



Contents lists available at JurnalIICET

Jurnal Konseling dan Pendidikan

ISSN: 2337-6740 (Print) ISSN: 2337-6880 (Electronic)

Journal homepage: <http://jurnal.konselingindonesia.com>



Pengembangan instrumen kualitas layanan dosen universitas swasta menggunakan pemodelan Rasch

Yuyun Elizabeth Patras¹, Rais Hidayat²

^{1,2}Universitas Pakuan, Indonesia

Article Info

Article history:

Received Des 18th, 2019

Revised Jan 21st, 2020

Accepted Feb 1st, 2020

Keyword:

Private university

Rasch model

Service quality

ABSTRACT

This research aims to get a research instrument, namely quality service lecturer at private university. The research method used mixed methods consists of literature reviews, interviews with students, expert validation, and a test use Rasch modeling. Literature review is carried out to obtain a description of the service quality according to the developing theory. Interviews with students aim to obtain items of review that fit the research context. Expert validation aims to obtain content validity. Collecting data through a questionnaire instrument containing 32 items. The number of respondents was 543 students at Pakuan University. The finding that 31 items of valid lecturer service quality instrument and one (1) item that is invalid must be discarded; This lecturer service quality instrument is very reliable because it has a reliability value above 0.9, both in terms of people reliability, alpha cronbach and item reliability. However, based on the item difficulty level test using item measure, item dimensionality and variable map, and Item: DIF test found several items that must be corrected.



© 2020 The Authors. Published by Indonesian Institute for Counseling, Education and Therapy (IICET). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

Corresponding Author:

Yuyun Elizabeth Patras,

Universitas Pakuan

Email: yuyunpatras64@gmail.com

Pendahuluan

Pendidikan merupakan jasa (service) yang diberikan lembaga pendidikan kepada masyarakat/konsumen (Rajab et al., 2011). Jasa pendidikan, termasuk jasa pendidikan tinggi yang diberikan universitas swasta maupun negeri, seiring dengan permintaan dari masyarakat saat ini sedang tumbuh (Abdullah & Mohamad, 2016). Pertumbuhan permintaan jasa ini harus diiringi dengan pemberian jasa pendidikan yang bermutu, bukan hanya sarana fisik namun yang utama adalah dari para pengajar/dosennya (Dursun, Oskayba, & Gökmen, 2013). Keadaan tersebut menyebabkan lembaga harus memberi perhatian penuh pada sejauhmana kualitas layanan (service quality) yang diberikan kepada pelanggannya/masyarakat (Esmacilpour & Ranjbar, 2016).

Pada sebuah universitas, agenda pemberian kualitas layanan pada pelanggannya menjadi sangat penting (Rajab et al., 2011). Kenapa penting, karena mahasiswa sebelum mendaftar pada sebuah universitas akan menjadikan kualifikasi pelayanan universitas tersebut sebagai pertimbangan utama (Dursun et al., 2013). Oleh karena itu mengimplementasikan pelayanan terbaik bagi mahasiswa merupakan solusi bertahan bagi universitas swasta maupun negeri (Stefan, 2010). Apalagi sejalan dengan perubahan terjadi di masyarakat, hubungan antara penyedia jasa dan pelanggannya semakin dinamis, sehingga lembaga penyedia jasa untuk terus menerus mengevaluasi jasa yang diberikannya (Kumar, Shekhar, Lassar, & Chen, 2018). Sebuah lembaga yang tidak memberikan kualitas layanan terbaik pada pelanggannya tidak akan mampu berkompetisi (Roy, 2019).

Berdasarkan pemikiran bahwa pemberian jasa, termasuk jasa dalam bidang akademik, antara lain pengajaran yang dilakukan dosen di sebuah universitas, maka jasa tersebut diperlukan evaluasi. Evaluasi

yang dimaksud adalah mempertanyakan sejauh mana jasa yang diberikan dosen sudah memenuhi harapan (expected) mahasiswa. Penelitian ini akan berfokus pada tema tersebut. Hal baru dalam penelitian ini yaitu metode menggunakan metode bauran (mixed method) dan analisisnya menggunakan pemodelan Rasch, sedangkan instrument dikembangkan dari 5 dimensi kualitas layanan, yaitu: tangibility, reliability, responsiveness, assurance, dan empathy (Roy, 2019) yang kemudian dikembangkan ke dalam 32 butir pernyataan.

Kualitas layanan (service quality) didefinisikan sebagai pemberian layanan nyata yang dilakukan sebuah organisasi untuk memenuhi harapan dan persepsi pelanggannya (Boon, Shukur, & Bassim, 2016). Kualitas layanan berarti organisasi mengerti dan memahami persepsi dan harapan pelanggan sehingga organisasi menjadikan hal itu sangat penting untuk memperkirakan pemberian kualitas layanan (Deb & Ahmed, 2018). Pemberian kualitas layanan terbukti dapat memuaskan pelanggan dan memengaruhi perilaku pelanggan (Oriade & Scho, 2019), kualitas layanan juga berkorelasi sangat signifikan pada kehendak untuk menggunakan jasa atau produk secara berulang oleh para pelanggannya (Prentice & Kadan, 2019).

Banyak ragam pengukuran kualitas layanan. Namun demikian terdapat 5 dimensi kualitas layanan yang sering digunakan, yaitu: tangibles, empathy, assurance, reliability dan responsiveness (Boon et al., 2016). Hal itu diperkuat peneliti lainnya bahwa kualitas layanan memiliki dimensi sebagai berikut: physical dimensions or tangibility, reliability, responsiveness, and assurance, and empathy (Esmailpour & Ranjbar, 2016). Lebih jauh peneliti lainnya menjabarkan setiap dimensi yaitu: (1) Tangibles meliputi fasilitas fisik, peralatan, dan penampakan dari sisi personal dan cara berkomunikasi; (2) Reliability meliputi kinerja pelayanan berdasarkan prinsip akurasi dan kejujuran; (3) Responsiveness meliputi kesiapan dan kesanggupan memberikan layanan terbaik bagi pelanggan; (4) Assurance meliputi kompetensi, kebaikan, integritas dan rasa aman yang ditunjukkan pemberi layanan; (5) Empathy meliputi akses, komunikasi dan pemahaman si pemberi layanan kepada yang dilayani (Dursun et al., 2013).

Pada lembaga layanan pendidikan seperti universitas, persepsi mahasiswa sebagai pelanggan didasarkan pada kualitas dan layanan pendidikan yang ditawarkan. Di lembaga pendidikan tinggi, kualitas layanan yang dirasakan yang meliputi belajar mengajar atau perkuliahan, fasilitas dan layanan pendukung lainnya (Azam, 2018). Jika kualitas layanan pendidikan diterima baik oleh mahasiswa, maka akan menimbulkan kepuasan pada pelanggan (Rajab et al., 2011). Berdasarkan paparan di atas untuk selanjutnya peneliti akan menyebut upaya pemberian layanan pada mahasiswa sebagai kualitas layanan dosen (lecturer's service quality), khususnya di universitas swasta.

Pemodelan Rasch (Rasch Model) merupakan sebuah pendekatan statistik untuk mengukur kinerja, sikap, dan persepsi manusia. Penamaan tersebut berdasarkan nama penemunya, Georg Rasch, berkebangsaan Denmark, ia menerbitkan teorinya pada tahun 1960 dan meninggal pada tahun 1980 (Tesio, 2003). Analisis data menggunakan pemodelan Rasch dapat diterapkan pada ilmu social (Bond & Fox, 2007), penggunaan pemodelan Rasch sangat direkomendasikan untuk meningkatkan kualitas pengukuran, baik dalam penelitian qualitative maupun quantitative (Stef van Buuren, 2015).

Pemodelan Rasch mengacu pada teori response butir atau aitem (item response theory) yang memberikan alternatif yang lebih baik untuk memeriksa kualitas penilaian skala psikometrik dan menginformasikannya dengan lebih baik (Zile-Tamsen, 2017). Pemodelan Rasch dapat digunakan dalam berbagai penelitian. Pemodelan Rasch digunakan untuk meneliti bidang perilaku (Bond & Fox, 2007) kesehatan mental (Tesio, 2003) dan kinerja karyawan (Jiao, Lissitz, & Macready, 2011). Dalam pendidikan, misalnya untuk membuat instrumen yang dapat mengukur kondusifitas dalam pembelajaran (Nguyen Thanh Dong, 2014).

Pemodelan Rasch dalam menganalisis data memiliki berbagai keunggulan, misalnya ia dapat menjelaskan aitem (butir) pernyataan dan orang (person) (Carvalho, Primi, & Meyer, 2012). Dalam pemodelan Rasch, unit pengukuran skala untuk kemampuan dan kesulitan aitem (butir) dikenal sebagai "logit", kontraksi "unit odds log". Adanya unit pengukuran tersebut membuat Model Rasch dapat membantu untuk mengatasi pengukuran aitem (butir) dengan cara yang benar (Wu & Adams, 2007). Pemodelan Rasch mampu mengatasi masalah integritas data melalui akomodasi transformasi logit dengan menerapkan logaritma pada rasio ganjil dari data mentah yang diperoleh dari responden (Linacre, 2010).

Metode

Penelitian ini bertujuan untuk membuat instrument penelitian, yaitu kualitas layanan dosen. Penelitian ini menggunakan metode kombinasi (mixed-methode) (Sugiyono, 2015) yang dimulai secara kualitatif dan

dilanjutkan dengan kuantitatif deskriptif (Zohrabi, 2013). Proses kualitatif dimulai dengan kegiatan kajian literature, wawancara mahasiswa, dan validasi ahli. Sedangkan proses kuantitatif (Klette, 2012), yaitu melakukan analisis data menggunakan pemodelan Rasch (Rasch model). Kajian literatur dilakukan untuk mendapatkan validitas logis (logical validity) (John I. Pollock, 1967), yaitu instrumen yang ditemukan berdasarkan pada teori kualitas layanan (Meerah, Osman, Zakaria, & Ikhsan, 2012). Wawancara pada mahasiswa bertujuan untuk mendapatkan butir-butir pernyataan yang sesuai dengan konteks penelitian. Validasi ahli bertujuan untuk mendapatkan validitas isi (content validity) (Ghazali, 2016). Pengumpulan data melalui instrumen kuisioner yang disebarakan melalui Google Form. Jumlah responden sebanyak 543 mahasiswa. Penelitian ini dilakukan di Universitas Pakuan, Jawa Barat, yang telah dikenal sebagai salah satu universitas swasta terbesar di Bogor, Jawa Barat, Indonesia. Adapun gambaran demografi responden terlihat pada table 1.

Tabel 1. Demografi Responden Penelitian

Jenis Kelamin	Status Sekolah Menengah	Tahun Masuk	Pekerjaan Orang Tua	Fakultas
Laki-laki, n=139	Negeri, n= 301	2018, n= 188	Wiraswasta, n=167	Ekonomi, n = 213
Perempuan, n=404	Swasta, n= 242	2017, n= 108	PNS, n=131	Pendidikan, n = 164
		2016, n= 92	Swasta, n=122,	Hukum, n= 56
		2015, n = 123	TNI/Polri, n=21	MIPA, n= 51
			Lainnya, n=102	Sosial Budaya, n = 36
				Teknik, n= 23

Kontruksi instrumen menggunakan validitas logis, yaitu sebuah instrumen yang memenuhi persyaratan valid berdasarkan hasil penalaran karena dirancang secara baik, mengikuti teori dan ketentuan yang ada (John I. Pollock, 1967). Berdasarkan kontruksi tersebut peneliti mendapatkan 5 dimensi kualitas layanan, yaitu: (1) *tangible*, (2) *reliability* (3) *responsiveness* (4) *assurance*, (5) *empathy*. Langkah berikutnya adalah melakukan wawancara terhadap mahasiswa untuk mendapatkan butir-butir pernyataan sesuai dengan dimensi kualitas layanan. Berdasarkan wawancara pada mahasiswa diperoleh 4 butir yang terkait dengan dimensi *tangible* (TGL), 6 butir yang terkait *reliability* (RLB), 6 butir yang terkait *responsiveness* (RSP), 11 butir yang terkait dengan *assurance* (ASR) , 5 butir yang terkait *empathy* (EMP). Jadi keseluruhan instrumen ini memiliki 32 butir pernyataan. Draf instrument selanjutnya dinilai oleh dua ahli. Rekomendasi keduanya berupa saran perbaikan, khususnya konten (Ghazali, 2016). Saran tersebut ditindaklanjuti oleh peneliti.

Butir instrument yang terkait *tangible* yaitu: (TGL1) Dosen berpenampilan rapi ketika mengajar; (TGL2) Dosen menyampaikan materi kuliah berdasarkan referensi yang relevan; (TGL3) Dosen terlihat percaya diri; (TGL4) Dosen menggunakan media pembelajaran yang memudahkan mahasiswa untuk memahami materi kuliah.

Butir instrument yang terkait *reliability*, yaitu: (RLB1) Dosen menggunakan buku ajar yang relevan dengan mata kuliah; (RLB2) Dosen menginformasikan sumber-sumber belajar (link buku, jurnal, materi pelajaran) kepada mahasiswa; (RLB3) Dosen mengajar sesuai dengan jadwal yang diatur oleh fakultas/program studi; (RLB4) Dosen mengisi daftar hadir mahasiswa dan dosen; (RLB5) Dosen memberikan nilai kepada mahasiswa sesuai ketentuan yang ditetapkan.

Butir instrument yang terkait *responsiveness*, yaitu: (RSP1) Dosen mengembalikan hasil ujian kepada mahasiswa; (RSP2) Dosen memberikan pengayaan kepada mahasiswa yang membutuhkan; (RSP3) Dosen menggunakan teknologi pembelajaran terbaru dalam perkuliahan; (RSP4) Dosen memeriksa tugas perkuliahan yang diberikan kepada mahasiswa; (RSP5) Dosen menyampaikan rencana pembelajaran kepada mahasiswa; (RSP6) Dosen memberikan petunjuk yang jelas ketika memberikan tugas kepada mahasiswa.

Butir instrument yang terkait *assurance*, yaitu: (ASR1) Dosen membahas presentasi mahasiswa; (ASR2) Dosen memberikan sanksi terhadap mahasiswa yang melanggar peraturan perkuliahan; (ASR3) Dosen mampu menyelesaikan masalah yang terjadi dalam perkuliahan dengan efektif; (ASR4) Dosen mampu merespon pertanyaan mahasiswa dengan jawaban yang disertai data; (ASR5) Dosen melakukan teguran

kepada mahasiswa sesuai aturan yang berlaku; (ASR6) Dosen memberikan bimbingan sesuai peraturan di kampus; (ASR7) Dosen memberikan rasa aman kepada mahasiswa; (ASR8) Dosen mampu membangun rasa percaya dari mahasiswa; (ASR9) Dosen menghindari melakukan kritik pada mahasiswa di depan teman-temannya; (ASR10) Dosen menyampaikan materi kuliah dengan bahasa yang mudah dipahami; (ASR11) Dosen memberikan perhatian kepada kemajuan belajar mahasiswa.

Butir instrument yang terkait *empathy* yaitu: (EMP1) Dosen berinteraksi kepada mahasiswa secara adil; (EMP2) Dosen melakukan perwalian sesuai ketentuan fakultas/program studi; (EMP3) Dosen dapat bekerjasama dengan mahasiswa; (EMP4) Dosen meluangkan waktu yang cukup untuk melayani kebutuhan akademik mahasiswa; (EMP5) Dosen meminta unpan balik (feed-back) dari mahasiswa tentang perkuliahan yang diampunya.

Pengukuran jawaban responden menggunakan skala *Likert* (Brinkman, 2009) dengan lima pilihan peringkat (*rating*), semakin besar skor jawaban responden berarti hal itu menunjukkan responden lebih menyetujui butir pernyataan dan sebaliknya. *Rating*nya mulai 1= sangat tidak setuju, 2 = tidak setuju, 3 = netral, 4 = setuju, 5 = sangat setuju (Hendriks et al., 2012). Pengujian instrument dan persepsi atas kualitas layanan menggunakan pemodelan Rasch. Proses pengolahan data dimulai dengan input data dari kuesioner yang berbentuk skala ordinal kemudian dikonversi ke skala interval menggunakan pemodelan Rasch melalui perangkat lunak Winsteps versi 3.73.

Pengujian dilakukan pada 4 hal, yaitu: (1) Pengujian validitas butir (*item*) instrument dengan menggunakan *Item (Column): Fit Order*, dimana Nilai *Outfit Mean Square* (MNSQ) yang diterima $.05 < \text{MNSQ} < 1.5$, Nilai *Outfit Z-Standard* (ZSTD) yang diterima $-2 < \text{ZSTD} < + 2$, Nilai *Point Measure Correlation* (Pt Mean Cor): $0.4 < \text{PtMean Cor} < 0.85$ dan terakhir menggunakan *rating (partial-credit) scale*, dengan kriteria bahwa jika semua *rating* (angka) memiliki titik puncak, maka instrument tersebut memiliki validitas (Bambang Sumintono, 2015); (2) Pengujian reliabilitas instrument menggunakan *summary statistic*; (3) Pengujian butir instrumen yang sulit dan mudah disetujui responden menggunakan *item measure* dan *item: dimensionality*; dan (4) Pengujian persepsi atau bias berdasarkan demografi responden menggunakan *Differential Item Functional* (DIF) *plot* (Bambang Sumintono, 2014). Kriteria adanya perbedaan persepsi yang signifikan menggunakan tabel 30.4 pada Winsteps software dengan kriteria bahwa perbedaan persepsi dikatakan signifikan jika nilai probabilitas menunjukkan lebih kecil dari 0.05 ($p < 0.05$) (Bambang Sumintono, 2015).

Hasil dan Pembahasan

Pengujian validitas butir instrument menggunakan *Item (Column): Fit Order* pada Winsteps seperti terdapat pada Table 2 dengan kriteria pengambilan keputusan bahwa butir pernyataan dinyatakan valid apabila OUTFIT MNSQ berada pada rentang $0.5 < \text{MNSQ} < 1.5$. Memperhatikan hal itu maka butir yang “tidak valid” pada instrument kualitas layanan dosen ini berdasarkan OUTFIT MNSQ ini adalah butir 10, yaitu RSP1 dimana nilai MNSQ menunjukkan 1.85 yang berarti lebih besar dari 1.5. Sisanya sebanyak 31 butir memenuhi kriteria OUTFIT MNSQ yang dimaknai bahwa 31 butir (*item*) sudah fit atau tepat mengukur.

Pengujian validitas butir instrument menggunakan *Item (Column): Fit Order* pada Winsteps dengan melihat nilai OUTFIT Z-STANDARD (ZSTD) dengan kriteria valid jika $-2 < \text{ZSTD} < + 2$, maka ditemukan bahwa terdapat 8 butir yang memenuhi kriteria validitas, yaitu butir: RSP4 (1.8), RSP3 (1.3), TGL4 (0.7), EMP2 (0.2), (RSP5 (1.3), TGL2 (-0.5), RLB2 (-0.9) dan RSP5 (-1.3). Ini berarti sisanya sebanyak 24 tidak memenuhi OUTFIT Z-STANDARD atau harus diperbaiki.

Pengujian validitas instrumen dengan melihat nilai *Point Measure Correlation* (Pt Mean Cor) dengan kriteria valid jika $0.4 < \text{PtMean Cor} < 0.85$, maka ditemukan bahwa polaritas *item* (*item polaritas*) memiliki nilai *Point Measure Correlation* yang positif dan memenuhi kriteria yang dipersyaratkan. Ini berarti semua butir instrument kualitas layanan dosen tidak ada konflik antara butir dan konstruk yang diukur. Dengan demikian dari 3 kriteria validitas (OUTFIT MNSQ, OUTFIT Z-STANDARD, dan *Point Measure Correlation*) hanya terdapat 1 butir (RSP1) yang tidak memenuhi 2 syarat lainnya sehingga harus didrop atau tidak digunakan.

Tabel 2. Hasil Uji Item (Column): Fit Order

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	EXACT EXP.	MATCH OBS%	Item		
10	1705	543	1.30	.06	1.71	9.8	1.85	9.9	A .63 .78	42.7	50.1	RSP1
1	2327	543	-1.31	.07	1.52	6.6	1.41	3.7	B .62 .67	52.4	61.9	TGL1
24	1859	543	.73	.06	1.47	6.7	1.47	6.5	C .69 .77	52.6	51.2	ASR9
31	1913	543	.52	.06	1.38	5.6	1.37	5.1	D .72 .76	59.3	52.3	EMP5
8	2227	543	-.80	.07	1.06	.9	1.32	3.6	E .69 .70	59.6	58.1	RLB4
17	2016	543	.12	.06	1.27	4.0	1.29	3.9	F .70 .74	57.1	53.9	ASR2
5	2037	543	.03	.06	1.05	.9	1.28	3.8	G .72 .74	57.3	54.2	RLB1
16	2018	543	.11	.06	1.21	3.2	1.27	3.7	H .71 .74	56.5	53.8	ASR1
11	1760	543	1.10	.06	1.26	4.0	1.26	3.8	I .73 .78	48.3	49.9	RSP2
3	2259	543	-.96	.07	1.19	2.7	1.26	2.8	J .66 .69	56.3	58.7	TGL3
7	2010	543	.14	.06	1.07	1.2	1.18	2.5	K .71 .74	54.8	53.8	RLB3
13	1997	543	.19	.06	1.00	.1	1.13	1.8	L .74 .75	56.1	53.8	RSP4
12	1905	543	.55	.06	1.10	1.5	1.08	1.3	M .74 .76	55.8	51.9	RSP3
4	1998	543	.19	.06	.98	-.3	1.05	.7	N .74 .75	56.3	53.8	TGL4
28	2178	543	-.58	.07	1.03	.5	1.01	.2	O .72 .71	60.4	56.9	EMP2
2	2113	543	-.29	.07	.85	-2.5	.96	-.5	P .74 .73	61.6	55.6	TGL2
6	2100	543	-.23	.07	.96	-.6	.93	-.9	Q .73 .73	59.6	55.3	RLB2
14	2080	543	-.15	.06	.95	-.7	.91	-1.3	R .75 .73	58.3	55.0	RSP5
20	2106	543	-.26	.07	.92	-1.2	.85	-2.1	S .75 .73	64.9	55.3	ASR5
9	2122	543	-.33	.07	.85	-2.2	.84	-2.2	T .75 .72	64.7	55.7	RLB5
27	2028	543	.07	.06	.83	-2.9	.79	-3.3	U .78 .74	64.7	54.1	EMP1
22	2051	543	-.03	.06	.80	-3.4	.74	-4.1	V .78 .74	64.7	54.5	ASR7
15	2017	543	.11	.06	.76	-4.2	.74	-4.1	W .78 .74	64.9	53.8	RSP6
25	2087	543	-.18	.07	.72	-4.8	.75	-3.8	X .78 .73	66.7	55.2	ASR10
30	1972	543	.29	.06	.75	-4.4	.74	-4.4	Y .79 .75	64.7	53.4	EMP4
26	2011	543	.14	.06	.72	-4.8	.74	-4.2	Z .79 .74	65.5	53.8	ASR11
23	2047	543	-.01	.06	.73	-4.8	.70	-4.9	AA .79 .74	65.9	54.5	ASR8
19	2100	543	-.23	.07	.70	-5.2	.69	-4.8	AB .78 .73	67.4	55.3	ASR4
21	2109	543	-.27	.07	.70	-5.3	.65	-5.4	AC .79 .73	67.3	55.6	ASR6
18	2022	543	.09	.06	.68	-5.7	.64	-6.0	AD .80 .74	67.4	54.0	ASR3
29	2049	543	-.02	.06	.68	-5.7	.67	-5.4	AE .80 .74	68.2	54.5	EMP3
MEAN	2039.5	543.0	.00	.06	1.00	-.4	1.02	-.1		60.1	54.5	
S.D.	125.3	.0	.52	.00	.27	4.1	.30	4.1		6.1	2.3	

Pengujian validitas kontrak dengan menggunakan *item: dimensionality* pada *Winstep*, dengan kriteria bahwa dinyatakan memiliki kemampuan mengukur range variable atau mengukur semua responden jika *Raw Variance Explained by measures* berada di atas 40% untuk data type Likerts (Bambang Sumintono, 2015), maka seperti pada Tabel 3 ditemukan bahwa nilai yang dimiliki oleh *Raw Variance Explained by measures* sebesar 61,6%. Berdasarkan hal itu maka instrument kualitas layanan dosen ini memiliki validitas konstruk atau mampu mengukur range variable atau mengukur semua responden. Namun demikian jika dilihat menggunakan *Wright Map*, terlihat jelas seperti pada Tabel 7 terdapat pengumpulan butir (*item*) yang terkumpul di tengah. Ini berarti terdapat butir-butir instrument yang harus diperbaiki.

Tabel 3. Hasil Pengujian *Item: Dimensionality*

Total raw variance in observations	=	-- Empirical --	80.8	100.0%	Modeled	100.0%
Raw variance explained by measures	=		49.8	61.6%		61.6%
Raw variance explained by persons	=		31.7	39.2%		39.2%
Raw Variance explained by items	=		18.1	22.4%		22.4%
Raw unexplained variance (total)	=		31.0	38.4%	100.0%	38.4%
Unexplned variance in 1st contrast	=		3.1	3.8%	10.0%	
Unexplned variance in 2nd contrast	=		2.8	3.5%	9.1%	
Unexplned variance in 3rd contrast	=		2.0	2.5%	6.5%	
Unexplned variance in 4th contrast	=		1.6	2.0%	5.2%	
Unexplned variance in 5th contrast	=		1.4	1.8%	4.6%	

Hasil pengujian validitas dengan melihat hasil *test rating (partial –credit) scale* sebagaimana terdapat pada Tabel 4 ditemukan bahwa setiap *rating* (1, 2, 3, 4, 5) mempunyai puncak yang terpisah; atau berarti probabilitas tiap *rating* terlihat jelas perbedaannya oleh responden. Berdasarkan Tabel 3 bahwa instrument kualitas layanan dosen mampu dibedakan *ratingnya* oleh responden.

Tabel 4. Hasil Uji *rating (partial –credit) scale*

Hasil pengujian menggunakan *summary statistic* untuk instrument kualitas layanan dosen sebagaimana terlihat pada Tabel 5 bahwa instrument kualitas layanan ini memiliki reliabilitas orang (*person*), alpha cronbach dan reliabilitas butir (*item*) mendekati 1. Artinya reliabilitas instrument ini baik karena instrument ini memiliki reliabilitas *person* 0.97, reliabilitas *alpha cronbach* 0.94 dan reliabilitas butir (*item*) 0.98. Ini berarti instrument kualitas layanan ini jika digunakan oleh peneliti lain maka hasilnya relatif ajeg atau akan memiliki kemiripan hasil.

Tabel 5. Hasil Uji Menggunakan *Summary Statistic*

REAL RMSE	.34	TRUE SD	1.68	SEPARATION	4.92	Person RELIABILITY	.96
MODEL RMSE	.30	TRUE SD	1.68	SEPARATION	5.66	Person RELIABILITY	.97
S.E. OF Person MEAN = .08							
Person RAW SCORE-TO-MEASURE CORRELATION = .93							
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .98							
REAL RMSE	.07	TRUE SD	.51	SEPARATION	7.57	Item RELIABILITY	.98
MODEL RMSE	.06	TRUE SD	.52	SEPARATION	7.97	Item RELIABILITY	.98
S.E. OF Item MEAN = .09							

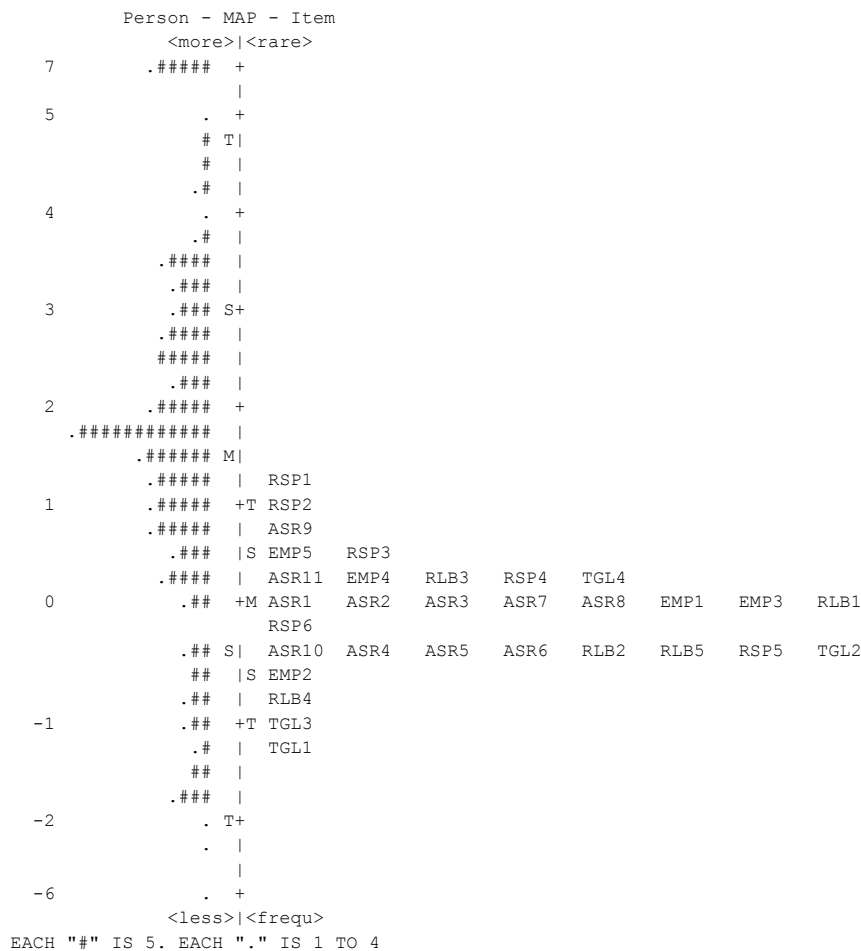
Hasil pengujian menggunakan *item measure* dan *variable map* untuk mengetahui butir (*item*) yang paling sulit disetujui dan paling mudah disetujui oleh responden ditemukan sebagaimana dalam Tabel 6 dan Tabel 7, dimana terdapat 10 butir (*item*) yang sulit disetujui reponden, yaitu: RSP1, RSP2, ASR9, EMP5, RSP3, ASR1, EMP4, RLB3, RSP4, dan TGL4. Sedangkan yang paling mudah disetujui yaitu TGL1. Namun demikian jika memperhatikan Tabel 7 ditemukan butir (*item*) instrument terlalu menumpuk di tengah atau kurang menyebar sesuai dengan responden (mahasiswa). Jika melihat *Standar Deviasi* (S.D) pada Tabel 6 terdapat S.D sebesar 0.52, maka ini berarti *range* tingkat kesulitan butir (*item*) tidak beragam sehingga perlu diperbaiki.

Table 6. Hasi Uji *Items Measure*

ENTRY	TOTAL	TOTAL	MODEL	INFIT	OUTFIT	PT-MEASURE	EXACT MATCH	
NUMBER	SCORE	COUNT	MEASURE	S.E. MNSQ	ZSTD MNSQ	ZSTD CORR.	EXP. OBS%	EXP% Item
10	1705	543	1.30	.06 1.71	9.8 1.85	9.9	.63 .78 42.7	50.1 RSP1
11	1760	543	1.10	.06 1.26	4.0 1.26	3.8	.73 .78 48.3	49.9 RSP2
24	1859	543	.73	.06 1.47	6.7 1.47	6.5	.69 .77 52.6	51.2 ASR9
12	1905	543	.55	.06 1.10	1.5 1.08	1.3	.74 .76 55.8	51.9 RSP3
31	1913	543	.52	.06 1.38	5.6 1.37	5.1	.72 .76 59.3	52.3 EMP5
30	1972	543	.29	.06 .75	-4.4 .74	-4.4	.79 .75 64.7	53.4 EMP4
13	1997	543	.19	.06 1.00	.1 1.13	1.8	.74 .75 56.1	53.8 RSP4
4	1998	543	.19	.06 .98	-.3 1.05	.7	.74 .75 56.3	53.8 TGL4
7	2010	543	.14	.06 1.07	1.2 1.18	2.5	.71 .74 54.8	53.8 RLB3
26	2011	543	.14	.06 1.27	-4.8 .74	-4.2	.79 .74 65.5	53.8 ASR11
17	2016	543	.12	.06 1.22	4.0 1.29	3.9	.70 .74 57.1	53.9 ASR2
15	2017	543	.11	.06 .76	-4.2 .74	-4.1	.78 .74 64.9	53.8 RSP6
16	2018	543	.11	.06 1.21	3.2 1.27	3.7	.71 .74 56.5	53.8 ASR1
18	2022	543	.09	.06 .68	-5.7 .64	-6.0	.80 .74 67.4	54.0 ASR3
27	2028	543	.07	.06 .83	-2.9 .79	-3.3	.78 .74 64.7	54.1 EMP1
5	2037	543	.03	.06 1.05	.9 1.28	3.8	.72 .74 57.3	54.2 RLB1
23	2047	543	-.01	.06 .73	-4.8 .70	-4.9	.79 .74 65.9	54.5 ASR8
29	2049	543	-.02	.06 .68	-5.7 .67	-5.4	.80 .74 68.2	54.5 EMP3
22	2051	543	-.03	.06 .80	-3.4 .74	-4.1	.78 .74 64.7	54.5 ASR7
14	2080	543	-.15	.06 .95	-.7 .91	-1.3	.75 .73 58.3	55.0 RSP5
25	2087	543	-.18	.07 .72	-4.8 .75	-3.8	.78 .73 66.7	55.2 ASR10
6	2100	543	-.23	.07 .96	-.6 .93	-.9	.73 .73 59.6	55.3 RLB2
19	2100	543	-.23	.07 .70	-5.2 .69	-4.8	.78 .73 67.4	55.3 ASR4
20	2106	543	-.26	.07 .92	-1.2 .85	-2.1	.75 .73 64.9	55.3 ASR5
21	2109	543	-.27	.07 .70	-5.3 .65	-5.4	.79 .73 67.3	55.6 ASR6
2	2113	543	-.29	.07 .85	-2.5 .96	-.5	.74 .73 61.6	55.6 TGL2
9	2122	543	-.33	.07 .85	-2.5 .84	-2.2	.75 .72 64.7	55.7 RLB5
28	2178	543	-.58	.07 1.03	.5 1.01	.2	.72 .71 60.4	56.9 EMP2
8	2227	543	-.80	.07 1.06	.9 1.32	3.6	.69 .70 59.6	58.1 RLB4
3	2259	543	-.96	.07 1.19	2.7 1.26	2.8	.66 .69 56.3	58.7 TGL3
1	2327	543	-1.31	.07 1.52	6.6 1.41	3.7	.62 .67 52.4	61.9 TGL1
MEAN	2039.5	543.0	.00	.06 1.00	-.4 1.02	-.1	60.1	54.5
S.D.	125.3	.0	.52	.00 .27	4.1 .30	4.1	6.1	2.3

Keterangan Tabel 4. TGL= *tangible*, RLB = *reliability*, RSP= *responsiveness*, ASR = *assurance*, EMP = *empathy*.

Tabel 7. Hasil Uji Variable Map



Hasil tes menggunakan *Differential Item Functional (DIF) plot* dan nilai probabilitas tabel 30.4 pada Winstep, dimana perbedaan persepsi dinyatakan signifikan jika nilai $p < 0.05$, maka seperti Gambar 1 dan Tabel 7 bahwa mahasiswa perempuan (garis 2 merah) dan mahasiswa laki-laki (garis 1 biru) memiliki perbedaan persepsi yang signifikan pada 8 butir instrument, yaitu: TGL1 ($p=0.0176$), RLB4 ($p=0.0216$), RLB5 ($p=0.0005$), RSP1 ($p=0.0176$) RSP2 (0.0036), RSP3 ($p=0.0065$), ASR2 ($p=0.0322$), dan ASR7 ($p=0.00185$). Ini berarti kedelapan butir yang dipersepsikan berbeda itu harus diperbaiki karena masih memiliki bias gender secara nyata.



Gambar 1 Hasil *Differential Item Functional (DIF) plot* Berdasarkan Jenis Kelamin

Keterangan: Mahasiswa laki-laki (biru) , 2 = Mahasiswa perempuan (merah)

Tabel 7. Hasil Uji D.F *Person Classes* Berdasarkan Jenis Kelamin

Person CLASSES	SUMMARY DIF			BETWEEN-CLASS		Item	
	CHI-SQUARE	D.F.	PROB.	MEAN-SQUARE	t=ZSTD	Number	Name
2	5.6328	1	.0176	2.3990	1.1898	1	TGL1
2	.8598	1	.3538	.3565	-.1457	2	TGL2
2	1.6018	1	.2056	.6786	.2142	3	TGL3
2	3.3403	1	.0676	1.3072	.6695	4	TGL4
2	2.2178	1	.1364	.8657	.3718	5	RLB1
2	.8357	1	.3606	.3226	-.1951	6	RLB2
2	.2092	1	.6474	.1022	-.6582	7	RLB3
2	5.2779	1	.0216	2.2069	1.1119	8	RLB4
2	11.9709	1	.0005	4.9607	1.9680	9	RLB5
2	5.6445	1	.0175	2.1791	1.1003	10	RSP1
2	8.4891	1	.0036	3.3107	1.5117	11	RSP2
2	7.4019	1	.0065	2.9111	1.3790	12	RSP3
2	3.4783	1	.0622	1.3618	.7014	13	RSP4
2	.0519	1	.8197	.0222	-1.0535	14	RSP5
2	.4653	1	.4952	.1770	-.4588	15	RSP6
2	3.7063	1	.0542	1.4550	.7539	16	ASR1
2	4.5859	1	.0322	1.8057	.9333	17	ASR2
2	.0000	1	1.0000	.0040	-1.3142	18	ASR3
2	.3019	1	.5827	.1276	-.5819	19	ASR4
2	.0796	1	.7779	.0543	-.8468	20	ASR5
2	2.2371	1	.1347	.9173	.4112	21	ASR6
2	5.5476	1	.0185	2.2503	1.1299	22	ASR7
2	1.8039	1	.1792	.7319	.2618	23	ASR8
2	.8408	1	.3592	.3220	-.1959	24	ASR9
2	3.1407	1	.0764	1.2806	.6537	25	ASR10
2	.0000	1	1.0000	.0009	-1.4419	26	ASR11
2	3.5649	1	.0590	1.4383	.7446	27	EMP1
2	.4185	1	.5177	.1789	-.4545	28	EMP2
2	.9813	1	.3219	.4005	-.0863	29	EMP3
2	.1433	1	.7050	.0849	-.7176	30	EMP4
2	.1765	1	.6744	.0873	-.7090	31	EMP5

Hasil tes menggunakan *Differential Item Functional (DIF) plot* dan nilai probabilitas tabel 30.4, dimana perbedaan persepsi dinyatakan signifikan jika nilai $p < 0.05$ seperti pada Gambar 2 dan Tabel 8, maka berdasarkan kriteria tersebut bahwa tidak ada perbedaan persepsi (bias) yang signifikan antara mahasiswa yang berasal dari sekolah menengah negeri (garis 1 biru) dan mahasiswa yang berasal dari sekolah menengah swasta (garis 2 merah). Ini berarti instrument kualitas layanan dosen dapat mengukur dengan baik mahasiswa yang berasal dari Sekolah Menengah (SMU/K) yang berstatus negeri maupun swasta.



Gambar 2: Hasil *Differential Item Functional (DIF) plot* Berdasarkan Status Sekolah Menengah

Keterangan: Gambar 2: 1 = Sekolah Negeri, 2 = Sekolah Menengah Swasta

Tabel 8. Hasil Uji D.F *Person Classes* Berdasarkan Status Sekolah Menengah

Person CLASSES	SUMMARY DIF			BETWEEN-CLASS		Item	
	CHI-SQUARE	D.F.	PROB.	MEAN-SQUARE	t=ZSTD	Number	Name
2	.8529	1	.3557	.4270	-.0525	1	TGL1
2	.3329	1	.5640	.1625	-.4923	2	TGL2
2	.4703	1	.4929	.2290	-.3520	3	TGL3
2	.0000	1	1.0000	.0050	-1.2877	4	TGL4
2	.0000	1	1.0000	.0003	-1.5011	5	RLB1
2	.5993	1	.4388	.2993	-.2310	6	RLB2
2	.0874	1	.7675	.0689	-.7804	7	RLB3
2	.7063	1	.4007	.3534	-.1501	8	RLB4
2	.1215	1	.7275	.0614	-.8131	9	RLB5
2	2.9999	1	.0833	1.4903	.7731	10	RSP1
2	.0000	1	1.0000	.0004	-1.4956	11	RSP2
2	.0000	1	1.0000	.0189	-1.0852	12	RSP3
2	.0000	1	1.0000	.0143	-1.1348	13	RSP4
2	.6053	1	.4366	.2968	-.2349	14	RSP5
2	.0825	1	.7739	.0646	-.7990	15	RSP6
2	.6702	1	.4130	.3294	-.1849	16	ASR1
2	.3004	1	.5836	.1471	-.5302	17	ASR2
2	.2164	1	.6418	.1057	-.6469	18	ASR3
2	1.3634	1	.2429	.6791	.2147	19	ASR4
2	.0000	1	1.0000	.0216	-1.0588	20	ASR5
2	1.6916	1	.1934	.8425	.3536	21	ASR6
2	.0000	1	1.0000	.0415	-.9153	22	ASR7
2	.4932	1	.4825	.2461	-.3205	23	ASR8
2	.0000	1	1.0000	.0003	-1.5156	24	ASR9
2	.0000	1	1.0000	.0064	-1.2565	25	ASR10
2	.1998	1	.6549	.0976	-.6731	26	ASR11
2	.0000	1	1.0000	.0076	-1.2321	27	EMP1
2	1.0036	1	.3165	.5010	.0349	28	EMP2
2	.0000	1	1.0000	.0014	-1.4148	29	EMP3
2	.0751	1	.7840	.0598	-.8203	30	EMP4
2	.0784	1	.7795	.0644	-.7995	31	EMP5

Pengujian menggunakan *Differential Item Functional (DIF) plot* dan nilai probabilitas tabel 30.4 pada Winstep, dimana perbedaan persepsi dinyatakan signifikan jika nilai $p < 0.05$, sebagaimana pada Gambar 3 dan Table 9, bahwa terdapat perbedaan persepsi mahasiswa di berbagai fakultas yang signifikan pada 9 butir, yaitu: TGL4 ($p = 0.0012$), RLB1 ($p = 0.0083$), RLB2 ($p = 0.0017$), RLB5 ($p = 0.0018$), RSP1 (0.0000), RSP6 ($p = 0.0437$), ASR1 ($p = 0.0000$), ASR2 ($p = 0.0094$) dan EMP5 ($p = 0.0127$). Ini berarti sembilan butir (*item*) tersebut harus diperbaiki karena masih mengandung bias berdasarkan fakultas.



Gambar 3: Hasil *Differential Item Functional (DIF) plot* Berdasarkan Fakultas

Keterangan: Gambar 3: 1 = FMIPA; 2 = FKIP; 3 = FISIB; 4 = FE; 5 = FH; 6 = FT

Tabel 7. Hasil Uji D.F *Person Classes* Berdasarkan Fakultas

Person CLASSES	SUMMARY DIF			BETWEEN-CLASS		Item	
	CHI-SQUARE	D.F.	PROB.	MEAN-SQUARE	t=ZSTD	Number	Name
6	10.1111	5	.0720	.2041	-1.7396	1	TGL1
6	5.4154	5	.3670	.1576	-1.9703	2	TGL2
6	10.0375	5	.0740	.2240	-1.6517	3	TGL3
6	20.1539	5	.0012	.9538	.1366	4	TGL4
6	15.5382	5	.0083	.3867	-1.0769	5	RLB1
6	19.3237	5	.0017	.5349	-.6822	6	RLB2
6	19.0970	5	.0018	.6045	-.5218	7	RLB3
6	9.7408	5	.0827	.2518	-1.5373	8	RLB4
6	19.1325	5	.0018	.5755	-.5870	9	RLB5
6	46.2855	5	.0000	1.9133	1.3561	10	RSP1
6	6.7106	5	.2428	.1266	-2.1507	11	RSP2
6	5.8067	5	.3252	.1277	-2.1438	12	RSP3
6	7.2724	5	.2009	.1453	-2.0391	13	RSP4
6	2.6606	5	.7520	.0743	-2.5387	14	RSP5
6	11.4126	5	.0437	.6261	-.4746	15	RSP6
6	30.3347	5	.0000	1.0801	.3342	16	ASR1
6	15.2401	5	.0094	.3973	-1.0454	17	ASR2
6	5.6165	5	.3450	.2864	-1.4059	18	ASR3
6	2.2276	5	.8167	.0296	-3.0654	19	ASR4
6	7.7081	5	.1728	.1153	-2.2237	20	ASR5
6	1.8282	5	.8723	.0261	-3.1247	21	ASR6
6	1.6188	5	.8989	.0542	-2.7375	22	ASR7
6	2.0872	5	.8369	.0791	-2.4961	23	ASR8
6	5.0614	5	.4081	.2067	-1.7281	24	ASR9
6	3.3650	5	.6437	.1511	-2.0063	25	ASR10
6	8.3094	5	.1397	.1580	-1.9685	26	ASR11
6	3.7065	5	.5922	.1281	-2.1412	27	EMP1
6	7.8028	5	.1672	.1671	-1.9202	28	EMP2
6	7.1052	5	.2126	.2432	-1.5718	29	EMP3
6	8.1729	5	.1467	.1585	-1.9656	30	EMP4
6	14.5028	5	.0127	.8546	-.0312	31	EMP5

Instrumen penelitian yang baik harus melalui uji terlebih dahulu (Brinkman, 2009), tujuannya agar instrument tersebut memenuhi syarat validitas (Ghazali, 2016). Valid berarti tingkat kepercayaan (Zohrabi, 2013). Valid berarti instrument yang digunakan memiliki kemampuan untuk mengukur apa yang akan diukur (Rahmawati, 2019). Instrument yang baik harus memenuhi beberapa kriteria validitas sehingga banyak istilah validitas instrument: validitas isi, validitas konstruk/bangun, validitas prediktif, validitas statistic/realibilitas (Baghaei, 2014). Jika peneliti menggunakan Model Rasch untuk menganalisis instrument penelitian, seperti halnya dalam penelitian ini, maka peneliti dapat menggunakan Model Rasch sebagai alternative (Zile-Tamsen, 2017). Selama ini pengujian validitas instrument menggunakan uji validitas *Pearson Product Moment* untuk instrumen *non-test*/kuesioner (Widi, 2011) dan uji *Point Biserial* untuk instrument *test* dan pilihan benar salah (Jesyca R.T. Muaju, Adi Setiawan, 2013). Penggunaan Model Rasch untuk pengujian validitas instrument hasilnya bisa lebih variatif karena validitas instrument bisa menggunakan berbagai kriteria (Othman, 2014). Penelitian ini jika menggunakan kriteria validitas *OUTFIT MNSQ* hasilnya 31 butir valid dan 1 butir tidak valid (RSP1); jika menggunakan *OUTFIT Z-STANDARD* hasilnya dari 8 butir valid dan 24 tidak valid; jika menggunakan kriteria validitas *Point Measure Correlation* hasilnya 32 butir valid; jika menggunakan *item: dimensionaly* untuk mengukur sejauh mana butir (*item*) dapat mengukur konstruk yang diukur, maka hasilnya beberapa butir harus diperbaiki karena terkonsentrasi pada satu dimensi saja; dan jika menggunakan *rating (partial –credit) scale test* hasilnya semua butir (*item*) instrumen memenuhi kriteria. Berdasarkan temuan ini terbukti bahwa kelebihan Rasch Model dalam menganalisis validitas instrument dapat dilakukan dari beberapa segi sehingga instrument yang dihasilkan dapat lebih dipercaya (Abdaziz, Jusoh, & Amlus, 2014). Analisis validitas menggunakan Rasch Model dapat dikatakan lebih baik karena kekonsistennya (Jusoh, 2018). Dengan demikian pemodelan Rasch dapat membantu untuk mengatasi pengukuran *item* (butir) dengan cara yang benar (Wu & Adams, 2007).

Selain memenuhi syarat validitas (Takaki, Taniguchi, & Fujii, 2014), sebuah instrument penelitian juga harus memenuhi syarat reliabilitas (Mohajan, 2017). Reliabilitas berarti instrument tersebut jika digunakan oleh peneliti lain akan memiliki hasil yang relative sama hasilnya (Ghazali, 2016). Keunggulan tes reliabilitas dalam Rasch model yaitu instrument bisa diketahui memiliki 3 reliabilitas, yaitu reliabilitas *person*, *alpha cronbach*, dan reliabilitas *item*/butir (Stef van Buuren, 2015). Pada penelitian ini menunjukkan bahwa instrument kualitas layanan dosen memiliki reliabilitas yang tinggi yaitu reliabilitas *person* 0.97, reliabilitas *alpha cronbach* 0.94 dan reliabilitas *item*/butir 0.98. Koefisien reliabilitas di atas 0.9 menunjukkan sangat baik sekali (Mohamad, Lisa, Sern, & Mohd, 2015). Ini berarti instrument kualitas layanan ini jika digunakan oleh peneliti lain maka hasilnya akan relatif ajeg atau memiliki kemiripin hasil. Instrumen yang memiliki reliabilitas tinggi merupakan salah satu ciri instrumen yang baik (Mohajan, 2017).

Instrumen yang baik harus pula memiliki kemampuan untuk menjangkau semua aspek yang diukur (Loubser, Casteleijn, & Bruce, 2015). Model Rasch merupakan aplikasi yang dapat digunakan untuk mengevaluasi apakah sebuah instrument memiliki kemampuan mengukur (*construct validity*) (Baghaei, 2014), karena Model Rasch mampu menunjukkan butir (*item*) instrumen yang sulit disetujui/dikerjakan oleh responden (Abdaziz et al., 2014) sekaligus disandingkan kemampuan respondennya (Bambang Sumintono, 2015). Hasil penelitian ini menunjukkan bahwa berdasar hasil analisis menggunakan *item measure* dan *variable map* serta melihat *Standar Deviasi* (S.D) diperoleh gambaran butir (*item*) yang paling sulit disetujui dan paling mudah disetujui oleh responden dan bagaimana sebaran butirnya. Penelitian ini menemukan 10 butir (*item*) yang sulit disetujui reponden, butir (*item*) terlalu menumpuk di tengah atau kurang menyebar sesuai dengan dengan responden (mahasiswa) dan *range* tingkat kesulitan butir (*item*) tidak beragam sehingga perlu diperbaiki. Analisis butir yang sulit dan bagaimana hubungannya dengan kemampuan responden merupakan ciri khas Model Rasch untuk mengembangkan instrument dengan lebih baik (Mohamad et al., 2015). Analisis butir intrumen yang dikaitkan dengan kemampuan respondennya sangat membatu pembuatan instrument yang menjangkau aspek yang mau diukur (Nurulhuda, Mohd, Fauzi, Yunus, & Ismail, 2018).

Instrument yang baik harus pula menghindari butir (*item*) yang ambigu (Aguilar, Cortes, Guerrero, Herrera, & Orozco, 2011). Instrument yang ambigu harus dihindari karena akan dimaknai secara berbeda sehingga hasilnya pun berbeda (Josephine & Mitan, 2014), banyak cara agar instrument terhindar dari ambigu (Bors, Gruman, & Shukla, 2010), penggunaan model Rasch secara baik dapat membantu agar terhindar dari instrument yang ambigu atau tidak mampu dipahami dengan baik oleh responden (Siti Eshah Mokshein, 2019). Dalam penelitian ini Model Rasch menunjukkan bahwa: terdapat 7 butir instrument dipersepsikan secara berbeda dengan signifikan oleh responden (mahasiswa) laki-laki dan perempuan; dan terdapat 9 butir instrument yang dipersepsikan berbeda dengan signifikan oleh mahasiswa dari fakultas yang berbeda-beda. Jika ditemukan butir instrument yang dipersepsikan berbeda, maka butir intrumen itu harus diperbaiki (Wibisono, 2014).

Kesimpulan

Instrumen kualitas layanan dosen yang memiliki validas dan reliabilitas tinggi sangat dibutuhkan untuk penelitian yang bertujuan meningkatkan kualitas layanan pada mahasiswa. Instrumen kualitas layanan yang diperoleh dalam penelitian ini berasal dari sebuah proses penelitian mulai dari tinjauan literatur, wawancara, validasi ahli dan analisis data menggunakan pemodelan Rasch. Berdasarkan hasil pengujian menggunakan Model Rasch ditemukan 31 butir instrument kualitas layanan dosen yang valid dan satu butir yang tidak valid sehingga harus dibuang. Berdasarkan uji uji reliabilitas menggunakan analisis Model rasch ditemukan bahwa instrument kualitas layanan dosen ini sangat terpercaya karena memiliki nilai reliabilitas di atas 0.9, baik dari segi reliabilitas *person*, *alpha cronbach* maupun reliabilitas butir (*items*). Namun demikian berdasarkan uji tingkat kesukaran butir menggunakan *item measure*, *item dimensionality* dan *variable map* serta uji *Item: DIF* ditemukan beberapa butir (*item*) yang harus diperbaiki karena masih mengandung bias (gender dan fakultas) dan kurang mampu menjangkau *range* variabel yang diteliti .

Referensi

- Abdaziz, A., Jusoh, M. S., & Amlus, H. (2014). Construct Validity : A Rasch Measurement Model Approaches. *Journal of Applied Science and Agriculture*, 9 (September), 7–12.
- Abdullah, N., & Mohamad, N. (2016). University Recreational Facilities Service Quality and Students ' Physical Activity Level. *Procedia - Social and Behavioral Sciences*, 224(August 2015), 207–212. <https://doi.org/10.1016/j.sbspro.2016.05.443>

- Aguilar, H. C., Cortes, P. P., Guerrero, A. V, Herrera, C. D. C., & Orozco, J. A. T. (2011). Construction and Validation of an Instrument To Evaluate Critical Reading of Research Papers. *2011 4th International Conference of Education, Research and Innovation*, (November), 5286–5294.
- Azam, A. (2018). Service Quality Dimensions and Students' Satisfaction : A study of Saudi Arabian Private Higher Education Institutions. *European Online Journal of Natural and Social Sciences*, 7(2), 275–284.
- Baghaei, P. (2014). The Rasch Model as a Construct Validation Tool. *Rasch Measurement Transactions*, 22(1).
- Bambang Sumintono, W. W. (2014). *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial*. Cimahi: Trim Komunikata.
- Bambang Sumintono, W. W. (2015). *Aplikasi Pemodelan Rasch pada Asesment Pendidikan*. Cimahi: Trim Komunikata.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model : Fundamental Measurement in the Human Sciences Second Edition University of Toledo*. New Jersey 07430: Lawrence Erlbaum Associates, Inc.
- Boon, T., Shukur, M., & Bassim, A. (2016). Evaluation of Service Quality of Private Higher Education using Service Improvement Matrix. *Procedia - Social and Behavioral Sciences*, 224(August 2015), 132–140. <https://doi.org/10.1016/j.sbspro.2016.05.417>
- Bors, D. A., Gruman, J. A., & Shukla, S. (2010). Measuring tolerance of ambiguity : Item polarity , dimensionality , and criterion validity. *Revue Europeenne de Psychologie Appliquee*, 60(4), 239–245. <https://doi.org/10.1016/j.erap.2010.07.001>
- Brinkman, W.-P. (2009). Design of a Questionnaire Instrument. *Handbook of Mobile Technology Research Methods*, (December), 31–57.
- Carvalho, L. D. F., Primi, R., & Meyer, G. J. (2012). Application of the Rasch model in measuring personality disorders. *Trends*, 34(2), 101–109.
- Deb, S., & Ahmed, M. A. (2018). Determining the service quality of the city bus service based on users' perceptions and expectations. *Travel Behaviour and Society*, 12(February), 1–10. <https://doi.org/10.1016/j.tbs.2018.02.008>
- Dursun, T., Oskayba, K., & Gökmen, C. (2013). The Quality Of Service Of The Distance Education. In *13th International Educational Technology Conference The* (Vol. 103, pp. 1133–1151). <https://doi.org/10.1016/j.sbspro.2013.10.441>
- Esmailpour, M., & Ranjbar, M. (2016). Investigating the Impact of Commitment , Satisfaction , and Loyalty of Employees on Providing High - Quality Service To Customer. *Journal of Modelling in Management*, 14(3), 355–388.
- Ghazali, N. H. M. (2016). A Reliability and Validity of an Instrument to Evaluate the School-Based Assessment System : A Pilot Study. *International Journal of Evaluation and Research in Education (IJERE)*, 5(2), 148–157.
- Hendriks, J., Fyfe, S., Styles, I., Skinner, S. R., Merriman, G., & Hendriks, J. (2012). Scale construction utilising the Rasch unidimensional measurement model : A measurement of adolescent attitudes towards abortion. *Australasian Medical Journal [AMJ 2012]*, 5(5), 251–261.
- Jesyca R.T. Muaju, Adi Setiawan, T. M. (2013). Uji Validitas dan Reliabilitas Menggunakan Metode Bootstrap pada Kuisioner Ripe Yes/No Questions. In *Prosiding Seminar Nasional Sains dan Pendidikan Sains VII* (p. 513).
- Jiao, H., Lissitz, R. W., & Macready, G. (2011). Exploring levels of performance using the mixture Rasch model for standard setting 1, *53(4)*, 499–522.
- John I. Pollock. (1967). Logical validity in modal logic. *The Monist*, 51(1), 128–135.
- Josephine, H., & Mitan, C. (2014). Studi Eksperimental : Mengurangi Bias Pengukuran Umum Balanced Scorecard dalam Penilaian Kinerja pada Mahasiswa S1 Program Akuntansi Pajak. *BUSINESS ACCOUNTING REVIEW*, 2(2), 32–41.
- Jusoh. (2018). Construct Validity for Measuring Entrepreneurial Readiness Among Malaysian Higher Education Students : A Stochastic Measurement Model Approach. In *MATEC Web of Conferences 150*, (Vol. 150, pp. 1–4).
- Klette, K. (2012). *Mixed Methods in Educational Research*. Oslo: The Research Council of Norway.
- Kumar, S., Shekhar, V., Lassar, W. M., & Chen, T. (2018). Journal of Retailing and Consumer Services Customer engagement behaviors : The role of service convenience , fairness and quality. *Journal of Retailing and Consumer Services*, 44(December 2017), 293–304. <https://doi.org/10.1016/j.jretconser.2018.07.018>
- Linacre, J. M. (2010). Predicting Responses from Rasch Measures. *JOURNAL OF APPLIED MEASUREMENT*, 11(1), 1–10.
- Loubser, H. J., Casteleijn, D., & Bruce, J. C. (2015). ScienceDirect The BETA ® nursing measure : Calibrating construct validity with Rasch analyses. *Health SA Gesondheid*, 20(1), 22–32.

- <https://doi.org/10.1016/j.hsag.2015.02.001>
- Meerah, T. S. M., Osman, K., Zakaria, E., & Ikhsan, Z. H. (2012). Developing an Instrument to Measure Research Skills. *Procedia - Social and Behavioral Sciences*, 60, 630–636. <https://doi.org/10.1016/j.sbspro.2012.09.434>
- Mohajan, H. K. (2017). Two Criteria for Good Measurements in Research: Validity and Reliability. *Annals of Spiru Haret University*, 17(3), 1–32.
- Mohamad, M. M., Lisa, N., Sern, L. C., & Mohd, K. (2015). Measuring the Validity and Reliability of Research Instruments. *Procedia - Social and Behavioral Sciences*, 204(November 2014), 164–171. <https://doi.org/10.1016/j.sbspro.2015.08.129>
- NGUYEN THANH DONG, D. N. F. S. (2014). Applying the Rasch Model to Investigate Singapore Principals' Instructional Leadership Practices. *Leading & Managing*, 20(2), 1–26.
- Nurulhuda, S., Mohd, T., Fauzi, M., Yunus, M., & Ismail, I. (2018). The use of rasch measurement model for the validity and reliability. *Journal of Counseling and Educational Technology*, 1(2), 22–27.
- Oriade, A., & Scho, P. (2019). An examination of the role of service quality and perceived value in visitor attraction experience. *Journal of Destination Marketing & Management*, 11(May 2018), 1–9. <https://doi.org/10.1016/j.jdmm.2018.10.002>
- Othman, N. B. (2014). Assessing Construct Validity and Reliability Of Competitiveness Scale Using Rasch Model Approach. In *The 2014 WEI International Academic Conference Proceedings* (pp. 113–120).
- Prentice, C., & Kadan, M. (2019). Journal of Retailing and Consumer Services The role of airport service quality in airport and destination choice. *Journal of Retailing and Consumer Services*, 47(July 2018), 40–48. <https://doi.org/10.1016/j.jretconser.2018.10.006>
- Rahmawati, E. (2019). Developing instruments of teacher's perception of critical thinking in elementary school. *Journal of Education and Learning (EduLearn)*, 13(4), 559–566. <https://doi.org/10.11591/edulearn.v13i4.13232>
- Rajab, A., Aishah, S., Rahman, A., Abdul, H., Shaari, R., & Saat, M. (2011). Service Quality in a Research University: A Post-Graduate Perspective. *Procedia - Social and Behavioral Sciences*, 29, 1830–1838. <https://doi.org/10.1016/j.sbspro.2011.11.431>
- Roy, S. (2019). Service quality versus service experience : An empirical examination of the consequential effects in B2B services. *Industrial Marketing Management*, XXX(February 2017). <https://doi.org/10.1016/j.indmarman.2019.02.017>
- Siti Eshah Mokshein, H. I. and H. A. (2019). The Use of Rasch Measurement Model In English Testing. *Cakrawala Pendidikan*, 38(1), 16–32. <https://doi.org/10.21831/cp.v38i1.22750>
- Stef van Buuren, D. J. D. W. (2015). Measuring Psychosocial Impact of CBRN Incidents by the Rasch Model. *JOURNAL OF APPLIED MEASUREMENT*, 16(3).
- Stefan, C. (2010). The quality consulting services management in the university educational system. *Procedia - Social and Behavioral Sciences*, 2(2), 5586–5589. <https://doi.org/10.1016/j.sbspro.2010.03.911>
- Sugiyono. (2015). *Metode Penelitian Kombinasi (Mix Methods)*. Bandung: Alfabeta.
- Takaki, J., Taniguchi, T., & Fujii, Y. (2014). Reliability, Validity, and Significance of Assessment of Sense of Contribution in the Workplace. *International Journal of Environmental Research and Public Health*, 11, 1594–1604. <https://doi.org/10.3390/ijerph110201594>
- Tesio, Il. (2003). *MEASURING BEHAVIOURS AND PERCEPTIONS: RASCH ANALYSIS AS A TOOL FOR REHABILITATION RESEARCH*. *J Rehabil Med*, 35, 105–115.
- Wibisono, S. (2014). *APLIKASI MODEL RASCH UNTUK VALIDASI FUNDAMENTALISME AGAMA BAGI RESPONDEN MUSLIM* Susilo Wibisono. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia*, III(3), 730–750.
- Widi, R. (2011). Uji Validitas dan Reliabilitas dalam Penelitian Epidemiologi Kedokteran Gigi. *Stomatognathic (J.K.G. Unej) Vol.*, 8(1), 27–34.
- Wu, M., & Adams, R. (2007). *Applying The Rasch Model To Psycho-Social Measurement A Practical Approach*. Melbourne.: Educational Measurement Solutions.
- Zile-Tamsen, C. Van. (2017). Using Rasch Analysis to Inform Rating Scale. *Research in Higher Education*. <https://doi.org/10.1007/s11162-017-9448-0>
- Zohrabi, M. (2013). Mixed Method Research: Instruments, Validity, Reliability and Reporting Findings. *Theory and Practice in Language Studies*, 3(2), 254–262. <https://doi.org/10.4304/tpls.3.2.254-262>