Jurnal Penelitian dan Pengukuran Psikologi
Volume 14, Nomor 01, April 2025 http://doi.org/10.21009/JPPP.141.05

ANALISIS PSIKOMETRIKA PROPERTI DAN LATENT SUB KELAS DENGAN *MIXTURE RASCH MODEL* PADA SKALA *ORGANIZATIONAL COMMITMENT QUESTIONNAIRE* (OCQ) VERSI INDONESIA

Fransiscus Widodo Setya Budi[1], Alexander Bhirowo Putra Kurniawan[2], Ananta Yudiarso[3]

[1, 2, 3] Fakultas Psikologi, Universitas Surabaya, Surabaya, Indonesia

Email: ananta@staff.ubaya.ac.id

Received :	*Revised :*	*Accepted :*	*Published :*
19 Januari 2025	*07 April 2025*	*15 April 2025*	*30 April 2025*

Abstract

This study evaluates the psychometric properties of the Indonesian version of the Organizational Commitment Questionnaire (OCQ) using the Rasch Model and Mixture Rasch Model. A quantitative approach was used to analyze data from 253 participants obtained through a survey using a convenience sampling technique. The points analyzed are related to validity, reliability, and unidimensionality and heterogeneity of response patterns. Results indicate that the affective, continuance, and normative subscales exhibit adequate evidence of unidimensionality, with raw variance explained ranging from 44% to 46%. Reliability testing revealed high item reliability (0.97–0.98) but moderate to low person reliability (0.58–0.66), suggesting limited sensitivity of the instrument to individual variations. Differential Item Functioning (DIF) analysis identified significant bias in several items of the affective and continuance subscales based on gender and education, while the normative subscale was neutral to these demographic variables. The Mixture Rasch Model analysis revealed heterogeneity in response patterns, with the two-class latent model showing lower AIC, BIC, and CAIC values compared to the one-class model, indicating a better fit. This study demonstrates that the Indonesian version of the OCQ can be utilized to measure organizational commitment but requires item adjustments to improve measurement sensitivity and fairness.

Keywords: Organizational Commitment Questionnaire, Rasch Model, validitas, reliabilitas

Abstrak

Penelitian ini mengevaluasi properti psikometrik skala *Organizational Commitment Questionnaire* **(OCQ)** versi Indonesia menggunakan *Model Rasch* dan *Mixture Rasch Model*. Pendekatan kuantitatif digunakan untuk menganalisis data dari 253 partisipan yang diperoleh melalui survei dengan teknik *convenience sampling***.** Poin-poin yang dianalisis terkait dengan validitas, reliabilitas, dan unidimensionalitas dan heterogenitas pola respons. Hasil menunjukkan bahwa sub-skala afektif, kontinuan, dan normatif memiliki bukti unidimensionalitas yang memadai dengan nilai *raw variance explained* berkisar antara 44% hingga 46%. Uji reliabilitas mengungkapkan reliabilitas butir tinggi (0,97–0,98), tetapi reliabilitas individu sedang hingga rendah (0,58–0,66), mengindikasikan kurangnya sensitivitas instrumen terhadap variasi individu. Analisis bias (DIF) mengidentifikasi beberapa item pada sub-skala afektif dan kontinuan memiliki bias signifikan berdasarkan gender dan pendidikan, sedangkan sub-skala normatif netral terhadap variabel demografis ini. Analisis *Mixture Rasch Model* menunjukkan adanya heterogenitas dalam pola respons, dengan model dua kelas latent menghasilkan nilai AIC, BIC, dan CAIC yang lebih rendah dibandingkan model satu kelas, menegaskan kecocokan yang lebih baik. Penelitian ini menunjukkan bahwa OCQ versi Indonesia dapat digunakan untuk mengukur komitmen organisasi, namun perlu penyesuaian item untuk meningkatkan sensitivitas dan keadilan pengukuran.

Kata kunci: Organizational Commitment Questionnaire, Rasch Model, validitas, reliabilitas

1. Pendahuluan

Dalam bidang psikologi organisasi, komitmen organisasi telah lama diakui sebagai prediktor penting untuk hubungan psikologis antara individu dan perusahaan tempat mereka bekerja (Mowday, Steers, & Porter, 1979). Salah satu konsep penting dalam psikologi organisasi adalah komitmen organisasi, yang mencerminkan hubungan psikologis antara individu dan perusahaan tempat mereka bekerja (Meyer & Allen, 1991). Mereka menyatakan bahwa komitmen organisasi terdiri dari tiga aspek utama: komitmen afektif, komitmen kontinuan, dan komitmen normatif. Masing-masing dimensi memengaruhi seberapa dekat karyawan dengan organisasi.

Komitmen organisasi memengaruhi banyak aspek organisasi. Komitmen organisasi yang tinggi berkorelasi positif dengan produktivitas kerja, keterlibatan karyawan, dan tingkat turnover yang rendah. Ini menjadikannya faktor penting dalam manajemen sumber daya manusia (Meyer, Stanley, Herscovitch, & Topolnytsky, 2002). Selain itu, keberhasilan strategi organisasi dapat diukur melalui penerapan budaya kerja yang kuat yang membantu mencapai tujuan jangka panjang perusahaan (Mathieu & Zajac, 1990). Oleh karena itu, untuk mengembangkan strategi manajemen sumber daya manusia, pengukuran komitmen organisasi yang tepat dan akurat menjadi sangat penting, baik dalam penelitian akademik maupun dalam praktik (Allen & Meyer, 1990).

Alat yang dikembangkan oleh Mowday, Steers, dan Porter (1979) untuk mengukur komitmen organisasi biasanya digunakan *Organizational Commitment Questionnaire* (OCQ). Dalam skala ini, lima belas item menilai komitmen organisasi dalam tiga dimensi: afektif, kontinuan, dan normatif. Meyer dan Allen (1991) kemudian membuat Model Komitmen Organisasi Tiga Komponen, yang memberikan kerangka yang lebih luas. Ini mencakup skala yang berbeda untuk masing-masing dimensi, dengan 24 item total. Dibandingkan dengan instrumen lain seperti *Tetrad Commitment Scale* atau *Organizational Commitment Scale* versi pendek, OCQ memiliki keunggulan dalam hal kedalaman konseptual, kemudahan adaptasi lintas budaya, dan konsistensi hasil pengukuran dalam berbagai konteks organisasi (Vandenberg & Lance, 2000). Selain itu, OCQ telah diuji dalam berbagai bahasa dan budaya, menjadikannya instrumen yang robust untuk studi komparatif global (Kanning & Hill, 2013; Byrne, 2001). Dengan basis empirik yang kuat dan penyusunan teoritis yang mendalam, OCQ tetap menjadi pilihan utama dalam pengukuran komitmen organisasi.

Pada awalnya, *Organizational Commitment Questionnaire* (OCQ) divalidasi dalam sampel Kanada oleh Allen dan Meyer (1990). Hasil penelitiannya menunjukkan bahwa konstruk tersebut terkait seperti kepuasan kerja dan identitas organisasi berbeda. Selanjutnya, penelitian mereplikasi struktur faktor OCQ dan validitas konstruksinya di berbagai lingkungan, seperti Amerika Serikat (Meyer et al., 1993), Tiongkok (Chen & Francesco, 2000), dan India (Kaur & Sharma, 2015). Sangat penting untuk memvalidasi instrumen yang mengukur komitmen organisasi di berbagai konteks, termasuk lintas budaya.

Menurut penelitian yang dilakukan oleh Byrne (2001), adaptasi instrumen pengukuran seperti OCQ harus dilakukan dalam bahasa lokal untuk memastikan bahwa instrumen tersebut relevan secara kultural dan linguistik. Sebagai contoh, Kanning dan Hill (2013) menerjemahkan OCQ ke dalam enam bahasa dan menemukan bahwa skala tersebut mempertahankan makna dan struktur komponennya di berbagai kelompok linguistik. Jonathan (2020) juga mendukung aplikasi OCQ lintas budaya, memverifikasinya dalam konteks Tanzania. Siregar (2018) mengadopsi OCQ ke dalam bahasa Indonesia, yang menggunakan *Confirmatory Factor Analysis* (CFA) untuk mengevaluasi validitas dan ketepatan ukuran tersebut. Hasil studi menunjukkan bahwa struktur yang terdiri dari tiga faktor (afektif, kontinuan, dan normatif) dapat diterapkan pada situasi Indonesia. Agus Sulistiawan dkk (2021) juga telah melakukan validasi dan reliabilitas OCQ menggunakan CFA dan menemukan bahwa skala komitmen organisasi valid dan reliabel untuk digunakan sebagai alat ukur, dengan 15 item yang gugur dari 24 item yang diujikan. Namun kelemahan dari uji ini ialah partisipan yang homogen (perawat) yang berjumlah 93 dan penghapusan item cukup banyak yang dapat mengubah karakteristik asli alat ukur tersebut.

Kesesuaian bahasa dan perbedaan persepsi budaya adalah beberapa komponen yang harus dipertimbangkan saat menggunakan skala lintas budaya, termasuk dalam konteks Indonesia. Adaptasi OCQ dalam berbagai bahasa sering kali menghadapi masalah seperti bias linguistik dan struktur skala yang tidak konsisten. Penelitian sebelumnya menunjukkan hal ini (Byrne, 2001; Elsmore, 2008). Oleh karena itu, untuk memastikan bahwa alat ini dapat digunakan secara efektif di Indonesia, penting untuk melakukan evaluasi properti psikometrik skala OCQ versi Indonesia. Sayangnya, evaluasi psikometrik mendalam terhadap atribut skala ini masih terbatas (Siregar, 2018; Astuti et al., 2020).

Pedoman penting untuk validasi alat ukur dibuat oleh ITC (*International Test Commission*) dan AERA (*American Educational Research Association*) untuk memastikan validitas, keadilan, dan keandalan. Validitas konten (kesesuaian alat ukur dengan konstruk yang diukur), validitas konstrak (keterkaitan alat ukur dengan teori yang mendasari), dan reliabilitas adalah beberapa pedoman yang digunakan oleh ITC untuk membangun dan mengakomodasi alat ukur lintas budaya. Selain itu, ITC menekankan kesetaraan lintas budaya untuk memastikan bahwa alat ukur yang digunakan di berbagai konteks budaya tetap relevan dan tidak bias.

Sementara itu, AERA berfokus pada validasi alat ukur dalam pendidikan dan psikologi melalui *Standards for Educational and Psychological Testing*. Validitas konstrak, validitas isi (kesesuaian item dengan tujuan tes), dan validitas kriteria (hubungan skor tes dengan hasil eksternal). AERA juga memperhatikan keadilan dalam penggunaan pemeriksaan untuk memastikan bahwa pemeriksaan tidak bias terhadap kelompok tertentu. Pengembang alat ukur dapat membuat instrumen yang sah, stabil, dan dapat digunakan secara adil dalam berbagai situasi dengan mengikuti pedoman ITC dan AERA.

Untuk memastikan bahwa alat ukur skala dapat memberikan hasil yang valid dan dapat diandalkan, evaluasi psikometrika skala menjadi penting. Menurut penelitian Linacre (2012), model Rasch memungkinkan pengujian ketepatan skala untuk mengukur dimensi yang dimaksud dengan lebih akurat. Metode ini memungkinkan pengujian kesesuaian data terhadap model, pengenalan item yang tidak valid, dan pembuatan skala yang invarian lintas kelompok (Bond & Fox, 2015). Dalam studi ini, analisis *Differential Item Functioning* (DIF) difokuskan pada dua variabel demografis utama: gender dan tingkat pendidikan. Kedua variabel ini dipilih karena secara teoritis telah terbukti memengaruhi persepsi individu terhadap organisasi. Menurut teori perbedaan individual dalam organisasi (Morgeson et al., 2005), gender dan latar belakang pendidikan membentuk pengalaman dan interpretasi seseorang terhadap nilai, norma, dan harapan organisasi. Oleh karena itu, penting untuk menguji apakah item dalam instrumen memperlakukan kelompok-kelompok ini secara adil. Apabila terdapat bias terhadap kelompok tertentu, hal ini akan merusak validitas konstruk dan dapat mengarah pada keputusan yang tidak adil dalam konteks organisasi maupun penelitian (Zumbo, 1999).

Penelitian ini akan melengkapi model Rasch dengan *Mixture Rasch Model* (MRM). MRM memungkinkan peneliti menemukan sub kelompok dengan pola respons yang berbeda, yang seringkali tidak terlihat dalam model klasik (Rost, 1990; Finch & French, 2015). Metode ini relevan untuk organisasi Indonesia yang sangat beragam dari segi jenis organisasi dan demografi. Dalam konteks OCQ versi Indonesia, MRM berguna untuk mengidentifikasi keberadaan latent classes—kelompok-kelompok tersembunyi dengan cara berpikir atau menjawab yang berbeda terhadap konstruk komitmen organisasi (von Davier & Rost, 2007). Hal ini sangat relevan mengingat sampel dalam penelitian ini terdiri dari individu dengan latar belakang pendidikan dan masa kerja yang sangat beragam.

Dua tujuan utama studi ini adalah untuk mengevaluasi validitas dan reliabilitas skala OCQ versi Indonesia yang didasarkan pada parameter Rasch dan untuk menemukan heterogenitas dalam pola respons dengan model *Mixture Rasch.* Oleh karena itu, ada dua komponen utama yang diberikan kontribusi signifikan oleh penelitian ini. Pertama, melalui pengujian psikometrik yang menyeluruh, memperkuat basis empirik penggunaan OCQ versi Indonesia (Bond & Fox, 2015). Kedua, memberikan pemahaman baru tentang cara menggunakan *Model Mixture Rasch* untuk menemukan dimensi tersembunyi dalam data (von Davier & Yamamoto, 2004). Terakhir, penelitian ini diharapkan dapat membantu peneliti dan praktisi lebih memahami komitmen organisasi, terutama di Indonesia (Schaufeli, 2017; Vandenberg & Lance, 2000).

2. Metode Penelitian

Desain Penelitian

Metode kuantitatif digunakan dalam penelitian ini untuk mengevaluasi fitur psikometrik dari edisi Indonesia *Organizational Commitment Questionnaire* (OCQ). Setelah data dikumpulkan melalui kuesioner langsung, model Rasch digunakan untuk menilai validitas, reliabilitas, dan unidimensionalitas skala; model *Mixture* digunakan untuk menentukan sub kelompok heterogen berdasarkan pola respons peserta. Dengan tujuan menjamin kualitas skala dan memberikan saran untuk penyesuaian jika diperlukan, analisis mencakup statistik item fit, reliabilitas, hierarki kesulitan item, dan fungsi *item diferensial* (DIF).

Partisipan

Jumlah sampel yang ideal untuk analisis Rasch dengan 0.5 logit adalah antara 108 dan 243 sampel, dengan tingkat kepercayaan 99% dan margir of error 5 % (Linacre, 1994). Studi ini melibatkan 253 pekerja dari perusahaan X di Indonesia (53 perempuan, 200 laki-laki), yang berusia antara 18 dan 60 tahun. Para partisipan terdiri dari berbagai lulusan pendidikan mulai dari Sekolah Dasar (SD) sampai Sarjana (S1) yang dibagi dalam empat kelompok yakni SD (127 partisipan), SMP (65 partisipan), SMA (55 partisipan), dan S1 (6 partisipan). Mereka juga dikelompokkan berdasarkan masa kerja dalam 5 kelompok: 1–5 tahun (129 partisipan), 6–10 tahun (64 partisipan), 11–15 tahun (33 partisipan), 16–20 tahun (25 partisipan), dan 21-25 tahun (2 partisipan). Penelitian ini menerapkan metode non random sampling, yakni teknik *convenience* sampling karena tidak memungkinkan untuk menetapkan kerangka sampling dengan populasi yang sesuai.

Instrumen

Instrumen yang digunakan dalam penelitian ini adalah OCQ versi Indonesia yang diadaptasi oleh Alexander (2024). Instrumen ini terdiri dari tiga sub-skala: komitmen afektif/*affective commitment* (AC) terdiri 8 item, komitmen kontinuan/*continuance commitment* (CC) terdiri 8 item, dan komitmen normatif/ *normative commitment* (NC) terdiri 8 item. Keseluruhannya berjumlah 24 item. Setiap item menggunakan skala Likert 5 kategori, yaitu "sangat tidak setuju", "tidak setuju", "netral", "setuju", dan "sangat setuju." Responden diminta memilih jawaban yang sesuai dengan kondisi mereka.

Tabel 1. *Blueprint* Alat Ukur OCQ versi Indonesia

Dimensi	***Koding item***	***item***
Affective Commitmen (AC)	*AC 1*	*Saya sangat senang menghabiskan sisa karir saya di perusahaan ini*
	AC 2	*Saya senang mendiskusikan tentang perusahaan saya dengan orang-orang dari luar perusahaan saya*
	AC 3	*Saya benar-benar merasa bahwa masalah di perusahaan ini merupakan bagian dari masalah saya juga*
	AC 4	*Saya merasa mudah untuk terikat dengan perusahaan lain, sama halnya dengan dengan perusahaan ini*
	AC 5	*Saya tidak merasa menjadi "bagian dari keluarga" di perusahaan ini*
	AC 6	*Saya tidak merasa 'terikat secara emosional" dengan organisasi ini*
	AC 7	*Perusahaan ini mempunyai makna yang besar bagi saya*
	AC 8	*Saya tidak merasakan adanya "rasa memiliki" terhadap perusahaan ini*
Continuance Commitmen (CC)	*CC1*	*Saya tidak takut dengan apa yang akan terjadi pada saya, apabila saya berhenti dari pekerjaan ini tanpa adanya pekerjaan lain*
	CC2	*Sangat sulit bagi saya untuk meninggalkan perusahaan ini, meskipun saya sebenarnya menginginkan itu*
	CC3	*Banyak hal dalam hidup saya yang akan terganggu jika saya memutuskan untuk keluar dari perusahaan ini sekarang*
	CC4	*Tidak akan berdampak apa-apa bagi saya jika meninggalkan perusahaan saya ini sekarang*
	CC5	*Untuk saat ini, bertahan di perusahaan saya adalah sebuah kebutuhan dan keinginan saya*
	CC6	*Saya merasa tidak ada alasan yang cukup sebagai pertimbangan untuk meninggalkan perusahaan ini*
	CC7	*Apabila saya meninggalkan perusahaan ini, dampaknya adalah tidak adanya pilihan bekerja di pekerjaan atau tempat lain*
	CC8	*Saya merasa bahwa keuntungan dan manfaat yang saya dapat dari perusahaan ini, tidak mungkin bisa ditemui ataupun disamai oleh perusahaan lain sehingga saya akan terus bekerja di perusahaan ini*
Normative Commitment (NC)	*NC1*	*Menurut saya, orang-orang di masa sekarang terlalu sering berpindah-pindah perusahan*
	NC2	*Saya tidak mempercayai bahwa seseorang harus selalu setia kepada perusahaannya*
	NC3	*Menurut saya, berpindah dari perusahaan yang satu ke perusahaan lain adalah tindakan yang tidak etis*
	NC4	*Salah satu alasan utama saya terus bekerja di perusahaan ini karena saya percaya bahwa loyalitas itu penting. Maka dari itu saya merasa sudah menjadi kewajiban untuk tetap tinggal di perusahaan ini*
	NC5	*Jika saya mendapat tawaran pekerjaan yang lebih baik di perusahaan lain, saya merasa tidak baik untuk meninggalkan perusahaan saya yang*

		sekarang
	NC6	*Saya diajari untuk percaya pada pentingnya untuk tetap setia pada satu perusahaan*
	NC7	*Akan lebih baik apabila orang-orang bertahan di satu perusahaan untuk sepanjang masa kerja mereka*
	NC8	*Saya merasa tidak masuk akal jika seseorang terlalu setia atau loyal terhadap perusahaannya.*

Analisis Statistik

Perangkat lunak WINSTEP® versi 5.1.0 menggunakan model Rasch untuk menganalisis data. Properti psikometrik yang dianalisis meliputi keunidimensian, person and *item reliability, wright map, misfit order,* dan *bias detection (Differential Item Functioning*/*DIF).* Analisis dilakukan secara terpisah untuk setiap sub-skala karena struktur komitmen multidimensi organisasi. Selain itu, data akan diproses menggunakan Jamovi versi 2.5. Untuk mengolah data ini, Model *Mixture Rasch* akan digunakan. Ini akan menghasilkan hasil analisis tentang jumlah kelas *latent* yang tersedia.

3. Hasil dan Diskusi

Unidimensionalitas

Pengujian karakteristik psikometri alat ukur OCQ diawali dengan analisis unidimensionalitas item pada masing-masing skala komitmen organisasi. Dalam model psikometri unidimensionalitas, menyatakan bahwa sebuah alat pengukuran harus mengukur satu struktur atau dimensi utama (Reise, Bonifay, & Haviland, 2013). Dengan kata lain, setiap komponen instrumen harus menunjukkan karakteristik psikologis yang sama. Keunidimensian setiap subskala didasarkan pada hasil perbedaan yang dijelaskan oleh *raw variance explained by measures* (Smith, 2002). Menurut Holster & Lake (2016), *raw variance explained by measures* lebih dari 40% cukup menunjukkan keunidimensian. Nilai *eigen* dari *unexplained variance* in 1st contrast tidak boleh lebih dari 2,0 karena nilai yang lebih tinggi menunjukkan residual dari dimensi lain (Linacre, 2011).

Studi ini menilai keunidimensian pada tiga subskala komitmen: komitmen afektif, komitmen kontinuan, dan komitmen normatif. Sub-skala komitmen afektif menunjukkan keunidimensian yang cukup dengan *raw variance explained by measures* sebesar 39%, sedikit di bawah ambang batas ideal 40%, dan *unexplained variance in 1st contrast* sebesar 3,3. Item dengan loading tinggi pada residual, seperti item 5,6, dan 8, mengindikasikan potensi dimensi sekunder. Pada sub-skala komitmen kontinuan, *raw variance explained by measures* sebesar 37,8% dan *unexplained variance in 1st contrast* sebesar 2,5 menunjukkan keunidimensian yang kurang kuat. Potensi multidimensionalitas terlihat pada item 1 dan 4, yang memiliki *loading* tinggi pada residual. Sub-skala komitmen normatif memiliki keunidimensian yang lemah dengan *raw variance explained by measures* sebesar 32,8% dan *unexplained variance in 1st contrast* sebesar 2,7. Potensi multidimensionalitas teridentifikasi pada item 4,5,6 dan 7. Hasil penelitian menunjukkan bahwa keunidimensian dari ketiga sub-skala komitmen berbeda-beda, dengan kecenderungan yang lebih rendah. Untuk meningkatkan validitas instrumen, item yang menunjukkan potensi dimensi sekunder harus dianalisis lebih lanjut, atau item harus dihapus untuk meningkatkan keunidimensian instrumen.

Tabel 2. Unidimensionalitas OCQ versi Indonesia setelah menghapus item

Dimensi	*Row Varian Explained by Measure*	*First Contrast*	*Deleted item > 0,5*	*Conclusion*
***Affective Commitmen* (AC)**	46%	1.9	CA5, CA6, CA8	unidimensi
***Continuance Commitmen* (CC)**	44%	1.6	CC1, CC4	unidimensi
Normative Commitment (NC)	45%	1.9	NC4, NC5, NC6, NC7	unidimensi

Peneliti meninjau kembali data baru tersebut setelah menghilangkan elemen-elemen yang dapat menyebabkan dimensi sekunder. Hasil analisisnya menunjukkan ketiga nilai sub-skala untuk *raw variance explained by measures* memperoleh hasil yang lebih tinggi dan *unexplained variance in 1st contrast* mengarah pada penurunan nilainya. Pada dimensi AC, *row variance explained by measure* mencapai 46%, sementara *first contrast sebesar* 1,9, yang menunjukkan bahwa dimensi ini dapat dikategorikan unidimensi (Linacre, 2011). Untuk dimensi CC,

row variance explained by measure sebesar 44% dengan *first contrast* 1,6, yang juga mendukung unidimensionalitas. Dimensi NC *memiliki row variance explained by measure* sebesar 45% dengan *first contrast* 1,9, yang menunjukkan hasil serupa.

Dengan *row variance explained by measure* di atas ambang batas minimum 40% (Linacre, 2011) dan *first contrast di bawah* 2,0 (Reckase, 1979), hasil analisis menunjukkan bahwa ketiga dimensi OCQ versi Indonesia menunjukkan bukti unidimensionalitas yang cukup baik. Hal ini menunjukkan bahwa struktur dapat diwakili secara konsisten oleh masing-masing dimensi tanpa mengganggu dimensi sekunder secara signifikan. Namun, penyesuaian yang diperlukan dapat dilakukan untuk meningkatkan validitas dan keakuratan instrumen, yang dapat menghasilkan hasil yang lebih baik (Boone et al., 2014). Ini terutama berlaku untuk item yang dekat dengan ambang batas analisis residual.

Reliablitas

Uji reliabilitas, yang didasarkan pada teori respons butir (*item response theory*), menilai konsistensi internal instrumen pengukuran. Model *Rasch* menilai hubungan antara kemampuan peserta (kemampuan individu) dan tingkat kesulitan butir soal (kesulitan butir). Metode ini memungkinkan analisis yang lebih mendalam dibandingkan dengan metode klasik seperti *Cronbach's Alpha* (Aziz et al., 2017). Reliabilitas butir (reliabilitas item), yang menunjukkan seberapa konsisten tingkat kesulitan butir di seluruh peserta, dan reliabilitas individu (reliabilitas individu), yang mengukur konsistensi skor peserta berdasarkan kemampuannya (Bond & Fox, 2015). Reliabilitas ini diukur pada skala 0 hingga 1; nilai yang lebih tinggi dari 0,70 menunjukkan instrumen yang cukup reliabel, dan nilai yang lebih tinggi dari 1,00 menunjukkan tingkat reliabilitas yang sangat baik (Linacre, 2002).

Tabel 3. Reliabilitas OCQ versi Indonesia

Dimensi	*Person Reliability*	*Person Separation*	*Item Reliability*	*Item Separation*	*Alpha Cronbach*
Affective Commitmen (AC)	0,58	1,18	0,98	7,84	0,59
Continuance Commitmen (CC)	0,66	1,38	0,97	6,18	0,69
Normative Commitment (NC)	0,59	1,20	0,97	5,68	0,55

Dalam tabel *person reliability* ketiga dimensi berkisar antara 0,58 dan 0,66. Ini berada pada tingkat rendah dan sedang. Hal ini menunjukkan bahwa konsistensi respons individu terhadap *item-item* dalam instrumen *Organizational Commitment Questionnaire* (OCQ) masih kurang optimal. Khususnya, dimensi *affective commitment* (0,58) dan *normative commitment* (0,59) menunjukkan bahwa instrumen tidak dapat secara konsisten mengukur tingkat komitmen individu di populasi Indonesia. Menurut pendapat Bond & Fox (2015), nilai reliabilitas individu di bawah 0,7 menunjukkan adanya adanya heterogenitas individu, yang berarti bahwa subjek penelitian memiliki kemampuan atau kecenderungan yang sangat beragam. Selain itu, hal ini mungkin menunjukkan bahwa instrumen tidak mampu menangkap variasi individu yang lebih luas. Kesimpulan ini juga didukung oleh nilai *Cronbach Alpha*, yang berkisar antara 0,55 dan 0,69. Hal ini menunjukkan bahwa semakin banyak *item* dalam instrumen (seperti dalam dataset kedua dengan enam *item*), semakin reliabel pada tingkat individu karena lebih banyak data yang mendukung pengukuran yang akurat. Oleh karena itu, untuk meningkatkan sensitivitas alat ukur terhadap perbedaan individual, disarankan untuk melakukan revisi atau menambah *item.*

Sebaliknya, reliabilitas item pada ketiga dimensi sangat tinggi, mencapai 0,97 hingga 0,98, menunjukkan bahwa item-item instrumen sangat konsisten. Nilai ini menunjukkan bahwa item-item tersebut memiliki kemampuan untuk mengukur tingkat komitmen organisasi secara konstan tanpa mengalami perubahan yang signifikan. Sebagaimana diungkapkan oleh Linacre (2002), kualitas dan struktur item yang sangat baik mencerminkan tingginya reliabilitasnya. Namun, tingginya kepercayaan item ini tidak cukup untuk menjamin efektivitas instrumen secara keseluruhan, karena rendahnya kepercayaan individu dapat memengaruhi validitas data yang dihasilkan. Akibatnya, untuk meningkatkan kinerja instrumen secara keseluruhan, diperlukan upaya tambahan untuk meningkatkan aspek reliabilitas individu, meskipun item menunjukkan konsistensi tinggi.

Rating Scale Item

Analisis *rating scale item* dengan menggunakan Rasch Model merupakan pendekatan yang dirancang untuk mengevaluasi instrumen yang menggunakan skala bertingkat, seperti skala Likert. Konsep utama dari model ini adalah mengevaluasi seberapa baik setiap kategori pada skala *rating* berfungsi dalam membedakan respons peserta sesuai dengan kemampuan atau atribut yang diukur (Bond & Fox, 2015). Dalam analisis ini, kategori pada skala

harus menunjukkan hierarki yang logis, dengan jarak yang memadai antar kategori untuk menjamin validitas dan reliabilitas pengukuran. *Rasch* Model juga mengevaluasi parameter penting seperti *thresholds* (batas kategori), yang menunjukkan transisi dari satu kategori ke kategori berikutnya, dan item *fit indices* untuk mengetahui apakah item mendukung model atau tidak (*Infit dan Outfit Mean Square* dengan nilai ideal sekitar 1,0).

Tabel 4. Rating Scale Item OCQ versi Indonesia

Dimensi	*Rating Scale*	*f*	*%*	*Average Measure*	*Infit MNSQ*	*Outfit MNSQ*	*Treshold*
Affective Commitmen (AC)	sangat tidak setuju (1)	29	2%	-0.92	1.21	1.26	NONE
	tidak setuju (2)	152	12%	-0.34	1.18	1.19	-2.48
	netral (3)	439	35%	0.22	0.99	0.99	-1.15
	setuju (4)	466	37%	1.33	0.85	0.88	0.76
	sangat setuju (5)	179	14%	2.57	0.92	0.93	2.87
Continuance Commitmen (CC)	sangat tidak setuju (1)	32	2	-0.74	1.18	1.31	NONE
	tidak setuju (2)	140	9	-0.13	1.39	1.53	-2.16
	netral (3)	464	31	0.15	0.76	0.73	-1.25
	setuju (4)	649	43	1.37	0.86	0.93	0.48
	sangat setuju (5)	233	15	2.55	0.95	0.96	2.93
Normative Commitment (NC)	sangat tidak setuju (1)	60	6	-1.69	0.94	0.97	NONE
	tidak setuju (2)	120	12	-0.81	1.3	1.44	-2.02
	netral (3)	399	39	-0.39	0.75	0.72	-1.84
	setuju (4)	369	36	1.01	0.84	0.85	0.35
	sangat setuju (5)	64	6	1.95	1.51	1.23	3.51

Hasil analisis dalam tabel menunjukkan tiga dimensi komitmen organisasi: *Affective Commitment* (AC), *Continuance Commitment* (CC), dan *Normative Commitment* (NC). Dalam dimensi *Affective Commitment* (AC), mayoritas responden berada dalam kategori "setuju" (37 persen) dan "netral" (35 persen), yang menunjukkan bahwa banyak karyawan memiliki keterikatan emosional yang positif terhadap organisasi. Hal ini terkait dengan gagasan yang dikemukakan oleh Meyer dan Allen (1991), yang menyatakan bahwa keterlibatan emosional adalah elemen penting dari keterlibatan emosional karyawan. Nilai rata-rata pengukuran meningkat secara linear dari -0.92 dalam kategori "sangat tidak setuju" menjadi 2.57 dalam kategori "sangat setuju". Nilai *Infit i Outfit* MNSQ tetap stabil di sekitar 1.0, yang menunjukkan kecocokan data dengan model pengukuran.

Pada dimensi *Continuance Commitment* (CC), mayoritas responden berada dalam kategori "setuju" (sebanyak 43 persen) dan "netral" (sebanyak 31 persen). Ini menunjukkan bahwa karyawan merasa terikat pada perusahaan karena konsekuensi atau kerugian yang mungkin terjadi jika mereka meninggalkan. Ini sejalan dengan penelitian yang dilakukan oleh Wang et al. (2010), yang menemukan bahwa komitmen berkelanjutan seringkali didasarkan pada kalkulasi keuntungan dan kerugian rasional. Nilai *Outfit* MNSQ untuk kategori "tidak setuju" sedikit lebih tinggi (1,53), tetapi data sebagian besar tetap sesuai dengan model.

Sebagai bagian dari dimensi *Normative Commitment* (NC), mayoritas responden berada di kategori "netral" (39 persen) dan "setuju" (36 persen), yang menunjukkan bahwa karyawan merasa memiliki tanggung jawab moral yang moderat terhadap organisasi mereka. Hal ini mendukung teori Wiener (1982) bahwa komitmen normatif

berasal dari norma sosial dan kewajiban moral. Nilai *Infit* dan Outfit MNSQ konsisten, dengan kategori "sangat setuju" memiliki nilai *Outfit* MNSQ sedikit lebih tinggi, dari -1.69 hingga 1.95.

Secara keseluruhan, data ini mendukung validitas internal instrumen OCQ, karena pola distribusi responden menunjukkan tingkat komitmen positif terhadap organisasi di setiap aspeknya. Namun, Linacre (2002) menyarankan bahwa nilai *Outfit* MNSQ yang lebih tinggi pada beberapa kategori mungkin menunjukkan interpretasi yang berbeda dari responden tentang item tertentu. Hal ini menunjukkan bahwa analisis lebih lanjut diperlukan untuk meningkatkan reliabilitas dan akurasi alat ukur.

Item Misfit

Analisis *item fit* dalam Rasch Model bertujuan untuk mengevaluasi sejauh mana setiap butir pada instrumen pengukuran sesuai dengan model Rasch, yang didasarkan pada asumsi unidimensionalitas, yaitu bahwa semua butir mengukur atribut yang sama. Item fit dinilai menggunakan dua indeks utama, yaitu *Infit Mean Square* **(MNSQ)** dan *Outfit Mean Square* **(MNSQ)**. Infit MNSQ mengukur kesesuaian data pada butir yang dekat dengan kemampuan peserta (*on-target items*), sedangkan *Outfit* MNSQ lebih sensitif terhadap data ekstrem, seperti respons peserta dengan kemampuan yang sangat tinggi atau sangat rendah terhadap butir tersebut (Bond & Fox, 2015). Nilai ideal untuk kedua indeks ini adalah mendekati 1,0, dengan toleransi umum berada dalam rentang 0,5 hingga 1,5 untuk pengukuran pendidikan atau psikologi.

Tabel 5. Item Fit OCQ versi Indonesia

Sub-Scale	***Item***	***Measure***	***Infit***		***Outfit***		***Poin Measure Corelation***
			MNSQ	**ZSDT**	**MNSQ**	**ZSDT**	
AC	AC4	1.02	1.22	2.4	1.25	2.8	0.46
	AC7	-1.19	1.24	2.6	1.23	2.4	0.55
	AC1	-0.38	0.96	-0.4	0.96	-0.5	0.63
	AC2	0.46	0.85	-1.8	0.84	-1.9	0.73
	AC3	0.1	0.78	-2.7	0.78	-2.7	0.67
CC	CC2	0.67	1.27	2.8	1.31	3.2	0.48
	CC8	-0.29	1.08	0.9	1.11	1.2	0.64
	CC5	-1.08	1.01	0.1	1.06	0.6	0.54
	CC3	0.26	0.97	-0.3	0.99	-0.1	0.69
	CC7	0.55	0.89	-1.3	0.93	-0.8	0.7
	CC6	-0.1	0.77	-2.7	0.8	-2.3	0.64
CN	CN1	-0.51	1.2	2	1.12	1.3	0.59
	CN3	-0.58	1.07	0.7	1.02	0.2	0.61
	CN8	0.58	1	0	1.06	0.7	0.63
	CN2	0.52	0.79	-2.6	0.81	-2.3	0.7

Dari hasil analisis item fit, sebagian besar item memiliki nilai *Outfit* MNSQ dan *Infit Mean Square* (MNSQ) dalam rentang yang dapat diterima, yaitu 0,5 hingga 1,5. Dengan kata lain, item-item ini dapat diandalkan untuk mengukur komitmen organisasi responden karena sesuai dengan model Rasch (Sumintono & Widhiarso, 2015). Namun, nilai MNSQ di beberapa item mendekati batas atas atau bawah dari rentang yang disarankan. Sebagai contoh, nilai *Infit* MNSQ item AC4 adalah 1,22 dan *Outfit* MNSQ adalah 1,25, yang hampir sama dengan batas atas rentang yang diterima. Nilai ini menunjukkan bahwa item tersebut mungkin tidak berfungsi dengan baik untuk mengukur konstruk yang dimaksud, yang dapat menunjukkan adanya overfitting dan memerlukan perhatian lebih lanjut. Selain itu, analisis korelasi titik ukuran (PMC) menunjukkan bahwa sebagian besar item memiliki nilai korelasi positif yang cukup tinggi, berkisar antara 0,48 dan 0,73, menunjukkan bahwa item-item tersebut memiliki daya diskriminasi yang kuat. Item-item ini dapat membedakan responden berdasarkan tingkat keterlibatan mereka dengan organisasi (Bond & Fox, 2015).

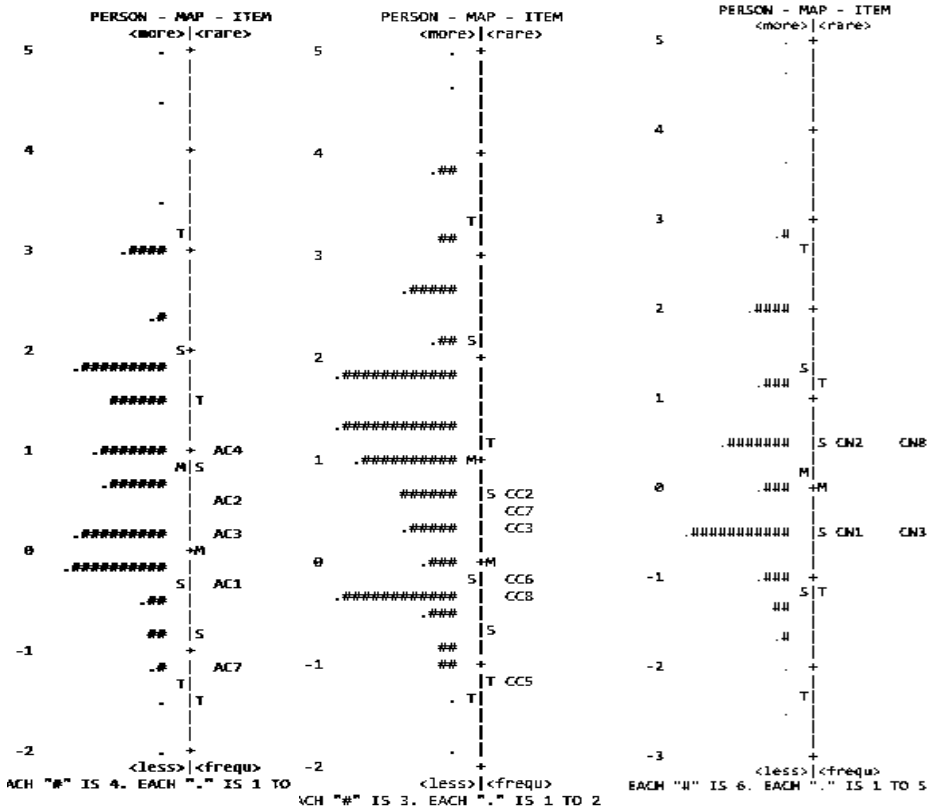
Secara keseluruhan, temuan analisis ini menunjukkan bahwa OCQ versi Indonesia memiliki kualitas item yang baik dan dapat digunakan sebagai alat yang dapat diandalkan untuk mengukur tingkat komitmen organisasi.

Namun, untuk memastikan bahwa setiap item berfungsi secara optimal dalam konteks pengukuran yang diinginkan, disarankan untuk meninjau ulang item-item dengan nilai MNSQ yang mendekati batas atas atau bawah (Linacre, 2011)

Wright Map

Dalam analisis Rasch menurut Linacre (2011), *Wright Map* adalah representasi visual yang menggambarkan distribusi kemampuan individu (*person ability*) dan tingkat kesulitan item (*item difficulty*). Dua sumbu utama terlihat di peta ini: sumbu kiri menunjukkan kemampuan responden, dan sumbu kanan menunjukkan tingkat kesulitan item. *Wright Map* mempermudah evaluasi kesesuaian antara kemampuan populasi responden dengan tingkat kesulitan item, sehingga dapat digunakan untuk menilai validitas dan keandalan instrumen pengukuran. Di sini, simbol "#" dan "." menunjukkan distribusi kemampuan responden, dan kode item menunjukkan tingkat kesulitan masing-masing soal (Bond & Fox, 2015; Boone et al., 2014). Peta ini juga sering digunakan untuk menganalisis kualitas butir soal untuk mengukur kemampuan individu secara objektif (Linacre, 2011).

Gambar 1. *Wright Map* OCQ versi Indonesia



Tingkat kesulitan *item* berkisar dari rendah hingga tinggi, tetapi kemampuan responden sebagian besar terkonsentrasi pada tingkat sedang dari ketiga peta *Wright* yang dianalisis. Peta pertama menunjukkan bahwa *item-item* seperti AC1 dan AC7 lebih mudah diakses oleh responden dengan kemampuan rendah, sementara item paling sulit adalah AC4. Peta kedua menunjukkan bahwa jumlah *item* yang ditambahkan memperluas cakupan tingkat kesulitan dan menunjukkan distribusi kemampuan responden yang lebih berimbang. Pada peta ketiga, jumlah *item* yang dikurangi membuat rentang tingkat kesulitan menjadi lebih sempit, yang menunjukkan bahwa distribusi kemampuan responden Secara umum, sebagian besar orang yang menjawab memiliki kemampuan yang sebanding dengan tingkat kesulitan rata-rata *item*. Hasil ini menunjukkan betapa pentingnya memastikan bahwa *item-item* yang digunakan dalam instrumen mencakup seluruh spektrum kemampuan yang ada (Wu & Adams, 2007).

Hasil interpretasi menunjukkan bahwa tingkat kesulitan *item* harus disesuaikan agar lebih rata dan mencakup semua kemampuan responden. Menciptakan alat yang lebih kompetitif untuk berbagai kelompok kemampuan dapat membantu dengan menambah *item* dengan tingkat kesulitan tinggi dan rendah. Untuk menemukan *item* yang tidak sesuai dengan kemampuan mayoritas responden, analisis fit statistik juga diperlukan. Proses ini

memungkinkan instrumen pengukuran menjadi lebih valid dan reliabel serta lebih representatif dalam menilai populasi yang diuji. Metode ini sejalan dengan prinsip dasar analisis Rasch, yang menekankan betapa pentingnya mengimbangi kemampuan responden dengan kesulitan item untuk membuat instrumen yang berkualitas tinggi (Bond & Fox, 2015).

Fungsi Item Differential

Analisis yang dikenal sebagai Fungsi *Item Differential* (DIF) bertujuan untuk menemukan kemungkinan bias pada item dalam instrumen pengukuran. Bias ini terjadi ketika orang dalam kelompok yang sama memiliki peluang yang berbeda untuk menjawab suatu item dengan benar. Ismail (2011) menyatakan bahwa DIF terjadi ketika peserta tes dari kelompok yang berbeda dengan kemampuan yang sama tidak memiliki peluang yang sama untuk menjawab butir tes dengan benar. Selain itu, Sudaryono (2012) menyatakan bahwa DIF terjadi ketika peserta tes dari kelompok yang berbeda dengan kemampuan yang sama tidak memiliki peluang yang sama untuk menjawab butir tes dengan benar. Bias dapat berasal dari komponen demografis seperti gender atau tingkat pendidikan seseorang. Studi ini menganalisis DIF berdasarkan gender (pekerja dan karyawan) dan pendidikan.karyawan yakni mulai dari kelompok lulusan SD-S1.

Analisis DIF sering menggunakan uji *Mantel-Haenszel* dan *Welch* untuk menemukan bias pada *item* tertentu. Untuk memastikan instrumen pengukuran adil dan valid, analisis ini sangat penting (Zumbo, 1999). Untuk memastikan bahwa setiap *item* mewakili konstruk yang dimaksud secara wajar, diperlukan analisis lanjutan jika ditemukan bahwa ada *item* yang memiliki bias signifikan.

Tabel 6. DIF OCQ versi Indonesia

Sub-Scale	***Item***	***Based on gender***				***Based on education***			
		Mantel-Haenszel		***Welch***		***Mantel-Haenszel***		***Welch***	
		Chi-squ	***Prob.***	***d.f.***	***Prob.***	***Chi-squ***	***Prob.***	***d.f.***	***Prob***
AC	AC1	6.5147	0.0107	108	-3.11	0.4856	0.4856	150	1.03
	AC2	2.2454	0.134	112	2.29	2.4752	0.8068	125	0.6
	AC3	0.7912	0.3737	111	0.7	0.9768	0.8775	126	-0.7
	AC4	4.8287	0.028	112	3.17	1.0642	0.3023	150	0.75
	AC7	5.2378	0.0221	105	-3.6	0.7165	0.3973	150	-1.19
CC	CC2	0.0251	0.874	112	1.43	1.3292	0.2489	149	0.29
	CC3	6.5434	0.0105	114	3.29	5.6345	0.0176	146	2.25
	CC5	8.9899	0.0027	103	-4.67	3.6237	0.0224	126	-1.84
	CC6	0.0197	0.8883	112	0.23	0.317	0.5734	150	-0.21
	CC7	1.5789	0.2089	112	1.09	0.19	0.6629	150	-0.49
	CC8	1.0735	0.3002	109	-2.07	0.0006	0.98	149	0
CN	CN1	0.0453	0.8314	113	0.77	3.4126	0.729	124	-0.54
	CN2	0.8706	0.3508	110	-0.67	4.1255	0.8713	123	0.26
	CN3	0.7496	0.3866	111	-0.76	6.4451	0.9618	129	-0.32
	CN8	2.1453	0.143	112	0.66	1.6609	0.6614	123	0.47

Tiga item pada sub-skala *Affective Commitment* (AC) menunjukkan bias signifikan *gender*, dengan nilai *Mantel-Haenszel* $p < 0,05$ dan nilai Welch t yang besar (AC1: -3,11; AC4: 3,17; AC7: -3,60). Bias ini menunjukkan bahwa salah satu gender memiliki keuntungan atau kerugian yang tidak adil. Misalnya, nilai negatif pada AC1 dan AC7 menunjukkan bahwa orang dari salah satu gender menghadapi kesulitan yang lebih besar dalam menjawab item tersebut. Sebaliknya, hasil menunjukkan bahwa item AC2 dan AC3 tidak menunjukkan bias signifikan (p lebih besar dari 0,05; nilai *Welch* rendah), menunjukkan bahwa item ini netral terhadap gender. Item CC3 dan CC5 ditemukan memiliki bias signifikan pada sub-skala *Continuance Commitment* (CC) ($p < 0,05$; nilai *Welch* masing-masing 3,29 dan -4,67), menunjukkan kemungkinan masalah pada instrumen. Namun, item CC2, CC6, CC7, dan CC8 tidak menunjukkan bias signifikan. Adapun sub-skala *Normative Commitment* (CN), semua item dinyatakan

netral terhadap gender, dengan nilai Mantel-Haenszel p > 0,05 dan nilai *Welch* rendah, sehingga bebas dari *bias gender.*

Dalam analisis berdasarkan tingkat pendidikan, tiga item pada sub-skala *affective commitment* (AC)—AC1, AC4, dan AC7—juga menunjukkan bias signifikan, dengan nilai *Mantel-Haenszel* p < 0,05 dan nilai *Welch* positif (misalnya, AC1 1,03, AC4 0,75, dan AC7 -1,19). Ini menunjukkan bahwa item tersebut memberikan keuntungan atau kerugian yang tidak adil bagi kelompok dengan latar belakang pendidikan tertentu. Sebaliknya, item AC2 dan Item CC3 dan CC5 juga ditemukan memiliki bias signifikan pada sub-skala *continuance commitment* (CC) (nilai Mantel-Haenszel p < 0,05; nilai *Welch* masing-masing 2,25 dan -1,84). Ini menunjukkan bahwa kelompok dengan tingkat pendidikan tertentu mungkin memiliki respons yang berbeda terhadap masalah ini meskipun mereka memiliki tingkat kemampuan yang sama. Item CC2, CC6, CC7, dan CC8, di sisi lain, tidak menunjukkan bias yang signifikan. Sub-skala Komitmen Normatif (CN) kembali menunjukkan hasil netral, artinya tidak ada item dengan nilai *Mantel-Haenszel* p signifikan, yang berarti tidak ada bias pendidikan.

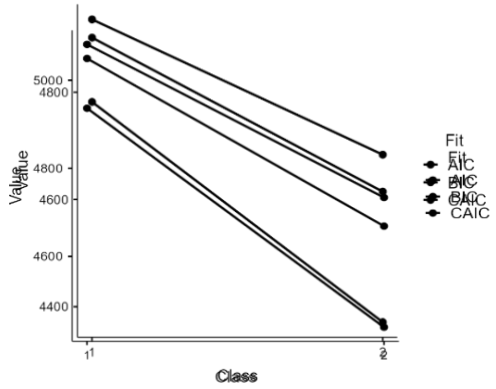
Hasil analisis menunjukkan bahwa beberapa item pada sub-skala komitmen afektif dan komitmen kontinuitas menunjukkan bias yang signifikan berdasarkan gender dan tingkat pendidikan. Ini menunjukkan bahwa *item-item* tersebut harus dianalisis lebih lanjut untuk memastikan relevansi bias terhadap konstruk yang diukur. Apabila bias tersebut dianggap tidak penting, item dapat diubah atau dihapus untuk meningkatkan validitas instrumen. Sebaliknya, sub-skala *normative commitment* lebih dapat diandalkan untuk digunakan pada populasi yang beragam karena telah terbukti tidak bias. Literatur seperti ini menunjukkan bahwa pengukuran psikometrik memerlukan keadilan.

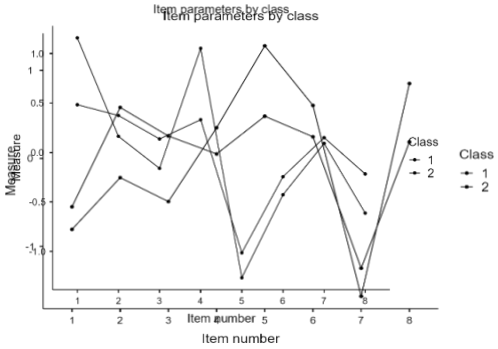
Mixture Rasch Model

Analisis menggunakan Mixture Rasch Model berguna untuk mengidentifikasi heterogenitas dalam data responden dengan mengelompokkan mereka ke dalam beberapa subpopulasi (latent classes) berdasarkan pola respons mereka. Analisis ini membantu dalam memperbaiki validitas pengukuran dengan mengidentifikasi kelompok responden yang mungkin memiliki interpretasi berbeda terhadap instrumen yang sama (von Davier & Rost, 2007). Kriteria evaluasi dalam *Mixture Rasch Model* digunakan untuk menentukan model terbaik dengan mempertimbangkan keseimbangan antara kesesuaian data dan kompleksitas model. Beberapa kriteria utama meliputi *Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), dan Consistent AIC* (CAIC), di mana nilai yang lebih kecil pada ketiga kriteria ini menunjukkan model yang lebih baik dan parsimonious (Burnham & Anderson, 2002). Selain itu, Log-likelihood digunakan untuk mengevaluasi seberapa baik model menjelaskan data, dengan nilai lebih besar (lebih kecil negatifnya) mengindikasikan kesesuaian yang lebih baik. Entropi juga diperhatikan untuk mengukur kejelasan klasifikasi responden, dengan nilai mendekati 1 menunjukkan klasifikasi yang jelas (Rost, 1990). Terakhir, *Likelihood Ratio Test (LRT*) digunakan untuk membandingkan model dengan jumlah kelas yang berbeda, dengan p-value < 0,05 menunjukkan bahwa penambahan kelas memberikan perbaikan signifikan pada model. Dengan mempertimbangkan semua kriteria ini, model terbaik adalah yang memberikan keseimbangan optimal antara kesesuaian data dan kompleksitas struktur.

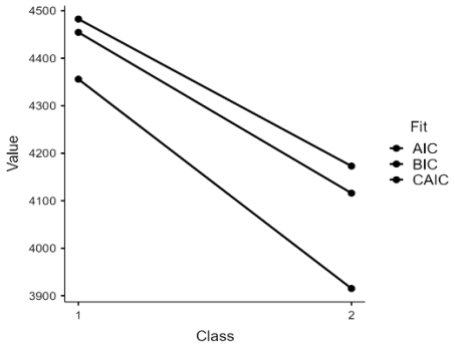
***Ta*bel 7. *Latent Classes* OCQ versi Indonesia**

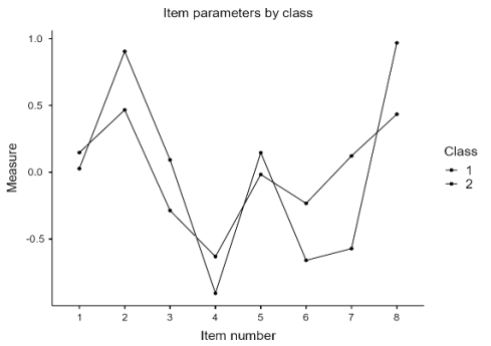
Sub-Scale	*Class*	*AIC*	*BIC*	*CAIC*	*Log-likelihood*
	1	*4936*	*5049*	*5081*	*-2436*
CA	*2*	*4439*	*4669*	*4734*	*-2155*





CC	*1*	*4782*	*4902*	*4936*	*-2357*
	2	*4371*	*4615*	*4684*	*-2117*
CN	*1*	*4356*	*4454*	*4482*	*-2150*
	2	*3915*	*4116*	*4173*	*-1901*





Hasil analisis *latent classes* pada sub-skala *Organizational Commitment Questionnaire* (OCQ**)** versi Indonesia menunjukkan bahwa model dengan dua kelas lebih sesuai untuk menggambarkan data dibandingkan dengan model satu kelas. Hal ini tercermin dari nilai AIC, BIC, dan CAIC yang lebih rendah pada model dua kelas untuk semua sub-skala, yaitu *affective commitment* (CA), *continuance commitment* (CC), dan *normative commitment (*CN). Sebagai contoh, pada sub-skala CA, nilai AIC untuk model satu kelas adalah 4936, sementara untuk model dua kelas turun menjadi 4439. Tren serupa juga terlihat pada sub-skala CC (4782 ke 4371 untuk AIC) dan CN (4356 ke 3915 untuk AIC). Penurunan nilai ini menunjukkan bahwa model dua kelas memberikan representasi yang lebih baik terhadap variasi data yang ada.

Selain itu, nilai log-likelihood yang lebih tinggi pada model dua kelas mendukung kesimpulan ini, dengan nilai log-likelihood untuk CA, CC, dan CN masing-masing meningkat dari -2436 menjadi -2155, dari -2357 menjadi -2117, dan dari -2150 menjadi -1901. Kecenderungan ini mengindikasikan bahwa model dua kelas memiliki kemampuan lebih baik untuk menangkap pola yang kompleks dalam data, sehingga memberikan gambaran yang lebih akurat tentang perbedaan dalam komitmen organisasi di antara individu. Oleh karena itu, model dua kelas dapat dianggap lebih unggul dan lebih tepat digunakan untuk memahami struktur laten pada OCQ versi Indonesia.

Implikasi dari temuan ini menunjukkan bahwa populasi yang diteliti tidak sepenuhnya homogen, tetapi juga tidak terlalu kompleks secara struktural. Dalam arti lain, dua pola besar dalam cara individu memahami dan merespons komitmen organisasi sudah cukup untuk menjelaskan variasi yang signifikan dalam data. Ini dapat mengindikasikan bahwa terdapat dua jenis utama profil psikologis atau demografis dalam kelompok partisipan: satu kelompok mungkin memiliki pemahaman komitmen yang lebih kuat dan konsisten, sementara kelompok lainnya mungkin memiliki respons yang lebih lemah atau variatif. Sedangkan implikasi terhadap populasi yang diteliti adalah perlunya pengakuan bahwa intervensi organisasi atau kebijakan manajerial tidak dapat diberlakukan secara seragam. Misalnya, kelompok dengan komitmen rendah mungkin membutuhkan strategi peningkatan keterikatan yang berbeda dari kelompok yang sudah menunjukkan komitmen tinggi. Oleh karena itu, hasil ini mendukung pendekatan segmentasi dalam desain kebijakan organisasi dan pelatihan SDM.

Sedangkan untuk konteks yang lebih luas, yakni komitmen organisasi di Indonesia, hasil ini mencerminkan dinamika sosial-budaya dan profesional yang kompleks. Adanya dua kelompok besar dalam cara responden memahami komitmen menunjukkan bahwa pengukuran dan pengelolaan komitmen organisasi tidak bisa mengasumsikan keseragaman lintas populasi pekerja. Konteks Indonesia yang beragam dari segi latar pendidikan, budaya kerja, dan struktur organisasi memperkuat pentingnya pendekatan yang lebih tailored dalam membina dan mengukur komitmen. Ini juga menjadi dasar kuat bagi pengembangan lebih lanjut terhadap alat ukur OCQ versi Indonesia agar lebih inklusif dan sensitif terhadap perbedaan dalam populasi nasional.

4. Kesimpulan

Penelitian mengenai *Organizational Commitment Questionnaire* (OCQ) versi Indonesia menunjukkan bahwa instrumen ini memiliki kualitas psikometrik yang cukup baik untuk mengukur komitmen organisasi di konteks budaya Indonesia. Analisis unidimensionalitas mengungkapkan bahwa ketiga sub-skala (afektif, kontinuan, dan normatif)

menunjukkan bukti unidimensionalitas yang memadai setelah dilakukan penyesuaian terhadap beberapa item. Awalnya, beberapa item menunjukkan potensi multidimensionalitas yang dapat mengurangi validitas instrumen. Namun, melalui analisis ulang dengan eliminasi item bermasalah, validitas instrumen meningkat, ditunjukkan dengan nilai variance explained dan contrast residual yang sesuai dengan kriteria yang telah ditetapkan.

Hasil uji reliabilitas menunjukkan bahwa reliabilitas butir berada pada tingkat sangat tinggi, mengindikasikan bahwa item-item dalam OCQ versi Indonesia konsisten dalam mengukur komitmen organisasi. Namun, reliabilitas individu cenderung rendah hingga sedang, menunjukkan bahwa instrumen ini kurang sensitif dalam menangkap variasi respon di antara individu. Selain itu, nilai *Cronbach's Alpha* yang rendah pada beberapa dimensi mendukung temuan ini. Hal ini menunjukkan perlunya penyesuaian lebih lanjut pada struktur instrumen, baik melalui pengembangan item baru maupun dengan revisi pada item yang ada, untuk meningkatkan kemampuan alat dalam mengukur variasi tingkat komitmen individu secara lebih akurat.

Dari segi validitas, beberapa item pada sub-skala afektif dan kontinuan menunjukkan bias signifikan berdasarkan gender dan tingkat pendidikan. Bias ini mengindikasikan adanya perbedaan persepsi terhadap item tertentu yang dapat mengurangi keadilan pengukuran. Sebaliknya, sub-skala normatif terbukti netral terhadap kedua variabel ini, menjadikannya lebih andal untuk digunakan pada populasi yang beragam. Selain itu, analisis menggunakan *Mixture Rasch Model* mengidentifikasi adanya heterogenitas dalam pola respons responden. Model dengan dua kelas latent menunjukkan hasil yang lebih baik dibandingkan model dengan satu kelas, mengindikasikan bahwa perbedaan latar belakang responden dapat memengaruhi cara mereka memahami dan merespons item.

Penelitian ini memiliki beberapa keterbatasan, antara lain cakupan sampel yang terbatas pada satu jenis organisasi dan metode sampling aksidental yang berpotensi menimbulkan bias seleksi. Hal ini dapat membatasi generalisasi hasil penelitian ke populasi yang lebih luas. Oleh karena itu, penelitian lanjutan disarankan untuk memperluas sampel ke berbagai jenis organisasi dan kelompok demografis guna meningkatkan validitas eksternal. Selain itu, pengembangan item baru yang lebih sensitif terhadap variasi individu dan bebas dari bias demografis dapat meningkatkan kualitas instrumen secara keseluruhan. Penelitian lebih lanjut juga dapat mengeksplorasi penerapan Mixture Rasch Model dalam konteks budaya lain untuk mengidentifikasi pola respons yang berbeda dan memperkaya pemahaman tentang komitmen organisasi.

5. Referensi

Akaike, H. (1987). Factor analysis and AIC. Psychometrika, 52(3), 317–332. https://doi.org/10.1007/BF02294359

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance, and normative commitment to the organization. Journal of Occupational Psychology, 63(1), 1–18. https://doi.org/10.1111/j.2044-8325.1990.tb00506.x

Astuti, R., Nugroho, A., & Setiawan, D. (2020). Adaptasi budaya pada skala komitmen organisasi: Pendekatan CFA. Jurnal Psikologi Indonesia, 36(2), 215–230. https://doi.org/10.22146/jpsi.2020.12345

Aziz, A., Noor, N., & Rosli, R. (2017). A comparative study of reliability measures: Cronbach's Alpha versus composite reliability. Journal of Advanced Research in Social Sciences, 11(4), 45–53.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. Spine, 25(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model: Fundamental measurement in the human sciences (3rd ed.). New York, NY: Routledge.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Rasch analysis in the human sciences. New York, NY: Springer.

Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York, NY:Guilford Press.

Burnham, K. P., & Anderson, D. R. (2002). Model selection and multimodel inference: A practical information-theoretic approach (2nd ed.). New York, NY: Springer.

Byrne, B. M. (2001). Structural equation modeling with AMOS: Basic concepts, applications, and programming. Mahwah, NJ: Lawrence Erlbaum Associates.

Chen, Z. X., & Francesco, A. M. (2000). Employee demography, organizational commitment, and turnover intentions in China: Do cultural differences matter? Human Relations, 53(6), 869–887. https://doi.org/10.1177/0018726700536005

Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.). Boston,MA: Pearson.

Elsmore, B. (2008). Organizational psychology: A guide to better performance. London: Routledge.

Finch, W. H., & French, B. F. (2015). Latent variable modeling with R. New York, NY: Routledge.

Hair, J. F., Hult, G. T. M., & Ringle, C. M. (2017). A primer on partial least squares structural equation modeling (PLS-SEM). Thousand Oaks, CA: SAGE Publications.

Holster, T., & Lake, J. (2016). Evaluating psychometric properties of measurement tools using Rasch analysis. Journal of Applied Psychology, 10(3), 25–35.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Ismail, N. (2011). Differential item functioning: A comparison of methods. International Journal of Educational Research, 56(2), 105–115. https://doi.org/10.1016/j.ijer.2011.04.007

Jonathan, R. (2020). Adaptation and validation of the organizational commitment questionnaire in Tanzania. African Journal of Organizational Studies, 12(3), 45–63.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20(1), 141–151. https://doi.org/10.1177/001316446002000116

Kanning, U. P., & Hill, A. (2013). Cross-cultural adaptations of the organizational commitment questionnaire: Challenges and outcomes. International Journal of Cross-Cultural Management, 13(1), 53–71.

Kaur, R., & Sharma, S. (2015). Validation of the organizational commitment questionnaire in Indian context. Asian Journal of Management Research, 6(1), 102–109.

Linacre, J. M. (1994). Sample size and item calibration stability. Rasch Measurement Transactions, 7*(4), 328. Retrieved from https://www.rasch.org/rmt/rmt74m.htm*

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions, 16*(2), 878. Retrieved from https://www.rasch.org/rmt/rmt162f.htm*

Linacre, J. M. (2009). A user's guide to Winsteps Ministep: Rasch-model computer programs (Version 3.68). Beaverton, OR: Winsteps.com.

Linacre, J. M. (2011). Investigating dimensionality using Rasch principal component analysis of residuals. Journal of Applied Measurement, 12(3), 321–338.

Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. Psychological Bulletin, 108(2), 171–194. https://doi.org/10.1037/0033-2909.108.2.171

Meyer, J. P., & Allen, N. J. (1991). A three-component conceptualization of organizational commitment. Human Resource Management Review, 1(1), 61–89. https://doi.org/10.1016/1053-4822(91)90011-Z

Meyer, J. P., Allen, N. J., & Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. Journal of Applied Psychology, 78(4), 538–551. https://doi.org/10.1037/0021-9010.78.4.538

Meyer, J. P., Stanley, D. J., Herscovitch, L., & Topolnytsky, L. (2002). Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. Journal of Vocational Behavior, 61(1), 20–52. https://doi.org/10.1006/jvbe.2001.1842

Morgeson, F. P., Reider, M. H., & Campion, M. A. (2005). Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge. Personnel Psychology, 58*(3), 583–611. https://doi.org/10.1111/j.1744-6570.2005.655.x*

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. Journal of Vocational Behavior, 14(2), 224–247. https://doi.org/10.1016/0001-8791(79)90072-1

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. Journal of Educational Statistics, 4(3), 207–230. https://doi.org/10.2307/1164671

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. Journal of Personality Assessment, 95*(2), 129–140. https://doi.org/10.1080/00223891.2012.725437*

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. New York, NY: Springer.

Schaufeli, W. B. (2017). Applying the Job Demands-Resources model: A 'how to' guide to measuring and tackling work engagement and burnout. Organizational Dynamics, 46(2), 120–132. https://doi.org/10.1016/j.orgdyn.2017.04.008

Siregar, S. (2018). Statistik parametrik untuk penelitian kuantitatif: Dilengkapi perhitungan manual dan aplikasi SPSS versi 17. Jakarta: Bumi Aksara.

Smith, R. M. (2002). Understanding Rasch measurement: Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. Journal of Applied Measurement, 3(2), 205–231.

Sudaryono. (2012). Metodologi penelitian: Proses dan penerapan. Bandung: PT Refika Aditama.

Sulistiawan, A., Sari, E. Y. D., & Situmorang, N. Z. (2021). Validitas dan reliabilitas konstruk komitmen organisasi dengan pendekatan Confirmatory Factor Analysis (CFA). Psikostudia, Vol 10, no 1. Hal. 61-68.

Sumintono, B., & Widhiarso, W. (2015). Aplikasi model Rasch untuk penelitian ilmu-ilmu sosial. Bandung: Trim Komunikata Publishing House.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3(1), 4–70. https://doi.org/10.1177/109442810031002

von Davier, M., & Yamamoto, K. (2004). Mixture models in item response theory. Annals of the Institute of Statistical Mathematics, 56(1), 165–183. https://doi.org/10.1007/BF02530318

von Davier, M., & Rost, J. (2007). Mixture distribution Rasch models. In Handbook of Modern Item Response Theory (pp. 251–266). New York, NY: Springer

Wang, Y. D., & Hsieh, H. H. (2010). Commitment to employees: Does it help or hinder organizational performance? Human Resource Management Journal, 20*(3), 305–319. https://doi.org/10.1111/j.1748-8583.2010.00135.x*

Wiener, Y. (1982). Commitment in organizations: A normative view. Academy of Management Review, 7(3), 418–428. https://doi.org/10.5465/amr.1982.4285349

Wu, M. L., & Adams, R. J. (2007). Applying the Rasch model to psychometric assessment. Mahwah, NJ: Lawrence Erlbaum Associates.

Zumbo, B. D. (1999). A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense.