mahasiswa. Sebagai industri jasa, kualitas produk berkaitan dengan kualitas proses pembelajaran. Kualitas produk inilah yang memerlukan analisis dan identifikasi elemen-elemen khusus yang harus dibayar. Barret (1996) mengusulkan suatu model integratif faktor-faktor kunci yang mempengaruhi kualitas belajar mahasiswa dan proses pendidikan. Di antara faktor-faktor tersebut, kualitas dosen termasuk kualitas jurusan/program-studi adalah faktor khusus yang penting.

Tujuan dari studi ini adalah untuk mengembangkan kuesioner Kualitas Layanan Tridharma Perguruan Tinggi (KLTP) dan mengevaluasi reliabilitas dan validitas skala pengukuran yang dikonstruksi dari butir-butir dalam kuesioner.

Tinjauan Pustaka

Holdford & Reinders (2001) mengemukakan bahwa kualitas layanan tridarma perguruan tinggi adalah sikap yang dihasilkan dari persepsi mahasiswa tentang kinerja layanan pendidikan tinggi tersebut. Sebagaimana diketahui, produk pendidikan tinggi dalam melayani mahasiswa adalah pengajaran, penelitian, dan pengabdian pada masyarakat. Jadi, kualitas produk yang dimaksud adalah kualitas pengajaran, kualitas penelitian, dan kualitas pengabdian. Pengajaran tercermin dalam kurikulum (termasuk kokurikuler dan ekstrakurikuler), meliputi aspek kognitif, afektif dan psikomotor, yang akan membekali mahasiswa untuk terjun ke pasar kerja. Penelitian memberi bekal mencari dan menghasilkan ilmu pengetahuan agar dapat meningkatkan, menyebarkan, menginterpretasikan, dan mengaplikasikan ilmu pengetahuan, serta membudayakan kreativitas ilmiah. Pengabdian masyarakat dapat berupa upaya pemberian kecakapan bermasyarakat dan aplikasi langsung kegiatan tersebut ke masyarakat, seperti kegiatan organisasi kampus, kuliah lapangan, menjadi pelatih, konsultan di berbagai bidang, bakti sosial seperti pemeliharaan lingkungan, kebersihan, dan sebagainya.

Pengukuran kualitas layanan yang paling populer dan hingga kini banyak dijadikan acuan dalam riset pemasaran adalah model pengukuran Servqual (*service quality*) yang dikembangkan oleh Parasuraman, Zeithaml, & Berry (1993) dalam serangkaian penelitian mereka dalam enam sektor jasa: reparasi peralatan rumah tangga, kartu kredit, asuransi, sambungan telepon jarak jauh, perbankan ritel, dan pialang sekuritas sejak tahun 1985 hingga 1994. Model pengukuran ini erat kaitannya dengan model pengukuran kepuasan pelanggan. (Fandy, 2005: 262).

Ada lima unsur yang menentukan kualitas jasa menurut Parasuraman, Zeithaml & Berry (1988) dalam memodifikasi model Servqual, yaitu *reliability* (kehandalan), *responsiveness* (daya tanggap), *assurance* (jaminan), *empaty* (empati), dan *tangible* (bukti fisik). *Reliability* dimaksudkan untuk mengukur aspek-aspek kehandalan sistem pelayanan yang diberikan oleh pemberi jasa, dalam hal ini apakah jasa yang diberikan sesuai dengan standar-standar umum atau bahkan standar internasional. *Responsiveness*

berkaitan dengan kecepatan tanggapan pemberi jasa dalam memberikan jasa sekaligus menangkap aspirasi yang muncul dari pelanggan. *Assurance* adalah jaminan bahwa jasa yang diberikan akan memberikan jaminan keamanan, kemampuan (kompetensi) sumberdaya dalam memberikan pelayanan sesuai dengan standar dan lain-lain, yang sifatnya memberikan jaminan bahwa seluruh unsur pemberi jasa sesuai dengan apa yang diharapkan. *Empaty* berkaitan dengan kemudahan dalam mendapatkan pelayanan, keramahan, komunikasi dan kemampuan memahami kebutuhan pelanggan. Terakhir *tangible* meliputi aspek-aspek nyata yang dapat diamati atau diraba dalam pelayanan. Kelima unsur tersebut sebagai acuan dimensi dalam mengembangkan indikator kualitas layanan pendidikan tinggi tingkat sarjana.

Konsumen menggunakan kelima dimensi ini dalam melakukan penilaian terhadap kualitas layanan, yang didasari oleh perbandingan antara layanan yang diharapkan dan layanan yang diterima. Gap di antara layanan yang diharapkan dan layanan yang dialami/diterima (*perceived*) adalah ukuran dari kualitas layanan.

Kitchroen (2004) mengemukakan bahwa model Serqual 5-dimensi tersebut dapat diadaptasi untuk studi kualitas layanan dalam industri pendidikan. Beberapa indikator telah digunakan dalam berbagai penelitian. Riportella, Couste, & Torres (2001) menggunakan indikator dalam mengevaluasi kualitas pengajaran yang dialami mahasiswa program MBA, yaitu: dosen mengajar dengan jelas; dosen mengajar dengan antusias; dosen mengajak untuk berpartisipasi dalam kelas; bermanfaat dan menarik untuk dibaca dan ada rujukan referensi materi yang dibawakan dosen; kedisiplinan dosen berkaitan dengan waktu; dan proses assesmen.

Metode Penelitian

Unit Analisis (Responden)

Responden yang menanggapi instrumen adalah mahasiswa program sarjana tingkat akhir (sedang memprogramkan matakuliah tugas akhir/skripsi). Selama di kampus, mahasiswa ditanya untuk melengkapi instrumen yang telah dikembangkan di enam perguruan tinggi negeri (Univ.

Hasanuddin, Univ. Negeri Makassar, Univ. Islam Negeri Alauddin) dan swasta (Univ. Muhammadiyah, Univ. Muslim Indonesia, Univ. Indonesia Timur) melalui survei. Tiga ratus sebelas mahasiswa diambil secara acak (multistage cluster random sampling) yang dapat menanggapi instrumen dari keenam universitas tersebut.

Pengembangan Kuesioner

Instrumen KLTPT dikembangkan dalam bentuk kuesioner. Kuesioner yang dikembangkan didasarkan pada prinsip berikut:

- Kuesioner menyertakan variabel laten, dan variabel manifes yang telah diidentifikasi mengikuti model konseptual.
- Format kuesioner dirancang untuk tidak begitu banyak menghabiskan waktu responden saat pengisian kuesioner.
- Menggunakan bahasa yang jelas dan bermakna sehingga dapat difahami dengan mudah oleh responden.

Pengembangan kuesioner dan skala pengukuran dalam studi ini, menggunakan prosedur pengembangan yang dikemukakan oleh Churchill (1979) dalam Churchill & Iacobucci (2002: 315 & 418). Fransi (2000) juga menggunakan prosedur ini dengan sedikit penambahan langkah sesuai dengan kondisi populasi pelanggan internet. Proses pengembangan kuesioner dan skala pengukuran mengikuti langkah-langkah sebagai berikut:

- Langkah 1. Spesifikasi domain konstruk
- Langkah 2. Mengidentifikasi dimensi/faktor pembentuk
- Langkah 3. Membangun Indikator dan butir pernyataan/pertanyaan dari masing-masing dimensi melalui tinjauan literatur
- Langkah 4. Reduksi butir (kalau perlu) melalui panel ahli dengan metode Delphi
- Langkah 5. Menentukan sekuensi butir termasuk karakteristik fisik kuesioner
- Langkah 6. Menentukan parameter pengukuran dan target sasaran responden yang akan mengisi kuesioner tersebut
- Langkah 7. Mengumpulkan data

Langkah 8. Menentukan model pengukuran yang lebih sesuai, reduksi butir kalau perlu melalui penilaian validitas

Langkah 9. Mendesain kuesioner definitif berdasarkan hasil analisis validitas dan reliabilitas.

Dalam kaitannya dengan pengukuran, maka konstruk kualitas layanan pengajaran, penelitian, dan pengabdian merupakan variabel laten dalam pengukuran.

Kuesioner KLTPPT terdiri dari tiga blok kuesioner yang berbeda sesuai dengan konstruksya yang diperhatikan dalam penelitian ini. Semua variabel tersebut didesain dalam perangkat kuesioner yang diharapkan dapat direspon dalam waktu tidak terlalu lama. Blok masing-masing kuesioner sebagai berikut:

- a. Kuesioner Kualitas Layanan Pembelajaran (KLPB)
- b. Kuesioner Kualitas Layanan Penelitian (KLPN)
- c. Kuesioner Kualitas Layanan Pengabdian (KLPD)

Butir pernyataan, skala, dan skoring

Skala Likert 7-titik diambil sebagai patokan pada semua butir pernyataan dalam kuesioner. Rensis Likert (1970) mengemukakan bahwa skala Likert akan memproduksi ukuran sikap yang dapat diinterpretasi sebagai pengukuran pada skala metrik sejati dengan syarat: (1) skala memuat lebih dari satu butir, (2) tingkat respon dijangka dengan bilangan bulat berurutan, dan (3) tingkat respon dijangka dengan lebel verbal dan simetrik pada bagian netral atau tengah. (Uebersax, 2006). Oleh karena itu, dalam Skala Likert 7-titik ini, responden akan diberikan butir pernyataan dengan pilihan antara 1 sampai 7. Angka 1 sampai 7 ini merupakan skor perolehan responden yang menyatakan bahwa 1 (sangat tidak setuju) hingga 7 (sangat setuju).

Menilai validasi dan reliabilitas konstruk

Untuk menilai validitas dan reliabilitas konstruk menurut berbagai penulis seperti Wang (2003), Rodgers dkk (2007), Deniz dkk. (2008) dapat menggunakan kriteria (1) unidimensional, (2) reliabilitas, (3) validitas isi, (4)

validitas terkait kriteria (prediktif dan konkuren), (5) validitas konvergen, (6) validitas dikriminan, dan (7) validitas nomologikal.

Analisis faktor konfirmatori (confirmatory factor analysis)

Sebagaimana dinyatakan oleh Churchill (1979), bahwa untuk menaikkan reliabilitas dan menurunkan *error* pengukuran, digunakan skala multi-butir sebagai pengganti skala satu-butir. Pendekatan Churchill (1979) untuk pengembangan skala telah diperluas oleh Anderson & Gerbing (1988) dalam Lages, dkk (2005) dengan menggunakan Analisis Faktor Konfirmatori (CFA). CFA dapat menilai sifat-sifat pengukuran mengenai keberadaan skala melalui *full-information maximum likelihood* (FIML) dalam AMOS 7.0. (Arbuckle, 2006).

Dalam analisis ini, dilakukan juga pemeriksaan atas asumsi yang diperlukan dalam metode *maximum likelihood* (ML). *Goodness-of-fit* model pengukuran dievaluasi melalui dengan kriteri bersama (atau indeks *overall fit*), yaitu *chi-square* (χ^2), relatif χ^2 / df , *root mean square error of approximation* (RMSEA), *comparative fit index* (CFI), dan *Tucker-Lewis index* (TLI). Model yang memiliki *overall fit* kurang baik dilakukan reduksi terhadap parameter melalui prosedur *modification indices* dengan pendekatan *building-trimming* yang tersedia dalam AMOS, dan model dinyatakan final (model final) bila indeks *overall fit* telah memenuhi kriteria (Garson, 2006). Kriteria tersebut atau biasa disebut *cut-off threshold* adalah χ^2 cukup kecil dengan nilai-p > alpha yang dipilih, $1 < \text{relatif } \chi^2 / df < 2$, RMSEA < 0.08, CFI > 0.90, dan TLI > 0.90. (Arbuckle, 2006)

Pada metode estimasi ML, uji- χ^2 biasanya digunakan untuk mengukur *goodness-of-fit* bila variabel observasi dianggap mengikuti distribusi normal multivariat. Pengujian normalitas dilakukan baik secara univariat maupun multivariat, selanjutnya transformasi data dilakukan bila diperlukan. Syarat analisis lainnya tetap diperhatikan dalam studi ini seperti linieritas suatu model, dan multikolinieritas.

Analisis dan Hasil Penelitian

Data cleaning dan pemeriksaan data

Data *cleaning* meliputi pengecekan konsistensi data dan perlakuan terhadap respon *missing*. Untuk itu dilakukan identifikasi data yang *out of range*, tidak konsisten secara logis, atau memiliki nilai ekstrim. Skor yang *out of range* adalah *inadmissible* dan harus dikoreksi (Malhotra, 2004: 409).

Dari 322 responden mahasiswa yang disurvei ada sebanyak 39 responden yang tidak dapat menanggapi atau mengisi kuesioner sehingga memerlukan pergantian sampel. Setelah melakukan pengambilan sampel pengganti. Instrumen yang tidak lengkap tidak diinput. Setelah data ditabulasi dilakukan pemeriksaan atas data tersebut termasuk data-data yang *out of range*, *missing* dan sebagainya. Data pencilan (*outlier*) dipertimbangkan dalam analisis. Berdasarkan ini maka ditetapkan ada 11 data yang tidak diikutkan dalam *data set*, sehingga total sampel penelitian sebanyak 311 unit dan dinyatakan *clean*. Sebagai tambahan, kebanyakan *software* komputer tidak dapat memproses adanya data *missing*. Oleh karena itu, data *missing* diestimasi, yaitu mengganti *missing* tersebut dengan rata-rata yang sesuai melalui prosedur *expectation-maximum* (EM). Jumlah nilai-nilai *missing* kurang dari 2% dari semua data, sehingga hasil analisis rata-rata tidak menyebabkan distorsi (atau pergeseran) variansi (Hair, dkk. 1998).

Kelayakan ukuran sampel

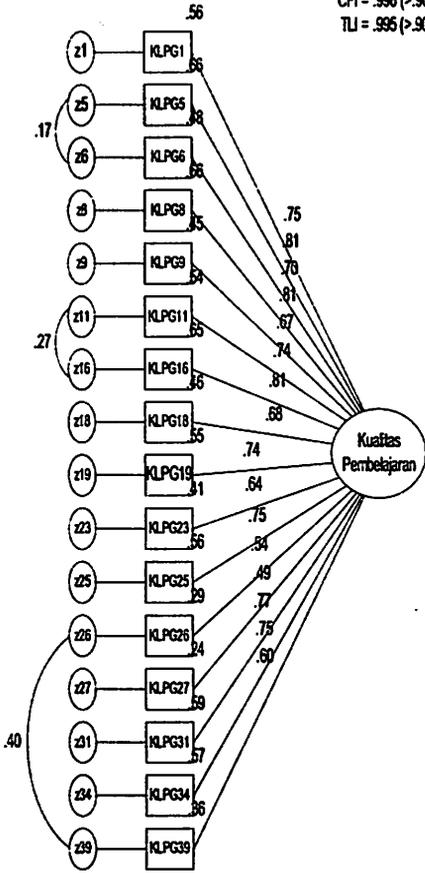
Sebagaimana diketahui bahwa dalam metode multivariat, ukuran sampel berperan penting dalam estimasi dan interpretasi hasil-hasil analisis. Hair dkk. (1998: 605) merekomendasikan ukuran sampel minimal untuk pemodelan analisis model pengukuran dalam *structural equation modeling* sebesar 200, dan yang lebih tipikal adalah paling sedikit 5-10 responden untuk setiap parameter yang diestimasi dipandang lebih sesuai. Jadi, ukuran sampel sebanyak 322 dipandang layak untuk analisis statistika berikutnya.

Model pengukuran untuk menilai instrumen KLTP

Model final dalam validasi empiris disajikan dalam Gambar 1. Hasil pengolahan data menunjukkan bahwa asumsi linieritas dapat dinyatakan telah dipenuhi untuk semua indikator terhadap konstruksya. Asumsi linieritas diambil berdasarkan matriks plot indikator dengan skor komposit konstruksya. Multikolinieritas (berakibat pada singularitas) dapat dijamin tidak terjadi karena matriks korelasi indikator semuanya memiliki koefisien korelasi $r < 0.85$.

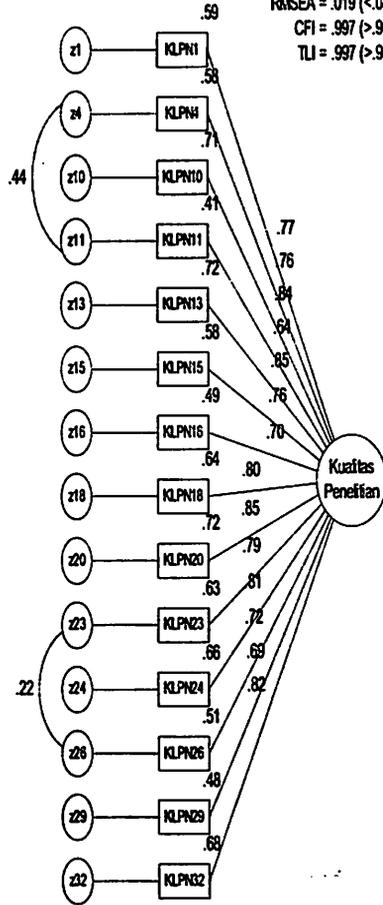
Hasil pemeriksaan univariat dan multivariat menunjukkan bahwa ada beberapa nonnormal multivariat dalam data. Masalah ini telah dikemukakan oleh Micceri (1989) (dalam Shimizu & Kano, 2006) bahwa variabel terobservasi dalam sains sosial hampir tidak pernah terdistribusi normal. Selanjutnya, dilakukan transformasi terhadap variabel yang memuat pencilan, namun data transformasi masih gagal untuk menampilkan normalitas multivariat dan lebih sukar untuk di "interpretasi" ketimbang data asli (Tabachnick & Fidell, 2001; Ullman, 2001 dalam Abramson, dkk. 2005). Meskipun nampak ketidaknormalan dalam data, namun tetap diputuskan untuk menggunakan estimasi ML, karena teknik ini cukup robust terhadap dilanggarnya normalitas (Engel dkk., 2003).

Chi-sqr. = 112.946 (Df = 101) P = .196
 Relative Chi-square = 1.118 (< 2)
 RMSEA = .020 (< .08)
 CFI = .996 (> .90)
 TLI = .995 (> .90)

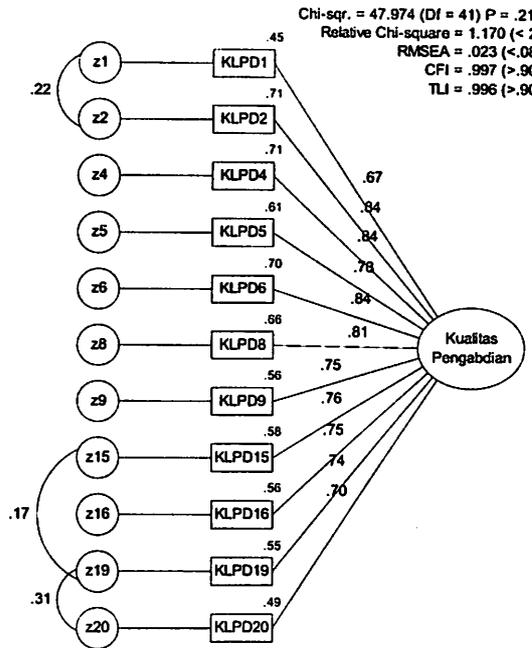


KUALITAS LAYANAN PEMBELAJARAN (KLPG)

Chi-sqr. = 83.056 (Df = 75) P = .245
 Relative Chi-square = 1.107 (< 2)
 RMSEA = .019 (< .08)
 CFI = .997 (> .90)
 TLI = .997 (> .90)



KUALITAS LAYANAN PENELITIAN (KLPN)



KUALITAS LAYANAN PENGABDIAN MASYARAKAT (KLPD)

Gambar 1. Model Final Pengukuran Skala KLTPT – Koefisien Terstandar CFA

Butir pernyataan instrumen KLPG, KLPN, dan KLPD yang mendukung model pengukuran dalam Gambar 1.

Hasil CFA model SEM final terhadap KLPG ditetapkan ada 16 indikator memberikan *goodness of fit* yang dapat diterima. Indeks *overall fit* yang diperoleh adalah $\chi^2(102) = 112.946$ dengan nilai- $p = 0.196 > 0.05$, $1 < \chi^2 / df = 1.118 < 2$, $RMSEA = 0.020 < 0.08$, $CFI = 0.996 > 0.95$, dan $TLI = 0.995 > 0.95$. Ini berarti model dapat digunakan dalam analisis lebih lanjut, sedangkan untuk instrumen KLPN ditetapkan sebanyak 14 indikator memberikan *goodness of fit* yang dapat diterima. Indeks *overall fit* yang

diperoleh adalah $\chi^2(75) = 83.056$ dengan nilai- $p = 0.245 > 0.05$, $1 < \chi^2/df = 1.107 < 2$, RMSEA = 0.019 < 0.08, CFI = 0.997 > 0.95, dan TLI = 0.997 > 0.95. Ini berarti model dapat digunakan dalam analisis lebih lanjut. Demikian pula dengan instrumen KLPD, diperoleh 11 indikator memberikan *goodness of fit* yang dapat diterima. Indeks *overall fit* untuk model final ini adalah $\chi^2(41) = 47.974$ dengan nilai- $p = 0.211 > 0.05$, $1 < \chi^2/df = 1.170 < 2$, RMSEA = 0.023 < 0.08, CFI = 0.997 > 0.95, dan TLI = 0.996 > 0.95. Ini berarti model dapat digunakan dalam analisis lebih lanjut.

Unidimensionalitas (unidimensionality)

Syarat yang diperlukan untuk melakukan analisis validitas dan reliabilitas adalah bahwa setiap indikator memiliki sifat unidimensional terhadap konstraknya (Hair, dkk. 1998). Unidimensinalitas dapat didefinisikan sebagai "*the existence of one construct underlying a set of items and has been recognised as one of the most critical and basic assumptions of measurement theory*" (Hattie, 1985: 139). Sederetan indeks *goodness-of-fit*, yaitu, GFI > 0.90, NFI > 0.90 dan RMR < 0.05, menyediakan petunjuk unidimensioalitas dari indikator konstruk (Hair dkk., 1998, Gefen, 2003).

Hasil CFA terhadap model SEM final diperoleh nilai GFI, NFI dan RMR untuk KLPG berturut-turut 0,959, 0,961, dan 0,055, untuk KLPN berturut-turut 0,965, 0,975, dan 0,035, dan untuk KLPD berturut-turut 0,973, 0,980, dan 0,036. Karena itu ketiga konstruk KLPG, KLPN, dan KLPD memiliki sifat unidimensional dari indikatornya.

Tabel 1. Hasil Analisis Faktor Konfirmatori dan Reliabilitas untuk KLPG

Indikator	Butir	Standardized Loading	Critical Ratio	p	Cronbach Alpha	
Tangible	KLPG1	0.750	15.275	0.000	0.82	0.94
	KLPG5	0.814	17.208	0.000		
	KLPG6	0.696	13.731	0.000		
Reliability	KLPG8	0.811	17.121	0.000	0.85	
	KLPG9	0.670	13.092	0.000		
	KLPG11	0.736	14.843	0.000		
Responsiveness	KLPG16	0.809	17.042	0.000		
	KLPG18	0.678	13.292	0.000	0.80	
	KLPG19	0.742	15.033	0.000		
	KLPG23	0.638	12.307	0.000		
Assurance	KLPG25	0.749	15.232	0.000		
	KLPG26	0.540	10.021	0.000	0.63	
	KLPG27	0.485	8.867	0.000		
Empaty	KLPG31	0.766	15.733	0.000		
	KLPG34	0.753	15.350	0.000	0.65	
	KLPG39	0.604	11.482	0.000		

CCR = 0.94 (>.70); AVE = 0.502 (>.50)

Reliabilitas (Reliability)

Reliabilitas dievaluasi dengan menggunakan penilaian konsistensi internal mengenai butir-butir yang merepresentasikan setiap faktor dari konstruk. Mengikuti Macdonald & Uncles (2007), maka reliabilitas yang digunakan dinilai melalui *Cronbach-alpha*, *Composite construct reliability* (CCR) dan *Average variance extracted* (AVE). Kriteria reliabilitas konstruk ini adalah Alpha > 0,70 (Nunnally, 1978), C-R > 0.70 dan AVE > 0.50. (Hair dkk. 1998: 612).

Berdasarkan tabel 2, 3, dan 4, nampak bahwa Cronbach alpha untuk keseluruhan butir dan untuk setiap indikator melewati *cut-off threshold*-nya, yaitu 0,70 baik untuk KLPG, KLPN, maupun KLPD. AVE merupakan ringkasan butir konvergensi dan juga sebagai rataan loading faktor kuadrat untuk setiap konstruk dalam model pengukuran. Skor untuk AVE pada

KLPG adalah 0,502, pada KLPN adalah 0,60, dan pada KLPD = 0,60 yang semuanya telah melewati *cut-off threshold*-nya. CCR juga terkait dengan model pengukuran. Semua skor CCR yang diperoleh pada KLPG, KLPN, dan KLPD melebihi threshold 0.7, yaitu berturut-turut 0.94, 0.95 dan 0.94.

Tabel 2. Hasil Analisis Faktor Konfirmatori dan Reliabilitas untuk KLPN

Indikator	Butir	<i>Standardized Loading</i>	<i>Critical Ratio</i>	<i>p</i>	<i>Cronbach Alpha</i>	
<i>Tangible</i>	KLPN1	0.767	15.841	0.000	0.75	0.95
	KLPN4	0.763	15.726	0.000		
<i>Reliability</i>	KLPN10	0.840	18.152	0.000	0.81	
	KLPN11	0.637	12.340	0.000		
	KLPN13	0.849	18.466	0.000		
<i>Responsiveness</i>	KLPN15	0.764	15.750	0.000	0.86	
	KLPN16	0.703	14.028	0.000		
	KLPN18	0.800	16.859	0.000		
	KLPN20	0.850	18.491	0.000		
<i>Assurance</i>	KLPN23	0.795	16.683	0.000	0.83	
	KLPN24	0.810	17.168	0.000		
	KLPN26	0.715	14.353	0.000		
<i>Empaty</i>	KLPN29	0.690	13.700	0.000	0.72	
	KLPN32	0.824	17.623	0.000		

CCR = 0.95 (>.70); AVE = 0.60 (>.50)

Validitas isi (Content validity)

Instrumen KLTPPT memenuhi kebutuhan struktur faktor yang reliabel dan konsisten. Akan tetapi, reliabilitas tinggi dan konsistensi internal merupakan syarat perlu untuk memiliki validitas konstruk tetapi belum cukup (Nunnally, 1978). Kriteria kualitatif yang mendasar berkaitan dengan validitas konstruk adalah validitas isi. Validitas isi menyaratkan bahwa instrumen tersebut harus menyajikan semua aspek-aspek konstruk yang mesti diukur. Churchill (1979) mengemukakan bahwa “*specifying the domain of the construct, generating items that exhaust the domain, and subsequently purifying the resulting scale should produce a measure which is content or face valid and*

reliable.” Oleh karena itu, prosedur yang digunakan dalam mengkonseptualisasi konstruk, membangun butir pernyataan, dan membersihkan ukuran konstruk sebagaimana langkah-langkah yang ditempuh dalam pengembangan kuesioner KLTPT menunjukkan bahwa kuesioner KLTPT, yaitu KLPG, KLPN, dan KLPD telah memiliki validitas isi yang kuat.

Tabel 3. Hasil Analisis Faktor Konfirmatori dan Reliabilitas untuk KLPD

Indikator	Butir	<i>Standardized Loading</i>	<i>Critical Ratio</i>	P	<i>Cronbach Alpha</i>	
<i>Tangible</i>	KLPD1	0.671	13.066	0.000	0.79	0.94
	KLPD2	0.843	18.171	0.000		
<i>Reliability</i>	KLPD4	0.844	18.202	0.000	0.86	
	KLPD5	0.783	16.237	0.000		
	KLPD6	0.837	17.961	0.000		
<i>Responsiveness</i>	KLPD8	0.813	17.163	0.000	0.74	
	KLPD9	0.746	15.136	0.000		
<i>Assurance</i>	KLPD15	0.764	15.657	0.000	0.76	
	KLPD16	0.752	15.305	0.000		
<i>Empaty</i>	KLPD19	0.739	14.911	0.000	0.80	
	KLPD20	0.697	13.777	0.000		

CCR = 0.94 (>.70); AVE = 0.60 (>.50)

Validitas terkait-kriteria (Criterion-related validity)

Validitas terkait-kriteria (prediktif dan konkuren) dinilai melalui korelasi antarskor total pada instrumen (jumlah dari semua butir) dan mengukur kriteria kevalidan (jumlah dari dua butir global). (Wang, 2003). Dalam studi ini validitas terkait-kriteria merujuk pada validitas konkuren di mana skor total terhadap instrumen (KLPG, KLPN, KLPD) dan skor terhadap kriteria kevalidan diukur pada waktu yang sama. Hubungan positif diharapkan terjadi antara skor total dan kriteria kevalidan jika instrumen itu mampu mengukur konstruk KLPG, KLPN, KLPD. Instrumen KLPG, KLPN, dan KLPD berturut-turut memiliki validitas terkait kriteria 0.81,

0,84, dan 0,84 pada taraf signifikansi 0.05. Ini menampilkan validitas terkait kriteria yang dapat diterima.

Validitas konvergen (Convergent validity)

Validitas konvergen dapat dinilai dari model pengukuran dengan menentukan apakah setiap butir/indikator yang diestimasi melalui *maximum likelihood loading* terhadap konstruk adalah signifikan. (Anderson & Gerbing, 1988). Faktor *loading* yang diperoleh dalam CFA merupakan informasi secara statistik untuk mengetahui tingkat validitas konvergen. Suatu butir/indikator menunjukkan validitas konvergen bila *standardized factor loadings* dalam CFA besar dan signifikan ($p < 0.05$) (Zumbo, 2005). Garson (2006) dalam April & Pather (2008) mendefinisikan *a rule-of-thumb* untuk nilai-nilai *standardized factor loadings* dalam CFA menjadi lemah bila nilai tersebut kurang dari 0.4. Validitas konvergen dan diskriminan yang diperiksa melalui analisis faktor tidak secara spesifik memeriksa unidimensionalitas karena faktor dibuat sebagai bobot jumlah dari semua butir pengukuran (Anderson & Gerbing, 1988, Hattie, 1985, Gefen, 2003). Berdasarkan hasil analisis yang ditunjukkan dalam table 2, 3, dan 4, memberikan keputusan bahwa semua butir pernyataan memiliki faktor *loading* yang lebih dari 0,4 dan signifikan ($p < 0.05$), sehingga semua butir tersebut memiliki validitas konvergen.

Validitas diskriminan (discriminant validity)

Validitas diskriminan (atau validitas divergen) terkait dengan kenyataan bahwa semua konstruk yang interkorelasi secara signifikan berbeda dari 1, dan varians yang terbagi antardua konstruk lainnya (yaitu kuadrat interkorelasinya) kurang dari rataan varians yang dijelaskan oleh konstruk dalam butir tersebut (MacKenzie dkk., 1999). Dapat pula dikatakan bahwa petunjuk adanya validitas diskriminan tersingkap oleh kemampuan suatu instrumen tidak terkorelasi dengan konstruk yang terkait. (Aaker dkk., 1998). (Lages dkk., 2005). Validitas diskriminan dikonfirmasi dengan menunjukkan signifikansi perbedaan chi-kuadrat antara model-model konstruk yang berpasangan untuk dikorelasikan

dengan model itu sendiri yang korelasinya dibatasi satu. (Bagozzi & Phillips, 1981 dalam Kim & Kim, 2005). Hasil analisis ditunjukkan dalam Tabel 4. Nampak dalam tabel tersebut $Vc_{df=1}^2$ semuanya signifikan ($p < 0.05$), sehingga dapat disimpulkan bahwa semua konstruk KLPG, KLPN, dan KLPD memiliki validitas diskriminan.

Tabel 4. Hasil Uji Perbedaan Chi-kuadrat untuk Validitas Diskriminan

No	Pasangan Konstruk	C^2 -statistik		$Vc_{df=1}^2$	P
		Model tak terbatas (d.f.)	Model terbatas (d.f.)		
1.	KJPG – KJPN	738.923 (399)	1637.224 (400)	898.30 (1)	0.000
2.	KJPG – KJPM	567.243 (317)	1598.174 (318)	1030.93 (1)	0.000
3.	KJPN – KJPM	429.767 (269)	1302.466 (270)	872.70 (1)	0.000

Validitas nomologikal (Nomological validity)

Suatu instrumen memiliki validitas nomologikal bila instrumen tersebut “*behaves as expected with respect to some other construct to which it is theoretically related*” (Churchill, 1995: 538). Papoutsakis (2008) dan Lages, dkk (2005) mengemukakan bahwa untuk menilai validitas nomologikal, pengukuran kita diuji secara statistika terhadap beberapa konstruk lain dengan referensi hasil penelitian terkait atau memiliki keterkaitan teoretis. Dalam Steenkamp & Trijp (1991) dikemukakan bahwa biasanya validitas nomologikal dinilai melalui analisis regresi atau korelasi sebagaimana dikemukakan pula oleh Westbrook (1980); Peter (1981); Comer (1984); Ruekert & Churchill (1984). Dalam studi pustaka, instrumen KLPB, KLPN, dan KLPD memiliki keterkaitan teoretis dalam satu paket instrumen KLTP. Validitas nomologikal dinilai melalui keterhubungan positif dari konstruk-konstruk tersebut yang saling terkait secara teoritis (Wang, 2003). Hasil analisis korelasi dalam model SEM antara KLPG & KLPN, KLPG & KLPD dan KLPN & KLPD, ketiganya memberikan korelasi yang signifikan ($p < 0,05$). Karena itu validitas nomologikal ketiga konstruk tersebut dipenuhi.

Simpulan dan Implikasi

Kualitas layanan perguruan tinggi, sebagaimana didefinisikan dalam studi ini, dinilai oleh mahasiswa. Walaupun penilaian mahasiswa tidak selalu mencerminkan realitas, tetapi mereka menolong menentukan *outcome* penting seperti partisipasi mahasiswa dalam ruang kuliah, keterlibatan dalam aktivitas ekstrakurikuler, dan pandangan tentang persekolahan yang mahasiswa ambil untuk dirinya guna kelulusannya dan kompetensi yang diperlukan dalam dunia kerja.

Jika mahasiswa memandang bahwa pendidikan adalah spekulasi kooperatif antara mahasiswa dan dosen, maka memahami persepsi mahasiswa tentang dosen dapat menjadi strategi untuk meningkatkan partisipasi mahasiswa dalam pendidikannya. Oleh karena itu, pendidikan tinggi seharusnya mengevaluasi kualitas pendidikan dengan menggunakan metode berganda termasuk kualitas layanan tridarma yang diberikan kepadanya. Pendidikan tinggi perlu menilai persepsi mahasiswa baik proses maupun *outcome* pengalaman pendidikannya. Penilaian sistematis sepanjang waktu dapat digunakan untuk mengidentifikasi dampak inisiasi utama dan promosi kampus dengan berdialog kepada mahasiswa tentang pendidikannya.

Hasil analisis data di atas menyimpulkan bahwa instrumen kualitas layanan pengajaran, layanan penelitian dan layanan pengabdian masyarakat semuanya memiliki validitas dan reliabilitas yang memadai berdasarkan model pengukuran yang diperoleh. Berdasarkan hasil kerja Parasuraman, Zeithaml, & Berry (1988) dalam pengembangan skala pengukuran kualitas layanan, yaitu Servqual 5-dimensi. Studi ini telah memberikan modifikasi servqual tersebut untuk layanan pendidikan tinggi, yang terdiri dari teachqual (kualitas layanan pengajaran), resqual (kualitas layanan penelitian), dan pubqual (kualitas layanan pengabdian masyarakat) dan masing-masing memiliki tingkat validitas dan reliabilitas yang memadai.

Daftar Pustaka

- Abramson, R., Rahman, S., & Buckley, P. 2005. Tricks and Traps in Structural Equation Modelling: a GEM Australia Example Using AMOS Graphics. *Proceedings of ABBSA Conference*. August: 558-599
- Anderson, J. C., & Gerbing, D. W. 1988. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103 (3): 411-423.
- April, G.D. & Pather, S. 2008. Evaluating Service Quality Dimensions within e-Commerce SMEs. *Electronic Journal Information Systems Evaluation* Volume 11, Issue 3: 109-124
- Arbuckle, J. L. 2006. *AMOS 7.0 User's Guide*. Amos Development Corporation. United States of America.
- Barret, R. 1996. Quality and the Abolition of Standards: Argument against some American Prescription for the Improvement of Higher Education. *Quality in Higher Education*, 2 (3): 201 - 210.
- Bowen & Brown, S. 1991. (Eds.), *Advance in services, marketing and management* 3: 1-67. Greenwich, CT: JAI.
- Buchari Alma. 2003. *Pemasaran Stratejik Jasa Pendidikan*. Alfabeta: Bandung.
- Carbonell, P & Rodriguez, A. I. 2006. The impact of market characteristics and innovation speed on perceptions of positional advantage and new product performance. *International Journal of Research in Marketing* 23: 1-12.
- Churchill Jr GA. 1979. A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research* 16 (Feb): 64-73.
- Churchill Jr GA. 1995. *Marketing research: methodological foundations*. Chicago: The Dryden Press.
- Churchill, G. A. & Iacobucci, D. 2002. *Marketing Research Mothodological Foundations. South-Western*. Eighth Edition. United States of America

- Deniz, ME., Kesici, S., dan Sumer, AS. 2008. The Validity and Reliability of The Turkish Version of The self-Compassion scale. *Social Behavior And Personality*, 36 (9): 1151-1160
- Dill, D. 2003. *Allowing the market to rule: the case of the United States*. The University of North Carolina: Chapel Hill.
- Duman, T. 2002. A Model Of Perceived Value For Leisure Travel Products. Disertasi. The Pennsylvania State University
- EFQM, European Foundation for Quality Management. 1995. Self-Assessment. Guidelines for public sector Education: *EFQM*, Bruselas.
- Engel, K.S., Moosbrugger, H, & Muller, H. 2003. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research* 8 (2): 23-74.
- Fandy T. 2005. *Pemasaran Jasa*. Bayumedia Publishing. Malang
- Fitzsimmons, J. A. & Fitzsimmons, M. J. 1998 *Service Management: Operations, Strategy and information Technology*. Singapore: McGraw-Hill.
- Fransi, E. C. 2000. Online Stores Service Quality: A Marketing Perspective. Departement of Business Management, University of Lleida: Spain. Online. <http://griho.udl.es/publicacions/2003/textonline%20stores%20service.pdf>. Diakses 4/6/2006.
- Garson, G. D. 2006. Structural Equation Modeling. North Carolina State University. Online: <http://www2.chass.ncsu.edu/garson/pa765/structur.htm>. Diakses 29/11/2006.
- Gefen, D. 2003. Assessing Unidimensionality Through Lisrel: An Explanation And Example. *Communications of the Association for Information Systems*. Volume 12: 23-47
- Hair, F.J., Anderson, R.E, Tatham, R.L., & Balck, W.C. 1998. *Multivariate Data Analysis*. 7th edition. Prentice Hall: New Jersey.

- Hattie, J.A. 1985. Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement* 9: 139-164.
- Holdford, D. & Wright, S.1997. Testing Commitment-Trust Theory in relationship beetwen pharmacy schools and students. *American Journal of Pharmaceutical Education* 61: 249-256
- Holdford, D., & Reinders, T. P. 2001. Development of an Instrument to Assess Student Perceptions of the Quality of Pharmaceutical Education. *American Journal of Pharmaceutical Education* 65: 125-131
- Jaworski, J. B. & Kohli, K, A. 1993. Market Orientation: Antecedents and Consequences. *Journal of Marketing*, 57(3): 53-70
- Kim, H-Y. & Kim, Y-K. 2005. Escapism, Consumer Lock-in, Attitude, and Purchase: An Illustration from an Online Shopping Context. *Journal of Shopping Center Research* 12 (2): 103-115.
- Kitchroen, K. 2004. Literature Review: Service Quality in Educational Institutions. *ABAC Journal* 4 (2). 14-25.
- Lages, C., Lages, C. R., dan Lages, L. F. 2005. The RELQUAL scale: a measure of relationship quality in export market ventures. *Journal of Business Research* 58, 1040-1048.
- Lee, S. Y. 2005. A Conceptual Model of The Roles of Price, Quality, and Intermediary Constructs in Determining Behavioral Intention to Visit A Festival. Dissertation. Texas A&M University.
- Malhotra, N. K. 2004. *Marketing Research: An Applied Orientation*. 4th-edition. Prentice Hall: New Jersey.
- Macdonald, E.K. dan Uncles, M.D. 2007. Consumer Savvy: Conceptualisation and Measurement. *Journal of Marketing Management* V. 23 No. 5-6: 497-517
- Nunnally, J.C. 1978. *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

- Papoutsakis, H. 2008. On Measuring Organizational Relationships: Threats to Validity in the Use of Key-Informants. *Electronic Journal of Knowledge Management* 6, Issue 2: 145-156.
- Parasuraman, A., Zeithaml V.A, & Berry, L.L., 1988. SERVQUAL: A Multiple Item Scale for Measuring Customer Perceptions of Service Quality. *Journal of Retailing*, vol 64.
- Pasuraman, A., Zeithaml, V.A., & Berry, L.L., 1993. More on Improving Service Quality Measurement. *Journal of Retailing*, 69(4) (winter).
- Riportella, C.C., Couste, N.L. & Torres, P.R. 2001. Measurement and Effects of Teaching Quality: An Empirical Model Applied to Masters Program. *Business Economics Series* 10: 01-31.
- Rodgers, S. Wang, Y. Rettie, R. dan Alpert, F. 2007. The Web Motivation Inventory: Replication, extension and application to internet advertising. *International Journal of Advertising*, 26(4): 447-476
- Shimizu, S. & Kano, Y. 2006. Use of nonnormality in structural equation modeling: Application to direction of causation. Online <http://homepage.mac.com/shoheishimizu/papers/JSPI05.pdf>. Diakses: 15/12/2006.
- Steenkamp, JB & Trijp, HC. 1991. The use of LISREL in validating marketing constructs. *Intern. J. of Research in Marketing* 8: 283-299.
- Uebersax J. S. 2006. Likert Scales: Dispelling the Confusion. Statistical Methods for Rater Agreement Website. Online: <http://ourworld.compuserve.com/homepages/jsuebersax/likert2.htm>. Diakses: 15/11/2006
- Wang, Y. 2003. Assessment of learner satisfaction with asynchronous electronic learning systems. *Information & Management* 41: 75-86
- Zhang, L. 2003. Do measures of College Quality Matter ? The Effect of College Quality on Graduates' Earnings. Online. http://www.ilr.cornell.edu/cheri/wp/cheri_wp64.pdf. Diakses 27/10/2006

- Zikmund, W.G. 1984. *Business Research Methods*. Fifth edition. The Dryden Press: United States of America.
- Zumbo, B.D. 2005. Structural Equation Modelling and Test Validation. *Encyclopedia of Statistics in Behavioral Science* 4: 1951-1958. John Wiley & Sons, Ltd, Chichester.