

Volume 14 No. 1, April 2025



Jurnal Penelitian & Pengukuran **Psikologi**

JPPP

Volume
14

Nomor
01

Hal.
1 - 96

ISSN
2337-4845

Diterbitkan Oleh
Fakultas Psikologi
Universitas Negeri Jakarta



**SUSUNAN DEWAN REDAKSI PERIODE 2025
JURNAL PENELITIAN PENGUKURAN PSIKOLOGI**

P-ISSN. 2337-4845

E-ISSN. 2620-7486

Penanggung jawab

Dekan Fakultas Psikologi, Universitas Negeri Jakarta

Editor in chief

Santi Yudhistira, M.Psi., Psikolog

Managing Editor

Karel Karsten Himawan, Ph.D., Psikolog

Gita Irianda Rizkyani Medellu, M.Psi., Psikolog

Rahmadianty Gazadinda, M.Sc.

Production Editor

Vinna Ramadhany Sy, M.Psi., Psikolog

Copy Editor

Adhissa Qonita, M.Psi., Psikolog

Dewi Fransiska Simajuntak M.Si

Kahfi Hizbullah, M.Psi, Psikolog

Reference Editor

Hermeilia Megawati, M.A.

Content Editor

Anggi Mayangsari, M.Psi., Psikolog

Layout Editor

Novitasari R. Damanik, M.Si.

SEKRETARIAT

Fakultas Psikologi

Jalan Rawamangun Muka

Kampus A Universitas Negeri Jakarta Gedung Dewi Sartika Lt. 7

Jakarta Timur 13220; Email: jppp@unj.ac.id



Jurnal Penelitian dan Pengukuran Psikologi

Jurnal yang terbit dua kali dalam satu tahun, pada bulan April dan Oktober, berisi tentang kajian dan hasil penelitian dan pengukuran di bidang psikologi.

Ketua Penyunting

Santi Yudhistira

Penyunting Pelaksana

Karel Karsten Himawan
Gita Irianda Rizkyani Medellu
Rahmadianty Gazadinda

Alamat Penyunting dan Tata Usaha: Program Studi Psikologi, Fakultas Psikologi, Universitas Negeri Jakarta, Jl. Halimun No.2 Kecamatan Setia Budi, Jakarta Selatan. Telp. (021) 4755115/29266297 Fax (021) 4897535. Email: santiyudhistira@unj.ac.id atau jppp@unj.ac.id

Jurnal Penelitian dan Pengukuran Psikologi diterbitkan oleh Fakultas Psikologi Universitas Negeri Jakarta. Terbit pertama kali pada bulan Oktober 2012.

Penyunting menerima tulisan yang belum pernah diterbitkan oleh media cetak lain. Naskah diketik dengan spasi 1 cm pada kertas ukuran A4 dengan panjang tulisan maksimal 10 Halaman. (Informasi detail dapat dilihat pada halaman akhir jurnal)

Daftar Isi

Nama	Judul Artikel	Halaman
<i>Adhissa Qonita</i>	Menilik Sikap Warga Sekolah Terhadap Siswa Berkebutuhan Khusus: Studi di Sekolah Menengah Inklusi Jakarta	1 – 7
<i>Hilda Egan, Iriani Indri Hapsari, Lussy Dwiutami Wahyuni, Gumgum Gumelar</i>	Analisis Psikometri Instrumen <i>Psychological Well-Being</i> pada Dokter: Adaptasi Skala Pengukuran	8 – 16
<i>Rizkiani Amalia & Devie Yudianto</i>	Adaptasi Instrumen <i>Depression Literacy (D-Lit) Questionnaire</i> Versi Bahasa Indonesia dengan Menggunakan Metode <i>Rasch Model</i>	17 – 29
<i>Muh Haritsah Hizbullah, Juliana Nur Afifah, Ulfah Sayidah Mutamimah, Ahmad Yasir Ibrahim, Alnaznabila Aghnii Hafnieva Imanda, Muhamad Arif Saefudin, Helli Ihsan</i>	Analisis Psikometri <i>Burnout Assessment Tool (BAT) For Student</i> Versi Bahasa Indonesia	30 – 38
<i>Fransiscus Widodo Setya Budi, Alexander Bhirowo Putra Kurniawan, Ananta Yudianto</i>	Analisis Psikometrika Properti dan Latent Sub Kelas dengan <i>Mixture Rasch Model</i> pada Skala <i>Organizational Commitment Questionnaire (OCQ)</i> Versi Indonesia	39 – 53
<i>Dyah Ayu Permatasari Gunawan, Lina Natalya, Nadia Sutanto</i>	Terapi Fobia Spesifik Sesi Tunggal dengan <i>The Heart Technique</i>	54 – 67
<i>Audita Izza Balqis</i>	<i>Validity And Reliability of The General Health Questionnaire-12 (GHQ-12) In Indonesia: Scale for The Detection Of Psychological Distress in Individuals Experiencing Quarter-Life Crisis</i>	68 – 75
<i>Maskuri Muhammad Cahyadi</i>	Modifikasi Skala <i>Self-Efficacy</i> : Analisis EFA dan CFA	76 – 81

Nama	Judul Artikel	Halaman
<i>Yasa Widayati, Zulfa Indira Wahyuni, Nia Tresniasari, Sitti Evangeline I. Suaidy, Ayunni Santicha, Niken Ayu Lestari, Yasmin Syahidah Al-Bisyri</i>	Adaptasi Alat Ukur <i>Learning Agility</i> pada Mahasiswa Baru: Versi Bahasa Indonesia	82 – 90
<i>Yasmin Siti Aulia, Aditya Jayadinata, Ally Allena Greene, Diah Zaleha Wyandini, Ghinaya Ummum Mukminin Hidayat</i>	Analisis Psikometris <i>The Moral Injury Scale</i>	91 – 96



MENILIK SIKAP WARGA SEKOLAH TERHADAP SISWA BERKEBUTUHAN KHUSUS: STUDI DI SEKOLAH MENENGAH INKLUSI JAKARTA

Adhissa Qonita

Prodi Psikologi, Fakultas Psikologi, Universitas Negeri Jakarta, Jakarta, Indonesia

Email: adhissa.qonita@gmail.com

<i>Received :</i>	<i>Revised</i>	<i>Accepted :</i>	<i>Published</i>
07 Maret 2025	17 April 2025	22 April 2025	30 April 2025

Abstract

Based on Peraturan Gubernur Nomor 32 Tahun 2021 Pasal 8 Ayat A, all public schools in DKI Jakarta must accept special needs students and conduct inclusive education. The implementation of inclusive schooling practices inextricably connected to the attitudes exhibited by school residents towards students with special needs. Unfortunately, the presense of students with special needs has not been readily accepted by the school community, indicating the underlying issues in the attitudes displayed toward. 105 teachers and 286 students in inclusive secondary schools in DKI Jakarta were included in this study. The measurement tools used were the Chedoke-McMaster Attitudes Toward Children with Handicaps (CATCH) for students and The Multidimensional Attitudes toward Inclusive Education Scale (MATIES) for teachers. The results of the study indicate that most students still exhibit a neutral attitude towards students with special needs. This suggests that while they do not display overtly negative attitudes, they also do not demonstrate clear positive acceptance. Interestingly, teachers' attitudes toward inclusive education appear to be relatively similar to those of the students, as their responses also tend to be neutral—neither explicitly positive nor negative. Although inclusive education in DKI Jakarta area has experienced significant development, the findings of this study indicate that most members of the school community have not yet presented positive attitudes toward students with special needs and inclusive education as envisioned by the government.

Keywords : Special needs students, inclusion, attitudes, inclusion teachers, secondary students

Abstrak

Peraturan Gubernur Nomor 32 Tahun 2021 Pasal 8 Ayat A menyebutkan bahwa semua sekolah negeri di wilayah DKI Jakarta harus menerima siswa berkebutuhan khusus dan menjalankan pendidikan inklusi. Berjalannya sekolah inklusi tidak lepas dari sikap yang ditunjukkan warga sekolah terhadap siswa berkebutuhan khusus. Sayangnya, keberadaan siswa berkebutuhan khusus tidak langsung diterima oleh warga sekolah sehingga mengindikasikan adanya isu tentang sikap yang ditunjukkan warga sekolah terhadap siswa berkebutuhan khusus. Studi ini dilakukan kepada 105 guru dan 286 siswa di sekolah menengah inklusi di Jakarta. Alat ukur yang digunakan adalah Chedoke-McMaster Attitudes Toward Children with Handicaps (CATCH) untuk siswa dan multidimensional attitudes toward inclusive education scale (MATIES) untuk guru. Hasil studi menunjukkan bahwa sebagian besar siswa menunjukkan sikap netral kepada siswa berkebutuhan khusus, yang artinya tidak menunjukkan sikap negatif terhadap siswa berkebutuhan khusus namun juga tidak menunjukkan sikap positif. Menariknya, sikap guru terhadap pendidikan inklusi pun tidak terlalu berbeda dengan siswa, yaitu sikapnya cenderung netral, tidak bersikap negatif maupun positif. Meskipun pendidikan inklusi di wilayah DKI Jakarta telah menunjukkan perkembangan pesat, hasil studi ini menunjukkan bahwa sebagian besar warga sekolah di DKI Jakarta belum menunjukkan sikap positif terhadap siswa berkebutuhan khusus ataupun terhadap pendidikan inklusi seperti yang diharapkan pemerintah.

Kata Kunci : Siswa Bekebutuhan Khusus, Inklusi, Sikap, Guru Inklusi, Siswa Menengah

1. Pendahuluan

Seperti yang sudah diketahui berdasarkan Pasal 31 ayat 1 UUD 1945 dan diperkuat oleh Pasal 5 Undang-undang Nomor 20 tahun 2003 bahwa setiap warga negara berhak mendapatkan pendidikan termasuk bagi siswa berkebutuhan khusus. Hal ini mendorong lebih banyak perubahan pandangan dan pemikiran warga Indonesia khususnya dalam segi pendidikan bagi anak berkebutuhan khusus. Salah satu perkembangan yang besar terhadap pendidikan anak berkebutuhan khusus juga tertuang dari Peraturan Menteri Pendidikan Nomor 70 Tahun 2009 yang memberikan upaya pendidikan untuk anak berkebutuhan khusus, atau yang dikenal sebagai pendidikan inklusi. Pendidikan inklusi membuka kesempatan pada siswa berkebutuhan khusus untuk mendapatkan pembelajaran bersama dengan siswa reguler di sekolah umum. Hal ini tentu sangat berbeda dengan sekolah khusus atau biasa disebut sekolah luar biasa (SLB) yang hanya menerima anak berkebutuhan khusus yang memiliki kondisi serupa satu sama lain.

Kehadiran Peraturan Menteri Pendidikan tersebut mendorong DPRD Provinsi DKI Jakarta mewajibkan seluruh sekolah negeri di wilayah DKI Jakarta untuk menjalankan pendidikan inklusi sejak tahun 2022. Kehadiran sekolah inklusi diharapkan dapat mengakomodasi anak-anak berkebutuhan khusus agar mendapatkan hak atas pendidikan yang layak. Berdasarkan Pasal 8 Ayat A Peraturan Gubernur Nomor 32 Tahun 2021 tentang petunjuk teknis penerimaan peserta didik baru di Jakarta, anak-anak berkebutuhan khusus bisa mendaftar di sekolah reguler melalui jalur afirmasi. DPRD Provinsi DKI Jakarta (2022) juga menegaskan bahwa setiap sekolah negeri harus memiliki satu kelas inklusi khusus di setiap tingkatnya untuk mengakomodasi anak-anak berkebutuhan khusus.

Dorongan aturan dan arahan yang diberikan di DKI Jakarta memberi harapan dalam pendidikan anak berkebutuhan khusus. Transisi sekolah reguler menjadi sekolah inklusi membawa tantangan bagi sekolah dan warganya untuk dapat mengakomodir kebutuhan siswa berkebutuhan khusus (ABK) di wilayah sekolahnya. Namun begitu, kebijakan ini juga menimbulkan kebingungan pada pihak sekolah untuk memenuhi tuntutan tersebut dikarenakan keterbatasan pengetahuan dan kurangnya persiapan yang terstruktur dalam mengelola sekolah inklusi (Dewi, 2017).

Lingkungan sekolah termasuk orang-orang di dalamnya baik guru, siswa, maupun tenaga pendidik diharapkan cukup terbuka dengan perubahan dan bergabungnya siswa ABK di sekolah. Guru diharapkan dapat mendidik, mengayomi, dan membantu siswa di kelas termasuk siswa berkebutuhan khusus (Haekal, 2018). Selain itu, siswa ABK, terutama di usia sekolah menengah, juga menginjak masa remaja sehingga juga memiliki keinginan untuk menghabiskan waktu dengan teman sebayanya (King, 2017), khususnya untuk mencari kebutuhan intimasi dan kepastian akan nilai diri mereka (Santrock, 2019). Oleh sebab itu, siswa ABK memerlukan respon yang positif atas kehadiran mereka baik dari guru dan siswa lain sebagai teman sebaya mereka di sekolah.

Hasanah, Ni'matuzaroh & Nurhamida (2016) menjelaskan bahwa siswa ABK membutuhkan sikap positif dari lingkungan sekolah berupa penerimaan dan kepedulian terhadap kehadiran mereka. Sikap positif yang diperlihatkan lingkungan belajar terhadap pendidikan inklusi dan siswa ABK dianggap sebagai hal yang sangat penting, mendukung, dan mendorong keberhasilan pendidikan inklusi (Heyder, Sudkamp & Steinmayr, 2020). Hanya saja saat ini respon yang diberikan kepada siswa ABK belum tentu positif. Salim, Jesslin & Rumatutur (2022) merangkum bahwa siswa di sekolah menengah masih muncul respon negatif kepada siswa ABK seperti prasangka dan stigma negatif, rendahnya penerimaan, penolakan, sikap menjauhi bahkan perundungan. Respon negatif ini dapat membuat siswa ABK menjadi tidak percaya diri, merasa malu, takut, marah, hingga dapat saja menarik diri dari pergaulannya sehari-hari (Sari & Kurnia, 2022). Respon positif dan negatif terhadap siswa ABK merupakan cerminan dari sikap yang mereka terima dari orang-orang di sekitarnya.

Armstrong (1986) menjelaskan sikap terhadap anak berkebutuhan khusus diartikan sebagai perasaan dan perlakuan individu terhadap anak berkebutuhan khusus. Sikap ini merupakan keyakinan yang dimiliki seseorang sebelum tertampilkan melalui perasaan dan perilaku (Myers & Twenge, 2016). Menurut Rosenberg & Hovland (dalam Avramidis et al., 2000), sikap merupakan gabungan dari reaksi emosional, pengetahuan, dan perilaku terhadap objek tertentu dan dapat berupa hal positif atau negatif terhadap objek tersebut. Menurut Mahat (2008), dimensi kognitif atau pengetahuan adalah persepsi yang dimiliki atau keyakinan terhadap pendidikan inklusi dan siswa ABK, dimensi afektif atau emosional adalah perasaan individu akibat persepsi tersebut, sedangkan dimensi behavior atau perilaku adalah kecenderungan respon perilaku terhadap pendidikan inklusi dan siswa ABK. Sikap positif yang diharapkan terhadap pendidikan inklusi dan siswa ABK dari lingkungan sekolah adalah penerimaan dalam kelas, keyakinan perbedaan karakteristik dan kebutuhan siswa ABK serta harapan dan dukungan terhadap berjalannya pendidikan inklusi (Subban & Sharma, 2006). Sikap positif ini diharapkan dapat dimiliki oleh sekolah inklusi, terutama di wilayah Jakarta yang sudah menerapkan sekolah inklusi sesuai Peraturan Gubernur Nomor 32 Tahun 2021. Sayangnya, keberadaan pendidikan inklusi dan siswa ABK masih cukup awam bagi sekolah reguler, sehingga sikap yang ditunjukkan warga sekolah atas siswa ABK dan pendidikan inklusi belum dapat dipastikan.

2. Metode Penelitian

Partisipan

Dalam studi ini, terdapat dua kelompok partisipan yang dilibatkan, yaitu siswa sekolah menengah inklusi dan guru sekolah menengah inklusi di wilayah DKI Jakarta. Partisipan studi ini terdiri dari 286 siswa sekolah menengah inklusi dan 105 guru sekolah menengah inklusi. Partisipan studi untuk siswa sekolah menengah memiliki karakteristik yaitu: 1) Siswa/siswi SMP/SMA/Sederajat; 2) Bersekolah di Provinsi DKI Jakarta 3) Bersekolah di sekolah inklusi. Partisipan dari kelompok guru yang dilibatkan pada penelitian ini memiliki karakteristik: 1) Guru di sekolah menengah SMP/SMA/Sederajat 2) Mengajar di wilayah Jakarta 3) Mengajar di sekolah inklusi yang menerima siswa ABK. Teknik pengambilan sampel yang digunakan adalah Teknik *non-probability sampling* dengan metode *convenience sampling*.

Instrumen

Variabel yang diukur dalam studi ini adalah sikap terhadap siswa berkebutuhan khusus dan pendidikan inklusi. Studi ini menggunakan dua alat ukur yang berbeda sesuai dengan konteks siswa dan guru sekolah menengah inklusi di wilayah Jakarta. Instrumen yang digunakan untuk mengukur sikap pada siswa ABK dalam studi ini adalah *Chedoke-McMaster Attitudes Toward Children with Handicaps (CATCH)* yang disusun oleh Armstrong (1986) berdasarkan teori Allport mengenai sikap yang terdiri dari komponen kognitif, afektif dan *behavioral/perilaku*. Instrumen ini diberikan untuk mengukur sikap anak dengan fisik normal kepada anak disabilitas. Instrumen ini memiliki 36 pertanyaan dengan lima pilihan jawaban (sangat tidak setuju-sangat setuju). Instrumen lalu diadaptasi dengan proses uji coba dengan *cronbach's alpha* sebesar 0.875 dan 24 item yang digunakan (masing-masing 8 item/komponen).

Instrumen yang digunakan untuk mengukur sikap guru adalah *Multidimensional Attitudes Toward Inclusive Education Scale (MATIES)* yang dikembangkan oleh Mahat (2008). MATIES terdiri dari 18 item dengan 5 pilihan jawaban yang dikelompokkan dalam tiga dimensi yaitu kognitif, afektif, dan *behavioral/perilaku*, masing-masing dimensi memiliki 6 item. Alat ukur ini lalu diadaptasi kedalam Bahasa Indonesia dan memiliki nilai *cronbach's alpha* sebesar 0,771 (kognitif), 0,764 (afektif), 0,778 (perilaku). Dari hasil uji coba terdapat dua item yang tidak digunakan sehingga jumlah item yang digunakan dalam instrumen adalah 16 item.

Analisa Data

Data yang dihasilkan dianalisis dengan menggunakan pendekatan kuantitatif dengan analisis statistika yang digunakan ialah analisis deskriptif. Penelitian ini menghasilkan gambaran perhitungan kategorikal yang didapatkan dari skor dari instrumen yang telah diberikan dan diisi oleh partisipan siswa atau guru di sekolah menengah inklusi di Jakarta.

3. Hasil dan Diskusi

Demografi

Tabel 1. Data Demografi

Data Demografi		N	Presentase	
Siswa	Usia			
		12	76	26,65
		13	59	20,6%
		14	33	11,5%
		15	14	4,9%
		16	14	4,9%
		17	77	26,9%
	18	13	4,5%	
Domisili	Jakarta Selatan	145	50,7%	
	Jakarta Timur	120	42%	
	Jakarta Barat	17	5,9%	
	Lainnya	4	1,4%	
Jenjang Sekolah	SMP	159	55,6%	
	SMA	127	44,4%	
Frekuensi Interaksi dengan anak berkebutuhan khusus	Tidak Pernah	36	12,6%	
	Jarang	70	24,5%	
	Terkadang	110	38,5%	
	Sering	70	24,5%	
Anak berkebutuhan khusus di kelas yang sama	Tidak Ada	102	35,7%	
	Ada	184	64,3%	

Guru				
Usia		20-30	26	24,8 %
		31-40	18	17,1 %
		41-50	21	20 %
		51-60	40	38,1 %
Jenis Kelamin		Laki-Laki	32	30,5%
		Perempuan	73	69,5%
Jenjang Mengajar		SMP	69	65,7%
		SMA	36	34,3%
Pengalaman mengajar berkebutuhan khusus Siswa		1-5 Tahun	79	75,2%
		6-10 Tahun	21	20%
		11-15 Tahun	4	3,8%
		16-20 Tahun	1	1 %
Pengalaman mengikuti Inklusi Pelatihan		Pernah	36	34,3%
		Tidak Pernah	69	65,7%

Berdasarkan penyebaran demografis partisipan siswa, hal yang memperlihatkan perbedaan signifikan terlihat dari frekuensi anak-anak yang berinteraksi dengan anak berkebutuhan khusus yaitu 38,5%. Selain itu di kelas mereka juga memiliki anak berkebutuhan khusus sebanyak 64,3%. Selain itu domisili terbanyak adalah Jakarta Selatan 50,7%. Saat ini siswa sekolah menengah yang mengikuti penelitian ini tidak terlalu banyak perbedaan antara siswa SMP dan SMA.

Sedangkan pada data demografis guru memperlihatkan perbedaan pada lebih banyak berusia 51-60 tahun 38,1%, guru Perempuan 69,5%, jenjang mengajarnya adalah SMP 65,7%, pengalaman mengajar 1-5 tahun 75,2% dan tidak pernah mengikuti pelatihan inklusi yaitu 65,75%.

Kategorisasi Sikap

Tabel 2. Kategorisasi Sikap Siswa terhadap Siswa ABK di sekolah inklusi

Kategorisasi	Jumlah	Presentase
Sikap Negatif	38	13,3%
Sikap Netral	215	75,2%
Sikap Positif	33	11,5%

Berdasarkan tabel 2 pada hasil sikap yang ditunjukkan oleh siswa terhadap siswa ABK diambil berdasarkan 3 kategori sikap menurut Doreen & Kurniawati (2018), pembagian rentang skor dilakukan dengan menggunakan mean empirik. Dari hasil tersebut diketahui bahwa kebanyakan sikap siswa sekolah menengah inklusi di Jakarta terhadap Pendidikan inklusi dan siswa ABK adalah netral 75,2%. Hal ini menunjukkan bahwa siswa sebenarnya tidak menentang ataupun sangat mendukung adanya Pendidikan inklusi dan adanya siswa ABK di kelas. Sikap netral juga menunjukkan sikap cenderung hanya menerima arahan sesuai dengan kebijakan sekolah inklusi.

Tabel 3. Kategorisasi Sikap Guru Terhadap Pendidikan Inklusi

Kategorisasi	Jumlah	Presentase
Sikap Positif	52	49,5%
Sikap Negatif	53	50,5%

Sikap yang ditunjukkan oleh guru terhadap Pendidikan inklusi dibagi menjadi dua kategori yaitu sikap positif dan sikap negatif seperti penelitian yang sebelumnya dilakukan oleh Muzdalifah & Billa (2017) sedangkan pembagian kategorisasi ini bagi dengan menggunakan mean empiric. Dari hasil tersebut terlihat bahwa presentase positif dan negatif hampir tidak memiliki perbedaan atau kemungkinan jika dibagi dengan kategori netral akan banyak guru berada di posisi netral. Perbedaan yang kecil ini juga cenderung menunjukkan bahwa masih adanya sikap yang negatif ditunjukkan oleh guru. Sikap negatif inilah yang sebenarnya diharapkan tidak muncul dalam pengajaran berbasis inklusi di sekolah menengah yang sudah diarahkan sesuai peraturan gubernur mengenai sekolah inklusi Jakarta.

4. Diskusi

Dalam penyelenggaraan sekolah inklusi, sikap siswa menjadi salah satu hal penting yang dapat menunjang dan membantu siswa ABK dalam belajar. Hasil studi yang dilakukan didapatkan bahwa kebanyakan siswa menunjukkan sikap yang netral (75,2%). Sikap netral ini menunjukkan bahwa siswa di sekolah inklusi

sebenarnya tidak menunjukkan sikap negatif terhadap siswa ABK, namun mereka juga tidak langsung menerima anak tersebut dalam pertemanan mereka (Hallberg & Klingberg, 2023; dalam Salsabila 2024). Di sisi lain masih adanya sikap negatif terhadap siswa ABK 13,3% menunjukkan bahwa masih ada siswa yang bersikap buruk terhadap kehadiran siswa ABK di sekolahnya. Hal ini menunjukkan bahwa dalam pelaksanaan Pendidikan inklusi di lapangan masih belum menunjukkan harapan atau iklim yang ramah terhadap ABK.

Menurut Wang, dkk (2021) frekuensi interaksi seseorang dengan individu berkebutuhan khusus berkorelasi positif dengan sikap orang tersebut terhadap individu berkebutuhan khusus tersebut. Dari hasil di atas dapat diketahui bahwa siswa tidak terlalu sering berinteraksi dengan siswa ABK (tidak pernah sampai jarang) meskipun berada di sekolah inklusi. Beberapa siswa yang mengikuti studi ini tidak memiliki siswa ABK di kelasnya sehingga mereka terkadang kurang menyadari sikap mereka, bahkan kurang menyadari bahwa sekolah mereka sebenarnya adalah sekolah inklusi yang menerima siswa berkebutuhan khusus. Selain interaksi dengan siswa ABK, sikap netral yang ditunjukkan oleh siswa disebabkan kurangnya pengetahuan mereka mengenai anak berkebutuhan khusus dan bagaimana seharusnya mereka bersikap terhadap mereka. Menurut (Fajrianti & Purwanti, 2021), informasi mengenai pendidikan inklusi kepada siswa masih belum lengkap sehingga siswa menjadi bingung apa yang diharapkan lingkungan terhadap mereka. Siswa cenderung mengetahui sikap yang buruk yang tidak boleh dilakukan terhadap siswa lain termasuk siswa berkebutuhan khusus, namun mereka tidak mengetahui sikap positif seperti apa dalam memperlakukan siswa ABK di sekolah mereka.

Dilihat dari sisi data demografi guru dapat disimpulkan bahwa munculnya siswa ABK di lingkungan sekolah regular sudah dialami oleh salah satu guru yang merasa sudah mengajar siswa ABK selama 16-20 tahun. Hal tersebut menegaskan bahwa keberadaan siswa ABK di sekolah regular sebenarnya sudah lebih lama dibandingkan munculnya peraturan penyelenggaraan Pendidikan inklusi. Hanya saja sampai saat ini, dari hasil pengambilan data yang dilakukan meski tidak berbeda jauh, namun guru masih sedikit lebih menunjukkan sikap negatif dibandingkan sikap positif terhadap Pendidikan inklusi yang terdapat siswa ABK. Hal ini diperkuat dengan kondisi di lapangan menunjukkan bahwa guru regular belum memiliki kompetensi untuk dapat melayani ABK dengan baik (Winarsih, 2022). Adanya pendidikan inklusi dan siswa ABK membuat bertambahnya peran guru di kelas. Sehingga guru regular memerlukan kompetensi tambahan seperti pendidikan dan pelatihan terkait anak berkebutuhan khusus agar dapat memenuhi tuntutan pelayanan pembelajaran yang dibutuhkan siswa ABK. Hanya saja di sekolah menengah inklusi di Jakarta kebanyakan guru belum pernah mendapatkan pelatihan terkait kebutuhan khusus siswa (65,7%).

Ditambah lagi hasil studi ini juga menunjukkan bahwa sebagian besar guru di berbagai sekolah menengah inklusi Jakarta baru berinteraksi dengan siswa ABK 1-5 tahun saja (75,2%). Hal tersebut menyebabkan guru regular mengalami kesulitan dalam memodifikasi kurikulum, menyesuaikan materi dan media pembelajaran yang dibutuhkan oleh siswa ABK (Pardede, 2020). Pengalaman yang belum begitu banyak dalam berhadapan dengan siswa ABK ditambah belum adanya bantuan tambahan kompetensi yang diberikan kepada guru terkait ABK membuat guru berada di posisi yang kurang menguntungkan guna menunjukkan sikap yang diharapkan pada siswa ABK. Hal inilah yang menjadi dasar kebingungan dari sikap guru terhadap pendidikan inklusi dan siswa ABK yang ada di lingkungan sekolah mereka, dimana sebagai pengajar mereka tidak memahami apa yang perlu dilakukan agar siswa ABK di kelas mendapatkan pembelajaran individual yang sesuai dengan karakteristik mereka. Ketidaktahuan dan ketidaksiapan dalam menjalankan pendidikan inklusi inilah yang mungkin membuat sikap guru terhadap pendidikan inklusi belum sepenuhnya positif dan masih cenderung negatif dibandingkan siswa dalam memandang siswa ABK. Hal tersebut belum sesuai dengan harapan penyelenggaraan Pendidikan inklusi dimana sikap yang diharapkan adalah sikap positif. Hal ini dikarenakan kehadiran sikap positif terhadap siswa ABK diperlukan guna membentuk model anti diskriminatif terhadap siswa ABK untuk siswa lain di kelasnya (Rahbini, 2012). Harapannya, guru dapat mengembangkan sikap positif terhadap siswa ABK sehingga pelaksanaan pendidikan inklusi dapat berjalan lebih baik).

5. Kesimpulan

Secara umum, sekolah yang melaksanakan pendidikan inklusi perlu menyesuaikan kebutuhan siswa ABK, termasuk menghadirkan sikap warga sekolah yang positif terhadap kehadirannya. Idealnya, seluruh warga sekolah baik guru, siswa lain, staff kependidikan dan berbagai pihak lain yang terlibat dengan sekolah perlu memiliki sikap positif terhadap keberadaan siswa ABK di sekolah yang telah ditetapkan sebagai sekolah inklusi. Hanya saja, dalam kenyataannya sikap yang ditunjukkan oleh siswa masih cenderung netral. Sedangkan sikap terhadap Pendidikan inklusi yang ditunjukkan oleh guru sedikit lebih negatif dibandingkan positif. Sikap netral pada siswa sendiri menimbulkan kebingungan dalam bersikap dan keterbukaan untuk masuk ke wilayah kedekatan satu sama lain.

Hasil sikap guru terhadap Pendidikan inklusi menunjukkan tidak banyak berbedanya sikap positif dan negatif, malah mengarah pada sikap negatif. Hal tersebut kemungkinan dipicu oleh kurangnya penambahan kompetensi guru dalam menghadapi siswa ABK yang sebenarnya menambah peran guru yang sejalan dengan penelitian dari Gaad & Khan (2007), guru merasa kurangnya keterampilan dan merasakan beban berat guna memenuhi kebutuhan siswa ABK. Dalam studi ini juga ditemukan bahwa sikap negatif juga masih cukup

ditunjukkan sehingga tidak jarang kasus mengenai siswa ABK mengalami hal kurang menyenangkan dan diskriminatif.

Meski demikian, dalam studi ini dapat dilihat juga bahwa sikap terhadap siswa ABK dan pendidikan inklusi mulai mengalami peningkatan terlihat dari sudah adanya guru atau siswa lain yang bersikap positif kepada siswa ABK. Kedepannya, sikap positif terhadap siswa ABK dan pendidikan inklusi diharapkan menjadi lebih meningkat. Siswa dan guru diharapkan lebih diberikan pengetahuan terkait dengan berbagai macam karakteristik siswa ABK dan bagaimana cara untuk bersikap dan memperlakukan mereka. Selain itu, guru juga diharapkan mendapatkan lebih banyak kompetensi untuk mengakomodir kebutuhan siswa ABK untuk dapat bersekolah.

6. Referensi

- Armstrong, R. W. (1986). *Attitudes Toward Disabled Children: A New Measure of Able-Bodied Children's Attitudes Toward Disabled Children*. [Doctoral's Thesis, McMaster University]. MacSphere. Retrieved from <https://macsphere.mcmaster.ca/bitstream/11375/5824/1/fulltext.pdf>
- Avramidis, E., Bayliss, P., & Burden, R. (2000). Student teachers' attitudes towards the inclusion of children with special educational needs in the ordinary school. *Teaching and teacher education*, 16(3), 277-293.
- Dewi, R. S. (2017). Pengaruh Pelatihan Efikasi Diri sebagai Pendidik terhadap Penurunan Burnout pada Guru di Sekolah Inklusi: Pengaruh Pelatihan Efikasi Diri sebagai Pendidik terhadap Penurunan Burnout pada Guru di Sekolah Inklusi. *Naturalistic: Jurnal Kajian dan Penelitian Pendidikan dan Pembelajaran*, 1(2), 155-167.
- Doreen, R. U., & Kurniawati, F. (2018). Gender differences in attitude toward students with disabilities among primary students in Jakarta. *2nd INDOEDUC4ALL-Indonesian Education for All (INDOEDUC 2018)*, 272, pp. 1-4. Atlantis Press.
- Fajrianti, S. P., & Purwanti, M. (2021). The role of knowledge and attitude of regular students to acceptance toward special educational needs. *Journal of Educational, Health and Community Psychology*, 10(2), 275-289.
- Gaad, E., & Khan, L. (2007). Primary Mainstream Teachers' Attitudes towards Inclusion of Students with Special Educational Needs in the Private Sector: A Perspective from Dubai. *International Journal of Special Education*, 22(2), 95-109.
- Haekal, T. M. (2018). Peran Wali Kelas Dalam Pengelolaan Anak Berkebutuhan Khusus (ABK) di SDN 1 Banda Aceh. (Doctoral dissertation, UIN Ar-Raniry Banda Aceh).
- Hasanah, U., Ni'matuzahroh, N., & Nurhamida, Y. (2016). Sikap siswa reguler terhadap siswa berkebutuhan khusus dengan kecenderungan *bullying* kelas inklusi. *Jurnal Psikologi Islam Al-Qalb*, 7(1), 43-58.
- Heyder, A., Suedkamp, A., & Steinmayr, R. (2020). How are teachers' attitudes toward inclusion related to the social-emotional school experiences of students with and without special educational needs?. *Learning and Individual Differences*, 77, 101776.
- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia. (2022). *Panduan Pelaksanaan Pendidikan Inklusif*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.
- King, L. (2017). *The Science of Psychology: An Appreciative View* (4th ed.). McGraw-Hill Education.
- Mahat, M. (2008). The Development of a Psychometrically-Sound Instrument to Measure Teachers' Multidimensional Attitudes toward Inclusive Education. *International journal of special education*, 23(1), 82-92.
- Myers, D. G., & Twenge, J. M. (2016). *Social Psychology* (12th ed.). McGraw-Hill Education.
- Pemerintah Indonesia. (1945). *Undang-Undang Dasar Negara Republik Indonesia 1945 Pasal 31 ayat 1 tentang hak mendapatkan Pendidikan*. Sekretariat Negara. Jakarta.
- Muzdalifah & Billah (2017) Pengaruh Efikasi pada Sikap Guru Terhadap Pendidikan Inklusif <https://www.researchgate.net/publication/321482206>
- Pemerintah Provinsi DKI Jakarta. (2021). *Peraturan Gubernur Daerah Khusus Ibukota Jakarta Nomor 32 Tahun 2021 tentang Petunjuk Teknis Penerimaan Peserta Didik Baru*. Sekretariat Daerah Provinsi DKI Jakarta.
- Pardede, P. F. D., & Fernandes, R. (2020). Implementasi Pendidikan Inklusif di SMA Negeri 3 Bukittinggi. *Jurnal Perspektif*, 3(2), 272-278. <https://doi.org/10.24036/perspektif.v3i2.237>
- Rahbini, R. (2012). Pendidikan Inklusif; Sebuah Konstruksi Pendidikan Anti Diskriminasi Bagi Difabel. *Jurnal Kependidikan Islam*, 2(1), 29-44. <https://jurnalftk.uinsby.ac.id/index.php/JKPI/article/download/1060/338>
- Santrock, J. W. (2019). *Life-Span Development* (17th ed.). McGraw-Hill Education.
- Salim, R. M., Jesslin, & Rumlatur, N. A. (2022). *Puppet book* membantu anak usia dini dalam memahami anak berkebutuhan khusus. *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*, 6(6), 5769-5781. doi:10.31004/obsesi.v6i6.2975

- Salsabila, Q.A (2024) Pengaruh Persepsi Iklim Inklusi di Sekolah terhadap Sikap Siswa pada Anak Berkebutuhan Khusus di Sekolah Menengah Jakarta. Skripsi : Universitas Negeri Jakarta.
- Sari, D. A., & Kurnia, I. (2022). Kenali pentingnya membangun pendidikan karakter pada anak berkebutuhan khusus (ABK) di sekolah reguler.
- Subban, P., & Sharma, U. (2006). Understanding educator attitudes toward the implementation of inclusive education. *Disability studies quarterly*, 25(2).
- Wang, Z., Xu, X., Han, Q., Chen, Y., Jiang, J., & Ni, G.-X. (2021). Factors associated with public attitudes towards persons with disabilities: a systematic review. *BMC Public Health*, 21(1058). Retrieved from <https://doi.org/10.1186/s12889-021-11139-3>
- Winarsih, M., Bintoro, T., & Nanik, U (2022). Pemberdayaan Guru Madrasah Ibtidaiyah dan Raudhatul Athfal dalam pelaksanaan Pendidikan Inklusif. *Jurnal Pengabdian Masyarakat*, 13(1), 120-125



ANALISIS PSIKOMETRI INSTRUMEN *PSYCHOLOGICAL WELL-BEING* PADA DOKTER: ADAPTASI SKALA PENGUKURAN

Hilda Egan¹, Iriani Indri Hapsari², Lussy Dwiutami Wahyuni³, Gumgum Gumelar⁴

^{1,2,3,4} Program Studi Magister Sains Psikologi, Fakultas Psikologi, Universitas Negeri Jakarta, Jakarta, Indonesia

Email: hildaegan14@gmail.com

Received :	Revised :	Accepted :	Published :
24 Juni 2024	06 April 2025	22 April 2025	30 April 2025

Abstract

This study involved 70 doctors in Indonesia with diverse demographic characteristics, consisting of 54.2% males and 45.8% females, with ages ranging from 25 to 60 years. Based on Rasch analysis using the *jMetrik* software version 4.1.1, the psychological well-being measurement instrument for doctors in Indonesia showed excellent reliability, with an item reliability value of 0.95 and person reliability of 0.86. These results indicate that the measurement tool used is valid and reliable for this study. The instrument also demonstrated excellent strata separation with a value of 6.77, as well as a range of item difficulty levels, from the easiest to the most difficult. Item number 18 was considered the easiest, while item number 8 was the most difficult. However, item number 26 was found to be unsuitable for the participants and was therefore excluded from the analysis. Overall, the quality of this instrument is rated as excellent, with only two items in the harmony dimension to be used in future research.

Keywords: *psychological well-being, medical doctor, adaptation, validation, psychometry properties*

Abstrak

Penelitian ini melibatkan 70 dokter di Indonesia dengan karakteristik demografis yang beragam, terdiri dari 54,2% laki-laki dan 45,8% perempuan, dengan usia berkisar antara 25 hingga 60 tahun. Berdasarkan analisis *Rasch* menggunakan perangkat lunak *jMetrik* version 4.1.1, instrumen pengukuran *psychological well-being* untuk dokter di Indonesia menunjukkan reliabilitas yang sangat baik, dengan nilai reliabilitas item sebesar 0.95 dan reliabilitas *person* sebesar 0.86. Teknik pengambilan *sample* pada penelitian ini yaitu *purposive sampling*. Hasil ini mengindikasikan bahwa alat ukur yang digunakan valid dan dapat diandalkan untuk penelitian ini. Instrumen ini juga menunjukkan pemisahan strata yang sangat baik dengan nilai 6.77, serta variasi tingkat kesulitan item yang mencakup rentang dari yang paling mudah hingga yang paling sulit. Item nomor 18 dianggap sebagai yang paling mudah, sementara item nomor 8 adalah yang paling sulit. Namun, item nomor 26 dinilai tidak cocok dengan partisipan, sehingga dikeluarkan dari analisis. Secara keseluruhan, kualitas instrumen ini dinilai sangat baik, dan hanya dua item dalam dimensi harmoni yang akan digunakan dalam penelitian selanjutnya.

Keywords: *psychological well-being, dokter, adaptasi, validasi, psikometri*

1. Pendahuluan

Uraian Rasio dokter per 1.000 penduduk merupakan indikator penting dalam menilai ketersediaan tenaga medis untuk memberikan layanan kesehatan yang memadai. Berdasarkan data dari World Health Organization (WHO) yang dihimpun oleh Index Mundi, Indonesia hanya memiliki 0,47 dokter per 1.000 penduduk, yang jauh di bawah standar minimum WHO yang seharusnya 1 dokter per 1.000 penduduk. Angka ini juga jauh lebih rendah dibandingkan dengan negara-negara tetangga seperti Malaysia yang memiliki rasio 2,2 dokter per 1.000 penduduk

dan Thailand yang memiliki 0,95 dokter per 1.000 penduduk. Ketimpangan rasio dokter ini mencerminkan adanya masalah serius dalam distribusi tenaga medis di Indonesia, di mana sebagian besar dokter terkonsentrasi di kota-kota besar, terutama di Pulau Jawa, sementara daerah terpencil dan kurang berkembang kekurangan tenaga medis yang memadai. Selain itu, faktor-faktor seperti terbatasnya jumlah lulusan pendidikan kedokteran, kurangnya insentif untuk dokter bekerja di daerah terpencil, serta masalah pengalokasian sumber daya kesehatan di daerah-daerah ini turut memperburuk ketimpangan ini. Dampak dari ketimpangan ini sangat signifikan terhadap kualitas layanan kesehatan yang diterima masyarakat, di mana akses terhadap dokter yang berkualitas terbatas, waktu tunggu untuk perawatan menjadi lebih lama, dan beban kerja yang tinggi bagi dokter yang ada dapat menurunkan kualitas layanan medis. Oleh karena itu, penting bagi pemerintah untuk meningkatkan jumlah dokter, memperbaiki distribusi tenaga medis, serta menyediakan insentif bagi dokter yang bekerja di daerah kurang berkembang agar dapat menciptakan sistem layanan kesehatan yang lebih merata dan berkualitas bagi seluruh lapisan masyarakat.

Krisis jumlah dokter di seluruh dunia, terutama di negara berkembang, telah menjadi perhatian utama dalam sektor kesehatan. Ketidakseimbangan antara jumlah dokter dengan populasi pasien yang terus berkembang dapat memperburuk kualitas pelayanan kesehatan. Dalam konteks ini, penting untuk mempertimbangkan berbagai faktor yang mempengaruhi *well-being* dokter serta dampaknya terhadap profesionalisme medis dan sikap dokter terhadap pasien. Krisis ini bukan hanya berdampak pada pasien, tetapi juga mempengaruhi kondisi emosional, psikologis, dan fisik dokter yang berhadapan dengan beban kerja yang tinggi dan tantangan profesinya. Kelelahan yang dialami oleh dokter merupakan sindrom terkait pekerjaan yang melibatkan kelelahan emosional, depersonalisasi dan perasaan berkurangnya pencapaian pribadi yang merupakan hal umum terjadi (West et al., 2018). Kelelahan profesional terdiri dari kelelahan fisiologis yang meliputi perasaan kehilangan energi dan kelelahan fisik serta mental akibat tuntutan pekerjaan yang berlebihan. Kelelahan sinisme, yang lebih dikenal sebagai depersonalisasi yang dialami ketika dokter mulai merasakan apatis atau tidak peduli terhadap pasiennya. Dan terakhir tidak mampunya dokter mencapai tujuan profesional meskipun sudah berusaha keras (Maslach et al., 2016). Hal tersebut menjadi dampak pada *psychological well-being* para dokter, meliputi *optimism*, kesejahteraan, kebahagiaan, kreatifitas, kebijaksanaan, kesehatan, pengambilan keputusan yang melibatkan kekuatan pribadi, dan *self-efficacy* (Gilman et al., 2014). Bahkan dapat mengakibatkan risiko buruk yang dialami oleh dokter seperti meningkatnya kecemasan, depresi, PTSD, masalah dengan tidur dan kualitas hidup yang rendah (Schneider et al., 2022).

Psychological well-being adalah sebuah kondisi di mana seseorang memiliki sikap positif terhadap diri sendiri, orang lain dan dapat membuat keputusan dan mengatur perilakunya. Peningkatan *well-being* pada seseorang juga didukung dengan peningkatan pendidikan dan status pekerjaan, hal ini dikarenakan mereka memiliki pandangan hidup yang optimis, menekankan pertumbuhan dan perkembangan diri (van Dierendonck, 2004). Pentingnya memprioritaskan dan melindungi *psychological well-being* dokter agar mereka dapat kompeten secara sosial dan berfungsi dengan baik di lingkungan tempat bekerja sosial mereka (Gómez-López et al., 2019). Secara keseluruhan, *psychological well-being* dokter tidak hanya memengaruhi kualitas pelayanan kesehatan yang mereka berikan tetapi juga *well-being* mereka sendiri. Mengadaptasi instrumen *psychological well-being* dari Fitria Abidin untuk dokter di Indonesia sangat penting karena profesi kedokteran menghadapi tekanan emosional dan fisik yang tinggi. Dokter perlu memiliki *psychological well-being* yang baik untuk dapat memberikan pelayanan medis yang berkualitas. Instrumen yang disesuaikan dengan konteks Indonesia, termasuk aspek budaya dan sosial, akan lebih akurat dalam menilai kondisi psikologis dokter. Hal ini memungkinkan identifikasi masalah kesejahteraan yang mungkin tersembunyi, seperti stres atau *burnout*, dan membantu merancang program dukungan yang tepat, sehingga meningkatkan kinerja dan *well-being* dokter. Dengan demikian, adaptasi ini mendukung dokter agar tetap kompeten dan berfungsi optimal di tempat kerja mereka. Dalam penelitian mengenai *psychological well-being*, Ryff (1995) mengidentifikasi enam dimensi utama yang mendukung kesejahteraan psikologis seseorang. Dimensi-dimensi ini saling terkait dan membentuk keseluruhan pengalaman kesejahteraan yang mendalam, yang mencakup penerimaan diri, hubungan positif dengan orang lain, otonomi, penguasaan lingkungan, pertumbuhan pribadi, dan tujuan hidup. Setiap dimensi ini memiliki peran penting dalam membentuk *psychological well-being* yang seimbang, terutama dalam konteks profesi yang penuh tantangan, seperti kedokteran. Penerimaan diri adalah dimensi pertama yang berfokus pada sikap positif terhadap diri sendiri, termasuk menerima kekuatan dan kelemahan pribadi. Hubungan positif dengan orang lain merupakan dimensi yang menekankan pentingnya hubungan yang mendalam dan bermakna dengan orang lain. Hal ini melibatkan empati, kasih sayang, dan dukungan sosial yang memperkuat rasa keterhubungan dengan orang-orang di sekitar mereka. Bagi seorang dokter, memiliki hubungan yang positif dengan kolega, pasien, dan keluarga sangat penting, tidak hanya dalam mendukung *psychological well-being* mereka sendiri, tetapi juga dalam menciptakan lingkungan kerja yang kolaboratif dan harmonis. Hubungan yang memuaskan ini memberikan rasa memiliki yang mendalam dan penting dalam menciptakan keseimbangan emosional di tengah tekanan pekerjaan.

Dimensi otonomi berkaitan dengan kemampuan untuk membuat keputusan secara mandiri dan memiliki kontrol atas hidupnya, menunjukkan pentingnya individu memiliki rasa kemandirian dalam bertindak dan mengatur hidupnya sesuai dengan nilai dan keyakinannya. Bagi seorang dokter, otonomi berarti memiliki kebebasan untuk membuat keputusan medis yang berdampak pada pasien, serta memiliki kontrol terhadap waktu dan cara kerja mereka. Hal ini sangat penting untuk menjaga kesejahteraan psikologis dokter agar tidak merasa terkekang atau tertekan oleh ekspektasi eksternal.

Dimensi penguasaan lingkungan adalah dimensi yang berfokus pada kemampuan seseorang untuk mengelola dan beradaptasi dengan tuntutan kehidupan sehari-hari. Individu yang memiliki penguasaan lingkungan yang baik mampu menangani tantangan dalam pekerjaan, keluarga, dan kehidupan sosial dengan efisien. Sebagai contoh, seorang dokter yang dapat mengelola waktu dan stres secara efektif akan lebih mampu mengatasi tantangan di tempat kerja, seperti beban pasien yang tinggi atau situasi darurat medis, serta tetap menjaga keseimbangan dengan kehidupan pribadi mereka. Pada dimensi pertumbuhan pribadi mengacu pada pengembangan diri secara berkelanjutan dan keinginan untuk mencapai potensi penuh seseorang. Dimensi ini melibatkan rasa berjuang untuk terus belajar, beradaptasi, dan memperbaiki diri, baik dalam konteks profesional maupun pribadi. Bagi seorang dokter, pertumbuhan pribadi adalah elemen penting dalam menjalani karir yang sukses dan memuaskan, yang mencakup pengembangan keterampilan medis, pengetahuan baru, serta kemampuan untuk beradaptasi dengan perubahan dalam dunia medis yang terus berkembang. Dimensi tujuan hidup merupakan dimensi terakhir yang mengacu pada adanya makna dan arah yang jelas dalam hidup seseorang. Individu yang memiliki tujuan hidup yang kuat merasakan bahwa setiap tindakan dan pilihan yang mereka buat memiliki dampak yang signifikan dan bermakna. Dalam konteks dokter, tujuan hidup seringkali terhubung dengan motivasi untuk membantu orang lain, memberikan kontribusi positif kepada masyarakat, dan berkomitmen terhadap penyembuhan dan perawatan pasien. Tujuan hidup yang jelas memberikan arah dan energi untuk terus berjuang meskipun menghadapi tantangan berat dalam profesi mereka.

Secara keseluruhan, dimensi-dimensi tersebut berkontribusi secara langsung pada *psychological well-being* dokter, yang pada gilirannya mempengaruhi kualitas layanan kesehatan yang mereka berikan. Penelitian tentang *psychological well-being* ini memberikan wawasan penting dalam menciptakan lingkungan kerja yang mendukung dan memungkinkan dokter untuk berkembang secara pribadi dan profesional, sehingga mereka dapat memberikan pelayanan medis yang terbaik untuk pasien dan masyarakat.

Alat ukur *psychological well-being* yang banyak digunakan pada orang dewasa dan diadaptasi di Indonesia dengan alat ukur terbaru adalah alat ukur Kesejahteraan Psikologis Remaja Indonesia Usia 12-15 Tahun (Abidin et al., 2023). Penggunaan alat ukur *Kesejahteraan Psikologis Remaja Indonesia Usia 12-15* menjadi pilihan yang relevan bagi dokter di Indonesia karena instrumen ini dikembangkan secara spesifik berdasarkan konteks budaya dan psikososial masyarakat Indonesia. Alat ukur ini telah melalui proses validasi dan reliabilitas yang ketat, sehingga memberikan hasil pengukuran yang akurat dan dapat dipercaya. Selain itu, instrumen ini menjawab kebutuhan akan alat asesmen lokal yang sesuai, yang sebelumnya masih didominasi oleh adaptasi dari instrumen luar negeri. Penelitian terdahulu menunjukkan bahwa kesejahteraan psikologis yang tinggi memiliki kesehatan mental dan fisik yang baik, hal tersebut sangat diperlukan agar dapat melakukan proses aktualisasi diri dan merasa berfungsi secara optimal (Ryff, 2013). Pada pembahasan ini peneliti lebih fokus pada alat tersebut dikarenakan dapat mengukur 10 dimensi: (1) afek positif (3 butir pernyataan), (2) afek negatif (3 butir pernyataan), (3) kepercayaan diri (3 butir pernyataan), (4) hubungan positif (3 butir pernyataan), (5) orientasi masa depan (2 butir pernyataan), (6) pengembangan diri (3 butir pernyataan), (7) pengelolaan emosi negatif (3 butir pernyataan), (8) tanggung jawab (3 butir pernyataan), (9) kemampuan menyelesaikan masalah (2 butir pernyataan), dan (10) harmoni (3 butir pernyataan).

Pada penelitian ini penggunaan model *Rasch* pada 70 dokter penting karena kemampuannya untuk menguji kualitas dan validitas instrumen secara objektif. Model ini dapat mengidentifikasi apakah aitem dalam instrumen berfungsi dengan baik pada sampel yang berbeda dan menangani potensi bias, baik dari sisi responden maupun item. Selain itu, *Rasch* memungkinkan analisis data ordinal menjadi skala interval yang lebih akurat, sehingga memberikan pengukuran yang lebih valid. Dengan sampel yang relatif kecil, model *Rasch* tetap efisien dan memberikan hasil yang mendalam mengenai reliabilitas dan validitas instrumen *psychological well-being* yang digunakan. Penelitian ini menggunakan *Rasch* model *rating scale model* (RSM), model ini mempermudah peneliti dalam mengolah data yang diperoleh dari skala *Likert* dan memungkinkan peneliti untuk menangani bias, menguji validitas dan memastikan pengukuran yang adil bagi semua responden.

Secara keseluruhan, penerapan model *Rasch* pada pengukuran *psychological well-being* dokter mengandalkan asumsi unidimensionalitas. *Psychological well-being* mengukur satu konstruk atau dimensi tunggal. Dalam hal ini, seluruh item dalam instrumen *psychological well-being* dokter seharusnya mengukur aspek yang sama, yaitu *psychological well-being* dokter. Jika instrumen mengukur lebih dari satu dimensi, hasil analisis *Rasch* mungkin tidak akurat.

2. Metode Penelitian

Uraian Dalam penelitian sebelumnya, Fitria Abidin mengadaptasi instrumen *psychological well-being* untuk memastikan relevansi dan validitas pengukuran yang dilakukan di Indonesia berdasarkan faktor-faktor lokal yang mempengaruhi *psychological well-being*. Pada tahap awal penelitian, peneliti meminta izin dari peneliti Alat Ukur Kesejahteraan Psikologis Remaja Indonesia Usia 12-15 Tahun (Abidin et al., 2023) untuk mengadaptasi dan memvalidasi instrumen, selanjutnya peneliti melakukan *expert review* yang dilakukan dengan delapan dokter (empat dokter umum dan empat dokter spesialis) berusia antara 25–60 tahun, untuk memastikan bahwa aitem-aitem dalam instrumen yang digunakan untuk mengukur *psychological well-being* dapat dipahami dengan benar oleh responden. Peneliti juga melakukan *peer review* yang melibatkan peneliti lain yang memiliki latar belakang dalam bidang yang sama, untuk memberikan umpan balik terhadap instrumen yang digunakan. Setelah validitas konten diuji, instrumen disebar dalam bentuk kuesioner online (*G-form*). Jenis penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif. Metode ini sebagai metode ilmiah atau saintifik karena memenuhi kaidah-kaidah ilmiah yaitu konkrit, objektif, terukur, rasional dan sistematis (Sugiyono, 2019).

Partisipan yang terlibat dalam penelitian ini adalah 70 dokter yang aktif berpraktek di Indonesia, 70 sampel tersebut mewakili populasi yang cukup heterogen dan mencakup variasi dalam skor atau dimensi yang ingin diukur. Terbagi dua kelompok yaitu dokter umum dan dokter spesialis. Sepanjang periode pengumpulan data dua minggu, partisipan yang secara sukarela setuju untuk berpartisipasi mengisi *informed consent* dan kuesioner penelitian.

Teknik pengambilan *sample* pada penelitian ini yaitu *purposive sampling*, dengan kriteria usia 25–60 tahun, aktif berpraktek di Indonesia. Dengan mengambil sampel dokter dalam rentang usia 25-60 tahun, peneliti berusaha untuk mendapatkan perspektif yang beragam dan mendalam dari berbagai dokter yang memiliki pengalaman medis yang relevan. Hal ini memungkinkan evaluasi instrumen tersebut valid secara konten dan dapat dipahami oleh target pengguna yang lebih luas, serta mencakup berbagai dimensi *psychological well-being* yang perlu diukur.

Teknik pengumpulan data dalam penelitian ini menggunakan kuesioner adaptasi Alat Ukur Kesejahteraan Psikologis Remaja Indonesia Usia 12-15 tahun (Abidin et al., 2023). Model skala yang digunakan penelitian ini menggunakan skala likert dengan 5 pilihan jawaban, yaitu dari Sangat tidak sesuai (1), Tidak sesuai (2), Netral (3), Sesuai (4), Sangat sesuai (5). Dalam model *Rasch*, WMS (*Weighted Mean Square*) dan UMS (*Unweighted Mean Square*) adalah dua statistik yang digunakan untuk mengevaluasi sejauh mana data yang dikumpulkan dari responden sesuai dengan model matematis yang diusulkan oleh *Rasch*. Kedua statistik ini berfokus pada kesesuaian data (*fit*), yang penting untuk memastikan bahwa instrumen yang digunakan untuk mengukur suatu konstruk psikologis atau karakteristik tertentu berfungsi dengan baik. Dengan panduan *Rating Scale Instrument Quality Criteria*, Fisher, W.P. Jr, menunjukkan bahwa *item reliability* sangat baik rentan .91 - .94, sedangkan *person reliability* tergolong buruk dengan rentan <.67. Hal ini menunjukkan adanya konsistensi partisipan dalam menjawab item-item pada instrumen, yang mencerminkan *person reliability* yang baik. *Person reliability* menggambarkan sejauh mana respons partisipan dapat diandalkan atau konsisten dalam konteks pengukuran. Jika nilai *person reliability* rendah, maka hal tersebut dapat mengindikasikan bahwa partisipan tidak merespons secara konsisten terhadap item yang disajikan, yang bisa disebabkan oleh ambiguitas aitem, kurangnya pemahaman, atau faktor eksternal lain. Dalam situasi tersebut, peneliti memerlukan penyesuaian terhadap instrumen, seperti merevisi item yang bermasalah atau mempertimbangkan penggunaan teknik analisis lanjutan, seperti analisis *Rasch* atau pemodelan *item response theory* (IRT), guna meningkatkan kualitas pengukuran (Boone et al., 2014). Aitem yang digunakan dikategorikan sangat baik dengan nilai pemisahan strata dari idealnya >3 dan dikategorikan aitem yang bervariasi dari paling mudah sampai tingkat kesukaran yang tinggi sampai. Pemisahan strata berfungsi untuk mengidentifikasi dan memetakan perbedaan kemampuan atau karakteristik individu (partisipan) serta kesulitan aitem (soal atau pernyataan) dalam pengukuran skala.

Uji reabilitas dalam penelitian ini menggunakan teknik alpha cronbarch. Menurut Sugiono (2016), jika nilai alfa cronbarch mencapai >.7 maka dikatakan *reliable*.

3. Hasil

Sample terdiri dari 70 dokter yang tersebar di Indonesia, (54,2% laki-laki dan 45,8% perempuan) yang usianya berkisar antara 25- 60 tahun ($SD = .80$). Sebagian besar responden merupakan dokter umum (71%), dan sisanya adalah dokter spesialis (29%). Pada penelitian ini menggunakan analisis *Rasch* dengan perangkat lunak (*software*) jMetrik version 4.1.1. Untuk mendapatkan butir soal yang fit untuk alat ukur *psychological well-being* pada dokter di Indonesia, perhitungan reabilitas pada alat ukur penelitian ini dengan mengadaptasi alat ukur Fitria Abidin, maka didapat nilai reabilitas sebesar .886, $SD = 16.66$ dengan $N=70$.

Tabel 1. Aitem *psychological well-being* pada dokter

Butir Soal	Pernyataan	Difficulty	WMS	UMS
1	Dalam satu bulan terakhir ini, saya merasa senang	.24	.67	.60
2	Dalam satu bulan terakhir, saya merasa bahagia	.21	.58	.59
3	Dalam satu bulan terakhir, merasa damai	-.10	.74	.70
4	Dalam satu bulan terakhir, saya merasa sedih*	-.58	.94	1.33
5	Dalam satu bulan terakhir, saya merasa marah*	-.41	.69	.73
6	Dalam satu bulan terakhir, saya merasa kesal*	-.69	.90	.90
7	Saya percaya dengan kemampuan yang saya miliki	.80	1.05	1.28
8	Saya percaya diri dengan apa yang saya lakukan	.91	.96	.78
9	Saya ragu dengan diri saya*	-.32	.72	.72
10	Saya adalah orang yang mudah mengerti perasaan rekan kerja saya	.06	.95	1.12
11	Saya adalah orang yang mau membantu rekan kerja	.67	.68	.60
12	Saya siap dimintai tolong oleh rekan kerja saya di jam kerja maupun di luar jam kerja	-.23	.92	.88
13	Saya memiliki rencana untuk kemajuan karier	.48	1.10	.93
14	Saya bingung jika ditanya mengenai rencana karier saya, karena belum terbayang sama sekali*	-.26	1.00	.91
15	Sampai dengan saat ini, saya belum tahu apa yang saya sukai untuk saya lakukan di luar jam kerja*	-.06	.82	.88
16	Saya tidak tertarik mengikuti kegiatan di luar jam kerja*	-.24	.96	.89
17	Saya berlatih untuk mengembangkan kemampuan di bidang yang saya minati (misal: kesenian, olahraga, pelajaran, dll)	.10	1.03	.96
18	Saat saya merasa kesal, saya tahu penyebabnya	-1.43	1.54	1.43
19	Saat merasa marah, saya tahu cara untuk mengatasinya	-1.36	1.93	1.86
20	Saya berusaha untuk menahan diri tidak membalas orang yang menyinggung perasaan saya	.30	1.03	1.21
21	Jika atasan saya tidak mengingatkan, berarti saya tidak perlu mengerjakan tugas di luar jam kerja	-.63	1.16	1.59
22	Saya melakukan tugas tanpa harus diingatkan	.41	1.22	1.14
23	Saya tidak ingat harus melakukan ibadah rutin	-.12	1.22	1.16
24	Saya tahu kapan saya meminta bantuan teman, kapan saya meminta bantuan rekan kerja/pihak lainnya untuk masalah saya	.36	1.16	1.32
25	Saya yakin memiliki kemampuan untuk menyelesaikan masalah yang saya hadapi	.63	.69	.64
26	Mengerjakan pekerjaan dan berteman sama-sama penting	.80	1.09	3.95
27	Saya berusaha untuk tidak tertinggal dalam pekerjaan maupun dalam pertemanan	.59	.90	1.05
28	Pasangan saya menilai saya lebih mementingkan teman dibanding keluarga	-.10	1.45	1.33

Unfavorable : 4, 5, 6, 9, 14, 15, 16, 18, 19, 21, 23, 28

Favorable : 1, 2, 3, 7, 8, 10, 11, 12, 13, 17, 20, 22, 24, 25, 26, 27

Tabel 2. Quality Statistics

SCALE QUALITY STATISTICS		
Statistic	Items	Persons
Observed Variance	0.3445	0.4262
Observed Std. Dev	0.5869	0.6529
Mean Square Error	0.0141	0.0558
Root MSE	0.1189	0.2361
Adjusted Variance	0.3304	0.3705
Adjusted Std. Dev.	0.5748	0.6087
Separation Index	4.8325	2.5775
Number of Strata	6.7766	3.7700
Reliability	0.9589	0.8692

Hasil dari penelitian ini yang disebarkan kepada partisipan dokter dengan nilai reabilitas *items* .95 dan reabilitas *persons* .86, menggambarkan bahwa partisipan dokter dikategorikan baik dan tepat pada alat ukur penelitian ini. Konsistensi jawaban dari partisipan penelitian dikategorikan baik sehingga kualitas butir aitem dalam instrumen aspek reabilitasnya baik. Aitem yang digunakan dikategorikan sangat baik dengan nilai pemisahan strata 6.77 dari idealnya >3 dan dikategorikan aitem yang bervariasi dari paling mudah sampai tingkat kesukaran yang tinggi sampai. Partisipan dalam penelitian ini dikategorikan memiliki variasi *psychological well-being* dengan nilai 3.77. Nilai *separation* menunjukkan bahwa kualitas instrumen dikategorikan sangat baik 4.83 dan dokter yang tersebar di Indonesia pada penelitian ini dianggap cukup mampu menjadi partisipan pada penelitian ini.

Tabel 3. Item Statistics

FINAL JMLE ITEM STATISTICS						
Item	Difficulty	Std. Error	WMS	Std. WMS	UMS	Std. UMS
pwb1	0.24	0.12	0.67	-2.18	0.60	-2.29
pwb2	0.21	0.12	0.58	-2.96	0.59	-2.43
pwb3	-0.10	0.11	0.74	-1.95	0.70	-1.91
pwb4	-0.58	0.09	0.94	-0.46	1.33	2.05
pwb5	-0.41	0.10	0.69	-2.63	0.73	-1.76
pwb6	-0.69	0.09	0.90	-0.79	0.90	-0.70
pwb7	0.80	0.16	1.05	0.31	1.28	1.09
pwb8	0.91	0.16	0.96	-0.11	0.78	-0.83
pwb9	-0.32	0.10	0.72	-2.17	0.72	-1.77
pwb10	0.06	0.11	0.95	-0.31	1.12	0.69
pwb11	0.67	0.15	0.68	-1.65	0.60	-1.91
pwb12	-0.23	0.10	0.92	-0.58	0.88	-0.71
pwb13	0.48	0.13	1.10	0.56	0.93	-0.26
pwb14	-0.26	0.10	1.00	0.03	0.91	-0.50
pwb15	-0.06	0.11	0.82	-1.16	0.88	-0.57
pwb16	-0.24	0.10	0.96	-0.23	0.89	-0.61
pwb17	0.10	0.11	1.03	0.26	0.96	-0.16
pwb18	-1.43	0.10	1.54	3.28	1.43	2.43
pwb19	-1.36	0.10	1.93	5.33	1.86	4.48
pwb20	0.30	0.12	1.03	0.21	1.21	1.03
pwb21	-0.63	0.09	1.16	1.28	1.59	3.42
pwb22	0.41	0.13	1.22	1.20	1.14	0.70
pwb23	-0.12	0.11	1.22	1.36	1.16	0.87
pwb24	0.36	0.13	1.16	0.92	1.32	1.45
pwb25	0.63	0.14	0.69	-1.68	0.64	-1.69
pwb26	0.80	0.16	1.09	0.49	3.95	6.98
pwb27	0.59	0.14	0.90	-0.48	1.05	0.29
pwb28	-0.10	0.11	1.45	2.57	1.33	1.64

Pada analisis *Rasch*, *Difficulty* dapat dianggap sebagai *Z score* dengan nilai normal antara -0.3 sampai +3. Disimpulkan bahwa aitem nomor 18 merupakan aitem paling mudah dan aitem nomor 8 merupakan aitem paling sulit bagi partisipan. Nilai WMS dan UMS pada Tabel 2 dapat dinyatakan sebagai aitem yang ideal dengan rentang *score* 0.5 – 1.5, aitem nomor 26 dengan nilai dengan tingkat aitem yang mudah dapat dinyatakan sebagai aitem yang tidak cocok dengan partisipan dokter, sehingga aitem tersebut *drop* pada penelitian ini. Pada dimensi harmoni, hanya dua butir aitem yang memenuhi kriteria analisis. Namun, jumlah ini belum mencukupi untuk menghasilkan pengukuran yang reliabel dan valid. Dengan hanya dua item, variasi respons individu cenderung terbatas, yang dapat mengurangi akurasi dan keandalan pengukuran. Hal ini juga membuat analisis *misfit* (ketidaksiharian) menjadi lebih sulit dilakukan karena kesulitan dalam mendeteksi pola respons yang tidak sesuai.

4. Pembahasan

Penelitian ini bertujuan untuk menguji validitas dan reliabilitas instrumen penelitian, mendapatkan butir aitem yang fit, analisis kualitas aitem, mengetahui tingkat abilitas dokter dan tingkat kesulitan soal, bahkan menguji kesesuaian *person* dan *item* secara simultan pada alat ukur *psychological well-being* pada dokter yang diadaptasi dari Alat Ukur Kesejahteraan Psikologis Remaja Indonesia (Abidin et al., 2023), yang kemudian diuji menggunakan metode analisis *Rasch* dengan software jMetrik version 4.1.1.

Hasil uji reliabilitas menunjukkan bahwa instrumen memiliki tingkat keandalan yang baik dengan nilai reliabilitas sebesar 0.886 (SD = 16.66, N = 70). Selain itu, hasil analisis menunjukkan bahwa reliabilitas aitem mencapai 0.95, sementara reliabilitas *person* sebesar 0.86. Nilai ini mengindikasikan bahwa konsistensi partisipan dalam menjawab tergolong baik, sehingga kualitas butir aitem dalam aspek reliabilitasnya dapat dikategorikan sangat baik.

Dari hasil penelitian ini, terdapat beberapa temuan yang perlu diperhatikan dalam penelitian ini. Hasil analisis *Rasch* menunjukkan bahwa aitem nomor 18 merupakan aitem paling mudah, sementara aitem nomor 8 adalah yang paling sulit bagi partisipan. Selain itu, nilai *Weighted Mean Square* (WMS) dan *Unweighted Mean Square* (UMS) menunjukkan bahwa sebagian besar aitem berada dalam rentang skor ideal (0.5 – 1.5), yang berarti instrumen memiliki kecocokan yang baik dengan model *Rasch*. Akan tetapi, aitem nomor 26 ditemukan tidak sesuai dengan karakteristik partisipan dokter dan akhirnya di-drop dari penelitian ini. Menyatakan bahwa tidak sesuai bagi dokter mengerjakan pekerjaan dan berteman adalah hal yang sama-sama penting pada dimensi harmoni yang ada pada Alat Ukur Kesejahteraan Psikologis Remaja Indonesia, hal ini didukung oleh beban pekerjaan dokter mengacu pada jumlah dan struktur pekerjaan dan sering diukur sebagai total jam kerja sehingga mengalami kesulitan dalam bersosialisasi (Pratama, 2023). Dokter dengan tingkat beban kerja yang tinggi dan dukungan sosial yang rendah cenderung mengalami kelelahan emosional, yang dapat berkontribusi pada penurunan kualitas komunikasi dan interaksi interpersonal (Smith et al., 2017). Selain itu, penelitian oleh Bakker & Demerouti (2007) dalam framework *Job Demands-Resources* (JD-R) juga menemukan bahwa beban kerja yang tinggi, dapat mengarah pada kelelahan dan *disengagement*, yang akhirnya berdampak pada kualitas hubungan profesional dan pribadi.

Hasil dari penelitian ini menunjukkan bahwa ketika individu menghadapi tuntutan pekerjaan yang tinggi tanpa diimbangi dengan kontrol kerja yang memadai serta dukungan sosial yang cukup, mereka berisiko mengalami stres yang tinggi dan *burnout*. Tuntutan pekerjaan yang tinggi pada dokter berhubungan dengan peningkatan risiko *burnout*, sementara kontrol kerja dan dukungan rekan kerja yang memadai berhubungan dengan penurunan risiko tersebut (Christiansen et al., 2024). Banyak aspek yang menunjang hal tersebut, diantaranya meningkatnya tanggung jawab pekerjaan dokter, diwajibkan untuk *multitasking*, perlunya penyesuaian dengan teknologi baru dalam pekerjaan dan tempat kerja yang kompetitif. Dokter juga menyadari bahwa pekerjaan mereka mengacu pada pertumbuhan pribadi mereka, memiliki tujuan yang jelas dan terarah di pekerjaannya (Muntean et al., 2022). Sehingga dokter tidak mau mengambil resiko yang berdampak pada pekerjaan dan jenjang karier mereka (Panagioti et al., 2018).

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa instrumen *psychological well-being* yang telah diadaptasi memiliki kualitas yang baik dalam mengukur *kesejahteraan psikologis* dokter di Indonesia. Meski demikian, beberapa perbaikan dan pengembangan lebih lanjut diperlukan untuk meningkatkan keakuratan dan validitas instrumen, khususnya dalam aspek tertentu seperti jumlah aitem dalam dimensi harmoni. Diharapkan pada penelitian selanjutnya menambah jumlah aitem untuk dimensi tersebut guna meningkatkan keandalan dan representasi dimensi secara lebih akurat. Jika penambahan aitem tidak memungkinkan, lakukan analisis *misfit* untuk memastikan kesesuaian aitem dengan model *Rasch* dan pertimbangkan untuk menggabungkan dimensi yang serupa. Dengan perbaikan ini, diharapkan instrumen yang dikembangkan dapat lebih optimal dalam mengukur *psychological well-being* di kalangan tenaga medis dan berkontribusi dalam penelitian-penelitian selanjutnya di bidang *psychological well-being* dokter di Indonesia.

menggunakan Times New Roman 10, spasi 1,0, rata kiri kanan. Bagian ini berisikan interpretasi dan paparan hasil analisis yang diperoleh dan dikaitkan dengan temuan atau teori. Penulis perlu memaparkan argumen kritis atas kemungkinan hal-hal yang dapat mempengaruhi hasil penelitiannya, seperti dikaitkan dengan karakteristik demografi partisipan penelitian atau dikaitkan dengan faktor lain diluar variabel yang diteliti, dan mengemukakan argumen secara logis, rasional, dan kritis. (isi dapat dituliskan font Times new roman 10, spasi 1,0, rata kanan kiri)

5. Kesimpulan

Penelitian ini berhasil mengadaptasi dan memvalidasi instrumen *psychological well-being* bagi dokter yang aktif berpraktik di Indonesia dengan menggunakan perangkat lunak (*software*) jMetrik version 4.1.1. Hasil analisis menunjukkan bahwa instrumen yang digunakan memiliki tingkat reliabilitas yang baik dengan nilai reliabilitas sebesar 0.886 serta nilai pemisahan strata sebesar 6.77, yang mengindikasikan daya diskriminasi yang baik dalam mengukur kesejahteraan psikologis dokter.

Meskipun demikian, terdapat beberapa aspek yang perlu diperbaiki, seperti jumlah aitem dalam dimensi harmoni yang masih terbatas dan perlu ditingkatkan untuk memastikan keakuratan pengukuran. Selain itu, beberapa aitem ditemukan kurang sesuai dengan karakteristik partisipan dan telah di-drop dari penelitian ini. Oleh karena itu, pengembangan lebih lanjut diperlukan untuk meningkatkan validitas dan representasi instrumen.

Secara keseluruhan, instrumen ini memiliki potensi besar untuk digunakan dalam penelitian dan praktik profesional guna memahami serta meningkatkan *psychological well-being* dokter di Indonesia. Perbaikan dan pengembangan lebih lanjut diharapkan dapat menjadikan instrumen ini lebih optimal dan aplikatif dalam berbagai konteks medis.

6. Referensi

- Abidin, F. A., Yudiana, W., & Fadilah, S. H. (2022). Parenting style and emotional well-being among adolescents: the role of basic psychological needs satisfaction and frustration. *Frontiers in psychology, 13*, 901646.
- Akhtar, H. (2019). Evaluasi Properti Psikometris Dan Perbandingan Model Pengukuran Konstruksi Subjective Well-Being. *Jurnal Psikologi, 18*(1), 29–40. <https://doi.org/10.14710/jp.18.1.29-40>
- Amalia, S. (2016). Analisa psikometrik alat ukur Ryff's Psychological Well-Being (RPWB) versi bahasa Indonesia: Studi pada lansia guna mengukur kesejahteraan dan kebahagiaan. In *Seminar Asean 2nd Psychology & Humanity* (pp. 430-437).
- Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of managerial psychology, 22*(3), 309-328.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences* (Vol. 10, pp. 978-94). Dordrecht: Springer.
- Christiansen, F., Gynning, B. E., Lashari, A., Zuberbühler, J. P., Johansson, G., & Brulin, E. (2024). Associations between job demand-control-support and high burnout risk among physicians in Sweden: a cross-sectional study. *Journal of Occupational Medicine and Toxicology, 19*(1), 42.
- Furlong, M. J., Gilman, R., & Huebner, E. S. (Eds.). (2014). *Handbook of positive psychology in schools*. New York, NY: Routledge.
- Gómez-López, M., Viejo, C., & Ortega-Ruiz, R. (2019). Psychological well-being during adolescence: Stability and association with romantic relationships. *Frontiers in psychology, 10*, 465708.
- Maslach, C., & Leiter, M. P. (2016). Burnout. In *Stress: Concepts, cognition, emotion, and behavior* (pp. 351-357). Academic Press.
- Muntean, L. M., Nireștean, A., Popa, C. O., Strete, E. G., Ghiga, D. V., Sima-Comaniciu, A., & Lukacs, E. (2022). The relationship between Emotional Stability, Psychological Well-being and life satisfaction of Romanian Medical doctors during COVID-19 period: a cross-sectional study. *International Journal of Environmental Research and Public Health, 19*(5), 2937.
- Panagioti, M., Geraghty, K., Johnson, J., Zhou, A., Panagopoulou, E., Chew-Graham, C., ... & Esmail, A. (2018). Association between physician burnout and patient safety, professionalism, and patient satisfaction: a systematic review and meta-analysis. *JAMA internal medicine, 178*(10), 1317-1331.
- Pratama, R., & Yufika, A. (2023). Physicians' Workload and Quality Healthcare in Indonesia. *Trends in Infection and Global Health, 3*(1), 43-55.
- Ryan, R., and Deci, E. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annu. Rev. Psychol. 52*, 141–166. doi: 10.1146/annurev.psych.52.1.141

- Ryff, C. D. (2017). Eudaimonic well-being, inequality, and health: Recent findings and future directions. *International Review of Economics*, 64(2), 159–178. <https://doi.org/10.1007/s12232-017-0277-4>
- Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of personality and social psychology*, 69(4), 719.
- Schneider, J., Talamonti, D., Gibson, B., & Forshaw, M. (2022). Factors mediating the psychological well-being of healthcare workers responding to global pandemics: A systematic review. *Journal of Health Psychology*, 27(8), 1875-1896.
- Smith, F., Goldacre, M. J., & Lambert, T. W. (2017). Adverse effects on health and wellbeing of working as a doctor: views of the UK medical graduates of 1974 and 1977 surveyed in 2014. *Journal of the Royal Society of Medicine*, 110(5), 198-207.
- Sugiyono, P. D. (2019). Buku sugiyono, metode penelitian kuantitatif kualitatif. *Revista Brasileira de Linguística Aplicada*, 5(1).
- Sumintono, B., & Widhiarso, W. (2015). Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Cimahi: Trim Komunikata.
- Sunardy, G. N., Abidin, F. A., & Qodariah, L. (2023). Adolescents' Psychological Well-being: Adaptation and Validation of the Brief Scale of Psychological Well-Being for Adolescents (BSPWB-A) in Indonesia. *Psychology Hub*, 40(3), 51-58.
- Van Dierendonck, D. (2004). The construct validity of Ryff's Scales of Psychological Well-being and its extension with spiritual well-being. *Personality and individual differences*, 36(3), 629-643.
- West, C. P., Dyrbye, L. N., & Shanafelt, T. D. (2018). Physician burnout: contributors, consequences and solutions. *Journal of internal medicine*, 283(6), 516-529.



ADAPTASI INSTRUMEN *DEPRESSION LITERACY (D-LIT) QUESTIONNAIRE* VERSI BAHASA INDONESIA DENGAN MENGGUNAKAN METODE *RASCH MODEL*

Rizkiani Amalia,¹ Devie Yundianto²

^{1,2}Fakultas Psikologi, Universitas Nahdlatul Ulama Indonesia, Jakarta, Indonesia

Email: rizkianiamalia932@gmail.com

Received :	Revised :	Accepted :	Published :
29 Desember 2024	30 Maret 2025	15 April 2025	30 April 2025

Abstract

Limited knowledge and misconceptions about psychological disorders, especially depression, often lead to negative attitudes towards sufferers. Efforts are needed to address this issue to reduce unwanted impacts. This study aims to determine how people's knowledge of depression and to adapt and analyze the psychometric properties of the *Depression Literacy (D-Lit) Questionnaire* instrument. This study analyzed the adapted *Depression Literacy (D-Lit) Questionnaire* instrument, with 225 participants aged 18-40 years, which later became 195 participants due to elimination. The analysis method used was the *Rasch Model*. This study looks at the reliability of persons and items and a rating scale based on the *Andrich Threshold*. The results of the analysis obtained were that the item reliability value reached 0.99 and the person reliability value with a value reached 0.64. Alpha Cronbach's reliability was found to be 0.68, indicating sufficient interaction between item and person. There were 30 persons that did not meet the invariance test criteria so the researcher eliminated 4 items because they did not meet the requirements of *Outfit MNSQ*, *ZTSD* and *Pt Mean Corr*. The raw variance contained in the analysis is 44.3%, where this value has met the unidimensionality requirement. The study conducted almost all met the requirements of the *Rasch Model*.

Keywords: *Depression Literacy, Rasch Model, Psychometric Properties*

Abstrak

Keterbatasan pengetahuan dan kesalahpahaman mengenai gangguan psikologis, khususnya depresi, seringkali menimbulkan sikap negatif terhadap penderitanya. Diperlukan upaya untuk mengatasi masalah ini guna mengurangi dampak yang tidak diinginkan. Penelitian ini memiliki tujuan untuk mengetahui bagaimana pengetahuan masyarakat terhadap depresi serta untuk mengadaptasi dan menganalisis properti psikometris instrumen *Depression Literacy (D-Lit) Questionnaire*. Penelitian ini menganalisis instrumen *Depression Literacy (D-Lit) Questionnaire* yang telah diadaptasi, dengan partisipan berusia 18-40 tahun sebanyak 225 orang yang kemudian menjadi 195 partisipan karena proses dieliminasi. Metode analisis yang digunakan adalah *Rasch Model*. Penelitian ini melihat reliabilitas *person* dan butir serta *rating scale* berdasarkan ambang batas *Andrich Threshold*. Hasil analisis yang didapat adalah nilai reliabilitas butir mencapai 0,99 dan nilai reliabilitas *person* dengan nilai 0,64. Reliabilitas Alpha Cronbach ditemukan sebesar 0,68 yang menunjukkan bahwa terdapat interaksi yang cukup antara butir dan person. Terdapat 30 person yang tidak memenuhi kriteria uji invariansi sehingga dieliminasi, serta 4 butir yang dieliminasi karena tidak memenuhi persyaratan *Outfit MNSQ* dan *ZTSD* serta *Pt Mean Corr*. *Raw variance* yang terdapat dalam analisis senilai 44,3% dimana nilai tersebut telah memenuhi persyaratan unidimensi. Analisis yang dilakukan hampir semua memenuhi persyaratan *Rasch Model*.

Kata Kunci: *Depression Literacy, Rasch Model, Properti Psikometris*.

1. Pendahuluan

Perkembangan teknologi dan ilmu pengetahuan terkini telah membawa perubahan dalam kehidupan manusia, khususnya dalam lingkup kesehatan jiwa (Annury et al., 2022). *World Health Organization* (WHO) menyatakan bahwa 1 dari 4 orang menderita gangguan psikologis seumur hidup mereka dan salah satu yang paling umum terjadi pada populasi dunia adalah depresi. Hal tersebut dikarenakan depresi memengaruhi orang-orang dari semua kalangan usia baik tua maupun muda, dan dapat menimbulkan dampak yang serius, seperti gangguan suasana hati yang akan berpengaruh ke kehidupan sosialnya, fungsional yang tidak normal, dan dalam kasus ekstrim dapat menyebabkan bunuh diri (Dhaliwal et al., 2022; Saleh et al., 2023).

Depresi adalah salah satu gangguan psikologis yang sering dijumpai. Gangguan ini ditandai dengan keadaan dimana seseorang merasa putus asa dan hilangnya minat terhadap aktivitas yang sebelumnya dianggap menyenangkan (Al-Azzam et al., 2013). Depresi juga merupakan kondisi dimana seseorang merasa sedih, kecewa saat mengalami suatu perubahan, kehilangan, merasa gagal, dan menjadi patologis ketika merasa tidak mampu untuk beradaptasi. Pengertian lainnya juga disebutkan bahwa depresi merupakan suatu keadaan yang dapat memengaruhi seseorang secara afektif, fisiologis, kognitif, dan perilaku, sehingga dapat mengubah pola dan respon tidak seperti biasanya (Hadi et al., 2017).

Menurut laporan WHO, 27% orang dengan depresi di seluruh dunia berada di Asia tenggara. Prevalensi nasional depresi pada generasi muda di Indonesia dilaporkan sebesar 29,3% untuk depresi sedang dan 8,0% untuk depresi berat. Depresi sekitar 29,0% lebih banyak terjadi pada pria dibandingkan wanita dengan persentase 25,5% (Purborini et al., 2021). Tidak hanya di Indonesia, Kasus depresi selalu mendapat angka yang tinggi hampir di setiap negara, namun sebenarnya depresi dapat diobati. Hanya saja, di banyak negara, tidak lebih dari 10% orang yang menderita depresi menerima perawatan dan dukungan yang mereka butuhkan. Kurangnya sumber daya, layanan kesehatan mental, stigma sosial, dan kurangnya literasi kesehatan mental menjadi hambatan dalam menerima pengobatan gangguan mental, khususnya depresi (Nguyen Thai & Nguyen, 2018).

Banyak penderita depresi yang tidak terdiagnosis dan tidak mendapat pengobatan yang tepat, meskipun mereka menunjukkan gejala-gejala depresi (Al-Azzam et al., 2013). Orang dengan penyakit mental mungkin tidak mencari bantuan atau menunda mencari bantuan dikarenakan berbagai hambatan, termasuk ketakutan akan penilaian dan reaksi dari orang-orang terdekatnya, serta diskriminasi yang berkaitan dengan depresi. Bahkan, beberapa orang berhenti mengonsumsi antidepresan karena alasan ini, karena sikap serta kurangnya dukungan dari orang-orang sekitar mereka (Nguyen Thai & Nguyen, 2018). Sikap masyarakat terhadap penyakit mental yang cenderung negatif ini, terutama ketika pengetahuan dan keyakinan mengenai penyakit mental terbatas, sehingga masyarakat terus mempunyai sikap negatif terhadap penderita penyakit mental dan dapat dengan cepat memengaruhi orang lain. Oleh karena itu, perlu adanya peningkatan kesadaran masyarakat untuk mengurangi stigma dan diskriminasi terhadap penderita gangguan mental, yang dalam kasus ini adalah depresi (Akinsulore et al., 2018).

Untuk meminimalisir sikap masyarakat yang dapat menimbulkan dampak buruk terhadap penderita depresi, diperlukannya pengetahuan atau literasi kesehatan mental (khususnya depresi) yang baik. Literasi kesehatan mental mencakup pengetahuan dan keterampilan yang membahas aspek biologis, psikologis, dan sosial dari kesehatan mental untuk meningkatkan pemahaman tentang kesehatan mental dan berbagai gangguan mental, mengurangi stigma, membantu mengenali dan mencegah gangguan mental, dan memfasilitasi perilaku pencarian bantuan. Literasi depresi merupakan perkembangan dari literasi kesehatan mental, dimana literasi ini meliputi pengetahuan dan keterampilan. Literasi depresi ini membahas seputar pengetahuan mengenai depresi seperti, pengertian, penyebab, gejala, penanganan, keterampilan untuk mengenali depresi, kemampuan membantu diri sendiri (*self-help*), dan pengetahuan mengenai bantuan profesional (Wahyuni & Nabila, 2022; Wei et al., 2013).

Sebuah penelitian menemukan bahwa lebih dari 50% masyarakat memiliki kesadaran rendah terhadap depresi (ALJadani et al., 2021). Kesadaran masyarakat merupakan faktor penting dalam pengelolaan depresi. Jika masyarakat memahami pengetahuan dan sikap yang semestinya terhadap penderita gangguan mental, maka stigma, khususnya terhadap orang dengan diagnosis depresi akan berkurang. Mayoritas orang, sadar atau tidak, akan berkontak langsung dengan penderita gangguan mental di dalam hidupnya. Oleh sebab itu, kesadaran dan pengetahuan tentang gangguan jiwa sangat diperlukan (Saleh et al., 2023). Hal tersebut terbukti dengan temuan penelitian bahwa individu dengan tingkat pendidikan yang lebih tinggi memiliki kesadaran masyarakat yang lebih tinggi terhadap depresi. Meskipun tidak dapat digeneralisasikan bahwa orang dengan tingkat pendidikan rendah memiliki kesadaran yang rendah tentang depresi, namun ada kemungkinan bahwa orang yang memiliki pengetahuan khusus tentang depresi dapat meningkatkan kesadarannya akan penyakit mental tersebut (Aletesh et al., 2021).

Sekalipun banyak informasi yang dapat diperoleh dari pengetahuan ilmiah mengenai depresi, namun tidak sedikit kepercayaan dan kesalahpahaman masyarakat tentang gangguan kesehatan mental berkontribusi terhadap

stigma yang terjadi (Kulwicka & Gasiorowska, 2023). Tidak sedikit pula orang-orang yang menyadari bahwa pengetahuan mengenai gangguan psikologis itu penting, namun masih belum memahami bagaimana kondisi individu saat menderita depresi. Anggapan-anggapan yang dipercaya masyarakat kurang baik seperti penderita depresi adalah orang yang lemah, harus bertanggung jawab atas kondisinya sendiri, dan tidak jarang masyarakat menganggap bahwa penderita depresi itu berbahaya. Perlakuan tersebut membuat penderita depresi enggan mencari bantuan kepada profesional. Rendahnya literasi mengenai penyakit mental, menimbulkan stigma di masyarakat yang membuat penderita depresi tetap berada dalam situasi tersebut tanpa mencari bantuan (Akinsulore et al., 2018).

Selain itu, intervensi yang tidak efektif juga menjadi salah satu penyebab rendahnya literasi depresi di kalangan masyarakat dan stigma yang masih ada di masyarakat. Selain itu juga, sumber informasi yang tersedia terbatas untuk mengetahui sumber informasi yang akurat (Kulwicka & Gasiorowska, 2023). Usia, jenis kelamin, pendidikan, budaya, serta kepribadian juga dapat menjadi faktor yang memengaruhi rendah atau tingginya literasi kesehatan mental, khususnya depresi (Nazira et al., 2022).

Oleh karena itu, diperlukan pengukuran untuk mengetahui tingkat pengetahuan masyarakat mengenai depresi dengan cara mengadaptasi alat ukur. Adaptasi alat ukur merupakan istilah yang merujuk pada penggunaan alat ukur yang biasanya berasal dari negara lain dengan bahasa dan budaya yang berbeda kemudian diadaptasi ke dalam bahasa dan budaya sesuai dengan daerah yang ditargetkan ("*ITC Guidelines for Translating and Adapting Tests (Second Edition)*," 2018). Sebelum melakukan adaptasi, peneliti harus memikirkan terlebih dahulu, apakah adaptasi alat ukur yang digunakan akan memiliki makna yang sama dalam suatu budaya yang akan dipelajari, atau apakah antara peneliti dan partisipan memiliki pemahaman yang sama mengenai konstruk tersebut. Proses adaptasi instrumen ini mengikuti panduan dari 18 pedoman yang dibagi menjadi 6 kategori untuk pengadaptasian tes menurut ITC. Sebuah studi oleh (Kulwicka & Gasiorowska, 2023) menemukan bahwa terdapat salah satu instrumen yang dapat digunakan untuk menilai literasi depresi dan mengukur pengetahuan dan keyakinan masyarakat tentang gangguan depresi, termasuk gejala, pengobatan dan pencegahan. Instrumen tersebut adalah *Depression Literacy (D-Lit) Questionnaire* yang dikembangkan oleh (Tehrani et al., 2022). Instrumen ini menunjukkan bahwa skor yang lebih tinggi menunjukkan tingkat literasi depresi yang lebih tinggi dan pemahaman yang lebih baik pula. Instrumen ini juga digunakan untuk mengevaluasi sejauh mana individu menyadari depresi dan memahami informasi yang relevan. Tujuan penelitian ini adalah untuk mengadaptasi dan menganalisis properti psikometris instrumen D-Lit dengan menggunakan *Rasch Model*.

2. Metode Penelitian

Partisipan

Penelitian ini menggunakan populasi pada orang dewasa dengan usia 18-40 tahun. Usia tersebut digunakan dalam sebuah penelitian depresi pada dewasa muda berusia 18-39 tahun (Brodaty et al., 1991). Peneliti memilih kelompok usia ini karena kelompok usia tersebut mewakili usia dewasa yang paling rentan mengalami gangguan kesehatan mental (Hakim & Aristawati, 2023). Penelitian ini menggunakan teknik *Convenience Sampling*, dimana metode pengambilan sampel tersebut memilih sampel yang aksesibilitasnya cukup mudah dijangkau oleh peneliti. Dengan itu, peneliti dapat memberikan kuesioner kepada orang-orang terdekatnya seperti teman, kerabat, keluarga, atau koleganya. Selain itu, teknik pengumpulan data yang digunakan dalam penelitian ini berupa pernyataan dalam bentuk kuesioner dengan menggunakan bantuan *google form* dan sosial media sebagai tempat untuk penyebaran kuesioner tersebut (Maidiana, 2021). Jumlah responden yang didapat dalam penelitian ini berjumlah berjumlah 225 orang dengan data demografis sebagai berikut.

Tabel 1. Data Demografis

Berdasarkan data dari tabel 1, mayoritas partisipan penelitian ini adalah wanita dengan persentase 82,6%

Kategori	Frekuensi	Persentase
Jenis Kelamin		
Pria	34	17,4%
Wanita	161	82,6%
Usia		
18-29	176	90,3%
30-40	18	9,2%
Tidak Menjawab	1	0,5%
Domisili		
Jabodetabek	85	43,6%
Luar Jabodetabek	109	55,9%
Tidak Menjawab	1	0,5%
Pendidikan Terakhir		
SMA/SMK	77	39,5%
S1/D3	115	59,0%
S2	3	1,5%
Masalah Psikologis		
Anxiety	78	40,0%
Depresi	53	27,2%
Lainnya	7	3,6%
Tidak Memiliki Masalah Psikologis	54	27,7%
Tidak Menjawab	3	1,5%
Mencari Informasi Depresi		
Tidak	16	8,2%
Ya	179	91,8%

dan pria 17,4%. Partisipan dengan usia paling banyak berada di usia 18-29 tahun sebanyak 90,3% dan usia 30-40 sebanyak 9,2%. Partisipan yang berdomisili di luar wilayah jabodetabek dengan persentase 55,9% lebih mendominasi dibandingkan yang berasal dari wilayah jabodetabek dengan jumlah 43,6% dalam penelitian ini. Mayoritas partisipan memiliki pendidikan terakhir D3/S1 sebesar 59%, kemudian SMA/SMK sebanyak 39,5% dan S2 sebanyak 1,5%. Sebagian partisipan melaporkan bahwa dirinya memiliki masalah psikologis, dengan rincian anxiety sebanyak 40%, depresi 27%, dan 3,6% mengalami masalah lain. Sementara itu, partisipan yang tidak memiliki masalah psikologis dengan data sebanyak 27,7% serta 3 partisipan tidak memberikan jawaban atas pertanyaan ini.

Instrumen

Peneliti mengadaptasi instrumen *Depression Literacy (D-Lit) Questionnaire* oleh Tehrani et al (2022). Kuesioner tersebut berisi 22 item dengan 5 subskala, yaitu pengetahuan tentang gejala psikologis, pengetahuan tentang gejala kognitif perilaku, pengetahuan tentang minum obat dan efek sampingnya, dan pengetahuan tentang tingkat keparahan penyakit. Skala likert yang akan digunakan berkisar antara 1-5, dengan keterangan Sangat tidak setuju (1), Tidak Setuju (2), Netral (3), Setuju (4), Sangat Setuju (5). Adaptasi instrumen ini mengikuti panduan ITC. Pada tahap awal, peneliti meminta izin kepada pemilik instrumen, lalu menerjemahkannya ke bahasa Indonesia dan melakukan *back translation*. Kemudian, *expert judgement* dilakukan untuk memastikan makna tetap sesuai, dengan revisi jika diperlukan. Selanjutnya, uji keterbacaan pada 5 responden lalu disebarluaskan melalui *Google Form*. Data yang terkumpul dianalisis dengan uji validitas, reliabilitas, unidimensionalitas, invariansi, independensi lokal, dan validitas skala peringkat. Terakhir pencatatan teknis atas perubahan yang terjadi selama proses analisis data.

Tabel 2. Instrumen D-Lit

Dimensi	No Item	Kode Item	Bunyi Item Adaptasi
Pengetahuan tentang gejala psikologis	1	PS1	Saya merasa orang yang mengalami depresi mungkin merasa bersalah meskipun mereka tidak melakukan kesalahan apapun.
	2	PS2	Saya merasa, hilangnya rasa kepercayaan diri dan rasa rendah diri merupakan tanda dari depresi.
	3	PS3	Saya merasa, kurang tidur atau tidur berlebihan merupakan gejala depresi.
	4	PS4	Saya merasa banyak makan atau kehilangan selera makan merupakan gejala depresi.
	5	PS5	Saya merasa orang mungkin bergerak lebih lambat atau menjadi gelisah karena depresi yang dialaminya.
Pengetahuan tentang efektivitas metode pengobatan yang tersedia	6	ET1	Saya merasa psikolog klinis dapat meresepkan obat antidepresan.
	7	ET2	Saya merasa banyak perawatan untuk depresi yang lebih efektif daripada obat antidepresan.
	8	ET3	Saya merasa efek konseling mirip dengan efek terapi perilaku kognitif untuk depresi.
	9	ET4	Saya merasa efek terapi perilaku kognitif sama dengan efek obat antidepresan untuk depresi tingkat ringan hingga tingkat sedang.
	10	ET5	Saya merasa saat seseorang mengalami depresi tingkat sedang dapat memengaruhi kehidupan seseorang seperti halnya penyakit yang berkaitan dengan gangguan saraf atau tuli.
Pengetahuan tentang gejala kognitif perilaku	11	CBS 1	Saya merasa orang dengan depresi sering berbicara tidak teratur dan relevan.
	12	CBS 2	Saya merasa perilaku ceroboh dan bodoh adalah gejala umum dari depresi.
	13	CBS 3	Saya merasa tidak berjalan di trotoar yang rusak bisa jadi gejala depresi.
	14	CBS 4	Saya merasa orang dengan depresi sering mendengar suara yang biasanya tidak nyata.
	15	CBS 5	Saya merasa depresi tidak memengaruhi daya ingat dan konsentrasi.
	16	CBS 6	Saya merasa, memiliki beberapa kepribadian yang berbeda dapat menjadi gejala depresi.
Pengetahuan tentang minum obat dan efek sampingnya	17	SE1	Saya merasa dari semua alternatif dan gaya hidup untuk mengatasi depresi, mengonsumsi vitamin adalah yang paling bermanfaat.
	18	SE2	Saya merasa penderita depresi harus berhenti minum obat antidepresan segera setelah mereka merasa lebih baik.
	19	SE3	Saya merasa obat antidepresan membuat ketagihan.
	20	SE4	Saya merasa obat antidepresan biasanya bekerja dengan cepat.
Pengetahuan tentang tingkat keparahan penyakit	21	SD1	Saya merasa sebagian besar penderita depresi perlu dirawat di rumah sakit (jiwa).
	22	SD2	Saya merasa banyak selebriti yang menderita depresi.

Terdapat total 22 item saat disebarkan pada responden dengan 5 subskala dan 5 opsi jawaban yang dimulai dari 1 (Sangat Tidak Setuju) hingga 5 (Sangat Setuju). Item-item di atas merupakan beberapa pernyataan tentang pengetahuan masyarakat tentang depresi. Instrumen tersebut dibagikan setelah melalui proses adaptasi menggunakan panduan ITC. Adapun item yang diadaptasi adalah PS1 yang item aslinya adalah “Saya merasa orang dengan depresi mungkin merasa bersalah ketika mereka tidak melakukan kesalahan apapun” menjadi “Saya merasa orang yang mengalami depresi mungkin merasa bersalah meskipun mereka tidak melakukan kesalahan apapun”, PS2 yang awalnya “Saya merasa hilangnya rasa percaya diri dan harga diri yang rendah dapat menjadi salah satu tanda depresi” menjadi “Saya merasa, hilangnya rasa kepercayaan diri dan rasa rendah diri merupakan tanda dari depresi”, PS3 dengan item aslinya “Saya merasa terlalu sedikit atau terlalu banyak tidur dapat menjadi gejala depresi” menjadi “Saya merasa, kurang tidur atau tidur berlebihan merupakan gejala depresi”, dan terakhir

PS4 dengan item asli “Saya merasa makan terlalu banyak atau kehilangan minat terhadap makanan bisa jadi merupakan depresi” menjadi “Saya merasa banyak makan atau kehilangan selera makan merupakan gejala depresi”. Item-item tersebut diadaptasi karena kalimatnya sulit untuk dipahami oleh orang awam. Adaptasi pada 4 item dilakukan atas saran saat melakukan *Expert Judgement* dan saat melakukan uji coba.

Desain Penelitian

Penelitian ini menggunakan teknik *Convenience Sampling*, dimana metode pengambilan sampel tersebut memilih sampel yang aksesibilitasnya cukup mudah dijangkau oleh peneliti. Dengan itu, peneliti dapat memberikan kuesioner kepada orang-orang terdekatnya seperti teman, kerabat, keluarga, atau koleganya. Selain itu, teknik pengumpulan data yang digunakan dalam penelitian ini berupa pernyataan dalam bentuk kuesioner dengan menggunakan bantuan *google form* dan sosial media sebagai tempat untuk penyebaran kuesioner tersebut (Maidiana, 2021)

Dalam proses adaptasi instrumen ini, digunakan panduan dari ITC, yaitu tahap prakondisi yang mana peneliti meminta izin untuk mengadaptasi ke dalam bahasa Indonesia kepada pemilik instrumen. Selanjutnya tahap pengembangan tes, di mana dalam tahap ini dilakukan penerjemahan dari bahasa Inggris (bahasa asli) ke Bahasa Indonesia (bahasa adaptasi) kemudian terjemahkan lagi ke dalam bahasa Inggris oleh ahli penerjemah (*back translate*). Setelah *back translate* selesai dilanjutkan dengan mengevaluasi pernyataan-pernyataan dalam instrumen melalui proses *expert judgement* untuk melihat makna-makna pernyataan asli dengan adaptasinya (Utami & Arfensia, 2024). Untuk pernyataan yg sulit dimengerti atau pun yang dapat menyebabkan bias, akan diberikan saran yang lebih baik agar menjadi pernyataan yang mudah dipahami dan tidak menyebabkan bias, serta sesuai dengan budaya di tempat yang akan dijadikan tujuan penelitian. Tahap selanjutnya adalah konfirmasi, yakni mengkonfirmasi bahwa instrumen adaptasi ini sudah sesuai dan dapat disebarakan kepada masyarakat yang menjadi target populasi. Sebelum disebarakan, instrumen adaptasi dilakukan uji terbaca terlebih dahulu kepada 5 orang yang sesuai dengan kriteria agar diketahui apakah item-item atau pernyataan-pernyataan dari adaptasi tersebut sudah dapat dipahami atau belum, jika sudah, instrumen adaptasi dapat disebarakan. Kemudian tahap administrasi, dilakukan dengan mempersiapkan materi, *informed consent*, dan instruksi dalam pengisian alat ukur untuk meminimalkan kesalahan. Alat ukur disebarakan melalui kuesioner dengan bantuan *google form* lalu dibagikan ke sosial media. Setelah tahap administrasi adalah tahap skala skor dan interpretasi, yang mana peneliti melakukan analisis data dari data yg didapatkan dari instrumen yang telah disebarakan. Analisis tersebut meliputi kategorisasi skor, uji unidimensionalitas, uji invariansi parameter, uji independensi lokal, uji validitas skala peringkat, dan uji reliabilitas. Tahap terakhir adalah dokumentasi, yaitu dokumentasi teknis mengenai perubahan apapun yang terjadi saat analisis data dilakukan.

Teknis Analisis

Penelitian ini menggunakan metode *Rasch Model* dengan teknik analisis data yang dilakukan adalah uji unidimensionalitas, yang dijelaskan dengan menggunakan nilai *raw variance* yang mana syarat nilai untuk memenuhi unidimensional adalah minimal 20%, uji invariansi parameter yang termasuk *wright map* untuk melihat butir yang paling sulit dan paling mudah bagi responden, uji fit/misfit responden dan uji fit/misfit butir yang dapat dilihat dari outfit MNSQ, outfit ZTSD, dan Pt Mean Corr. Untuk *Outfit Mean Square* (MNSQ), nilai yang diterima adalah sebesar $0,5 < MNSQ < 1,5$, nilai *Outfit Z-Standard* (ZTSD) yang diterima sebesar $-2,0 < ZTSD < +2,0$, dan nilai *Point Measure Correlation* (Pt Mean Corr) yang diterima sebesar $0,4 < Pt Measure Corr < 0,85$ (Sumintono & Widhiarso, 2014). Kemudian uji independensi lokal yang memiliki syarat untuk memenuhi kriteria apabila korelasi butirnya $< 0,3$ (Christensen et al., 2017), lalu uji validitas skala peringkat, dan uji reliabilitas yang mencakup *separation person* dan item, reliabilitas person dan item, serta Alpha Cronbach. Data yang didapat kemudian ditabulasikan dalam Ms. Excel dan dikonversikan serta dianalisis menggunakan software Winstep 3.73.

3. Hasil

Analisis Rasch

1) Uji Unidimensionalitas

Unidimensionalitas dianggap sebagai alat untuk menghitung reliabilitas dari suatu model yang menunjukkan bahwa dalam model tersebut memiliki tingkat kesesuaian dimensi yang baik. Asumsi unidimensional dalam ruang laten banyak digunakan oleh konstruktor, karena unidimensional dapat meningkatkan interpretasi berbagai hasil pengujian (Aprilia et al., 2021; Bond & Fox, 2015; Hambleton & Swaminathan, 1985) . Dalam penelitian ini, uji asumsi ini dipilih karena mampu memberikan penilaian apakah suatu instrumen mampu mengukur konstruk secara konsisten atau tidak. Uji unidimensionalitas dijelaskan menggunakan nilai raw variance. Data memenuhi syarat unidimensional jika minimal nilai yang didapatkan

adalah 20% (Sumintono & Widhiarso, 2014). Hasil pengujian unidimensionalitas pada instrumen D-Lit ada pada tabel 3 sebagai berikut.

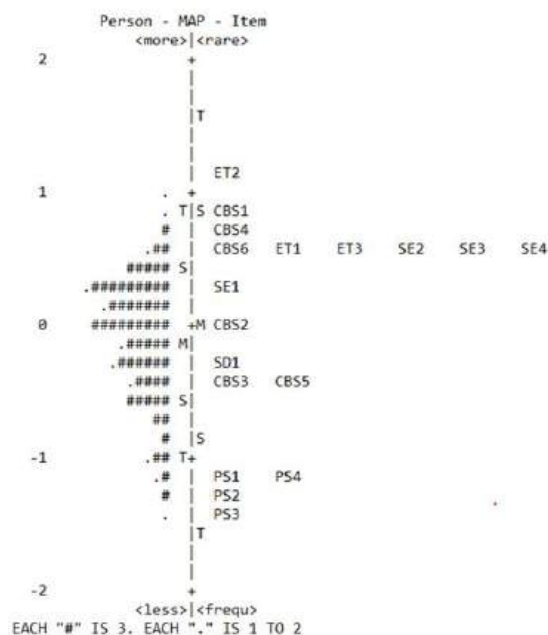
Tabel 3. Uji Unidimensionalitas Instrumen

		Empirical	Modeled
Total raw variance in observations	32.3	100.0%	100.0%
Raw variance explained by measures	14.3	44.3%	43.4%
Raw variance explained by persons	2.3	7.2%	7.1%
Raw Variance explained by items	12.0	37.1%	36.3%
Raw unexplained variance (total)	18.0	55.7%	100.0%
Unexplained variance in 1st contrast	2.7	8.5%	15.3%
Unexplained variance in 2nd contrast	1.8	5.6%	10.1%
Unexplained variance in 3rd contrast	1.6	5.1%	9.1%
Unexplained variance in 4th contrast	1.5	4.7%	8.4%
Unexplained variance in 5th contrast	1.1	3.5%	6.2%

Berdasarkan Tabel 3, hasil pengukuran *raw variance* data sebesar 44.3% dan nilai tersebut tidaklah jauh dari nilai ekspektasinya yaitu 43.4%. Dengan nilai tersebut dapat menunjukkan terpenuhinya persyaratan unidimensionalitas yang minimal 20%, nilai yang terdapat dalam *raw variance* adalah 44.3%, dimana nilai tersebut lebih dari 40% yang artinya kualitasnya lebih bagus.

2) Uji Invariansi

Uji invariansi ini dilakukan untuk menguji responden serta butir yang tidak sesuai dengan persyaratan, atau dengan kata lain responden dan butir tersebut misfit, sehingga responden atau butir tersebut harus dieliminasi. Uji asumsi ini melihat apakah antara individu dan butir item memiliki kestabilan dengan skor tes berdasarkan model pengukurannya yang nantinya akan membantu peningkatan kualitas pengukuran. Berikut hasil pengujian invariansi (item dan person) yang tidak sesuai dengan persyaratan pada gambar 1.



Gambar 1. Wright Map

Berdasarkan gambar 1, diketahui bahwa butir yang paling sulit disetujui oleh responden, adalah ET2, CBS1, CBS4, CBS6, ET1, ET3, SE2, SE3, SE4, dan SE1. Item dan person ditentukan berdasarkan nilai mean logit yaitu 0,00. Butir ET2 memiliki tingkat kesulitan tertinggi dengan bunyi item “Saya merasa banyak perawatan untuk depresi yang lebih efektif daripada obat antidepresan”. Terdapat butir-butir lain yang terlalu mudah untuk disetujui oleh responden, di antaranya PS3, PS2, PS1, dan PS4. Hal tersebut menunjukkan bahwa ada

kemungkinan bahwa butir-butir tersebut mengandung bias normatif sehingga responden cenderung setuju. Dari gambar 4.1 juga diketahui bahwa variasi tingkat kesukaran butir tidak sama dengan abilitas respondennya sehingga responden dengan abilitas rendah memiliki kecenderungan tidak terukur dengan baik karakteristiknya.

Tabel 4. Uji Fit/Misfit Respondent

Responden	INFIT	OUTFIT	ZTSD	PT. MEASURE
	MNSQ	MNSQ		CORR
122	4.69	5.24	8.1	A-.23
184	4.18	4.44	7.2	B-.85
074	3.74	4.43	6.5	C-.35
166	3.04	3.49	5.7	D-.73
070	3.15	3.27	5.4	E-.09
009	2.77	2.76	4.4	F .01
196	2.66	2.74	4.5	G-.28
006	2.42	2.58	4.2	I-.58
046	2.16	2.44	3.9	K-.73
038	2.27	2.37	3.7	L .05
182	2.35	2.37	3.8	M .06
014	2.17	1.87	2.4	O .86
069	2.17	1.87	2.4	P .86
034	1.92	2.13	3.0	R .14
050	1.88	1.93	2.7	T .08
055	1.88	1.67	2.0	X .14
219	1.65	1.77	2.3	Y-.07
109	1.68	1.74	2.2	Z .21
218	1.62	1.62	2.0	.07
210	1.61	1.61	1.8	.36
165	1.61	1.54	1.8	.14
208	1.52	1.60	2.0	.26
222	1.49	1.55	1.8	.06
101	.65	.70	-1.2	-.09
042	.61	.66	-1.3	-.02
029	.47	.51	-2.2	.00
063	.47	.51	-2.2	.00
048	.42	.45	-2.5	.34
220	.24	.21	-4.3	e .86
007	.17	.17	-4.9	a .87
Mean	1.01	1.01	-.3	
S.D	.67	.71	2.1	

Berdasarkan tabel 4 di atas terlihat bahwa dari 225 responden, sebanyak 30 responden tidak memenuhi persyaratan fit. Penentuan misfit responden dilihat dari kriteria outfit MNSQ, outfit ZSTD, dan Pt Mean Corr yang mengacu pada nilai *Outfit Mean Square* (MNSQ) yang diterima sebesar $0,5 < MNSQ < 1,5$, nilai *Outfit Z-Standard* (ZTSD) yang diterima sebesar $-2,0 < ZTSD < +2,0$, dan nilai *Point Measure Correlation* (Pt Mean Corr) yang diterima sebesar $0,4 < Pt Measure Corr < 0,85$. Selain itu, nilai korelasi *Rasch* (*point measure correlation*) atau daya diskriminasi *Rasch* yang memiliki kriteria sebesar 0,4 hingga 0,85 atau tidak memiliki nilai korelasi yang negatif. Implikasi dari nilai korelasi yang tidak sesuai kriteria dapat dikatakan bahwa responden tersebut memberikan jawaban yang tidak konsisten atau terdapat adanya indikasi *guessing* (menebak-nebak jawaban) pada responden. Oleh karena itu, responden (*person*) tersebut tidak fit karena tidak sesuai kriteria *outfit* dan *correlation* sehingga harus dieliminasi. Responden tersebut tidak menjawab butir soal dengan konsisten atau abilitasnya berubah-ubah karena terpengaruh dengan karakteristik butir yang berbeda.

Tabel 5. Uji Fit/Misfit Butir

Butir	INFIT	OUTFIT		PT. MEASURE	ITEM
	MNSQ	MNSQ	ZTSD	CORR	
22	1.30	1.31	2.8	A-.05	SD2
9	1.13	1.14	1.4	E-.22	ET4
5	.96	.95	-.4	i-.07	PS5
8	.69	.71	-3.2	b .29	ET3
Mean	.98	.98	-.3		
S.D	.16	.16	1.7		

Berdasarkan tabel 5 di atas, hasil pengujian didapatkan informasi bahwa butir SD2, ET4, PS5, dan ET3 memiliki nilai outfit MNSQ, outfit ZSTD, dan *Pt Measure Correlation* yang tidak memenuhi kriteria. Kriteria di sini sama dengan yang ada pada person misfit. Dari hasil tersebut, *point measure correlation* dari keempat butir di atas tidak mampu mendiskriminasi peserta dengan abilitas rendah dan tinggi dalam menjawab soal dengan kesulitan tertentu. Oleh karena itu, butir SD2, ET4, PS5, dan ET3 tidak memenuhi persyaratan sehingga harus dieliminasi. Jumlah butir yang digunakan pada instrumen D-lit berjumlah 18 butir dengan jumlah awal 22 butir. Dari total butir 22, terdapat beberapa person yang tidak memenuhi persyaratan outfit MNSQ, ZTSD, dan Pt Mean Corr sehingga 4 butir tersebut harus dieliminasi. Tersisa 18 butir yang dipertahankan dan telah memenuhi persyaratan, berikut adalah item akhirnya.

3) Uji Independensi Lokal

Uji independensi lokal bertujuan untuk mengetahui apakah suatu butir berkorelasi dengan butir lainnya, karena jika saling berkaitan akan memengaruhi estimasi parameter per individu sehingga akan menimbulkan masalah dengan validitas konstruk suatu pengukuran. Terdapat hal yang menjadi syarat untuk memenuhi kriteria uji independensi lokal adalah apabila korelasi butirnya $< 0,3$ (Christensen et al., 2017) yang bermakna bahwa jawaban peserta tes atau suatu butir tidak berpengaruh terhadap jawaban pada butir yang lain. Hasil pengujian independensi lokal dapat dilihat pada tabel 6 berikut.

Tabel 6. Hasil Uji Independensi Lokal

CORRELATION	ENTRY NUMBER ITEM	ENTRY NUMBER ITEM
.35	2 PS2	3 PS3
.35	3 PS3	4 PS4
.33	2 PS2	4 PS4
.29	1 PS1	2 PS2
-.34	2 PS2	10 CBS3
-.30	3 PS3	13 CBS6
-.30	2 PS2	9 CBS2
-.30	1 PS1	10 CBS3
-.29	5 ET1	18 SD1
-.28	4 PS4	7 ET3

Dapat dilihat bahwa terdapat butir soal yang memiliki korelasi rendah sebanyak 6 pasang butir, yaitu dari skor -.34 antara butir PS2 dan CBS3, skor -.30 antara butir PS3 dengan CBS6, skor -.30 antara butir PS2 dengan CBS2, skor -.30 antara butir PS1 dengan CBS3, skor 0.29 antara butir ET1 dengan SD1, dan skor -.28 antara butir PS4 dengan ET3. Butir-butir tersebut yang memiliki korelasi yang negatif yaitu butir PS2 dengan CBS3, butir PS3 dengan CBS6, butir PS2 dengan CBS2, butir PS1 dengan CBS3, butir ET1 dengan SD1, dan butir PS4 dengan ET3 dapat diartikan bahwa butir-butir tersebut tidak terdapat dependensi lokal di dalamnya. Terdapat 1 pasang butir yang memiliki korelasi $< 0,3$ yaitu butir PS1 dan PS2. Selain itu, terdapat 3 pasang butir lain yang memiliki nilai korelasi tertinggi sehingga tidak memenuhi persyaratan independensi lokal, yaitu nilai positif tidak melebihi 0,3. Butir-butir tersebut memiliki nilai korelasi di atas 0,3. Terdapat dampak nyata dari antarbutir, butir-butir menjadi saling terkait, sehingga kualitas satu butir berkaitan dengan kualitas butir lainnya. Hal tersebut menjelaskan bahwa 3 pasang butir tersebut saling memengaruhi jawaban responden satu butir ke butir lainnya.

Butir-butir tersebut memberikan petunjuk jawaban pada butir soal lainnya sehingga dapat merusak pengukuran instrumen. Oleh sebab itu, 3 pasang butir tersebut dapat dieliminasi. Jika tidak ingin mengeliminasi butir tersebut, dapat diatasi dengan cara mengubah kata-kata pada butir-butir tersebut menjadi kata yang berbeda namun memiliki makna yang sama, dengan tujuan mengalihkan perhatian antar butir. Alternatif ini dapat dipertimbangkan karena didukung dengan terpenuhinya asumsi item fit.

4) Uji Validitas Skala Peringkat

Tujuan pengujian validitas skala peringkat adalah untuk memverifikasi apakah pilihan jawaban yang diberikan kepada responden membuat responden bingung atau tidak. *Rating Scale* ini juga berfungsi dengan mengestimasi nilai abilitas responden dan tingkat kesulitan butir pada skala pengukuran interval yang disebut sebagai logit. *Rating Scale* ini digunakan karena untuk mengukur kualitas instrumen tidak hanya dari validitas atau reliabilitas saja, oleh karena itu *rating scale* dipakai dalam analisis ini (Ardiyanti & Dinni, 2018). *Rasch Model* dapat memberikan verifikasi mengenai skala peringkat yang diberikan dalam instrumen. Opsi jawaban yang diberikan berupa skala *Likert*. Berikut hasil uji validitas skala peringkat pada tabel 7.

Tabel 7. Hasil Uji Validitas Skala Peringkat

Category	Threshold	Observed Average	Observed count (%)	Infit	Outfit
Sangat tidak Setuju	NONE	-0,96	456 (13)	1,00	1,03
Tidak Setuju	-1,56	-0,61	1010 (29)	0,91	0,93
Netral	-0,19	-0,13	852 (24)	0,98	0,98
Setuju	0,21	0,64	819 (23)	0,74	0,73
Sangat Setuju	1,54	0,79	373 (11)	1,30	1,26

Berdasarkan tabel 7, terlihat bahwa rata-rata observasi dimulai dari logit dengan nilai -0,96 untuk opsi 1 (Sangat Tidak Setuju) dan meningkat secara konsisten menjadi logit 0,79 untuk opsi 5 (Sangat Setuju). Hal tersebut menjelaskan bahwa skala peringkat 1 – 5 tidak membuat responden bingung dan juga berarti rentang skala yang ditentukan sudah tepat untuk instrumen ini. Selain rata-rata observasi, ambang batas *Andrich Threshold* juga berfungsi sebagai kriteria untuk menguji apakah nilai politomi yang digunakan sudah benar atau belum. *Andrich Threshold* dimulai dari *NONE*, menuju ke nilai negatif, lalu menuju ke nilai positif secara berurutan. Artinya, kelima pilihan yang diberikan tidak membuat responden bingung saat memberikan jawaban/tanggapan, dan mereka dapat melihat perbedaan opsi yang tersedia. Dalam penelitian ini, *Andrich* dimulai dari *NONE* hingga 1,54.

5) Uji Reliabilitas

Adapun hasil dari pengujian reliabilitas instrumen dengan menggunakan *software* Winstep akan dijelaskan dalam tabel 8 berikut.

Tabel 8. Hasil Uji Reliabilitas

	Separation	Reliability	Alpha Cronbach
Person	2.10	.64	.68
Item	13.17	.99	

Hasil uji reliabilitas instrumen di atas menunjukkan nilai *separation* 2,10 dan reliabilitas *person* sebesar 0.64 yang berarti termasuk ke dalam kategori cukup dan nilai *separation* 13,17 serta reliabilitas item sebesar 0.99 yang menunjukkan nilai tersebut masuk ke dalam kategori istimewa. Selain itu, nilai Alpha Cronbach sebesar 0.68 yang artinya cukup antara interaksi responden dan item. Berikut adalah instrumen final *Depression Literacy (D-Lit) Questionnaire* setelah dieliminasi.

Tabel 9. Instrumen Final D-Lit

No Item	Item	Redaksi Item
1	PS1	Saya merasa orang yang mengalami depresi mungkin merasa bersalah meskipun mereka tidak melakukan kesalahan apapun.
2	PS2	Saya merasa, hilangnya rasa kepercayaan diri dan rasa rendah diri merupakan tanda dari depresi
3	PS3	Saya merasa, kurang tidur atau tidur berlebihan merupakan gejala depresi.
4	PS4	Saya merasa banyak makan atau kehilangan selera makan merupakan gejala depresi.
5	ET1	Saya merasa psikolog klinis dapat meresepkan obat antidepresan.
6	ET2	Saya merasa banyak perawatan untuk depresi yang lebih efektif daripada obat antidepresan.
7	ET5	Saya merasa saat seseorang mengalami depresi tingkat sedang dapat memengaruhi kehidupan seseorang seperti halnya penyakit yang berkaitan dengan gangguan saraf atau tuli.
8	CBS1	Saya merasa orang dengan depresi sering berbicara tidak teratur dan relevan.
9	CBS2	Saya merasa perilaku ceroboh dan bodoh adalah gejala umum dari depresi.
10	CBS3	Saya merasa tidak berjalan di trotar yang rusak bisa jadi gejala depresi
11	CBS4	Saya merasa orang dengan depresi sering mendengar suara yang biasanya tidak nyata
12	CBS5	Saya merasa depresi tidak memengaruhi daya ingat dan konsentrasi.
13	CBS6	Saya merasa, memiliki beberapa kepribadian yang berbeda dapat menjadi gejala depresi.
14	SE1	Saya merasa dari semua alternatif dan gaya hidup untuk mengatasi depresi, mengonsumsi vitamin adalah yang paling bermanfaat.
15	SE2	Saya merasa penderita depresi harus berhenti minum obat antidepresan segera setelah mereka merasa lebih baik.
16	SE3	Saya merasa obat antidepresan membuat ketagihan.
17	SE4	Saya merasa obat antidepresan biasanya bekerja dengan cepat.
18	SD1	Saya merasa sebagian besar penderita depresi perlu dirawat di rumah sakit (jiwa).

Setelah melalui proses analisis untuk mengetahui butir yang fit dan tidak fit berdasarkan outfit MNSQ, outfit ZSTD, dan *Pt Measure Correlation*, ditemukan 4 butir yang tidak memenuhi kriteria sehingga harus dieliminasi dan mendapatkan hasil akhir sebanyak 18 butir yang bertahan dengan hanya memiliki satu dimensi atau unidimensionalitas.

4. Pembahasan

Berdasarkan hasil dari pengadaptasian instrumen D-Lit, uji unidimensional yang diperoleh dengan *raw variance* data sebesar 44.3%. Nilai minimum untuk terpenuhinya persyaratan unidimensionalitas adalah 20%, sehingga *raw variance* dalam penelitian ini memiliki kualitas yang lebih bagus. Bagian uji variansi, terdapat 4 item yang dieliminasi karena tidak memenuhi persyaratan Outfit MNSQ, Outfit ZTSD, dan *Pt. Measure Correlation*. Item-item tersebut adalah SD2, ET4, PS5, dan ET3, sehingga item final pada instrumen D-Lit yang awalnya 22 butir menjadi 18 butir. Selain itu, hasil analisis instrumen *Depression Literacy* yang diadaptasi, diperoleh informasi bahwa nilai reliabilitas item sebesar 0,99 yang termasuk ke dalam kategori istimewa dan nilai reliabilitas person sebesar 0,64 yang termasuk ke dalam kategori cukup. Kedua reliabilitas tersebut dapat dilihat dalam tabel 8. Sedangkan nilai reliabilitas Alpha Cronbach sebesar 0,68 yang berarti terdapat interaksi yang cukup antara item dan person.

Untuk melengkapi penelitian ini, disertakan beberapa penelitian yang membahas tentang literasi depresi yaitu penelitian yang dilakukan oleh Kulwicka & Gasirowska (2023), ditemukan bahwa ada sebuah instrumen yang bisa digunakan untuk mengukur pemahaman masyarakat tentang depresi, termasuk pengetahuan dan pandangan mereka mengenai gejala, pengobatan, dan pencegahannya. Instrumen tersebut adalah *Depression Literacy (D-Lit) Questionnaire* yang dikembangkan oleh Tehrani et al. (2022). Kemudian, penelitian oleh Raiisi et al. (2024) menyebutkan bahwa dari 22 pertanyaan kuesioner, terdapat 17 item yang dipertahankan. Hasil koefisien korelasi

internal Spearman menunjukkan bahwa semua aspek literasi depresi berhubungan langsung dengan skor total kuesioner ini dan hasilnya signifikan. Selain itu, penelitian yang dilakukan oleh Pruksarungruang & Rhein (2022) menunjukkan bahwa literasi depresi di Thailand berada pada tingkat sedang hingga tinggi. Item-item individual pada skala literasi depresi menunjukkan beberapa area dalam mengidentifikasi depresi yang masih banyak orang yang salah paham, terutama mengenai adanya halusinasi dan beberapa kepribadian yang berbeda.

5. Kesimpulan

Penelitian ini menjelaskan tentang adaptasi instrumen *Depression Literacy (D-Lit) Questionnaire* dengan menggunakan metode *Rasch Model* untuk menilai pengetahuan masyarakat mengenai gangguan psikologis khususnya depresi, seperti yang telah disebutkan pada bagian rumusan masalah dan tujuan penelitian. Penelitian ini memberikan informasi tentang beberapa properti psikometri tambahan yang pada instrumen D-Lit versi asli, karena penelitian ini menggunakan metode *Rasch Model* sedangkan pada instrumen asli menggunakan metode *Confirmatory Factor Analysis (CFA)* dan *Exploratory Factor Analysis (EFA)*, serta digunakan untuk sampel dari bahasa dan budaya yang berbeda.

Dari hasil yang didapat, penelitian ini menunjukkan bahwa reliabilitas item dalam pengolahan data sudah masuk ke dalam kategori istimewa, dan interaksi antara responden dan item cukup baik. Untuk validitas, mulai dari uji unidimensional, uji invariansi, validitas skala peringkat, dan yang lainnya memenuhi persyaratan *Rasch Model*. Adapun keterbatasan yang ada dalam penelitian ini di antaranya karena menggunakan pendekatan *convenience sampling* yang siapapun dapat menjadi responden berdasarkan kesediaan responden sehingga populasi yang ada tidak mewakili secara umum, seperti domisili yang mayoritas berada di luar jabodetabek, usia lebih banyak di angka 18-23 tahun. Sedangkan adaptasi instrumen *Depression Literacy (D-Lit)* ini ditujukan untuk masyarakat umum dengan usia 18-40 secara umum.

Melihat dari hasil penelitian yang telah dilakukan, untuk penelitian selanjutnya yang tertarik mendalami tema yang sama, diharapkan dapat mengambil sampel yang lebih luas dan beragam tidak hanya terpaku pada satu generasi, tetapi kepada generasi lainnya. Selanjutnya, instrumen D-Lit yang sudah diadaptasi ini berjumlah 18 butir perlu diuji lebih lanjut pada sampel klinis dan non klinis agar mengetahui perbedaan hasil dari kedua sampel tersebut dan dapat divalidasi dalam konteks lain.

6. Referensi

- Akinsulore, A., Esimai, O. A., Mapayi, B. M., & Aloba, O. O. (2018). Public Awareness and Attitude Towards Depression: A Community Based Study Among an Adult Population in Ile-Ife South-Western Nigeria. *Community Mental Health Journal*, 54(6), 866–874. <https://doi.org/10.1007/s10597-017-0222-8>
- Al-Azzam, S., Alzoubi, K., Ayoub, N., Khader, Y., Karasneh, R., Al-Husein, B., & Al-Horani, M. (2013). An audit on public awareness of depression symptoms in Jordan. *International Journal of Occupational Medicine and Environmental Health*, 26(4). <https://doi.org/10.2478/s13382-013-0128-9>
- Aletesh, N., Alamrani, S., Alshreef, O., Alsharif, Z., Alshehri, M., Falah, H., Ali, R., Al-Etеш, Y., & Almubarak, S. (2021). Public awareness towards depression in Al-Jouf, Saudi Arabia. *International Journal of Medicine in Developing Countries*, 876–884. <https://doi.org/10.24911/IJMDC.51-1610991870>
- ALJadani, A. H., Alshammari, S. N., Alshammari, K. A., Althagafi, A. A., & AlHarbi, M. M. (2021). Public awareness, beliefs and attitude towards depressive disorders in Saudi Arabia. *Saudi Medical Journal*, 42(10), 1117–1124. <https://doi.org/10.15537/smj.2021.42.10.20210425>
- Annury, U. A., Yuliana, F., Suhadi, V. A. Z., & Karlina, C. S. A. (2022). *Dampak Self Diagnose Pada Kondisi Mental Health Mahasiswa Universitas Negeri Surabaya*.
- Aprilia, M., Lidinillah, D. A. M., & Giyartini, R. (2021). Pengembangan Instrumen Penilaian Kreativitas Siswa melalui Analisis Rasch Model di Sekolah Dasar. *Jurnal Basicedu*, 5(4), 2302–2310. <https://doi.org/10.31004/basicedu.v5i4.1199>
- Ardiyanti, D., & Dinni, S. M. (2018). Aplikasi Model Rasch dalam Pengembangan Instrumen Deteksi Dini Postpartum Depression. *Jurnal Psikologi*, 45(2), 81. <https://doi.org/10.22146/jpsi.29818>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd edition). Routledge, Taylor & Francis Group.
- Brodaty, H., Peters, K., Boyce, P., Hickie, I., Gordon, P., Mitchell, P., & Wilhelm, K. (1991). Age and Depression. *Journal of Affective Disorders*
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical Values for Yen's Q_3 : Identification of Local Dependence in the Rasch Model Using Residual Correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>

- Dhaliwal, R., Attri, I., Dhillon, S. S., Dev, K., Mehmi, N., & Attri, P. (2022). *Public Awareness of Depression – An Informal Clinical Study*.
- Hadi, I., Devianty, R., & Rosyanti, L. (2017). GANGGUAN DEPRESI MAYOR (MAYOR DEPRESSIVE DISORDER) MINI REVIEW. *Jurnal Penelitian*, 9(1).
- Hakim, Moh. A., & Aristawati, N. V. (2023). Mengukur depresi, kecemasan, dan stres pada kelompok dewasa awal di Indonesia: Uji validitas dan reliabilitas konstruk DASS-21. *Jurnal Psikologi Ulayat*, 10(2), 232–250. <https://doi.org/10.24854/jpu553>
- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory*. Springer Netherlands. <https://doi.org/10.1007/978-94-017-1988-9>
- ITC Guidelines for Translating and Adapting Tests (Second Edition). (2018). *International Journal of Testing*, 18(2), 101–134. <https://doi.org/10.1080/15305058.2017.1398166>
- Kulwicka, K., & Gasiorowska, A. (2023). Depression literacy and misconceptions scale (DepSter): A new two-factorial tool for measuring beliefs about depression. *BMC Psychiatry*, 23(1), 300. <https://doi.org/10.1186/s12888-023-04796-8>
- Maidiana. (2021). Penelitian Survey. *Journal of Education*
- Nazira, D., Mawarpury, M., Afriani., Kumala, I. D. (2022). Literasi Kesehatan Mental Pada Mahasiswa di Banda Aceh. *Jurnal Psikologi Unsyiah*
- Nguyen Thai, Q. C., & Nguyen, T. H. (2018). Mental health literacy: Knowledge of depression among undergraduate students in Hanoi, Vietnam. *International Journal of Mental Health Systems*, 12(1), 19. <https://doi.org/10.1186/s13033-018-0195-1>
- Pruksarungruang, J. & Rhein, D. (2022). Depression literacy: An analysis of the stigmatization of depression in Thailand. *Sage Open*, 12(4). <https://doi.org/10.1177/21582440221140376>
- Purborini, N., Lee, M.-B., Devi, H. M., & Chang, H.-J. (2021). Associated factors of depression among young adults in Indonesia: A population-based longitudinal study. *Journal of the Formosan Medical Association*, 120(7), 1434–1443. <https://doi.org/10.1016/j.jfma.2021.01.016>
- Raiisi, F., Ebrahimi, M., Hosseini, F. G., Jafari, K., & Rahmati, F. (2024). Translation, validity, and reliability of Depression Literacy Questionnaire in Iranian young adults. *International Journal of Preventive Medicine*, 15. <https://doi.org/10.4103/ijpvm.ijpvm.311.23>
- Saleh, M. S. N., Al-Hor, F. S., Aseel, M. T., Al-Ali, D. K., Al-Maraghi, S. A., Saeed, S. S., Bawazir, A. O., Seyam, N. I., & Emadi, N. M. (2023). Public Awareness and Prevalence of Depression Symptoms among Patients Attending Primary Health Care Centers in Qatar. *Asian Journal of Medicine and Health*, 21(6), 20–28. <https://doi.org/10.9734/ajmah/2023/v21i6818>
- Sumintono, B., & Widhiarso, W. (2014). Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial (Edisi Revisi)
- Tehrani, H., Nejatian, M., Moshki, M., & Jafari, A. (2022). Psychometric properties of Persian version of depression literacy (D-Lit) questionnaire among general population. *International Journal of Mental Health Systems*, 16(1), 40. <https://doi.org/10.1186/s13033-022-00550-x>
- Utami, G., & Arfensia, D. S. (2024). Adaptasi Binge-Watching Engagement Scale Questionnaire (BWESQ) Dalam Bahasa Indonesia. *Jurnal Penelitian Dan Pengukuran Psikologi: JPPP*, 13(2), 94–110. <https://doi.org/10.21009/JPPP.132.03>
- Wahyuni, E., & Nabila, F. (2022). Literasi Depresi: Gambaran dan Kebutuhan Self-help Book Untuk Meningkatkan Pemahaman Depresi. *INSIGHT: Jurnal Bimbingan Konseling*, 10(1), 23–29. <https://doi.org/10.21009/INSIGHT.101.04>
- Wei, Y., Hayden, J. A., Kutcher, S., Zygumt, A., & McGrath, P. (2013). The effectiveness of school mental health literacy programs to address knowledge, attitudes and help seeking among youth. *Early Intervention in Psychiatry*, 7(2), 109–121. <https://doi.org/10.1111/eip.12010>



ANALISIS PSIKOMETRI *BURNOUT ASSESSMENT TOOL (BAT) FOR STUDENT* VERSI BAHASA INDONESIA

Muh Haritsah Hizbullah¹, Juliana Nur Afifah², Ulfah Sayidah Mutamimah³, Ahmad Yasir Ibrahim⁴, Alnaznabila Aghnii Hafnieva Imanda⁵, Muhamad Arif Saefudin⁶, Helli Ihsan⁷

^{1,2,3,4,5,6,7} Fakultas Ilmu Pendidikan, Universitas Pendidikan Indonesia, Bandung, Indonesia

Email: harishizbullah03@upi.edu

Received :	Revised :	Accepted :	Published :
29 Desember 2024	30 Maret 2025	15 April 2025	30 April 2025

Abstract

This study aims to test the validity of the Indonesian version adaptation of the Burnout Assessment Tool (BAT) for students at the Universitas Pendidikan Indonesia. This scale consists of 6 dimensions and 33 items. The research methods included content validity testing with expert judgment and Rasch model analysis of 202 students. The results showed an Alpha Cronbach reliability of 0.913 (high) and raw variance explained by measures of 40.4% (good). Item fit analysis showed 31 of 33 items met the criteria. The Rasch model showed that the answer choices were not confusing and the person and item reliability were very good. However, 2 items were misfit and 26 people were misfit. The Indonesian version adaptation of BAT has good potential to measure student burnout, but further research with a larger sample is needed.

Keywords: burnout, adaptation, assessment, Rasch model

Abstrak

Penelitian ini bertujuan untuk menguji validitas adaptasi *Burnout Assessment Tool (BAT) for Student* versi bahasa Indonesia untuk mahasiswa di Universitas Pendidikan Indonesia. Skala ini terdiri dari 6 dimensi dan 33 item. Metode penelitian meliputi uji validitas isi dengan *expert judgement* dan analisis model Rasch terhadap 202 mahasiswa. Hasil menunjukkan reliabilitas Alpha Cronbach 0,913 (tinggi) dan *raw variance explained by measures* 40,4% (baik). Analisis *item fit* menunjukkan 31 dari 33 item memenuhi kriteria. Model Rasch menunjukkan pilihan jawaban tidak membingungkan dan reliabilitas person serta item sangat baik. Namun, 2 item *misfit* dan 26 orang *misfit*. Adaptasi BAT versi Indonesia memiliki potensi baik mengukur *burnout* mahasiswa, namun perlu penelitian lebih lanjut dengan sampel lebih besar.

Kata Kunci: burnout, adaptasi, asesmen, model Rasch

1. Pendahuluan

Dalam era modern yang ditandai dengan tekanan tinggi di berbagai aspek kehidupan, *burnout* telah menjadi perhatian utama di banyak kalangan, termasuk dalam konteks akademik. *Burnout*, sebagaimana didefinisikan dalam Klasifikasi Penyakit Internasional (ICD-11) oleh *World Health Organization (WHO)*, adalah sindrom yang terjadi akibat kegagalan dalam mengelola stres kronis di tempat kerja atau dalam aktivitas berbasis tugas lainnya. Sindrom ini terdiri atas tiga dimensi utama: kelelahan, jarak mental terhadap pekerjaan, dan penurunan efikasi profesional (Popescu dkk., 2024).

Walaupun awalnya dianggap terbatas pada dunia kerja, penelitian menunjukkan bahwa *burnout* juga dapat terjadi dalam konteks non-kerja, seperti di lingkungan pendidikan tinggi (Destyanto dkk., 2023). Berdasarkan penelitian Amarsa dkk (2023) sebanyak 73,3% mahasiswa Bimbingan dan Konseling mengalami burnout akademik dalam kategori sedang, dimana aspek *exhaustion* dan *cynicism* berada pada kategori tinggi dengan proporsi *burnout* yang lebih besar pada mahasiswa angkatan lebih senior di Universitas Pendidikan Indonesia. Selain itu penelitian Putri dkk (2023) menunjukkan sebanyak 78,4% mahasiswa kedokteran mengalami *burnout* pada tingkat sedang, dengan mahasiswa berusia 19 tahun (24,9%) dan angkatan 2021 (30,8%) mengalami burnout

lebih banyak dibandingkan kelompok lainnya. Dalam konteks akademik, mahasiswa sering menghadapi tuntutan yang serupa dengan pekerja profesional, seperti menghadiri kelas, menyelesaikan tugas, dan memenuhi tenggat waktu, yang semuanya dapat dianggap sebagai "pekerjaan". Tingginya tuntutan akademik, tekanan untuk mencapai hasil, dan beban tugas yang berlebihan seringkali memicu kelelahan emosional, sikap sinis terhadap aktivitas belajar, dan penurunan efikasi pribadi, yang secara kolektif dikenal sebagai *burnout* akademik (Sofa dkk., 2024).

Burnout dapat menyebabkan penurunan kesehatan fisik seperti menurunnya kualitas tidur dan konsentrasi menurun serta berkurangnya energi. Sedangkan pada psikis *burnout* dapat menyebabkan penurunan emosi dan mengalami depresi (Orpina & Praha, 2019). Selain itu, *burnout* juga dapat menyebabkan penurunan prestasi akademik karena berkurangnya konsentrasi, sehingga mahasiswa mengalami kesulitan dan kehilangan motivasi dalam menyelesaikan tugas perkuliahan (Aini & Sulaiman, 2024). Menurut Yang (dalam Orpina & Praha, 2019) *burnout* akademik ini tidak hanya berdampak pada kesehatan mental mahasiswa tetapi juga menurunkan keterlibatan akademik, prestasi belajar, dan meningkatkan risiko putus kuliah.

Untuk mengatasi permasalahan ini, penting untuk menggunakan alat ukur *burnout* yang valid dan andal. Popescu dkk. (2024) telah mengembangkan alat ukur *burnout* untuk mahasiswa yang diadaptasi dari *Burnout Assessment Tool* (BAT) oleh Schaufeli dkk. (2020). BAT dikembangkan untuk mengatasi keterbatasan dari alat ukur *burnout* sebelumnya, seperti *Maslach Burnout Inventory* (MBI), yang dinilai kurang komprehensif dalam mengukur gejala inti *burnout*. BAT dirancang dengan mencakup dimensi inti *burnout*, yaitu kelelahan, jarak mental, gangguan emosional, dan gangguan kognitif, serta dimensi tambahan seperti keluhan psikologis dan psikosomatik (Popescu dkk., 2024). Selain itu, BAT memiliki keunggulan dalam menghasilkan skor global yang mencerminkan sindrom *burnout* secara keseluruhan, dibandingkan hanya mengukur sub-komponen secara terpisah. Meskipun berdasarkan penelusuran telah banyak penelitian mengenai *burnout academic* di Indonesia, namun penelitian yang mengadaptasi alat ukur BAT *for student* dari Popescu dkk masih terbatas. Oleh karena itu, tujuan penelitian ini adalah mengadaptasi skala yang dikembangkan oleh Popescu et al. (2024) yaitu *Burnout Assessment Tool* (BAT) *for student* dalam budaya dan bahasa Indonesia sehingga penelitian terkait *burnout academic* di Indonesia dapat menggunakan alat ukur terbaru yang mampu mengatasi keterbatasan-keterbatasan alat ukur *burnout* sebelumnya.

2. Metode Penelitian

Burnout Assessment Tool (BAT) *for student* merupakan alat ukur yang disusun oleh Popescu dkk., (2024) sebagai hasil adaptasi dari *Burnout Assessment Tool* yang disusun oleh Schaufeli, dkk (2020). *Burnout* adalah kombinasi dari ketidakmampuan dan keengganan untuk tidak lagi mengeluarkan energi yang diperlukan di tempat kerja untuk menyelesaikan tugas dengan baik. Menurut mereka, "ketidakmampuan" terwujud dalam kurangnya energi dan "keengganan" dalam peningkatan resistensi, berkurangnya komitmen, kurangnya minat, dan ketid. dengan adanya penelitian ini dapat memberikan representasi yang lebih akurat.

Responden dalam penelitian ini adalah mahasiswa Universitas Pendidikan Indonesia. Pemilihan partisipan ini dilakukan karena populasi target dari alat ukur yang diadaptasi adalah mahasiswa, sehingga uji validitas dan reliabilitas perlu dilakukan pada kelompok yang sesuai. Universitas Pendidikan Indonesia dipilih sebagai lokasi pengambilan data karena memiliki populasi mahasiswa yang cukup besar dan beragam dari berbagai program studi. Selain itu, penelitian terdahulu (Amarsa dkk., 2023) menemukan bahwa mahasiswa Universitas Pendidikan Indonesia menunjukkan tingkat *burnout* akademik yang cukup tinggi, sehingga menjadikan populasi ini relevan dalam konteks pengukuran *burnout*. Dengan demikian, partisipan dari Universitas Pendidikan Indonesia dianggap dapat memberikan gambaran awal yang representatif untuk menguji kualitas psikometrik alat ukur hasil adaptasi ini, sebelum digunakan dalam konteks yang lebih luas. Tingkat *burnout* pada mahasiswa Universitas Pendidikan Indonesia akan diukur menggunakan aspek dari empat dimensi primer yaitu *exhaustion*, *mental distance*, *emotional impairment*, dan *cognitive impairment*; dan juga menggunakan dua dimensi lain yang bersifat sekunder yaitu *psychological complaints* dan *psychosomatic complaints*.

Dalam melakukan pengembangan alat ukur, terdapat tahapan-tahapan sehingga *item* tersebut berhasil diadaptasi dan dapat digunakan pada populasi yang lain dalam hal ini mahasiswa Universitas Pendidikan Indonesia. Tahapan-tahapan dalam mengembangkan alat ukur di antaranya adalah membuat *blueprint* dari alat ukur yang asli dengan cara menerjemahkan semua *item*, menguji validitas isi dengan menggunakan metode *expert judgement*, mengumpulkan data/menguji coba alat ukur, serta melakukan analisis *Rasch Model*. *Rasch Model* merupakan suatu pendekatan statistik yang memanfaatkan data mentah untuk membentuk skala pengukuran dengan interval yang setara (Sumintono & Widhiarso, 2015). Dalam *Burnout Assessment Tool*, pada awalnya terdapat sebanyak 33 *item* yang terbagi ke dalam enam dimensi *burnout* sebagai berikut.

Tabel 1. Daftar Item BAT

Dimensi	Item
<i>Exhaustion</i>	1. Ketika di kampus saya merasa lelah secara mental. 2. Segala sesuatu yang saya lakukan di kampus membutuhkan banyak upaya. 3. Setelah seharian di kampus, saya merasa sulit untuk memulihkan energi saya. 4. Ketika di kampus saya merasa lelah secara fisik. 5. Ketika saya bangun di pagi hari, saya merasa malas untuk memulai hari baru di kampus. 6. Saya ingin aktif di kampus, tapi entah kenapa saya tidak bisa mengaturnya. 7. Ketika saya memaksakan diri di kampus, saya cepat lelah. 8. Di akhir hari perkuliahan, saya merasa lelah secara mental dan kehabisan energi. 9. Saya kesulitan dalam menemukan gairah untuk berkuliah.
<i>Mental Distance</i>	10. Di kampus, saya tidak banyak berpikir tentang apa yang saya inginkan. 11. Saya merasa sangat tidak suka terhadap perkuliahan saya. 12. Saya merasa acuh tak acuh terhadap perkuliahan saya. 13. Saya merasa ragu menyampaikan maksud kuliah saya kepada orang-orang di sekitar saya.
<i>Cognitive Impairment</i>	14. Ketika di kampus saya sulit untuk tetap fokus. 15. Ketika di kampus saya kesulitan berpikir secara logis. 16. Saya cenderung pelupa dan mudah teralihkannya saat berada di kampus.
<i>Emotional Impairment</i>	17. Saat di kampus, saya merasa kesulitan untuk berkonsentrasi. 18. Saya kerap melakukan kesalahan di kampus karena pikiran saya teralihkannya pada hal-hal lain. 19. Di kampus, saya merasa kesulitan mengendalikan emosi saya. / Saya. 20. Saya tidak mengenali diri saya sendiri ketika saya bereaksi secara emosional di kampus. 21. Selama jam perkuliahan, saya menjadi mudah tersinggung jika sesuatu tidak berjalan sesuai dengan keinginan saya. 22. Saya menjadi kesal atau sedih saat di kampus tanpa alasan yang jelas. 23. Di kampus, saya secara tidak sadar memberikan reaksi yang berlebihan.
<i>Psychological Distress</i>	24. Saya mempunyai masalah tidur. 25. Saya cenderung khawatir. 26. Saya merasa tegang dan stress. 27. Saya merasa gelisah dan atau mengalami serangan panik. 28. Suara yang bising dan suasana keramaian membuat saya merasa terganggu.
<i>Psychosomatic Complaints</i>	29. Jantung saya berdebar-debar atau terasa nyeri. 30. Saya merasa tidak enak di bagian perut atau di bagian yang berkaitan dengan pencernaan. 31. Saya merasa sakit kepala. 32. Saya merasakan nyeri di bagian otot seperti di leher, pundak, atau punggung. 33. Saya seringkali jatuh sakit.

Uji validitas isi dengan *expert judgement* dilakukan dengan meminta 2 penilai yang merupakan dosen dari mata kuliah Pengembangan Alat Ukur Psikologi Program Studi Psikologi Universitas Pendidikan Indonesia untuk memberikan penilaian pada setiap aitem dari skala tersebut. Setelah proses *expert judgement*, pengambilan data dilakukan dengan menyebarkan kuesioner *online* melalui *Google Form* pada responden dengan teknik *purposive sampling* dengan kriteria 1) mahasiswa aktif Universitas Pendidikan Indonesia 2) berusia 18 - 25 tahun. Responden yang diperoleh adalah sebanyak 202 responden.

3. Hasil

Analisis data yang dilakukan menghasilkan temuan berdasarkan rasch model dengan mempertimbangkan beberapa aspek penting. Aspek-aspek tersebut meliputi *unidimensionalitas*, *item measure*, *item fit order*, *petal wright*, *person measure*, *person fit order*, *rating scale*, dan *summary statistic*. Penjelasan berikut akan menguraikan aspek-aspek penting tersebut secara lebih rinci.

Analisis Unidimensionalitas

Analisis unidimensionalitas adalah metode yang digunakan untuk memastikan bahwa suatu alat ukur hanya mengukur satu konstruk atau dimensi tertentu. Analisis unidimensionalitas menggunakan aplikasi Winstep, dengan hasil berikut. Uji dimensionalitas dilakukan untuk menilai sejauh mana suatu alat ukur mampu merepresentasikan satu konstruksi secara akurat (Widodo & Chotimah, 2023). *Raw variance explained by measures* digunakan untuk mengukur seberapa besar variasi dalam data yang dapat dijelaskan oleh instrumen yang digunakan. Semakin tinggi nilainya, semakin baik instrumen tersebut dalam menjelaskan variasi dalam respons peserta. Sementara *unexplained variance in 1 st to 5th contrast* digunakan untuk mengukur variasi yang tidak dapat dijelaskan oleh kontras pertama hingga kelima dalam rasch model. Peningkatan nilai *unexplained variance* menunjukkan ketidaksesuaian antara respons yang diamati dengan respons yang diprediksi oleh rasch model. Untuk memastikan terpenuhinya persyaratan pengukuran unidimensionalitas *raw variance explained by measures* sebaiknya melebihi 20%. Secara umum, interpretasi kategori nilai berkisar antara 20%-40% dianggap cukup, 40%-60% dianggap baik, dan lebih dari 60% dianggap sangat baik. Sedangkan, nilai *unexplained variance in 1 st to 5th contrast* idealnya kurang dari 15% (Sumintono & Widhiarso, 2014).

Tabel 2. Analisis Unidimensionalitas

Output	Hasil
Raw variance explained by measures	55.4%
Unexplained variance in 1st contrast	7.7%
Unexplained variance in 2nd contrast	4.6%
Unexplained variance in 3rd contrast	4.2%
Unexplained variance in 4th contrast	4.0%
Unexplained variance in 5th contrast	3.3%

Hasil analisis unidimensionalitas menunjukkan bahwa *raw variance explained by measures* sebesar 40,4%, yang termasuk dalam kategori “Baik”. Sementara itu, nilai *unexplained variance 1st to 5th contrast of residual* memiliki rincian sebagai berikut: *unexplained variance 1st contrast* sebesar 7,7%; *unexplained variance 2nd contrast* sebesar 4,6%; *unexplained variance 3rd contrast* sebesar 4,2%; *unexplained variance 4th contrast* sebesar 4,0%; dan *unexplained variance 5th contrast* sebesar 3,3%. Hal ini menunjukkan bahwa nilai *unexplained variance 1st to 5th contrast* dapat dikategorikan sebagai ideal.

Item Fit

Analisis *item fit* dilakukan dengan rasch model menggunakan aplikasi jMetrik dengan hasil yang diperoleh sebagai berikut.

Tabel 3. Item Statistic

Item	Difficulty	Std. Error	WMS	d. WMS	UMS	Std. UMS	Ket
b1	-0.44	0.07	0.70	-3.69	0.74	-3.12	<i>fit</i>
b2	-0.92	0.08	0.90	-1.15	0.98	-0.22	<i>fit</i>
b3	-0.29	0.07	1.06	0.65	1.11	1.24	<i>fit</i>
b4	-0.63	0.07	0.99	-0.11	1.12	1.25	<i>fit</i>
b5	-0.52	0.07	0.90	-1.09	0.97	-0.33	<i>fit</i>
b6	-0.41	0.07	1.22	2.28	1.20	2.11	<i>fit</i>
b7	-0.69	0.08	0.90	-1.09	0.89	-1.24	<i>fit</i>
b8	-0.76	0.08	0.80	-2.37	0.89	-1.24	<i>fit</i>
b9	0.18	0.08	0.77	-2.73	0.79	-2.42	<i>fit</i>
b10	0.19	0.08	0.96	-0.45	0.97	-0.34	<i>fit</i>
b11	1.02	0.09	0.99	-0.08	0.99	-0.02	<i>fit</i>
b12	1.17	0.09	0.94	-0.57	1.03	0.31	<i>fit</i>
b13	0.63	0.08	1.17	1.77	1.15	1.49	<i>fit</i>
b14	-0.08	0.07	0.68	-3.93	0.69	-3.81	<i>fit</i>

Item	Difficulty	Std. Error	WMS	d. WMS	UMS	Std. UMS	Ket
b15	0.42	0.08	0.81	-2.16	0.82	-1.98	<i>fit</i>
b16	-0.37	0.07	0.88	-1.33	0.88	-1.41	<i>fit</i>
b17	-0.14	0.07	0.67	-4.08	0.68	-3.93	<i>fit</i>
b18	0.05	0.08	0.73	-3.21	0.73	-3.24	<i>fit</i>
b19	0.76	0.08	0.86	-1.54	0.86	-1.48	<i>fit</i>
b20	0.82	0.08	1.06	0.69	1.05	0.51	<i>fit</i>
b21	0.65	0.08	1.07	0.75	1.11	1.12	<i>fit</i>
b22	0.44	0.08	1.62	5.62	1.69	6.05	<i>misfit</i>
b23	0.46	0.08	1.00	0.08	1.01	0.11	<i>fit</i>
b24	-0.28	0.07	1.56	5.39	1.62	5.85	<i>misfit</i>
b25	-0.93	0.08	0.89	-1.21	0.95	-0.50	<i>fit</i>
b26	-0.46	0.07	0.77	-2.71	0.79	-2.50	<i>fit</i>
b27	0.11	0.08	1.09	0.95	1.07	0.78	<i>fit</i>
b28	-0.62	0.07	1.32	3.23	1.37	3.70	<i>fit</i>
b29	0.65	0.08	1.07	0.76	1.06	0.61	<i>fit</i>
b30	0.22	0.08	1.36	3.61	1.35	3.48	<i>fit</i>
b31	-0.31	0.07	0.86	-1.67	0.90	-1.10	<i>fit</i>
b32	-0.37	0.07	1.29	2.99	1.39	3.88	<i>fit</i>
b33	0.44	0.08	1.01	0.14	0.99	-0.11	<i>fit</i>

Analisis *item fit* dilakukan untuk mengetahui item-item yang tidak sesuai atau inkonsisten dengan variabel yang diukur (Efandi & Putri, 2022). Berdasarkan Aduce, et. al. (2017) suatu *item* dianggap *fit* dan layak digunakan jika skor *Weighted Mean Squares* (WMS) dan *Unweighted Mean Squares* (UMS) berada pada rentang 0.5 - 1.5. Berdasarkan analisa yang dilakukan, diperoleh 2 dari 33 *item* termasuk dalam kategori *misfit*, yaitu *item* 22 (WMS = 1.62, UMS = 1.69) dan *item* 24 (WMS = 1.56, UMS = 1.62) sehingga *item* tersebut perlu dihapus.

Item Measure

Item measure digunakan untuk mendapatkan informasi mengenai instrumen yang paling mudah dan paling sulit disetujui oleh responden. Dengan demikian, *item measure* mencerminkan tingkat kesulitan suatu *item*. Klasifikasi tingkat kesulitan instrumen ditentukan oleh kombinasi nilai rata-rata logit dan standar deviasi, menurut Sumintono dan Widhiarso (2015 dalam Lestari dkk., 2022). Kategorisasi *item measure* adalah sebagai berikut.

Tabel 4. Kategorisasi Item Measure

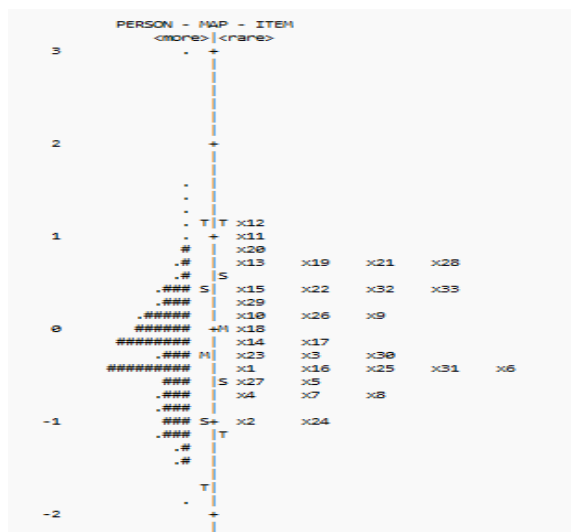
Item measure	Kategori	Nomor Item
> 0,73	Sangat sulit	b11, b12, b19, b20
0,0 sampai 0,73	Sulit	b9, b10, b13, b15, b18, b21, b22, b23, b27, b29, b30, b33
0,0 sampai - 0,73	Mudah	b1, b3, b4, b5, b6, b7, b14, b16, b17, b24, b26, b28, b31, b32
< -0,73	Sangat mudah	b2, b8, b25

Berdasarkan kategorisasi tersebut hasil analisa *item* menunjukkan bahwa *item* yang paling sulit disetujui oleh responden adalah *item* 12 dengan nilai logit 1.17. Item ini berbunyi “*Saya merasa acuh tak acuh terhadap perkuliahan saya*”. Sedangkan *item* yang paling mudah disetujui oleh responden adalah *item* 25 dengan nilai logit -0.93 yang berbunyi “*Saya cenderung khawatir*”.

Analisis Peta, Wright

Menurut Gambar Peta *Wright*, item X12 memiliki nilai logit yang menunjukkan tingkat kesukaran tinggi, sementara item X2 dan X24 memiliki nilai logit yang menunjukkan tingkat kesukaran rendah.

Gambar 1. Peta Wright



Analisis Person Measure

Person measure bertujuan untuk mengevaluasi sejauh mana karakteristik individu dalam satu dimensi tunggal yang diukur oleh instrumen, berdasarkan respons mereka terhadap *item* dalam instrumen tersebut. Nilai logit paling tinggi adalah subjek 130 dengan nilai logit 3.04 dan subjek dengan nilai logit terendah adalah subjek 109 dengan nilai logit -3.08.

Analisis Person Fit Order

Person fit order digunakan untuk mengidentifikasi individu yang memiliki respons tidak sesuai dengan yang diharapkan. Individu yang tidak sesuai ini dapat membantu peneliti dalam menemukan penyebab ketidaksesuaian tersebut (Febiana et al., 2024). Kriteria untuk mengetahui *person fit order* adalah dengan nilai *Infit* MNSQ dari setiap individu. Nilai rata-rata dan standar deviasi dijumlahkan terlebih dahulu, kemudian dibandingkan dengan nilai *Infit* MNSQ individu. Jika nilai logit seseorang lebih besar dari nilai tersebut, maka individu tersebut dikategorikan sebagai *misfit* (Sumintono & Widhiarso, 2014). Berdasarkan hasil yang telah didapat nilai MEAN dan SD adalah $1.02 + .58 = 1.6$. Berdasarkan kriteria tersebut terdapat 26 orang yang misfit karena nilai logitnya melebihi 1.6.

Analisis Rating Scale

Tabel 5. Analisis Rating Scale

Label Kategori	Observed Average	Andrich Threshold
1	-1.16	NONE
2	-.60	-1.46
3	-.11	-.43
4	.29	.52
5	.88	1.37

Analisis rating scale bertujuan untuk mengetahui apakah pilihan jawaban yang diberikan membingungkan bagi responden atau tidak (Febiana et al., 2024). Dalam instrumen ini terdapat lima pilihan jawaban yaitu sangat setuju, setuju, netral, tidak setuju dan sangat tidak setuju. Berdasarkan tabel di atas, nilai *observed average* mengalami peningkatan pada setiap pilihan jawabannya. Kemudian untuk nilai *Andrich Threshold* bergerak dari NONE kemudian negatif dan positif. Hal ini berarti bahwa opsi yang diberikan sudah valid (Sumintono & Widhiarso, 2014). Sehingga dapat disimpulkan bahwa pilihan jawaban yang diberikan tidak membingungkan bagi para responden.

Analisis Summary Statistik

Tabel 6. Rating Scale Instrument Quality Criteria

Criterion	Poor	Fair	Good	Very Good	Excellent
Targeting*	>2 errors	1-2 errors	< 1 error	< .5 error	< .25 error
Item Model Fit Mean-Square Range Extremes	<.33 - >3.0	.34-2.9	.5-2.0	.71-1.4	.77-1.3
Person and Item Measurement Reliability	<.67	.67-.80	.81-.90	.91-.94	>.94
Person and Item Strata Separated	2 or less	2-3	3-4	4-5	>5
Ceiling effect: % maximum extreme scores	>5%	2-5%	1-2%	.5-1%	<.5%
Floor effect: % minimum extreme scores	>5%	2-5%	1-2%	.5-1%	<.5%
Variance in data explained by measures**	??	??	??	??	??
Unexplained variance in contrasts 1-5 of PCA of residuals	>15%	10-15%	5-10%	3-5%	<.3%

This table has been developed by William P. Fisher, Jr. based on the Rasch literature and his many years of experience conducting Rasch analysis in different settings.

Tabel 7. Hasil Analisis Summary Statistic

	Output	Hasil
<i>Person</i>	<i>Mean</i>	92.0
	Reliabilitas	0.91
	Indeks Separasi	3.09
	Pemisah Strata (H)	4.45
<i>Item</i>	<i>Mean</i>	566.2
	Reliabilitas	0.98
	Indeks Separasi	7.16
<i>Person</i>	Pemisah Strata (H)	9.88
	<i>Infit</i> MNSQ	1.02
	<i>Outfit</i> MNSQ	1.03
	<i>Infit</i> ZSTD	-0.3
	<i>Outfit</i> ZSTD	-0.2
<i>Item</i>	<i>Infit</i> MNSQ	1.00
	<i>Outfit</i> MNSQ	1.03
	<i>Infit</i> ZSTD	-0.2
	<i>Outfit</i> ZSTD	0.2

Nilai *alpha cronbach* (KR20) merupakan nilai keseluruhan interaksi antara *person* dan *item*. Berdasarkan tabel, nilai *alpha cronbach* menunjukkan 0.92 yang termasuk pada kategori “*very good*”. Hal ini menunjukkan adanya kesesuaian antara responden dengan instrumen yang digunakan (Wibisono, 2016). Berdasarkan tabel diatas *person reliability* menunjukkan nilai 0.91 yang termasuk pada kategori “*very good*” dan *item reliability* menunjukkan nilai 0.98 yang termasuk pada kategori “*excellent*”. Hal ini berarti responden memberikan jawaban pada seluruh item dengan serius dan tidak secara acak dan keseluruhan *item* teridentifikasi memiliki ketepatan dan konsisten. Hasil analisis menunjukkan bahwa indeks separasi *person* adalah 3.09 dan indeks separasi itemnya adalah 7.16. Artinya terdapat 3 kelompok item dan 7 kelompok responden. Kemudian untuk mengetahui pemisahan strata menggunakan $H = \{(4 \times \text{indeks separasi}) + 1\}/3$. Hasilnya, untuk responden $H = 4.45$ (dibulatkan menjadi 5). Hal ini menunjukkan bahwa responden dapat dibagi ke dalam lima kelompok. Sementara itu, untuk *item* $H = 9.88$ (dibulatkan menjadi 10). Artinya, item-item yang digunakan dalam instrumen dapat dibagi ke dalam 10 level berdasarkan tingkat kesulitannya untuk disetujui responden.

Kemudian untuk ketentuan idealnya INFIT MNSQ dan OUTFIT MNSQ adalah 1.00 atau semakin mendekati 1.00 maka semakin baik. Berdasarkan tabel di atas hasil INFIT MNSQ *person* menunjukkan nilai 1.02 dan OUTFIT MNSQ *person* menunjukkan nilai 1.03, yang berarti keduanya masih dekat dengan nilai ideal, sehingga bisa dikatakan fit. Sedangkan untuk INFIT MNSQ *item* menunjukkan nilai 1.00 yang berarti termasuk baik dan OUTFIT MNSQ menunjukkan nilai 1.03 yang yang berarti masih dekat dengan nilai ideal, sehingga bisa dikatakan fit.

Kemudian untuk nilai INFIT ZSTD dan OUTFIT ZSTD person adalah -0.3 dan -0.02, keduanya mendekati 0 sehingga dapat dikatakan baik. Sedangkan untuk nilai INFIT ZSTD dan OUTFIT ZSTD item adalah -0.2 dan 0.2, keduanya mendekati 0 sehingga dapat dikatakan baik.

4. Pembahasan

Penelitian ini bertujuan untuk merancang dan mengembangkan versi Indonesia dari alat ukur *Burnout Assessment Tool* (BAT) for student yang diadaptasi dari instrumen BAT milik Popescu dkk., (2024). Pengujian dilakukan dengan *Rasch Model* agar dapat mengukur burnout pada mahasiswa Universitas Pendidikan Indonesia. Alat ukur ini terdiri dari 33 aitem dan 6 faktor yaitu *exhaustion, mental distance, emotional impairment, cognitive impairment, psychological distress, dan psychosomatic complaints*.

Hasil analisis unidimensionalitas menunjukkan *raw variance explained by measures* sebesar 40,4%, yang termasuk dalam kategori “Baik”. Hasil analisis juga menunjukkan bahwa nilai *unexplained variance 1st to 5th contrast* dapat dikategorikan sebagai ideal karena nilai kurang dari 15%. Berdasarkan hasil analisa uji *item fit*, terdapat 31 *item* yang memenuhi kriteria *fit*. Temuan ini menunjukkan bahwa 31 *item* tersebut memiliki kualitas yang baik. Hasil analisis *item measure* menunjukkan bahwa *item* yang paling sulit dijawab oleh responden adalah *item* 12 dengan nilai logit sebesar logit 1.17 dan *item* yang paling mudah dijawab oleh responden adalah *item* 25 dengan nilai logit sebesar -0.93. Analisis person measure menunjukkan bahwa nilai logit paling tinggi adalah subjek 130 dengan nilai logit 3.04 dan subjek dengan nilai logit terendah adalah subjek 109 dengan nilai logit -3.08. Selanjutnya hasil analisis *person fit order* menunjukkan bahwa terdapat 26 orang yang termasuk kategori misfit. Hasil analisis *rating scale* menunjukkan bahwa pilihan jawaban yang diberikan tidak membingungkan responden karena nilai *observed average* mengalami peningkatan pada setiap pilihan jawabannya dan nilai *Andrich Threshold* bergerak dari *NONE* kemudian negatif dan positif.

Berdasarkan hasil analisis *summary statistics* didapatkan nilai reliabilitas *alpha cronbach* 0.92 dan reliabilitas *person* 0.91 yang termasuk dalam kategori *very good*, serta item reliabilitas dengan nilai 0.98 yang termasuk dalam kategori *excellent*. Nilai indeks separasi person adalah 3.09 dan indeks separasi itemnya adalah 7.16. Kemudian dilakukan penghitungan pemisahan strata menggunakan $H = \{(4 \times \text{indeks separasi}) + 1\}/3$. Hasilnya menunjukkan bahwa responden dapat dibagi ke dalam lima kelompok sedangkan item-item yang digunakan dalam instrumen dapat dibagi ke dalam 10 level berdasarkan tingkat kesulitannya untuk disetujui responden. Nilai INFIT MNSQ (1.02) dan OUTFIT MNSQ person (1.03) masih mendekati ideal, sehingga *fit*. INFIT MNSQ item (1.00) tergolong baik, dan OUTFIT MNSQ item (1.03) juga *fit*. Sementara itu, nilai INFIT ZSTD dan OUTFIT ZSTD person (-0.3 dan -0.2) serta INFIT ZSTD dan OUTFIT ZSTD item (-0.2 dan 0.2) mendekati 0, sehingga dapat dikatakan baik.

Sejalan dengan hasil tersebut, alat ukur ini menunjukkan potensi yang baik dalam mengukur tingkat *burnout* mahasiswa di Universitas Pendidikan Indonesia. *BAT for Student* memungkinkan pengukuran yang akurat terhadap tingkat *burnout* di kalangan mahasiswa sehingga membuka peluang bagi penelitian lebih lanjut. Dengan alat ini, dosen dan pembimbing akademik dapat melakukan intervensi lebih awal untuk mahasiswa yang menunjukkan tanda-tanda *burnout*. Ini penting untuk mencegah dampak negatif yang lebih serius pada kesehatan mental dan akademik mereka

5. Kesimpulan

Penelitian ini bertujuan untuk mengembangkan versi Indonesia dari alat ukur *Burnout Assessment Tool* (BAT) untuk mahasiswa, yang diadaptasi dari instrumen BAT milik Popescu dkk. (2024). Pengujian yang dilakukan menggunakan *Rasch Model* menunjukkan hasil yang cukup menjanjikan. Sebanyak 31 item dinyatakan *fit*, dengan *raw variance explained by measures* sebesar 40,4% dan *unexplained variance pada 1st to 5th contrast* di bawah 15%, yang mengindikasikan unidimensionalitas yang memadai. Selain itu, reliabilitas alat ukur menunjukkan hasil yang baik, dengan nilai *Cronbach's alpha* 0,92, *person reliability* 0,91, dan *item reliability* 0,98. Indeks separasi person (3,09) dan item (7,16) menunjukkan bahwa responden dapat dikelompokkan ke dalam lima strata, sedangkan item dapat dibagi ke dalam 10 level berdasarkan tingkat kesulitannya. Analisis *fit statistics* menunjukkan nilai INFIT MNSQ dan OUTFIT MNSQ person serta item mendekati ideal (1,00–1,03), dan nilai INFIT ZSTD serta OUTFIT ZSTD mendekati 0, yang mengindikasikan kualitas alat ukur yang cukup baik.

Secara keseluruhan, alat ukur ini menunjukkan potensi yang baik untuk digunakan dalam mengukur tingkat *burnout* pada mahasiswa di Universitas Pendidikan Indonesia. Namun, diperlukan penelitian lebih lanjut dengan sampel yang lebih besar dan beragam untuk memastikan validitas dan reliabilitas alat ukur ini secara lebih menyeluruh. Bagi peneliti yang tertarik untuk mengembangkan atau menggunakan alat ukur ini, disarankan untuk melakukan analisis psikometrik ulang. Salah satu langkah yang dapat diambil adalah dengan meningkatkan jumlah responden dalam pengujian untuk mendapatkan hasil yang lebih valid dan reliabel. Selain itu, evaluasi yang lebih mendalam terhadap struktur item serta kemungkinan revisi desain alat ukur juga perlu dilakukan untuk memastikan alat ukur ini memenuhi standar validitas dan reliabilitas yang diharapkan.

6. Referensi

- Adruce., S. A. Z., Stephen, D., Suhaili, D. S. A. A., John, A. N., Othman, A. M. H. A., & Latif, N. A. (2017). Validation of Perceived Community Safety Instrument Using Polytomous Rasch Measurement Model. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(3-11), 125-129.
- Aini, N., & Sulaiman, S. (2024). Gambaran Academic Burnout Mahasiswa Tingkat Akhir Yang Bekerja. *Journal Of Public Health Science*, 1(2), 181-186.
- Amarsa, R. R., Ramdhani, R. N., Taufiq, A., & Saripah, I. (2023). Kecenderungan academic burnout pada mahasiswa bimbingan dan konseling universitas pendidikan indonesia. *G-Couns: Jurnal Bimbingan dan Konseling*, 7(03), 395-405.
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. Library of Congress Cataloging in Publication Data
- Destyanto, T. Y. R., Anindyajati, A., & Putri, I. S. (2023, October). Perilaku Workaholism sebagai Mekanisme Preventif Terhadap Kejadian Burnout Mahasiswa Teknik Industri. In *Prosiding Seminar Nasional Teknik Industri (SENASTI)* (Vol. 1, pp. 345-354).
- Efandi, S., & Putri, R. N. (2022). Pengujian Skala Efikasi Diri Perawat di RS X Menggunakan Model Rasch. *Psyche 165 Journal*, 170-175
- Febiana, R., Malay, M. N., & Wahyuni, C. (2024). ADAPTASI ALAT UKUR DIFFERENTIATED PROCESS SCALE OF SELF-FORGIVENESS. *Jurnal Psikologi*, 17(2), 340-359.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis: Pearson new international edition PDF eBook*. Pearson Higher Ed.
- Lestari, W., Wigati, I., Sholeh, M. I., & Pramita, D. (2022). Instrumen Literasi Digital Guru Menggunakan Model Rasch. *Orbital: Jurnal Pendidikan Kimia*, 6(2), 104-113.
- Orpina, S., & Prahara, S. A. (2019). Self-efficacy dan burnout akademik pada mahasiswa yang bekerja. *Indonesian Journal of Educational Counseling*, 3(2), 119-130.
- Popescu, B., Maricuțoiu, L. P., & De Witte, H. (2024). The student version of the Burnout assessment tool (BAT): psychometric properties and evidence regarding measurement validity on a romanian sample. *Current Psychology*, 43(3), 2037-2051.
- Putri, T. H., Masitoh, Z., & Khalid, F. (2023). Faktor yang mempengaruhi academic burnout pada mahasiswa keperawatan. *Jurnal Keperawatan Jiwa*, 11(2), 379-386.
- Schaufeli, W. B., Desart, S., & De Witte, H. (2020). Burnout Assessment Tool (BAT)—Development, validity, and reliability. *International Journal of Environmental Research and Public Health*, 17(24), 9495. <https://doi.org/10.3390/ijerph17249495>
- Sofa, P. N., Andini, D. R. A., Dya, D. A., Luthfiah, K., Fathiarafa, M., Aditiya, Y., & Ramdani, Z. (2024). Pengembangan Alat Ukur Skala *Self Regulated Learning* pada Mahasiswa yang Mengalami *Burnout Akademik*. *IBERS: Jurnal Pendidikan Indonesia Bermutu*, 3(1), 13-25.
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial (Edisi Revisi)*. Trim Komunikata Publishing House. *Cimahi City, Indonesia*.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. *Cimahi* : Trim Komunikata Publishing House.
- Wibisono, S. (2016). Aplikasi model rasch untuk validasi instrumen pengukuran fundamentalisme agama bagi responden muslim. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia*, 5(1), 1-29.
- Widodo, W., & Chotimah, C. (2023). Adaptasi dan Analisis Psikometri Skala Kompetensi Multikultural Calon Guru Menggunakan Pemodelan Rasch. *Jurnal Pendidikan Dan Kebudayaan*, 8(2), 153-172.



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

Fransiscus Widodo Setya Budi¹, Alexander Bhirowo Putra Kurniawan², Ananta Yudiarso³

^{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 konstruk (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 konstruk, 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 margin 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

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

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

Analisis Statistik

Perangkat lunak WINSTEP® versi 5.1.0 menggunakan model Rasch untuk menganalisis data. Properti psikometrik yang dianalisis meliputi keunidimensionalitas, 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. Keunidimensionalitas 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 keunidimensionalitas. 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 keunidimensionalitas pada tiga subskala komitmen: komitmen afektif, komitmen kontinuan, dan komitmen normatif. Sub-skala komitmen afektif menunjukkan keunidimensionalitas 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 keunidimensionalitas yang kurang kuat. Potensi multidimensionalitas terlihat pada item 1 dan 4, yang memiliki loading tinggi pada residual. Sub-skala komitmen normatif memiliki keunidimensionalitas 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 keunidimensionalitas 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 keunidimensionalitas instrumen.

Tabel 2. Unidimensionalitas OCQ versi Indonesia setelah menghapus item

<i>Dimensi</i>	<i>Row Varian Explained by Measure</i>	<i>First Contrast</i>	<i>Deleted item > 0,5</i>	<i>Conclusion</i>
<i>Affective Commitmen (AC)</i>	46%	1.9	CA5, CA6, CA8	unidimensi
<i>Continuance Commitmen (CC)</i>	44%	1.6	CC1, CC4	unidimensi
<i>Normative Commitment (NC)</i>	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.

Reliabilitas

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	<i>Person Reliability</i>	<i>Person Separation</i>	<i>Item Reliability</i>	<i>Item Separation</i>	<i>Alpha Cronbach</i>
<i>Affective Commitmen (AC)</i>	0,58	1,18	0,98	7,84	0,59
<i>Continuance Commitmen (CC)</i>	0,66	1,38	0,97	6,18	0,69
<i>Normative Commitment (NC)</i>	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 *affektive 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 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	Threshold
<i>Affective Commitmen (AC)</i>	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
<i>Continuance Commitmen (CC)</i>	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
<i>Normative Commitment (NC)</i>	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 Correlation
			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

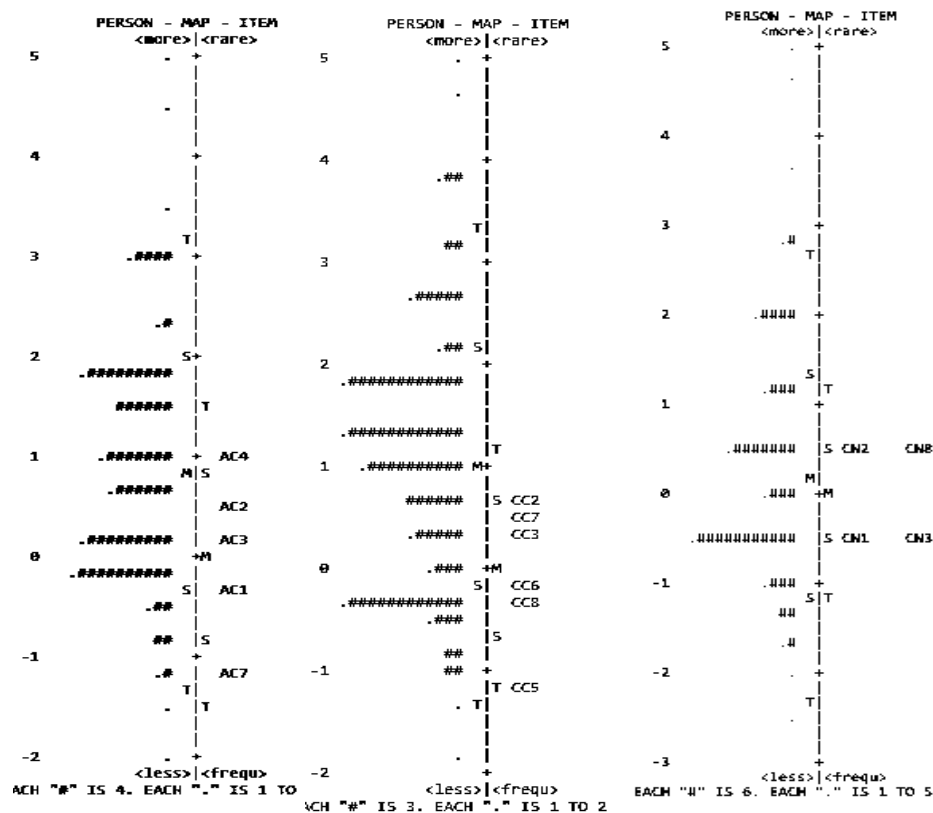
Dari hasil analisis item fit, sebagian besar item memiliki nilai *Outfit* MNSQ dan *Infit Mean Square (MNSQ)* dalam rentang yang dapat diterima, yaitu 0,5 hingga 1,5. Dengan kata lain, item-item ini dapat diandalkan untuk mengukur komitmen organisasi responden karena sesuai dengan model Rasch (Sumintono & 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.

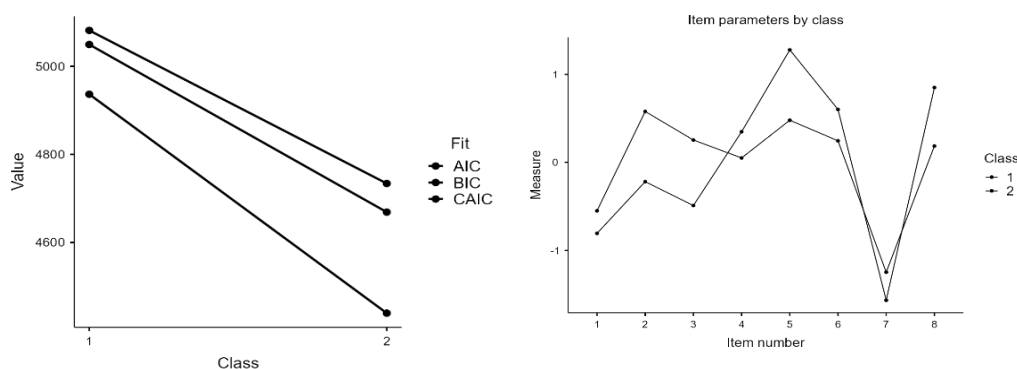
Hasil analisis menunjukkan bahwa beberapa item pada sub-skala komitmen afektif dan komitmen kontinuitas menunjukkan bias yang signifikan berdasarkan gender dan tingkat pendidikan. Ini menunjukkan bahwa *item-item* tersebut harus dianalisis lebih lanjut untuk memastikan relevansi bias terhadap konstruk 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

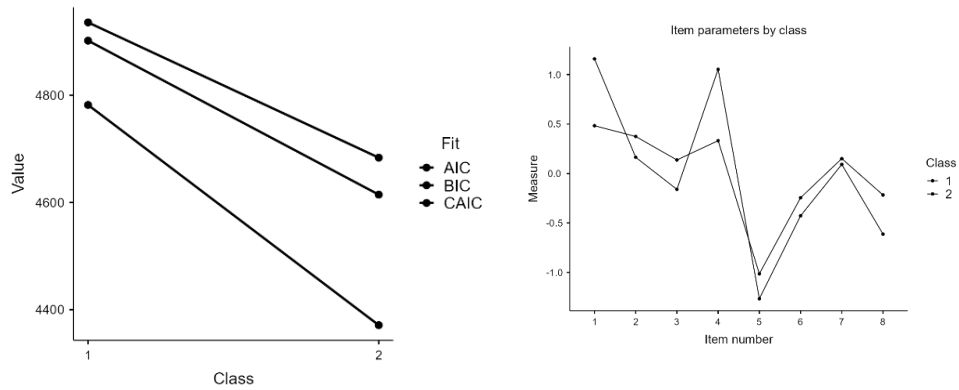
Analisis menggunakan Mixture Rasch Model berguna untuk mengidentifikasi heterogenitas dalam data responden dengan mengelompokkan mereka ke dalam beberapa subpopulasi (latent classes) berdasarkan pola respons mereka. Analisis ini membantu dalam memperbaiki validitas pengukuran dengan mengidentifikasi 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.

Tabel 7. Latent Classes OCQ versi Indonesia

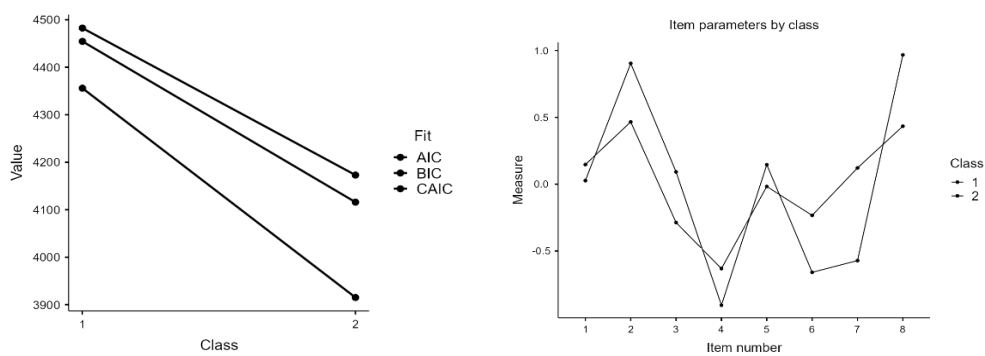
<i>Sub-Scale</i>	<i>Class</i>	<i>AIC</i>	<i>BIC</i>	<i>CAIC</i>	<i>Log-likelihood</i>
<i>CA</i>	1	4936	5049	5081	-2436
	2	4439	4669	4734	-2155
<i>CC</i>	1	4782	4902	4936	-2357
	2	4371	4615	4684	-2117
<i>CN</i>	1	4356	4454	4482	-2150
	2	3915	4116	4173	-1901



Grafik 1. Penurunan nilai AIC, BIC, CAIC dan Parameter Item Laten Class pada Dimensi CA



Grafik 2. Penurunan nilai AIC, BIC, CAIC dan Parameter Item Laten Class pada Dimensi CC



Grafik 3. Penurunan nilai AIC, BIC, CAIC dan Parameter Item Laten Class pada Dimensi CN

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](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](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.

TERAPI FOBIA SPESIFIK SESI TUNGGAL DENGAN *THE HEART TECHNIQUE***Dyah Ayu Permatasari Gunawan¹, Lina Natalya², Nadia Sutanto³**^{1,2,3} Prodi Psikologi, Fakultas Psikologi, Universitas Surabaya, Surabaya, IndonesiaEmail: ayupgunawan@gmail.com

Received :	Revised	Accepted :	Published
20 Oktober 2024	10 Maret 2025	18 April 2025	30 April 2025

Abstract

The Heart Technique is claimed by its users and creator to be effective in overcoming specific phobias in just 20 to 30 minutes. This study aims to test the effectiveness of *The Heart Technique* in online settings toward specific phobias. The research used two-group pretest–posttest design in individual and group sessions. The participants were 130 people aged 18-58 years, domiciled in Indonesia, selected using purposive sampling. The measuring instruments are Subjective Unit of Distress (SUD) and Severity Measure for Specific Phobia-Adult. The results of the Wilcoxon Signed-Rank Test analysis showed that there was a significant decrease in the measurement of the two measuring instruments of specific phobia ($p < 0.05$). Individual and group sessions were able to significantly reduce the severity of the phobia to a neutral or non-phobic category, however this technique was more effective in individual sessions as seen from the Mann Whitney-U test ($p < 0.05$). The results showed that *The Heart Technique* can be used as an alternative solution to overcome psychological problems, especially phobias.

Keywords: Emotional Problem, Specific Phobia, *The Heart Technique*, Therapeutic Effectiveness**Abstrak**

The Heart Technique diklaim oleh pengguna dan penciptanya, efektif mengatasi fobia spesifik dalam waktu 20 hingga 30 menit saja. Penelitian ini bertujuan menguji keefektifan *The Heart Technique* secara daring terhadap fobia spesifik. Metode penelitian menggunakan desain *two-group pretest–posttest design* pada sesi individu dan kelompok. Partisipan berjumlah 130 orang berusia 18-58 tahun, berdomisili di Indonesia, dipilih menggunakan *purposive sampling*. Alat ukur yang digunakan adalah Subjective Unit of Distress (SUD) dan Severity Measure for Specific Phobia-Adult. Hasil analisis Wilcoxon Signed-Rank Test menunjukkan penurunan signifikan pada pengukuran kedua alat ukur fobia spesifik ($p < 0.05$). Sesi individu dan kelompok mampu menurunkan secara signifikan tingkat keparahan fobia hingga menjadi netral atau kategori bukan fobia, namun teknik ini lebih efektif dalam sesi individu yang dapat dilihat dari uji Mann Whitney-U ($p < 0.05$). Hasil penelitian menunjukkan *The Heart Technique* dapat digunakan sebagai solusi alternatif untuk mengatasi permasalahan psikologis, terutama fobia.

Kata Kunci: Fobia Spesifik, Keefektifan Terapi, Permasalahan Emosi, *The Heart Technique*.**1. Pendahuluan**

Fobia merupakan sebuah program pikiran di bawah sadar, terdiri atas dua komponen, yaitu: kumpulan memori dan emosi negatif yang melekat dengan skala intensitas minimal menengah hingga tinggi. Program pikiran tersebut menjadi aktif ketika terpicu stimulus yang dapat berbentuk seperti objek atau situasi asli, cerita, gambar, ataupun imajinasi individu. Kondisi fobia ini menyebabkan individu secara aktif menghindari objek atau situasi fobia tersebut (American Psychiatric Association, 2013). Fobia spesifik memiliki lima kategori stimulus, yaitu: hewan, situasi, darah-injeksi-luka, lingkungan alam, dan objek atau situasi lainnya yang tidak termasuk pada keempat kategori sebelumnya.

Survei *World Health Organization* (WHO) memberikan hasil penelitian prevalensi fobia pada sampel sebanyak 124,902 orang dari 22 negara (Wardenaar et al., 2017). Survei tersebut menunjukkan sebanyak 5.5% individu masuk dalam kategori mengalami fobia dengan prevalensi 12 bulan dan sebanyak 7.4% individu masuk pada kategori prevalensi seumur hidup. Berdasarkan data prevalensi tersebut diketahui dari 124,902 individu, terdapat 6,869 hingga 9,242 individu yang hidupnya terganggu oleh fobia spesifik, sehingga perlu dipikirkan cara yang efektif dan efisien untuk membantu individu dengan gangguan fobia spesifik.

Melanjutkan, survei menyatakan bahwa wanita memiliki peluang lebih tinggi mengalami fobia spesifik dibandingkan pria (*American Psychiatric Association*, 2013; Wardenaar et al., 2017). Peneliti melakukan survei awal dan didapatkan data yang serupa dengan penelitian Wardenaar et al. dan *American Psychiatric Association*. Survei awal ini dilakukan dengan menyebarkan survei di media sosial yang ditujukan pada responden yang memiliki fobia spesifik. Responden yang mengisi survei diberikan penjelasan singkat tentang fobia spesifik dan kemudian bila responden merasa penjelasan yang diberikan sesuai dengan kondisinya maka dapat melanjutkan mengisi data diri, menjawab pertanyaan fobia apa saja yang dimiliki dan sudah berapa lama mengalaminya, perasaan yang dirasakan terhadap objek atau situasi fobia, dan apakah pernah mencoba mengatasi fobia yang dialami.

Survei oleh Wardenaar et al. (2017) juga memberikan data prevalensi fobia spesifik 12 bulan dan seumur hidup di mana pada wanita lebih tinggi 7.7% (9,617 orang) dan 9.8% (12,241 orang), dibandingkan pada pria 3.3% (4,122 orang) dan 4.9% (6,120 orang). Hasil survei ini sejalan dengan hasil penelitian *American Psychiatric Association* (2013) dan hasil survei awal oleh peneliti di mana sebanyak 75.76% (75 orang) wanita menyatakan mengalami fobia dan pria sebanyak 24.24% (24 orang). Data *American Psychiatric Association* (2013) juga melaporkan dari individu yang mengalami fobia, 75% memiliki lebih dari satu macam fobia. Data survei awal penelitian ini juga memberikan hasil 53.53% responden menyatakan memiliki lebih dari satu macam fobia dan di antaranya, pada kategori fobia spesifik yang berbeda.

DSM V menyatakan individu yang mengalami fobia spesifik akan merasakan takut dan cemas (*American Psychiatric Association*, 2013), namun hasil survei awal peneliti menunjukkan emosi negatif yang dirasakan oleh individu tidak hanya kedua emosi itu saja. Hasil survei awal menyatakan para responden juga merasakan jijik, tertekan, tidak mampu, geli, capek, sedih, muak, benci, dan masih banyak emosi negatif lainnya terhadap fobianya. Selain itu, emosi-emosi ini juga memberikan pengaruh pada tubuh fisik mereka. Respon fisik ketika berhadapan dengan stimulus fobia yang dilaporkan antara lain adalah jantung berdebar kencang, gemetar, panik, pusing, gatal-gatal, mual, dan masih banyak lagi.

Kerugian-kerugian yang dialami individu dengan fobia tidak berhenti pada kondisi fobia mempengaruhi pikiran, emosi, dan tubuh secara negatif saja, namun kondisi ini juga mengganggu kehidupan individu. Fobia spesifik dapat menjadi pencetus munculnya gangguan emosi lainnya seperti gangguan kecemasan umum, panik, dan depresi mayor (Lieb et al., 2016; Sancassiani et al., 2019). Salah satu contoh ekstrim individu yang mengalami fobia adalah seorang individu dengan fobia kecoak, setiap kali melihat kecoak di jalan, individu tersebut dapat memberikan respon panik dan berlari ke tengah jalan (Gunawan, 2020). Meskipun tipe gangguan kecemasan ini umum dialami masyarakat, data survei awal juga mengungkapkan 98.98% responden tidak berupaya secara aktif mencari bantuan untuk sembuh dari kondisi ini. Rata-rata responden melaporkan memilih untuk menghindari (*active avoidance*) atau menahan perasaan negatif (contoh: rasa takut dan cemas) ketika berhadapan dengan stimulus dan/ atau memberanikan diri dengan melakukan paparan diri (*exposure*). Salah satu alasan yang dapat menjelaskan individu dengan fobia, tidak mencari bantuan profesional adalah karena dibutuhkan banyak sesi dengan durasi yang lama pada setiap sesi untuk menyembuhkan fobia.

Penelitian-penelitian sebelumnya menunjukkan terdapat beberapa teknik psikoterapi yang sering digunakan untuk mengatasi fobia spesifik. Teknik *cognitive-behavioral therapy* (CBT) melalui 12 sesi yang masing-masing sesi berdurasi 60 menit (Riddle-Walker et al., 2016) dan teknik *time intensive* CBT dilakukan dalam 8 sesi di mana durasinya memiliki variasi dari 90 menit hingga 4 jam per sesi (Keyes et al., 2020). Teknik *exposure therapy* dilakukan dalam durasi 3 jam selama 1 sesi (Botella et al., 2016; Miloff et al., 2019). Penelitian-penelitian psikoterapi di atas mampu menyembuhkan fobia, namun ini dicapai setelah partisipan melalui bersesi-sesi intervensi dan/atau durasi intervensi yang lama dengan derajat capaian yang berbeda juga. Selain itu, belum ditemukan penelitian yang menguji efektivitas teknik terapi terhadap seluruh kategori fobia spesifik. Oleh karenanya, penelitian ini memiliki tujuan menguji efektivitas *The Heart Technique* pada semua kategori fobia spesifik.

The Heart Technique merupakan teknik terapi yang menggabungkan teori dan pendekatan *ego personality*, hipnoterapi, motivasi pikiran bawah sadar, trauma dan tubuh, sistem meridian, beserta neurosains. *The Heart Technique* bekerja dengan menghilangkan emosi yang melekat pada memori yang menjadi penyebab fobia. *The Heart Technique* dapat menyembuhkan beragam masalah emosi, termasuk fobia (Gunawan, 2018b). Teknik ini memiliki keunggulan yaitu teknik ini sederhana, mudah untuk dipelajari dan dipraktikkan pada orang lain ataupun pada diri sendiri karena teknik ini diciptakan untuk tujuan swaterapi, dan dapat mengatasi masalah emosi, fobia, dalam waktu relatif singkat yaitu berkisar 20 hingga 30 menit. Pada dasarnya teknik terapi ini

dilakukan oleh individu terhadap diri mereka sendiri dengan mengikuti panduan protokol terapi yang tertera di buku atau dengan bantuan tuntunan terapis yang mengikuti protokol di buku *The Heart Technique*. Dengan demikian, individu memiliki peran aktif dalam proses menyembuhkan dirinya sendiri.

Ribuan orang telah membeli buku *The Heart Technique* dan mempraktikkan teknik ini sejak tahun 2018. Selain itu juga terdapat banyak pelatihan *The Heart Technique* yang telah diselenggarakan untuk masyarakat umum seperti di Jakarta dengan 1,100 peserta dan Medan dengan 1,300 peserta (Gunawan, 2018a), Palembang (Tribun Sumsel, 2018), di sekolah dengan peserta murid dan guru (Rumah Sejahtera, 2018; 2019b); di sekolah tinggi (Rumah Sejahtera, 2019a), dan masih banyak lagi. Oleh karena itu, peneliti tertarik untuk membandingkan keefektifan *The Heart Technique* ketika dilakukan secara individu dan kelompok. Penelitian ini dilakukan secara daring agar dapat menjangkau partisipan di seluruh Indonesia melalui teknologi internet. Alasan lain adalah dengan pertimbangan alasan kesehatan yang pada saat penelitian ini dilaksanakan masih pada masa pandemi COVID-19.

Penelitian ini memiliki dua hipotesis, yaitu: 1) *The Heart Technique* yang dilakukan secara daring efektif mengatasi fobia spesifik dan 2) Tidak ada perbedaan efektivitas terapi *The Heart Technique* secara kelompok dan individu.

2. Metode Penelitian

Penelitian eksperimen ini menggunakan *two group pretest-posttest design* yang dilakukan secara daring. Tahapan penelitian ini dimulai dengan, pertama, merancang kuesioner *online* menggunakan media *Google Form*. Kuesioner *online* terdiri atas penjelasan penelitian, *informed consent*, identitas diri partisipan, pertanyaan tentang fobia spesifik yang dialami, berapa lama partisipan mengalami fobia spesifik tersebut, dan dua alat ukur sebagai pertanyaan *pre-test*, yang juga digunakan lagi pada *post-test*.

Penelitian ini menggunakan dua alat ukur untuk melihat perubahan kondisi partisipan terhadap fobia yang dialaminya, yaitu *Subjective Unit of Distress* (SUD) dan *Severity Measure for Specific Phobia-Adult*. SUD merupakan sistem penilaian ketidaknyamanan emosional individu terhadap suatu peristiwa secara subjektif. SUD menggunakan skala 0 (tenang) hingga 100 (kondisi cemas intens dan terburuk) (Wolpe & Lazarus, 1966 dalam Tanner, 2012). SUD ini juga merupakan alat asesmen yang menjadi bagian dari protokol *The Heart Technique* yang digunakan sebelum, saat, dan sesudah proses intervensi dilakukan. SUD yang digunakan pada *The Heart Technique* menggunakan skala 0 hingga 10. Pengukuran SUD diambil langsung dari buku *The Heart Technique* (Gunawan, 2018b). Pertanyaan dimulai dengan kalimat “Coba Anda menutup mata dan ingat kembali objek atau situasi yang membuat Anda mengalami fobia, yang ingin Anda atasi. Berapakah intensitas emosi yang Anda rasakan? Intensitas ini dinyatakan dengan angka 0 hingga 10. Angka 0 artinya Anda sama sekali tidak merasakan apapun, sementara angka 10 artinya emosi ini sangat intens (kuat) Anda rasakan.”

Severity Measure for Specific Phobia-Adult merupakan alat ukur yang dibuat oleh Craske et al. (2013) berdasarkan DSM V dan dipublikasi oleh *American Psychiatric Association*. Berbeda dengan alat ukur lainnya yang hanya mengukur secara spesifik satu jenis atau kategori fobia, alat ukur ini dirancang untuk mengukur seluruh jenis fobia spesifik yang dimiliki individu. Peneliti melakukan *forward translation* (dari bahasa Inggris ke bahasa Indonesia) dan modifikasi pada alat ukur ini sehingga alat ukur ini menggunakan 9 butir pertanyaan dan 10 pilihan respon, dengan pilihan skoring 1 (Sangat tidak menggambarkan diri saya) hingga 10 (Sangat menggambarkan diri saya). Reliabilitas alat ukur yang memiliki unidimensi ini pada setting non-klinis adalah $\alpha = 0.93$ (tinggi) (LeBeau et al., 2012) dan $\alpha = 0.90$ (tinggi) (Knappe et al., 2014) dan pada setting klinis yaitu $\alpha = 0.96$ (tinggi) (LeBeau et al., 2012). Peneliti juga melakukan uji reliabilitas pada alat ukur dan mendapatkan hasil konsisten dengan hasil reliabilitas LeBeau et al. (2012) dan Knappe et al. (2014), dengan Alpha Cronbach > 0.7 yaitu $\alpha = 0.925$ (tinggi) dan nilai CITC > 0.3 (CITC = 0.468 - 0.894). Hasil pengukuran konsistensi butir menandakan bahwa alat ukur *Severity Measure for Specific Phobia-Adult* siap digunakan dan hasilnya dapat dipercaya memberikan pengukuran yang akurat. Pertanyaan pada alat ukur ini dimulai dengan pernyataan “Saat saya mengingat atau membayangkan bertemu atau berhadapan dengan objek, atau berada di situasi yang membuat saya mengalami fobia, maka saya...”, dilanjutkan dengan butir alat ukur, dan pemberian pilihan jawaban menggunakan skoring 1 hingga 10.

Pada tahap kedua, pengambilan data penelitian, proses ini dimulai dengan menyebarkan informasi penelitian beserta *pre-test* sesi individu di media sosial (Facebook, Instagram, dan Line). Setelah data terkumpul, peneliti melakukan *screening* atau seleksi partisipan penelitian berdasarkan hasil *self report* calon partisipan. Pemilihan partisipan menggunakan teknik *purposive sampling* di mana partisipan terpilih setelah memenuhi kriteria inklusi yang ditentukan berdasarkan hasil *pre-test* yang disebarkan melalui berbagai media sosial. Kriteria inklusi penelitian ini adalah: (1) Individu kategori dewasa, berusia minimal 18 tahun; (2) Memenuhi kriteria diagnostik DSM V yaitu mengalami fobia spesifik selama minimal 6 bulan terakhir atau lebih; (3) Memiliki minimal skala 6 pada *Subjective Unit of Distress* (SUD) terhadap fobia yang dimilikinya; (4) Memiliki tingkat minimal kategori sedang pada *Severity Measure for Specific Phobia-Adult*; (5) Tidak memiliki gangguan mental lainnya; (6) Mengikuti sesi intervensi ini atas kehendak diri sendiri dan bersedia mengikuti hingga selesai.

Partisipan penelitian eksperimen ini berjumlah total 130 orang, di mana 30 orang (5 laki-laki dan 25 perempuan) merupakan partisipan penelitian sesi individu dan 100 orang (27 laki-laki dan 73 perempuan) merupakan partisipan sesi kelompok. Rata-rata partisipan penelitian sesi individu berusia 35 tahun dengan SD = 10,813. Rata-rata partisipan penelitian sesi klasikal berusia 34 tahun dengan SD = 9,449. Para partisipan penelitian ini berasal dari berbagai pulau di Indonesia. Partisipan sesi kelompok, terbanyak berasal dari Jawa dengan jumlah 73 orang, Sumatera 11 orang, Bali 7 orang, Kalimantan 4 orang, Sulawesi 2 orang, dan masing-masing 1 orang untuk pulau Papua, Sumbawa, dan Timor. Partisipan sesi individu yang berasal dari Jawa terdapat 20 orang, masing-masing 2 orang dari pulau Kalimantan, Sulawesi, dan Riau. Adapun dari pulau Sumatera, Bali, Lombok, dan Papua masing-masing 1 orang. Kategori fobia spesifik tertinggi yang dimiliki oleh partisipan penelitian, sebesar 34.62% (45 orang) adalah binatang. Diikuti dengan kategori situasi, sebesar 33.08% (43 orang). Selanjutnya, kategori kondisi natural lingkungan sebesar 18.46% (24 orang), situasi lain sebesar 9.23% (12 orang), dan terakhir darah-injeksi-luka sebesar 4.62% (6 orang).

Seluruh partisipan penelitian menggunakan teknik intervensi *The Heart Technique*. Teknik ini terdiri atas lima tahapan (1) Identifikasi Masalah, (2) *Over Energy Correction*, Reset Jalur Energi, & *Tracing*, (3) *Priming the Subconscious for Change*, (4) *Reset Jalur Energi & Tracing*, dan (5) Pelepasan Emosi. Penelitian ini mengikuti sesuai dengan tahapan protokol terapi yang tertera dalam buku *The Heart Technique* (Gunawan, 2018b).

Seluruh sesi penelitian dilakukan secara daring dan sesi intervensi menggunakan aplikasi Zoom. Penggunaan Zoom memiliki kelebihan partisipan dapat mengikuti sesi intervensi dari mana saja, tanpa batasan geografi. Di sisi lain, kelemahan dari penggunaan Zoom adalah peneliti harus membayar biaya langganan aplikasi dan partisipan harus menyiapkan kuota internet yang cukup untuk mengikuti sesi intervensi. Sesi intervensi individu, yaitu satu sesi *one on one* antara terapis dan partisipan, dilakukan selama 60 menit dengan pembagian waktu 30 menit sesi perkenalan, penjelasan fobia dan 5 tahapan *The Heart Technique*. Selanjutnya 30 menit sesi intervensi yaitu partisipan melakukan 5 tahapan *The Heart Technique* yang dipandu oleh terapis, tanya-jawab, dan penutup. Terapis sesi intervensi individu adalah pencipta *The Heart Technique* dan psikolog klinis aktif.

Adapun sesi kelompok dilakukan dalam bentuk satu sesi webinar yang dibawakan oleh pencipta *The Heart Technique*. Sesi kelompok ini berdurasi 90 menit dengan pembagian waktu 60 menit sesi perkenalan, penjelasan fobia dan 5 tahapan *The Heart Technique*. Selanjutnya 30 menit sesi intervensi yaitu partisipan melakukan 5 tahapan *The Heart Technique* yang dipandu oleh pencipta *The Heart Technique*, tanya-jawab, dan penutup. Sesi kelompok diberikan durasi di awal yang lebih lama dibandingkan sesi individu dengan tujuan memberikan waktu para partisipan untuk bergabung sebelum sesi intervensi dimulai. Ini dilakukan dengan tujuan seluruh partisipan dapat mengikuti rangkaian terapi dengan lengkap. Setelah sesi intervensi selesai, peneliti menyebarkan tautan *post-test* secara langsung.

Setelah proses pengambilan data selesai, peneliti melanjutkan pada tahapan analisis data dan terakhir, menulis simpulan penelitian. Hasil uji normalitas data menggunakan Kolmogorov-Smirnov mendapatkan hasil Sig. = 0.036 ($p < 0.05$), yang berarti data tidak berdistribusi normal sehingga analisis data penelitian menggunakan teknik non-parametrik. Analisis data menggunakan uji beda Wilcoxon signed-rank test untuk hipotesis 1, yang membandingkan hasil nilai rata-rata *pre-test* dan *post-test*. Hipotesis 2 menggunakan uji mann whitney-u untuk membandingkan dua nilai rata-rata pada dua kelompok independen, yaitu sesi individu dan kelompok. Analisis data menggunakan program SPSS for Mac versi 25. Hipotesis diterima dan ada perbedaan jika nilai *p-value* kurang dari sama dengan 0.05 ($\text{sig} \leq 0.05$).

Penelitian yang dilakukan tidak mengajukan kaji etik dikarenakan penelitian ini dilakukan saat peneliti melakukan studi S1 Psikologi di Universitas Surabaya, di mana para mahasiswa S1 diperbolehkan melanjutkan penelitian setelah lulus ujian Seminar Proposal Penelitian. Penelitian ini memiliki *Ethical Eligibility Declaration - Ethics Clearance for Conducting Research* yang ditandatangani oleh dosen pembimbing sebagai bukti penelitian ini tidak menyalahi kaidah etik dan disupervisi secara ketat oleh dosen pembimbing.

3. Hasil dan Diskusi

Hasil pengukuran penelitian ini terdiri atas uji pada skor SUD dan skor alat ukur *Severity Measure for Specific Phobia-Adult*. Pada Tabel 1, dapat dilihat data *pre-test* dan *post-test* dari sesi individu, kelompok, dan data gabungan. Dapat dilihat berdasarkan selisih skor antara skor *pre-test* dan *post-test* di kedua alat ukur, terdapat penurunan skor tingkat keparahan fobia spesifik secara signifikan baik pada sesi individu dan kelompok setelah menjalani sesi intervensi dengan *The Heart Technique*. Dengan demikian dapat disimpulkan *The Heart Technique* efektif digunakan untuk menyembuhkan fobia. Dilakukan uji analisis statistik Wilcoxon Signed-Rank Test untuk melihat perbedaan *mean pre-test* dan *post-test* baik pada sesi individu dan kelompok yang dapat dilihat pada Tabel 2. Pada Tabel 3 dilakukan uji analisis statistik Mann Whitney-U untuk membandingkan perbedaan *mean* antara kedua kelompok.

Tabel 1. Statistik deskriptif

Alat ukur	Tipe Data	Sesi intervensi	Min	Max	Mean	SD
SUD	Pre-Test	Gabungan	6	10	8.4154	1.42396
		Individu	8	10	9.3667	0.80872
		Kelompok	6	10	8.1300	1.44708
	Post-Test	Gabungan	0	5	0.3385	0.83112
		Individu	0	2	0.0667	0.36515
		Kelompok	0	5	0.4200	0.91210
<i>Severity Measure for Specific Phobia-Adult</i>	Pre-Test	Gabungan	42	90	67.9385	13.66984
		Individu	73	90	80.5667	5.47523
		Kelompok	42	90	64.1500	13.11054
	Post-Test	Gabungan	9	54	13.1846	8.47686
		Individu	9	15	9.6333	1.51960
		Kelompok	9	54	14.2500	9.38016

Tabel 2. Hasil analisis statistik Wilcoxon Signed-Rank Test

Sesi Individu				
Alat Ukur	Asymp. Sig.	Z/ Wilcoxon	Mean Pre-test	Mean Post-test
SUD	0.001	-4.898	9.3667	0.667
<i>Severity Measure for Specific Phobia-Adult</i>	0.001	-4.787	80.5667	9.6333
Sesi Kelompok				
Alat Ukur	Asymp. Sig.	Z/ Wilcoxon	Mean Pre-test	Mean Post-test
SUD	0.001	-8.718	8.1300	0.4200
<i>Severity Measure for Specific Phobia-Adult</i>	0.001	-8.683	64.1500	14.2500

Tabel 3. Hasil analisis statistik Mann Whitney-U

Alat ukur	Asymp. Sig.	U (Mann-Whitney U)	Mean Sesi Individu	Mean Sesi Kelompok
SUD	0.014	1194.500	55.32	68.56
<i>Severity Measure for Specific Phobia-Adult</i>	0.003	1020.500	49.52	70.30

Pada Tabel 2 dapat dilihat hasil uji analisis Wilcoxon *Signed-Rank Test*, yaitu terdapat perbedaan signifikan antara hasil *pre-test* dan *post-test* di mana kedua alat ukur memiliki $p = 0.001$ ($p < 0.05$). Pada sesi individu, alat ukur SUD menunjukkan perbedaan signifikan skor antara pengukuran *pre-test* ($M = 9.3667$) dengan pengukuran *post-test* ($M = 0.667$) ($Z = -4.898$). Hasil uji statistik Wilcoxon *Signed-Rank Test* pada alat ukur *Severity Measure for Specific Phobia-Adult* sesi individu, memiliki hasil konsisten dengan SUD, di mana terdapat perbedaan signifikan skor antara pengukuran *pre-test* ($M = 80.5667$) dengan pengukuran *post-test* ($M = 9.6333$) ($Z = -4.787$).

Pada sesi kelompok, alat ukur SUD memberikan hasil perbedaan signifikan skor pengukuran *pre-test* ($M = 8.1300$) dengan pengukuran *post-test* ($M = 0.4200$) ($Z = -8.718$). Hasil ini konsisten dengan hasil tes alat ukur *Severity Measure for Specific Phobia-Adult* yang mengalami penurunan skor secara signifikan dari *pre-test* ($M = 64.1500$) ke *post-test* ($M = 14.2500$) ($Z = -8.683$).

Pada Tabel 3 dapat dilihat hasil uji Mann Whitney-U pada kedua alat ukur memiliki hasil $p < 0.05$ yang dapat diartikan terdapat perbedaan signifikan skor *post-test* antara sesi individu dan sesi kelompok. Hasil uji pada SUD *The Heart Technique*, ditemukan ada perbedaan signifikan skor antara sesi individu ($Mdn = 55.32$) dengan sesi kelompok ($Mdn = 68.56$) ($U = 1194.500$, $p = 0.014$). Adapun hasil uji pada *Severity Measure for Specific Phobia-Adult* juga memberikan hasil konsisten dengan SUD, yaitu ada perbedaan signifikan skor antara sesi individu ($Mdn = 49.52$) dengan sesi kelompok ($Mdn = 70.30$) ($U = 1020.500$, $p = 0.003$, $p < 0.05$). Dengan demikian, hipotesis tidak ada perbedaan keefektifan terapi *The Heart Technique* secara kelompok dan individual, ditolak. Dilihat dari nilai *mean*, sesi individu lebih efektif dibandingkan dengan sesi kelompok.

Penelitian ini juga membandingkan frekuensi emosi partisipan terhadap objek atau situasi fobia (Tabel 4), frekuensi wilayah tubuh merasakan emosi (Tabel 5), dan frekuensi simptom fisik (Tabel 6) yang dapat dilihat pada lampiran. Tabel 4, 5, dan 6 tertera di lampiran. Pada setiap tabel dapat dilihat penurunan frekuensi yang signifikan yang dilaporkan para partisipan dan bahkan banyak tipe emosi, sensasi di wilayah tubuh, maupun simptom fisik yang sebelumnya dirasakan karena fobia, sesuai sesi intervensi sudah tidak lagi dirasakan dan dialami oleh para partisipan.

4. Diskusi

Hasil uji *Wilcoxon Signed-Rank Test* menerima hipotesis pertama ($p = 0.001$; $p < 0.05$). Hasil pengujian menunjukkan terdapat penurunan skor tingkat keparahan fobia spesifik secara signifikan ketika membandingkan skor *pre-test* dan *post-test*.

Pada sesi individu, skor rata-rata *pre-test* SUD adalah 9.3667, di mana skor *pre-test* SUD tersebut termasuk kategori sangat parah. Kemudian skor rata-rata pada *post-test* menjadi 0.667 yang mendekati batas skor terendah yaitu 0, termasuk dalam kategori bukan fobia atau netral. Hasil perubahan skor SUD sesi individu konsisten dengan alat ukur *Severity Measure for Specific Phobia-Adult*. Skor rata-rata *pre-test* sesi individu alat ukur *Severity Measure for Specific Phobia-Adult* adalah 80.5667, termasuk dalam kategori sangat parah, mendekati nilai skor tertinggi yaitu 90. Kemudian skor rata-rata partisipan menurun, dengan perubahan rata-rata *Severity Measure for Specific Phobia-Adult* di *post-test* menjadi 9.6333, yang mendekati batas skor terendah yaitu 9, termasuk dalam kategori bukan fobia atau netral.

Pada sesi kelompok, skor rata-rata *pre-test* SUD adalah 8.1300, termasuk kategori sangat parah karena skor tersebut mendekati skor 10 yang merupakan batas skor tertinggi. Kemudian setelah sesi intervensi, perubahan skor dapat terlihat secara langsung, dengan perubahan rata-rata SUD di *post-test* menjadi 0.4200. Skor ini mendekati batas skor terendah yaitu 0. Sama dengan pada sesi individu, hasil alat ukur SUD konsisten dengan alat ukur *Severity Measure for Specific Phobia-Adult*. Skor rata-rata *pre-test* *Severity Measure for Specific*

Phobia-Adult sesi kelompok adalah 64.1500, termasuk dalam kategori parah. Skor rata-rata pada *post-test* menjadi 14.2500, mendekati skor terendah yaitu 9.

Hasil skor *post-test* sesi individu dan sesi kelompok, pada alat ukur SUD dan *Severity Measure for Specific Phobia-Adult* menunjukkan hasil rata-rata partisipan termasuk dalam kategori bukan fobia. Hasil pengukuran ini dapat memberikan kesimpulan yaitu *The Heart Technique* yang dilakukan secara daring efektif mengatasi fobia spesifik.

The Heart Technique efektif dilakukan secara daring (*online*), dapat dijelaskan oleh tiga alasan: Pertama, di awal sesi intervensi dimulai, terapis memberikan penjelasan mengenai fobia, *The Heart Technique*, dan cara *The Heart Technique* bekerja untuk menyelesaikan fobia. Penjelasan ini membantu partisipan untuk memahami dan membangun cara pandang yang tepat terkait fobia, penyebab fobia, proses yang akan dijalani, dan hasil yang akan dicapai dengan menggunakan *The Heart Technique*. Terapis menganalogikan fobia sebagai sebuah beban yang dapat dilepaskan dengan mudah di sesi intervensi ini. Kedua, terapis memberikan instruksi yang jelas dan contoh-contoh secara langsung untuk diikuti. Ketiga, terapis menanyakan kondisi diri partisipan selama sesi intervensi berlangsung. Terakhir, terapis juga memberikan sesi tanya-jawab untuk memastikan partisipan memahami instruksi dan perihal terkait *The Heart Technique*.

Edukasi tentang kondisi fobia yang dialami dan juga bagaimana *The Heart Technique* dapat membantu kondisi peserta bersifat terapeutik. Pencapaian edukasi yang telah dilakukan pada penelitian ini dapat dijelaskan oleh

manfaat dari psikoedukasi. Psikoedukasi dapat meningkatkan secara positif kondisi yang dialami individu, menurunkan stigma yang dibuat pada diri sendiri, membuat individu mau mengikuti rangkaian intervensi, dan meningkatkan keterlibatan individu dalam penanganan kondisi yang dialaminya (Motlova et al., 2017). Partisipan yang mendapatkan pengetahuan baru dapat meningkatkan ekspektasi dirinya untuk sembuh.

Keberhasilan *The Heart Technique* menurunkan tingkat keparahan fobia baik pada sesi individu maupun kelompok disebabkan proses yang terjadi pada setiap partisipan adalah sama. Proses ini berlangsung di pikiran bawah sadar, di mana *The Heart Technique* mengaktifkan memori beserta emosi yang menjadi penyebab fobia ketika partisipan membayangkan atau mengingat objek atau situasi fobia yang dialaminya, dan kemudian emosi tersebut dinetralisir. Ketika emosi negatif dinetralisir, memori yang tadinya membuat individu merasakan emosi negatif (fobia) berubah menjadi netral. Selain itu, teknik terapi ini sesungguhnya dilakukan partisipan pada diri mereka sendiri dengan mengikuti tuntunan terapis, bukan dilakukan terapis terhadap partisipan, sehingga partisipan memiliki peran aktif untuk menyembuhkan dirinya sendiri.

Keberhasilan *The Heart Technique* tidak berhenti pada penurunan skor yaitu dengan membandingkan nilai *pre-test* dan *post-test*. Efek keberhasilan teknik terapi ini juga dapat dilihat dari partisipan yang melaporkan perubahan tipe emosi (Tabel 4), wilayah tubuh yang merasakan emosi (Tabel 5), dan simtom fisik (Tabel 6) ketika membayangkan objek atau situasi fobia yang dialaminya saat *pre-test* dan *post-test*. Pada Tabel 4 dapat dilihat dari 28 jenis emosi negatif yang dirasakan menjadi tersisa 8 jenis pada *post-test* dengan jumlah partisipan yang melaporkan menurun drastis. Selain itu, hasil *post-test* juga menunjukkan terdapat 117 partisipan yang melaporkan merasa netral terhadap objek atau situasi fobia mereka dan disertai emosi positif. Sebagai contoh pada hasil *pre-test* para partisipan melaporkan tipe emosi yang paling banyak dirasakan adalah cemas (98 orang) dan di *post-test* menjadi tersisa hanya 8 orang yang merasakannya, emosi ngeri dari 75 partisipan menjadi tersisa 1 partisipan, merasa tidak berdaya dari 57 partisipan menjadi 0 partisipan, dan masih banyak lagi perubahan yang dialami partisipan.

Perubahan pada aspek emosi dari negatif menjadi netral terhadap fobia dapat dijelaskan menggunakan teori-teori pembentuk *The Heart Technique*. Pertama, teori *ego personality*, instruksi di dalam protokol *The Heart Technique* mampu dengan tepat mengaktifkan *ego personality* yang memegang informasi, memori kejadian, dan seluruh emosi partisipan terhadap fobia yang dialaminya. Dengan *ego personality* tersebut aktif, maka terapis mengetahui fobia yang ingin disembuhkan dan diselesaikan telah tepat sasaran (Daharnis et al, 2021). Kedua, instruksi *The Heart Technique* untuk mengingat kejadian fobia tersebut, membuat reseptor yang terletak di sistem meridian menjadi lebih rentan untuk diintervensi. Dengan reseptor menjadi reseptif dan titik akupuntur K27 (*kidney-27*) diketuk berulang kali, memberikan efek depotensiasi terhadap reseptor pada jalur saraf yang memunculkan emosi negatif dan mengakibatkan emosi negatif ternetralisir (Hui et al, 2000; Eden & Feinstein, 2006; Feinstein, 2008).

Ketiga, teknik ini memanfaatkan fungsi proteksi pikiran bawah sadar yaitu bekerja untuk memproteksi diri individu dari segala hal yang dipikirkan, dipersepsikan, dirasa, dipercaya, diasumsikan, dan diyakini merugikan atau membahayakan diri individu baik dari sisi pikiran maupun fisik (Banyan & Kein, 2001; Erickson, 2005; Gunawan, 2018b; Tebbetts, 1987). Melalui penyusunan pertanyaan dalam protokol *The Heart Technique* yang sedemikian rupa, yang merupakan bagian dari Tahapan (3) *Priming the Subconscious for Change*, teknik ini meningkatkan motivasi pikiran bawah sadar untuk merubah situasi individu, yaitu sembuh dari fobia yang dialaminya. Terakhir pada Tahapan (5) Pelepasan Emosi, menggunakan perpaduan metafora dan sistem meridian. *The Heart Technique* menetralisir seluruh emosi baik pada lokasi primer maupun sekunder (Gunawan, 2018b). Lokasi primer adalah lokasi wilayah tubuh yang merasakan emosi, yang dilaporkan di *pre-test* oleh para partisipan. Adapun lokasi sekunder merupakan lokasi emosi yang sebelumnya tidak disadari dimiliki oleh partisipan.

Hasil intervensi dengan *The Heart Technique* juga memberikan perubahan pada wilayah tubuh yang merasakan sensasi emosi oleh partisipan, yang dapat dilihat pada Tabel 5. Sebanyak 101 dari 130 partisipan penelitian ini melaporkan tidak merasakan sensasi di wilayah tubuh manapun setelah menjalani sesi intervensi. Sebanyak 29 partisipan lainnya masih melaporkan merasakan sensasi di beberapa wilayah tubuh. Perbandingan data *pre-test* dan *post-test* adalah sebagai berikut: 76 partisipan melaporkan merasakan sensasi di wilayah dada saat *pre-test* dan data *post-test* menyatakan hanya tersisa 7 partisipan. Area kepala dari 60 partisipan menjadi tersisa 9 partisipan, area kaki dari 47 partisipan menjadi tersisa 3 partisipan, dan masih banyak lagi perubahan pada wilayah tubuh.

Selanjutnya, melihat pada Tabel 6, perubahan jumlah partisipan yang melaporkan simtom fisik adalah sebagai berikut: Jantung berdebar dari 93 partisipan menjadi tersisa 10 partisipan, gemetar dari 60 partisipan menjadi tersisa 3 partisipan, napas memburu dari 49 partisipan menjadi 0 partisipan, mau menangis dari 40 partisipan menjadi 0 partisipan, dan masih banyak tipe simtom fisik lainnya yang partisipan menyatakan tidak merasakannya setelah melalui sesi intervensi di *post-test*.

Simtom fisik, atau sensasi fisik, adalah bentuk dari keberadaan emosi, dengan intensitas tertentu, yang tinggal di wilayah tubuh tertentu (Levine & Frederick, 1997; Nummenmaa et al., 2014). Intensitas sensasi yang dirasakan

di tubuh berjalan sebanding dengan intensitas emosi yang dirasakan oleh individu (Nummenmaa et al., 2018). Semakin intens emosi yang dirasakan, semakin kuat sensasi yang dirasakan di tubuh individu. Semakin rendah intensitas emosi, maka semakin pudar sensasi yang dirasakan di tubuh. Maka, ketika partisipan melaporkan intensitas emosi sudah menurun hingga menjadi nol, partisipan tidak akan merasakan sensasi di tubuh mereka (Nummenmaa et al., 2014).

Terdapat dua alasan yang menjelaskan mengapa partisipan masih melaporkan merasakan sensasi di tubuh, meskipun intensitas emosi terhadap fobia telah menjadi nol. Pertama, bagian tubuh partisipan sedikit lebih lambat untuk melepas “sisa” sensasi di tubuh fisik, dan ini wajar tergantung keunikan setiap individu. Solomon dan Kohn (2014) menjelaskan durasi adaptasi, adaptor merekrut *classical receptive field* dan sinyal normalisasi, antar individu dapat berbeda. Maka, “sisa” sensasi fisik tersebut akan pudar dengan sendirinya dalam beberapa menit atau jam kemudian. Alasan kedua adalah karena di lokasi fisik tersebut tinggal lebih dari satu macam emosi (Nummenmaa et al., 2014), baik yang berasal dari fobia atau dari pengalaman traumatik lainnya. *Nesting* (bersarang) adalah kondisi di mana beberapa emosi tinggal di satu tempat di wilayah tubuh. Ketika partisipan mengaktifkan emosi di wilayah tersebut, partisipan dapat secara tidak sengaja mengaktifkan emosi lain yang tinggal di wilayah tubuh yang sama. Emosi-emosi lain tersebut yang menyebabkan partisipan masih merasakan sensasi fisik walaupun intensitas emosi terhadap fobia sudah dinetralisir dengan *The Heart Technique*.

Keefektifan *The Heart Technique* ketika diaplikasikan secara daring terhadap fobia spesifik, menjadi salah satu nilai kelebihan teknik ini. *The Heart Technique* merupakan teknik intervensi yang efektif, cepat dan mudah dipraktikkan secara daring, bagi pengguna. Hasil penelitian *The Heart Technique* efektif diaplikasikan secara daring ini dapat menjadi referensi untuk pembahasan teknik intervensi berbasis daring untuk mengatasi fobia, karena selama ini tidak banyak ditemukan data dan pembahasan terkait. Hasil penelitian ini juga sejalan dengan penelitian-penelitian sebelumnya (Mor et al., 2021; Peñate & Fumero, 2015; Rondung et al., 2018) yang menyatakan teknik terapi berbasis psikologi berbasis internet dapat dengan efektif membantu menyelesaikan kondisi fobia spesifik.

Meneliti lebih jauh, berikut penjelasan keefektifan *The Heart Technique* pada partisipan sesi individu dan kelompok. Analisis uji statistik deskriptif pada Tabel 1 menunjukkan rata-rata skor sesi individu dan kelompok pada *pre-test* ketika dibandingkan dengan hasil *post-test* menurun secara signifikan. Dengan demikian dapat disimpulkan, *The Heart Technique* efektif membantu dalam menyembuhkan fobia yang dialami para partisipan baik dari sesi individu maupun kelompok.

Keefektifan *The Heart Technique* ini dapat dilihat pada sesi individu dan kelompok, akan tetapi hasil uji *Mann Whitney U* memberikan hasil uji hipotesis $p < 0.05$, baik untuk alat ukur SUD ($p = 0.014$) maupun *Severity Measure for Specific Phobia-Adult* ($p = 0.003$). Hasil uji hipotesis ini menolak hipotesis kedua, yang berarti terdapat perbedaan keefektifan terapi *The Heart Technique* pada sesi individu dan kelompok. Meskipun hipotesis kedua penelitian ini ditolak, *The Heart Technique* terbukti efektif menyelesaikan dan/ atau menurunkan tingkat keparahan fobia pada sesi kelompok.

Perbedaan hasil sesi individu dan kelompok dapat dijelaskan oleh tantangan yang muncul pada sesi kelompok. Peneliti telah mencoba meminimalisir situasi yang berada di luar kontrol peneliti, seperti dengan memberikan pemberitahuan proses intervensi membutuhkan ruangan yang tenang dan tidak terganggu, sebelum sesi intervensi dilakukan. Namun, terdapat tiga tantangan yang dihadapi pada sesi kelompok. Ketiga tantangan ini sesuai dengan penjelasan Weinberg (2020) yang menyatakan sesi kelompok yang dilakukan secara daring memiliki tantangan tersendiri.

Pertama, setiap partisipan memiliki situasi tempat tinggal yang bervariasi. Kedua, partisipan merasa kurang mendapatkan *eye-to-eye contact* atau kontak mata antara partisipan dan terapis. Terakhir, peserta cenderung tidak berpartisipasi atau tidak aktif dalam sesi intervensi. Dampak dari tantangan ini adalah partisipan terdistraksi karena kondisi rumah yang kurang kondusif, beberapa partisipan tidak berpartisipasi aktif selama sesi intervensi berlangsung yang dapat dilihat dari kamera yang tidak dinyalakan ataupun tidak menjawab pertanyaan terapis selama sesi intervensi berlangsung. Berdasarkan laporan, partisipan terdistraksi karena lingkungan di rumah, membuat mereka tidak dapat mengikuti dengan penuh protokol *The Heart Technique*. Partisipan tidak aktif selama proses intervensi memberikan dampak terapis tidak dapat membantu partisipan jika mereka tidak mengikuti dengan benar rangkaian sesi intervensi ataupun jika partisipan salah mengerti instruksi yang diberikan.

Terakhir, beberapa partisipan melaporkan merasa ragu terhadap *The Heart Technique* yang dilakukan secara singkat namun efektif mengatasi fobia yang dialaminya. Para partisipan menyatakan rasa ragu terhadap *The Heart Technique* ini disebabkan para partisipan telah membawa beban emosi, fobia ini, selama bertahun-tahun, sedangkan hanya dalam waktu singkat *The Heart Technique* dapat membantu menghilangkannya. Perasaan ragu tersebut memberikan dampak partisipan berpikir kalau mereka masih merasakan emosi negatif terhadap fobia yang dialaminya. Sebagai contoh, salah satu partisipan menyatakan ia merasa nyaman dengan kondisi fobia dialaminya dan ingin menurunkan SUD-nya hanya hingga skala SUD 2 dari 10. Meskipun terapis memberikan penilaian bahwa ini dapat diturunkan hingga skala SUD 0 (netral), namun terapis menghargai keinginan partisipan. Peneliti melakukan konfirmasi pada para partisipan yang menyatakan ragu dan mendapatkan hasil dari

aspek simtom fisik maupun sensasi yang dirasakan di wilayah tubuh tertentu pada *pre-test*, sudah tidak dirasakan lagi setelah partisipan menjalani sesi intervensi. Peneliti mengambil kesimpulan bahwa para partisipan bukan ragu terhadap hasil intervensi, namun lebih termasuk bingung dan/ atau sulit percaya dan menerima kenyataan fobia yang dialami bertahun-tahun dapat disembuhkan dalam waktu 20 hingga 30 menit saja.

Oleh karena itu, *The Heart Technique* memiliki kelemahan yaitu pengguna teknik ini bila tidak percaya dapat mengatasi masalahnya, merasa ketika menggunakan teknik ini terlalu mudah untuk membawakan perubahan, atau teknik ini terlalu mudah untuk mengatasi permasalahan, akan melaporkan tidak sepenuhnya sembuh dari permasalahan yang dimilikinya. Begitu juga dengan pengguna teknik ini yang tidak memiliki niat dan keinginan untuk mengatasi permasalahan yang dialami, maka perubahan atau proses penyembuhan yang diterima tidak akan maksimal. Terakhir, kelemahan dari *The Heart Technique* adalah apabila pengguna teknik ini memiliki kendala kesehatan pada area tangan, seperti cedera, terkilir, keseleo, atau patah, maka mereka akan kesulitan mengikuti gerakan tertentu yang menjadi bagian dari teknik ini, sebagai contoh di Tahap (2) *OverEnergy Correction*, Reset Jalur Energi, & *Tracing*.

5. Kesimpulan

The Heart Technique adalah teknik terapi alternatif yang dapat diaplikasikan untuk mengatasi permasalahan emosional, termasuk kelima kategori fobia spesifik, dengan cepat dan mudah. Aplikasi *The Heart Technique* efektif digunakan secara *online* (daring) dan menyembuhkan fobia hanya dengan satu sesi intervensi. Kesimpulan ini dapat dilihat dari hasil uji statistik perbandingan rata-rata skor partisipan di *pre-test* dengan skor *post-test* setelah diberikan teknik intervensi *The Heart Technique* menurun secara drastis ($p = 0.001$; $p < 0.05$). Penurunan skor ini terjadi pada seluruh partisipan sesi individu dan sesi kelompok, baik pada alat ukur SUD dan juga *Severity Measure for Specific Phobia-Adult*. Mengikuti lima tahap *The Heart Technique*, emosi-emosi negatif yang dirasakan ketika berhadapan dengan objek atau situasi fobia yang tersimpan dalam pikiran bawah sadar individu diaktifkan dan kemudian dinetralisir. Dengan emosi telah menjadi netral, maka sensasi fisik yang dulunya juga muncul ketika berhadapan dengan objek atau situasi fobia, tidak akan dirasakan lagi. Inti cara kerja *The Heart Technique* adalah partisipan berperan aktif dalam proses terapi. Oleh karena itu, teknik terapi ini dapat diaplikasikan secara efektif pada proses daring baik pada sesi individu maupun kelompok. *The Heart Technique* dapat digunakan sebagai pilihan teknik terapi yang efektif untuk membantu permasalahan emosi, seperti fobia spesifik.

Saran pada penelitian selanjutnya adalah untuk mengambil data *follow up* agar mengetahui konsistensi hasil terapi. Penelitian selanjutnya juga dapat mengaplikasikan *The Heart Technique* ini pada gangguan emosional lainnya, baik terhadap emosi negatif maupun terhadap emosi positif. Keterbatasan penelitian ini adalah penelitian ini tidak melakukan *follow-up* terhadap kondisi diri para partisipan seusai sesi intervensi selesai, sehingga tidak ada data terkait konsistensi hasil terapi antar waktu dan uji efek terapeutik. Selain itu, penelitian ini dilakukan secara *online* atau daring sehingga peneliti tidak dapat mengontrol situasi ruangan setiap partisipan dan ini memberikan pengaruh pada hasil terapi.

6. Referensi

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
- Banyan, C. D & Kein, G. F. (2001). *Hypnosis and Hypnotherapy Basic to Advanced Techniques for the Professional*. Abbot Publishing House
- Botella, C., Pérez-Ara, M. Á., Bretón-López, J., Quero, S., García-Palacios, A., & Baños, R. M. (2016). In vivo versus augmented reality exposure in the treatment of small animal phobia: a randomized controlled trial. *PloS one*, 11(2), e0148237
- Craske, M., Wittchen, U., Bogels, S., Stein, M., Andrews, G., & Lebeu, R. (2013). *Severity Measure for Specific Phobia—Adult*. https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM5_Severity-Measure-For-Specific-Phobia-Adult.pdf
- Daharnis, D., Ildil, I., Amalianita, B., Zola, N., & Putri, Y. (2021). The Effectiveness of Ego-state Therapy in Reducing Trypanophobia. *Addictive Disorders & Their Treatment*, 20(1), 61-65. DOI:10.1097/ADT.0000000000000210
- Feinstein, D. (2008). Energy psychology: A review of the preliminary evidence. *Psychotherapy: Theory, Research, Practice, Training*, 45(2), 199–213. <https://doi.org/10.1037/0033-3204.45.2.199>
- Eden, D., and Feinstein D. (2006). *Energy medicine: basic 5 day taining* [DVD]. Innersource
- Erickson, M. H. (2005, May 1). *The Wisdom of Milton H. Erickson: The Complete Volume*. Crown House Publishing

- Gunawan, A. W. (2018a, April). *Teknologi di Balik The Heart Technique*®. <https://adiwgunawan.id/news/teknologi-di-balik-the-heart-technique>
- Gunawan, A. W. (2018b, Desember). *The Heart Technique* (2nd ed.). Penerbit Transformasi Insan Mulia.
- Gunawan, A. W. (2020, April 30). *Fobia: Apa dan Bagaimana Mengatasinya? #1* [Video]. Youtube. [youtube.com/watch?v=ROtKY_xxVDM&ab_channel=AdiWGunawan](https://www.youtube.com/watch?v=ROtKY_xxVDM&ab_channel=AdiWGunawan)
- Hui, K. K. S., Liu, J., Makris, N., Gollub, R. L., Chen, A. J. W., I. Moore, C. I., Kennedy, D. N., Rosen, B. R., & Kwong, K. K. (2000). Acupuncture modulates the limbic system and subcortical gray structures of the human brain: evidence from fMRI studies in normal subjects. *Human brain mapping*, 9(1), 13-25. [https://doi.org/10.1002/\(sici\)1097-0193\(2000\)9:1<13::aid-hbm2>3.0.co;2-f](https://doi.org/10.1002/(sici)1097-0193(2000)9:1<13::aid-hbm2>3.0.co;2-f)
- Keyes, A., Deale, A., Foster, C., & Veale, D. (2020). Time intensive cognitive behavioural therapy for a specific phobia of vomiting: A single case experimental design. *Journal of behavior therapy and experimental psychiatry*, 66, 101523. <https://doi.org/10.1016/j.jbtep.2019.101523>
- Knappe, S., Klotsche, J., Heyde, F., Hiob, S., Siegert, J., Hoyer, J., Strobel, A., LeBeau, E. T., Craske, M. G., Wittchen, H., & Beesdo-Baum, K. (2014). Test-retest reliability and sensitivity to change of the dimensional anxiety scales for DSM-5. *CNS spectrums*, 19(3), 256-267. <https://doi.org/10.1017/S1092852913000710>
- LeBeau, R. T., Glenn, D. E., Hanover, L. N., Beesdo-Baum, K., Wittchen, H. U., & Craske, M. G. (2012). A dimensional approach to measuring anxiety for DSM-5. *International Journal of Methods in Psychiatric Research*, 21(4), 258-272. <https://doi.org/10.1002/mpr.1369>
- Levine, P. A. & Frederick A. (1997). *Waking the tiger: Healing trauma*. North Atlantic Books
- Lieb, R., Miché, M., Gloster, A. T., Beesdo-Baum, K., Meyer, A. H., & Wittchen, H. U. (2016). Impact of specific phobia on the risk of onset of mental disorders: A 10-year prospective-longitudinal community study of adolescents and young adults. *Depression and Anxiety*, 33(7), 667-675. <https://doi.org/10.1002/da.22487>
- Miloff, A., Lindner, P., Dafgård, P., Deak, S., Garke, M., Hamilton, W., Heinsoo, J., Kristoffersson, G., Rafi, J., Sindemark, K., Sjölund, J., Zenger, M., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2019). Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial. *Behaviour research and therapy*, 118, 130-140. <https://doi.org/10.1016/j.brat.2019.04.004>
- Mor, S., Grimaldos, J., Tur, C., Miguel, C., Cuijpers, P., Botella, C., & Quero, S. (2021). Internet-and mobile-based interventions for the treatment of specific phobia: A systematic review and preliminary meta-analysis. *Internet Interventions*, 26, 100462. <https://doi.org/10.1016/j.invent.2021.100462>
- Motlova, L. B., Balon, R., Beresin, E. V., Brenner, A. M., Coverdale, J. H., Guerrero, A. P., ... & Roberts, L. W. (2017). Psychoeducation as an opportunity for patients, psychiatrists, and psychiatric educators: why do we ignore it?. *Academic Psychiatry*, 41, 447-451. DOI 10.1007/s40596-017-0728-y
- Nummenmaa, L., Glerean, E., Hari, R., & Hietanen, J. K. (2014). Bodily maps of emotions. *Proceedings of the National Academy of Sciences*, 111(2), 646-651. <https://doi.org/10.1073/pnas.1321664111>
- Nummenmaa, L., Hari, R., Hietanen, J. K., & Glerean, E. (2018). Maps of subjective feelings. *Proceedings of the National Academy of Sciences*, 201807390. doi:10.1073/pnas.1807390115
- Peñate, W., & Fumero, A. (2015). A meta-review of Internet computer-based psychological treatments for anxiety disorders. *Journal of Telemedicine and Telecare*, 22(1), 3-11. doi:10.1177/1357633x15586491
- Riddle-Walker, L., Veale, D., Chapman, C., Ogle, F., Rosko, D., Najmi, S., Walker, L. M., Maceachern, P., & Hicks, T. (2016). Cognitive behaviour therapy for specific phobia of vomiting (Emetophobia): A pilot randomized controlled trial. *Journal of anxiety disorders*, 43, 14-22. <https://doi.org/10.1016/j.janxdis.2016.07.005>
- Rondung, E., Ekdahl, J., Hildingsson, I., Rubertsson, C., & Sundin, Ö. (2018). Heterogeneity in childbirth related fear or anxiety. *Scandinavian journal of psychology*, 59(6), 634-643. <https://doi.org/10.1111/sjop.12481>
- Rumah Sejahtera. (2018, Oktober). The Heart Technique di SMA Swasta Budi Agung. https://rumahsejahtera.id/gallery/main/4_sma_swasta_budi_agung
- Rumah Sejahtera. (2019a, November). The Heart Technique di STIE EKA PRASETYA. https://rumahsejahtera.id/gallery/main/12_stie_eka_prasetya
- Rumah Sejahtera. (2019b, November). The Heart Technique di SEKOLAH TK - SD - SMP BODHISATTA. https://rumahsejahtera.id/gallery/main/10_sekolah_tk_sd_smp_bodhisatta
- Sancassiani, F., Romano, F., Balestrieri, M., Caraci, F., Di Sciascio, G., Drago, F., Hardoy, M. C., Moro, M.F., Roncone, R., Piras, M., Preti, A., Dell'Osso, L., Faravelli, C., & Carta, M.G. (2019). The prevalence of specific phobia by age in an Italian Nationwide Survey: how much does it affect the quality of life?. *Clinical practice and epidemiology in mental health: CP & EMH*, 15, 30-37. doi: 10.2174/1745017901915010030

- Solomon, S. G., & Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Current Biology*, 24(20), <https://doi.org/10.1016/j.cub.2014.09.001>.
- Tebbetts, C. (1987). *Miracles on demand: The Radical Short-Term Hypnotherapy 2nd Edition*. Westwood Publishing
- Tribun Sumsel.com. (2018). Mengenal The Heart Technique, Cara Mengatasi Anak Mudah Marah dan Takut Bicara di Depan Umum. <https://sumsel.tribunnews.com/2018/10/28/mengenal-the-heart-technique-cara-mengatasi-anak-mudah-marah-dan-takut-bicara-di-depan-umum>.
- Wardenaar, K. J., Lim, C. C., Al-Hamzawi, A. O., Alonso, J., Andrade, L. H., Benjet, C., Bunting, B., de Girolamo, G., Demyttenaere, K., Florescu, S. E., Gureje, O., Hisateru, T., Hu, C., Huang, Y., Karam, E., Kiejna, A., Lepine, J. P., Navarro-Mateu, F., Oakley Browne, M., Piazza, M., Posada-Villa, J., Ten Have, M. L., Torres, Y., Xavier, M., Zarkov, Z., Kessler, R. C., Scott, K. M., & de Jonge, P. (2017). The cross-national epidemiology of specific phobia in the World Mental Health Surveys. *Psychological medicine*, 47(10), 1744. doi:10.1017/S0033291717000174
- Weinberg, H. (2020). Online group psychotherapy: Challenges and possibilities during COVID-19—A practice review. *Group Dynamics: Theory, Research, and Practice*, 24(3), 201. <https://doi.org/10.1037/gdn0000140>

Lampiran

Tabel 4. Frekuensi emosi partisipan terhadap objek atau situasi fobia

No.	Emosi	<i>Pre-Test</i>		<i>Post-Test</i>	
		Frekuensi	Persentase	Frekuensi	Persentase
1	Cemas	98	17.10%	8	3.64%
2	Ngeri	75	13.09%	1	0.45%
3	Merasa tidak berdaya	57	9.95%	0	0%
4	Tertekan	53	9.25%	2	0.91%
5	Jijik	45	7.85%	9	4.09%
6	Frustrasi	38	6.63%	1	0.45%
7	Takut	37	6.46%	3	1.36%
8	Kesal	27	4.71%	0	0%
9	Jengkel	24	4.19%	0	0%
10	Sedih	23	4.01%	0	0%
11	Benci	22	3.84%	1	0.45%
12	Capek	19	3.32%	3	1.36%
13	Sebal	17	2.97%	0	0%
14	Muak	11	1.92%	0	0%
15	Marah	7	1.22%	0	0%
16	Panik	7	1.22%	0	0%
17	Malu	3	0.52%	0	0%
18	Kaget	2	0.35%	0	0%
19	Berdosa	1	0.17%	0	0%
20	Curiga	1	0.17%	0	0%
21	Gelisah	1	0.17%	0	0%
22	Ingin menghancurkannya	1	0.17%	0	0%
23	Kecewa	1	0.17%	0	0%
24	Menyesal	1	0.17%	0	0%
25	Tegang	1	0.17%	0	0%
26	Waspada	1	0.17%	0	0%

27	Lega/ plong/ tidak ada beban lagi	0	0%	43	19.55%
28	Senang/ bahagia/ hati enak	0	0%	12	5.45%
29	Damai/ nyaman/ relax/ tenang	0	0%	8	3.64%
30	Ragu	0	0%	5	2.27%
31	Percaya diri/ berani menghadapi objek atau situasi fobia	0	0%	5	2.27%
32	Semangat	0	0%	1	0.45%
33	Terharu	0	0%	1	0.45%
34	Tidak merasakan emosi negatif	0	0%	117	53.18%
Total		573	100%	220	100%

Tabel 5. Frekuensi wilayah tubuh merasakan sensasi emosi

No.	Wilayah Tubuh	<i>Pre-Test</i>		<i>Post-Test</i>	
		Frekuensi	Persentase	Frekuensi	Persentase
1	Dada	76	20.82%	7	4.61%
2	Kepala	60	16.44%	9	5.92%
3	Kaki	47	12.88%	3	1.97%
4	Tangan	46	12.60%	7	4.61%
5	Perut	45	12.33%	7	4.61%
6	Leher	38	10.41%	7	4.61%
7	Pundak	29	7.95%	7	4.61%
8	Punggung	13	3.56%	2	1.32%
9	Hati	2	0.55%	1	0.66%
10	Hidung	2	0.55%	0	0%
11	Telinga	2	0.55%	0	0%
12	Lengan	1	0.27%	1	0.66%
13	Lutut	1	0.27%	0	0%
14	Mulut	1	0.27%	0	0%
15	Tenggorokan	1	0.27%	0	0%
16	Tumit	1	0.27%	0	0%
17	Tidak merasakan apapun	1	0%	101	66.45%

Total	365	100%	152	100%
-------	-----	------	-----	------

Tabel 6. Frekuensi simtom fisik

No.	Simtom Fisik	<i>Pre-Test</i>		<i>Post-Test</i>	
		Frekuensi	Persentase	Frekuensi	Persentase
1	Jantung berdebar	93	24.09%	10	7.09%
2	Gemetar	60	15.54%	3	2.13%
3	Tangan/ kaki/ badan terasa dingin	51	13.21%	5	3.55%
4	Napas memburu	49	12.69%	0	0%
5	Mau menangis	40	10.36%	0	0%
6	Keringat dingin	34	8.81%	0	0%
7	Mual/ mau muntah	28	7.25%	2	1.42%
8	Merinding	8	2.07%	2	1.42%
9	Gatal	4	1.04%	0	0%
10	Lemas	4	1.04%	0	0%
11	Sesak Napas	4	1.04%	0	0%
12	Pusing	3	0.78%	2	1.42%
13	Perut tidak nyaman	2	0.52%	1	0.71%
14	Tangan/ kaki/ badan terasa kaku	2	0.52%	1	0.71%
15	Tegang	2	0.52%	0	0%
16	Konsentrasi berkurang	1	0.26%	0	0%
17	Mata bergetar	1	0.26%	0	0%
18	Badan ringan & enak	0	0%	2	1.42%
19	Terharu/ Menangis lega	0	0%	2	1.42%
20	Tidak ada simtom fisik	0	0%	111	78.72%
Total		386	100%	141	100%



VALIDITY AND RELIABILITY OF THE GENERAL HEALTH QUESTIONNAIRE-12 (GHQ-12) IN INDONESIA: SCALE FOR THE DETECTION OF PSYCHOLOGICAL DISTRESS IN INDIVIDUALS EXPERIENCING QUARTER-LIFE CRISIS

Audita Izza Balqis

Fakultas Psikologi, Universitas Muhammadiyah Malang, Malang, Indonesia

Email: auditaizza@umm.ac.id

<i>Received :</i>	<i>Revised</i>	<i>Accepted :</i>	<i>Published</i>
07 Januari 2025	29 Maret 2025	18 April 2025	30 April 2025

Abstract

This study aims to test the validity and reliability of the General Health Questionnaire-12 (GHQ)-12 for the Indonesian version, specifically for individuals who experience a quarter-life crisis. Individuals who experience a quarter-life crisis are identical in the form of psychological distress. The development of the GHQ-12 instrument was carried out to facilitate professionals in detecting levels of psychological distress as well as as an instrument in providing appropriate treatment. 300 participants (M= 23.70; SD = 2.64) were involved in this study. The reliability of Cronbach's Alpha ($\alpha=0.903$) showed GHQ-12 as an acceptable instrument. The GHQ-12 was found a one-dimensional model. CFA with goodness of-fit was found GHQ-12 was accepted, value of $X^2= 153.893$; RMSEA=0.079; GFI=0.921; TLI= 0.918; and CFI=0.933. This study's results indicate satisfactory validity and reliability for GHQ-12 as an instrument in detecting psychological distress for individuals experiencing a quarter-life crisis. The implications of this study can be used as a reference in providing appropriate interventions for overcoming psychological distress.

Keywords: General Health Questionnaire (GHQ-12), psychological distress, mental health, Quarter-life crisis

Abstrak

Penelitian ini bertujuan untuk melakukan uji validitas dan reliabilitas pada skala *General Health Questionnaire-12* (GHQ-12) untuk populasi di Indonesia, spesifik pada individu yang mengalami *quarter-life crisis*. Individu yang mengalami *quarter-life crisis* identik dengan adanya bentuk distres psikologis. Pengembangan instrumen GHQ-12 dilakukan untuk memudahkan tenaga profesional dalam melakukan deteksi tingkat distres psikologis serta sebagai instrumen dalam memberikan penanganan yang tepat. 300 partisipan (M=23.70; SD=2.64) terlibat dalam penelitian ini. Reliabilitas *Cronbach's Alpha* ($\alpha=0.903$) menunjukkan GHQ-12 sebagai instrumen yang dapat diterima. Hasil uji EFA menunjukkan GHQ-12 memiliki satu faktor. CFA dengan metode *goodness-of-fit* menunjukkan GHQ-12 diterima dengan nilai $X^2=153.893$; RMSEA=0.079; GFI=0.921; TLI= 0.918; dan CFI=0.933. Hasil penelitian ini menunjukkan validitas dan reliabilitas yang memuaskan untuk GHQ-12 sebagai instrumen dalam mendeteksi distres psikologis untuk individu yang mengalami *quarter-life crisis*. Implikasi penelitian ini dapat digunakan sebagai referensi dalam memberikan intervensi yang tepat dalam mengatasi distres psikologis.

Kata kunci: *General Health Questionnaire* (GHQ-12), distres psikologis, kesehatan mental, *quarter-life crisis*

1. Pendahuluan

Mental health problems in Indonesia today need special attention. The 2018 Basic Health Research (Kemenkes RI, 2018) found the prevalence of mental health cases in individuals older than 15 years, more than 19 million people experience mental-emotional disorders, and more than 12 million people experience depression. The results supported by Kaligis et al. (2021), the transitional age (N=393), 95.4% experienced anxiety, and 88% experienced depression. In Indonesia, it is also known that there is a gap, namely the still low proportion of people receiving adequate mental health services. In comparison, mental disorders in adulthood are known. Therefore a

need for the right strategy to overcome these problems. In addition to looking for the causative factors, the development of appropriate methods for early detection of individuals experiencing psychological distress is also carried out.

Various studies have found factors that contribute to psychological distress. The phenomenon of a quarter-life crisis can be categorized as one of the predictors of individuals experiencing psychological distress. Quarter-life crisis is a popular term related to problems regarding career, education, and romantic relationships in their 20s. Quarter-life Crisis are usually experienced by individuals facing the transition from adolescence to adulthood (Robinson et al., 2017). The characteristics of the quarter-life crisis occur in individuals during the transition period of 18-29 years (Arnett & Mitra, 2018). This transition period requires good adaptability. If individuals are unable or hampered in adapting, they tend to experience a quarter-life crisis and are at risk of experiencing psychological distress. Several studies have found that individuals who experience a quarter-life crisis are characterized by feelings of anxiety about the future, life goals, failure, or difficulty in making decisions (Robinson & Wright, 2013). If the quarter-life crisis occurs over a long period, it can cause individuals to be prone to experiencing psychological distress such as anxiety and depression (Arnett & Mitra, 2018; Arnett et al., 2014; Rosen et al., 2014) as well as problems in social relations.

Based on the problems above, it is necessary to detect individuals experiencing a quarter-life crisis accurately. If the detection is carried out correctly, it can assist in identifying, intervening, or taking preventive measures against the psychological distress of individuals experiencing a quarter-life crisis. Supported by Costello (2016) stated that early detection of mental health disorders is an effort to prevent them by considering the risk factors involved. Therefore, it is necessary to develop an appropriate measuring tool as a screening tool for individuals experiencing a quarter-life crisis. Previous studies on psychological distress measurement tools have widely utilized instruments such as the DASS-21 or the Kessler Psychological Distress Scale (K-10), both of which assess psychological distress. This study focuses on the GHQ-21 as a variation of measurement tools that relate to the characteristics of the quarter-life crisis, such as life uncertainty, major decisions, and identity changes, without a clinical orientation. Compared to the DASS-21 or K-10, which are more relevant for clinically significant disorders (Peixoto et al., 2021; Davies et al., 2015).

The General Health Questionnaire (GHQ) is a measuring tool developed by Goldberg to detect psychological distress and mental disorders. The original version of the GHQ questionnaire had 60-item, later developed into several versions (30-item, 28-item, 20-item, and 12-item). In comparison, the 12-point version is used as a screening instrument in various countries by WHO in research on mental health disorders. This is because the 12-point version of the GHQ is a useful measuring tool among similar instruments (Goldberg et al., 1997). General Health Questionnaire-12 (GHQ-12) is a measuring tool intended to screen the general population (nonclinical) for mental health problems (Goldberg et al., 1997). Various studies have assigned different names to dimensions depending on the type of model used. However, the common dimensions identified include general anxiety and social dysfunction. There are three scoring models for this measuring instrument: a bimodal scale (0-0-1-1) recommended in clinical research; a Likert scale with 4 points (1-2-3-4) most commonly used in various kinds of research; and the C-GHQ (0-0-1-1) or (0-1-1-1) models. GHQ-12 has been translated into 38 countries and cultures (Jackson, 2007).

Although the original version of the GHQ-12 measuring instrument developed by Goldberg (1988), based on research by Hystad & Johnsen (2020), has a unidimensional model, many other researchers doubt the internal structure of the measuring instrument. Therefore, more and more research is being carried out to develop and explore further the model and internal structure of GHQ-12. For example, a study conducted by Graetz (1991) and Abubakar and Fischer (2012) succeeded in forming the GHQ-12 model in a multidimensional version, namely anxiety/depression, social dysfunction, and loss of confidence. Besides of three-factor model, several other studies (Politi et al., 1994) found that GHQ-12 has a two-model version (general dysphoria and social dysfunction). Although many versions of the GHQ-12 model exist, the unidimensional version is still being researched. For example, Anjara et al. (2020) and Hystad and Johnsen (2020) have confirmed that the unidimensional version of the GHQ-12 study also showed satisfactory validity and reliability for the unidimensional model. Because of this diversity and discussion, this study explores the model factors for GHQ-12 in Indonesia.

In Indonesia, research on the validity and reliability of GHQ-12 has been carried out, including a screening tool for primary care patients who are at high risk for mental disorders (Anjara et al., 2020); screening for patients with somatoform disorders (Salma & Hidayat, 2013); and screening for adjustment disorders (Primasari & Hidayat, 2016). These studies found that GHQ-12 has good validity and reliability. For example, the study of Anjara et al. (2020) GHQ-12 has reliability ($\alpha=0.863$) with a Likert scale model scoring. In addition, the calculation of CFAs shows results under the theoretical model. Meanwhile, this study aimed to evaluate the psychometric GHQ-12 for the nonpsychiatric population. Specifically for the adult population who experiencing a quarter-life crisis in Indonesia. This research was conducted as a psychometric evaluation of GHQ-12, which is appropriate to be used as an appropriate measuring tool for the population in Indonesia, especially for individuals experiencing a quarter-life crisis.

2. Methods

Respondents

The criteria for subjects in this study were individuals aged 18-29 Indonesian people who were experiencing a quarter-life crisis. The recruitment process for subjects was conducted entirely online using Google Forms through various social media platforms. The sampling method in this study employed a non-probability sampling technique. Participants underwent a screening process to identify individuals experiencing a quarter-life crisis. The instrument used for screening was the Quarter-Life Crisis Screening Instrument (Petrov et al., 2022). From the 357 participants screened, 300 met the criteria for experiencing a quarter-life crisis and were further analyzed by completing the GHQ-12 scale, which was then evaluated for reliability and validity.

Instrument

General Health Questionnaire (GHQ-12). The General Health Questionnaire (GHQ-12) developed by David Goldberg (1970) is used to detect general psychological distress that can be applied to both community and nonclinical populations. This measuring instrument consists of 12 items that measure psychological distress over the past few weeks. This study uses a unidimensional model with a Likert scale (0-1-2-3-4).

Quarter-life crisis screening Instrument. A scale of quarter-life crisis screening adapted and modified from Petrov et al. (2022) was used to screen individuals who had a quarter-life crisis. This measuring instrument has a 12-point model scale (yes or no). 'Yes'=0, and 'no'=1. Screening is carried out using a cut-off-score of 6 points for the 'Yes' answer will be analyzed further.

Design

Research was conducted to test the validity and reliability of the General Health Questionnaire-12 (GHQ-12) scale. Before conducting validity and reliability tests, this study translated the scale from English to Indonesian. The original scale is translated by translators who are experts in English and Indonesian. The results of the measuring instrument, translated into Indonesian, are translated back into English (back translation), which is then compared with the original version of the measuring instrument. Experts in the relevant field also provide suggestions on the GHQ-12 and the screening scale for the quarter-life crisis.

Analysis Technique

The psychometric test aims to determine the validity and reliability of the GHQ-12 measuring instrument in the Indonesian population, specifically for individuals with a quarter-life crisis, carried out through several stages of analysis. The research data were collected by analyzing demographic data and testing assumptions using Skewness and Kurtosis. Researchers calculated scale reliability using Cronbach's Alpha estimation. Pearson's Product Moment was to analyze the item correlation. The factor analysis test using Exploratory Factor Analysis (EFA) was carried out to build a structural model. Confirmatory Factor Analysis (CFA) test to determine the extent to which the correlation between items and components of the measuring instrument followed the primary construct. CFA calculations were carried out to obtain evidence of the internal structure of GHQ-12 in Indonesia. The accuracy of the model is carried out using goodness-fit with parameters in the form of chi-square (χ^2), Root Mean Square Error of Approximation (RMSEA), Tucker-Lewis index (TLI), goodness-of-fit index (GFI), and Comparative Fit Index (CFI).

3. Result

A total of 300 participants (M= 23.70, SD= 2.64) were aged 18 and 30 years old (Table 1). Male participants comprised 83 people (27.7%), and women consisted of 217 people (72.3%). The occupational status of the participants was employed (40.7%), students (34.7%), employed at the same time as students (10%), and unemployed (14.7%). Meanwhile, 88.3% of marital statuses stated that they were not married, and 11.7% were married.

Table 2. shows the average score of each GHQ-12 Item. The highest average of all items was M=1.87 (SD=0.93), while the lowest average was M=1.22 (SD=1.10). The skewness of all items shows a range of -0.40 to 0.81, while Kurtosis shows -0.8 to 1.08. The standard error of Kurtosis is 0.281. Based on these results, it is determined to meet the normality assumption.

Table 1. Participants Demographic

Characteristics	N	%
Age M=23.70 (2.64)		
18-20	31	10
21-30	79	90
Education		
High School	38	12.7
Diploma	28	9.3
University	232	77.3
Others	2	0.7
Employment status		
Student	104	34.7
Employed	122	40.7
Student + Employed	30	10
Unemployed	44	14.7
Marital status		
Unmarried	265	88.3
Married	35	11.7

Table 2. The average score and item-total correlation of GHQ-12

Item	Mean	SD	Min.	Maks.	Skewness	Kurtosis	Item-total (r)
Tidak mampu berkonsentrasi	1.29	0.98	0.00	3.00	0.22	-0.8	.626
Sulit tidur karena cemas	1.52	1.11	0.00	3.00	0.02	-1.2	.548
Merasa terbebani dalam banyak hal	1.87	0.93	0.00	3.00	-0.40	-0.52	.558
Sulit dalam membuat keputusan	1.30	0.97	0.00	2.00	0.13	-1.08	.609
Merasa terus menerus tertekan	1.48	1.03	0.00	3.00	0.81	-1.00	.631
Ketidakmampuan mengatasi kesulitan	1.31	1.06	0.00	4.00	0.33	-0.73	.720
Tidak menikmati aktivitas sehari-hari	1.37	1.11	0.00	3.00	0.15	-0.28	.602
Tidak mampu menghadapi masalah	1.22	1.10	0.00	3.00	0.35	-0.12	.664
Merasa tidak bahagia dan tertekan	1.38	1.11	0.00	3.00	0.17	-1.22	.675
Kehilangan kepercayaan diri	1.44	1.03	0.00	3.00	0.71	-1.07	.697
Merasa menjadi orang yang tidak berguna	1.56	1.16	0.00	4.00	-0.23	-1.25	.571
Merasa tidak bahagia dengan segala hal	1.34	1.13	0.00	2.00	0.14	-1.4	.606

Internal Consistency

Cronbach's Alpha for internal consistency GHQ-12, ($\alpha=0.903$) with the coefficient item-total correlation between 0.548 to 0.720 (Table 2.). Internal consistency is declared adequate if the reliability value is above 0.7. Cronbach's Alpha if an item deleted shows a smaller number on each indicator, it can be stated that all GHQ-12 items have an accepted reliability value. Internal consistency is also done by distinguishing male participants ($\alpha=0.90$) and female participants (0.903).

Factor Analyses

Exploratory Factor Analysis (EFA) with varimax rotation was used to determine the construct validity using the maximum likelihood method. Assumption tests were carried out using Kaiser-Meyer-Olkin (KMO) and Barlett's sphericity tests. The measure of sampling adequacy (MSA) showed a score of 0.925 and Barlett's Test of Sphericity ($p < 0.001$). These results indicate that the assumptions are met for factor analysis because the KMO score is above 0.60 and Barlett's test ($p < 0.001$) (Steed & Coakes, 2003). The factor loadings for each item on the GHQ-12 scale indicate a value greater than 0.5. While the variance produces one factor that can explain 48.61% of the total variance. The Pearson coefficients for each Item on the GHQ-12 scale are presented from coefficient correlation from 0.287 to 0.627 (Table 3).

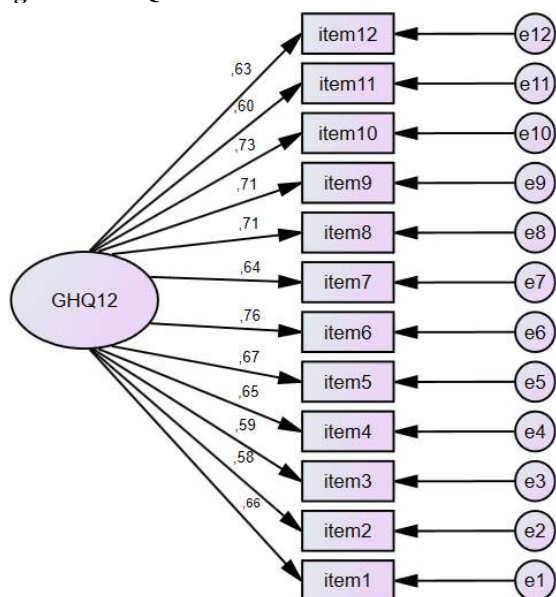
Table 4. Pearson's Correlation

	GHQ 2	GHQ 3	GHQ 4	GHQ 5	GHQ 6	GHQ 7	GHQ 8	GHQ 9	GHQ 10	GHQ 11	GHQ 12
GHQ1	.412	.382	.543	.420	.544	.423	.419	.452	.438	.409	.379
GHQ2	1.00	.400	.358	.366	.481	.331	.390	.412	.403	.321	.390
GHQ3		1.00	.374	.547	.416	.376	.378	.373	.487	.287	.320
GHQ4			1.00	.494	.519	.380	.442	.397	.489	.373	.343
GHQ5				1.00	.529	.360	.433	.468	.522	.398	.343
GHQ6					1.00	.453	.627	.494	.537	.467	.407
GHQ7						1.00	.455	.522	.480	.381	.472
GHQ8							1.00	.568	.493	.413	.453
GHQ9								1.00	.510	.385	.560
GHQ10									1.00	.485	.474
GHQ11										1.00	.490
GHQ12											1.00

Table 5. Goodness-of-Fit Index

Index Goodness of Fit	Index Score
Chi-square (X ²)	153.893
Probability Scaled Chi-Square (p-value)	.000
Root Mean Square Error of Approximation (RMSEA)	.079
Tucker-Lewis Index (TLI)	.918
The goodness of Fit Index (GFI)	.921
Comparative Fit Index (CFI)	.933

Confirmatory Factor Analysis (CFA) with the maximum likelihood of testing the internal structure of each GHQ-12 Item. Index measurement model Goodness of Fit (Table 4). The parameters used are chi-square (χ^2), Root Mean Square Error of Approximation (RMSEA), Tucker-Lewis index (TLI), goodness of fit index (GFI), and Comparative Fit Index (CFI). Figure 1. presents the model of the unidimensional factor model of the GHQ-12. The results of construct validity analysis on GHQ-12 showed chi-square (χ^2)= 153.893, df=54, p-value (<0.001), RMSEA=0.079, TLI= 0.918, GFI=0.921, and CFI=0.933. The items on the GHQ-12 scale show a factor loading ranging from 0.58 to 0.73. These results meet the recommended assumptions (Hair et al., 2010). GFI, CFI, and TLI values are above 0.90, while RMSEA is less than 0.08.

Figure 1. GHQ-12 Factor Model.

4. Discussion

The reliability of the GHQ-12 measuring instrument using the Likert scoring method shows Cronbach's Alpha was 0.903. This value indicates that the GHQ-12 has an excellent internal consistency for the population in Indonesia. Another good reliability value for GHQ-12 research in Indonesia is research by Primasari & Hidayat (2016) with a population of individuals with adjustment disorders using the same scoring method, having a Cronbach's Alpha of 0.863 (N=243). Research by Rahmandani et al. (2021) GHQ-12 (N=552) has Cronbach's Alpha ($\alpha=0.836$), and Anjara et al. (2020) with ($\alpha=0.76$). In addition, research on GHQ-12 from various other countries, such as Qin et al. (2018), found the GHQ-12 reliability value of 0.9. Padrón et al. (2012), GHQ-12 research with a population of paramedics in Brazil has a Cronbach's Alpha value of 0.89.

The test was conducted for Exploratory Factor Analysis first because EFA analysis was rarely found in subjects experiencing a quarter-life crisis. Exploratory Factor Analysis (EFA) resulted in one factor with a total variance of 48.61%. These results support the original version of the Goldberg and Williams (1988) scale. At the same time, different from other researchers from various countries found that GHQ-12 has three factors. GHQ-12 has three other factors, such as research by Gao et al. (2012); and Suzuki et al. (2011), consisting of anxiety or depression, distress, and loss of positive emotions. Research GHQ-12, 3 factors with other alternative models, such as anxiety, social dysfunction, and loss of self-confidence (Padrón et al., 2012; Abubakar & Fischer, 2012; Gelaye et al., 2015). The loading factor for this study's EFA test on GHQ-12 is around 0.287-0.627.

The model fit test from Confirmatory Factor Analysis (CFA) was conducted to evaluate the model using the goodness-of-fit method. With the chi-square of 153.893 with $p < 0.001$, it is interpreted that the measurement model is not following the empirical data. The RMSEA value in this study shows a value of 0.079, indicating a value of less than 0.08, so the model meets the model suitability (Hair et al., 2010). The RMSEA value in this study is similar to the research of Abubakar and Fisher (2012) with adult participants who produced an RMSEA of 0.80. However, the RMSEA value in this study is known to be relatively higher when compared to other studies such as research (Molina et al., 2014) with worker participants in Spain with an RMSEA of 0.067. In addition, the GFI and CFI values show values above 0.90, indicating the model's suitability to be interpreted (Hair et al., 2010). In addition, the TLI value is above 0.8 with a value of 0.918, it is stated that the accuracy of the model is good. The CFI value of 0.93 in this study is higher than the research of Abubakar and Fisher (2012), with a CFI of 0.89. The one-factor CFA test showed that the loading ranged from 0.58 to 0.73 (Figure 1). It can be concluded that the factor loading for GHQ-12 in the Indonesian version is quite good.

Although several studies using GHQ-12 partially support Graetz's (1991) theory that uses three dimensions some researchers also suggest conducting research using the unidimensional version of the GHQ-12 to obtain varying scores (Padrón et al., 2012). Different factors from various studies indicate that although the original scale of the GHQ-12 is unidimensional, the internal structure and dimensions of this measuring instrument are still a matter of debate. Different results from this comparative research show that unidimensional models that are influenced by the effect of statement sentences on items produce measurement accuracy compared to two and three-dimensions (Hystad & Johnsen, 2020). On the other hand, research on the comparison of internal structure and dimensions by Hystad and Johnsen (2020) concluded that GHQ-12 is flexible and multidimensional because of the influence on the use of sentences in each Item. In addition to using GHQ-12 with multidimensionality, it is not recommended to interpret subscale scores because the GHQ-12 measuring tool reflects more on general mental health than different constructs (Gnambs, 2018).

The proper screening instrument will be able to help professionals manage efficiency in the field of prevention and treatment of mental health. Especially in Indonesia, it is necessary to prioritize the development of screening measurement tools regarding mental health issues. This is because, based on previous research, it is known that Indonesia's health screening system is still little included in the routine procedures of public health services (Anjara et al., 2020). Especially for the target population, individuals in the transitional age from adolescence to early adulthood (Kaligis et al., 2021) are prone to experiencing psychological distress as well as a quarter-life crisis. Therefore, the implementation of screening for mental health problems needs special attention.

This study supported the original theory that the GHQ-12 was designed as a unidimensional measuring instrument. EFA and CFA tests also support existing constructs. However, this research is not without its weaknesses. The GHQ-12 validity and reliability test for the population in Indonesia does not compare the test with the bimodal scale form. In addition, this study also did not examine the differences between unidimensional and multidimensional. Another weakness is that the results of this study indicate the influence of demographic data on mental health. It is known that the demographic data for this study are women. The facts reveal that women are more prone to experiencing psychological distress than men (Malik et al., 2014). The implications of using this measuring instrument can be used to measure psychological distress in general.

Professionals can use this measurement tool for assessments to identify aspects requiring intervention, ensuring practical application for both practitioners and researchers. It can assist professionals in screening and developing support programs such as counseling, interventions, or institutional support services. The GHQ-12, proven valid and reliable for measuring psychological distress in individuals experiencing a quarter-life crisis,

helps differentiate between typical developmental transitions and significant psychological distress requiring clinical referral. For researchers, this tool aids in further exploration of the quarter-life crisis concept. Future studies can expand research on GHQ-12 with a more diverse population.

5. Conclusion

GHQ-12 shows adequate reliability and validity values for use by the population in Indonesia, especially for individuals experiencing a quarter-life crisis. This study's internal structure for the GHQ-12 supports the original theory, which is included in a unidimensional measuring instrument. However, GHQ-12 is flexible and can be used with multidimensional shapes. This study confirms that GHQ-12 can be used as a measurement tool in studies to detect general or nonclinical psychological distress.

6. References

- Abubakar, A., and Fischer, R. (2012). The factor structure of the 12-item general health questionnaire in a literate Kenyan population. *Stress Health* 28, 248–254. <https://doi.org/10.1002/smi.1420>
- Anjara, S.G., Bonetto, C., Van Bortel, T., & Brayne, C. Using the GHQ-12 to screen for mental health problems among primary care patients: Psychometrics and practical considerations. (2020). *International Journal of Mental Health Systems*, 14(62). <https://doi.org/10.1186/s13033-020-00397-0>
- Arnett, J. J., & Mitra, D. (2018). Are the features of emerging adulthood developmentally distinctive? A Comparison of ages 18–60 in the United States. *Emerging Adulthood*. <https://doi.org/10.1177/2167696818810073>
- Arnett, J. J., Žukauskienė, R., & Sugimura, K. (2014). The new life stage of emerging adulthood at ages 18–29 years: implications for mental health. *The Lancet. Psychiatry*, 1(7), 569–576. [https://doi.org/10.1016/S2215-0366\(14\)00080-7](https://doi.org/10.1016/S2215-0366(14)00080-7)
- Costello, E. J. (2016). Early detection and prevention of mental health problems: Developmental epidemiology and systems of support. *Journal of Clinical Child & Adolescent Psychology*, 24(6), 710–717. <https://doi.org/10.1080/15374416.2016.1236728>
- Davies, G., Caputi, P., Skarvelis, M., & Ronan, N. (2015). The depression anxiety and stress scales: Reference data from a large psychiatric outpatient sample. *Australian Journal of Psychology*, 67(2), 97–104. <https://doi.org/10.1111/ajpy.12069>
- Gnambs, T., & Staufienbiel, T. (2018). The structure of the general health questionnaire (GHQ-12): Two meta-analytic factor analyses. *Health Psychology Review*, 1–38. <https://doi.org/10.1080/17437199.2018.1426484>
- Gao, W., Stark, D., Bennett, M. I., Siegert, R. J., Murray, S., and Higginson, I. J. (2012). Using the 12-item general health questionnaire to screen psychological distress from survivorship to end-of-life care: dimensionality and item quality. *Psycho-Oncology*, 21 (9), 954–961. <https://doi.org/10.1002/pon.1989>
- Gelaye, B., Tadesse, M. G., Lohsoonthorn, V., Lertmeharit, S., Pensuksan, W. C., Sanchez, S. E., Lemma, S., Berhane, Y., Vélez, J. C., Barbosa, C., Anderade, A., & Williams, M. A. (2015). Psychometric properties and factor structure of the General Health Questionnaire as a screening tool for anxiety and depressive symptoms in a multi-national study of young adults. *Journal of Affective Disorders*, 187, 197–202. <https://doi.org/10.1016/j.jad.2015.08.045>
- Graetz, B. (1991). Multidimensional properties of the general health questionnaire. *Social, Psychiatry, and Epidemiology*, 26, 132–138. <https://doi.org/10.1007/BF00782952>
- Goldberg, D.P., & Blackwell, B. (1970). Psychiatric illness in general practice: a detailed study using a new method of case identification. *British Medical Journal*, 2(5707), 439–443. <https://doi.org/10.1136/bmj.2.5707.439>
- Goldberg D. & Williams P. (1988). *A Users' Guide To The General Health Questionnaire*. London: GL Assessment
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T.B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27(1), 191–7. <https://doi.org/10.1017/S0033291796004242>
- Hair, J.F., Black, W.C., Babin, B.J. (2010). *Multivariate Data Analysis: A Global Perspective, 7th ed.*, Pearson Education International, Prentice Hall.
- Hystad, S.W. & Johnsen, B.H. (2020). The dimensionality of the 12-item general health questionnaire (GHQ-12): Comparisons of factor structures and invariance across samples and time. *Frontiers Psychology*, 11, 1300. <https://doi.org/10.3389/fpsyg.2020.01300>
- Jackson, C. (2007). The general health questionnaire. *Occupation Medicine*, 57(1), 79. <https://doi.org/10.1093/occmed/kql169>

- Kaligis, F., Ismail, R.I., Wiguna, T., Prasetyo, S., Indriatmi, W., Gunardi, H., Pandia, V., & Magdalena, C.C. (2021). Mental health problems and needs among transitional-age youth in Indonesia. *International Journal of Environmental Research and Public Health*, 18, 4046. <https://doi.org/10.3390/ijerph18084046>
- Kementerian Kesehatan Republik Indonesia (Kemenkes RI). (2018). *Riset kesehatan dasar: Hasil utama riskesdas 2018*. Layanan Data Kementerian Kesehatan RI. <https://repository.badankebijakan.kemkes.go.id/id/eprint/3514/1/Laporan%20Riskesdas%202018%20Nasional.pdf>
- Malik, S & Farooqi, Y.N. (2014). General and sexual harassment as predictors of posttraumatic stress symptoms among female health professionals. *World Journal of Medical Sciences*, 10(1): 43-49. <https://doi.org/10.5829/idosi.wjms.2014.10.1.81128>
- Molina, J.G., Rodrigo, M.F., Losilla, J.M., & Vives, J. (2014). Wording effects and the factor structure of the 12-item general health questionnaire (GHQ-12). *Psychological Assessment*, 26(3), 1031–1037 <http://dx.doi.org/10.1037/a0036472.supp>
- Padrón, A., Galán, I., Durbán, M., Gandarillas, A., and Rodríguez-Artalejo, F. (2012). Confirmatory factor analysis of the general health questionnaire (GHQ- 12) in Spanish adolescents. *An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*, 21(7), 1291–1298. <https://doi.org/10.1007/s11136-011-0038-x>
- Peixoto, E. M., Zanini, D. S., & de Andrade, J. M. (2021). Cross-cultural adaptation and psychometric properties of the Kessler Distress Scale (K10): An application of the rating scale model. *Psicologia, reflexao e critica : revista semestral do Departamento de Psicologia da UFRGS*, 34(1), 21. <https://doi.org/10.1186/s41155-021-00186-9>
- Petrov, N., Robinson, O.C., & Arnett, J.J. (2022). The developmental crisis questionnaire (DCQ-12): Psychometric development and validation. *Journal of Adult Development*. <https://doi.org/10.1007/s10804-022-09403-w>
- Politi, P. L., Piccinelli, M., and Wilkinson, G. (1994). Reliability, validity, and factor structure of the 12-item general health questionnaire among young males in Italy. *Acta Psychiatrica Scandinavica*, 432–437. <https://doi.org/j.1600-0447.1994.tb01620.x>
- Primasari, I., & Hidayat, R. (2016). General Health Questionnaire-12 (GHQ-12) sebagai instrumen skrining gangguan penyesuaian. *Jurnal Psikologi*, 43(2), 121-134. <https://doi.org/10.22146/jpsi.9155>
- Qin, M., Vlachantoni, A., Evandrou, M., & Falkingham, J. (2018). General health questionnaire-12 reliability, factor structure, and external validity among older adults in India. *Indian Journal of Psychiatry*, 60(1), 56-59. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_112_17
- Rahmandani, A., La kahija, Y F. & Salma, S. (2021). Will self-compassion relieve distress?: A correlational study among Indonesian undergraduate students. *Journal of Medical Science*, 64 (2), 192-199. <https://doi.org/10.33160/yam.2021.05.013>
- Robinson, O. C., Demetre, J. D., & Litman, J. A. (2017). Adult life stage and crisis as predictors of curiosity and authenticity: Testing inferences from Erikson's lifespan theory. *International Journal of Behavioral Development*, 41(3), 426–431. <https://doi.org/10.1177/0165025416645201>
- Robinson, O. C., Wright, G. R. T., & Smith, J. A. (2013). The holistic phase model of early adult crisis. *Journal of Adult Development*, 20(1), 27–37. <https://doi.org/10.1007/s10804-013-9153-y>
- Rosen, N. O., Ivanova, E., & Knäuper, B. (2014). Differentiating intolerance of uncertainty from three related but distinct constructs. *Anxiety, Stress & Coping*, 27(1), 55-73. <https://doi.org/10.1080/10615806.2013.815743>
- Salma, & Hidayat, R. (2013). *Validasi klinik general health questionnaire-12 sebagai instrumen skrining gangguan somatoform di Puskesmas*. [Unpublished thesis]. Universitas Gadjah Mada, Yogyakarta, Indonesia.
- Steed, L., & Coakes, S.J. (2003). *SPSS analysis without anguish. Version 11 for Windows*. Milton, Queensland: John Wiley and Sons Australia
- Suzuki, H., Kaneita, Y., Osaki, Y., Minowa, M., Kanda, H., Suzuki, K., et al. (2011). Clarification of the factor structure of the 12-item general health questionnaire among Japanese adolescents and associated sleep status. *Psychiatry Research*, 188, 138–146. <https://doi.org/10.1016/j.psychres.2010.10.025>



MODIFIKASI SKALA *SELF-EFFICACY* : ANALISIS EFA DAN CFA

Maskuri Muhammad Cahyadi

Fakultas Ilmu Pendidikan dan Psikologi, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

Email: maskurymuhammadcahyadi@gmail.com

<i>Received :</i>	<i>Revised</i>	<i>Accepted :</i>	<i>Published</i>
07 Januari 2025	09 Maret 2025	21 April 2025	30 April 2025

Abstract

This research aims to determine the dimensions of the self-efficacy scale based on Bandura's theory using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) analysis for 151 junior high school (SMP) students. Based on the EFA analysis, it produces three dimensions, namely generality, level and strength. The researchers created 37 scales. However, after carrying out the EFA test, 25 items remained. The researchers dropped the items due to factor loading values being <0.4. The factors formed from EFA analysis were tested using CFA and three factors were formed. From the CFA analysis, the Goodness of Fit results were found to be less than fit, so it was necessary to modify the indicators eight times to improve the fit of the model so that it could be said to be fit based on the chi-square, CFI, TLI, SRMR and RMSEA criteria. The reliability results in the medium category indicate the need for other reliability tests. The results of EFA and CFA show that self-efficacy measuring instruments can help measure self-confidence in early adolescents or junior high school students.

Keywords: *Self-efficacy, EFA, CFA*

Abstrak

Penelitian ini memiliki tujuan untuk mengetahui dimensi-dimensi pada skala *self-efficacy* berdasarkan teori Bandura dengan menggunakan analisis *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA) pada siswa Sekolah Menengah Pertama (SMP) berjumlah 151. Berdasarkan analisis EFA menghasilkan tiga dimensi yaitu *generality*, *level* dan *strength*. Skala yang dibuat peneliti berjumlah 37. Namun, setelah dilakukan uji EFA tersisa 25 item. Item yang gugur disebabkan nilai *factor loading* <0,4. Faktor yang terbentuk dari analisis EFA diuji menggunakan CFA dan membentuk tiga faktor. Dari analisis CFA didapatkan hasil *Goodness Of Fit* yang kurang fit sehingga diperlukan *modification indices* sebanyak delapan kali untuk memperbaiki fit model sehingga didapatkan dikatakan fit berdasarkan kriteria *chi-square*, CFI, TLI, SRMR, dan RMSEA. Hasil reliabilitas berkategori sedang menunjukkan diperlukannya uji reliabilitas yang lain. Dari hasil EFA dan CFA menunjukkan bahwa alat ukur *self-efficacy* dapat membantu mengukur keyakinan diri pada remaja awal atau siswa Sekolah Menengah Pertama.

Kata Kunci: *Self-efficacy, EFA, CFA*

1. Pendahuluan

Siswa Sekolah Menengah Pertama memasuki masa remaja awal (11-13 tahun) yang merupakan periode transisi dari anak-anak menuju dewasa. Pada masa ini tentu akan mendapatkan perubahan dari aspek emosi, kognitif, dan fisik. Masa tersebut akan banyak menimbulkan gejala masalah atas perubahan yang terjadi. Perubahan dari Sekolah Dasar ke Sekolah Menengah Pertama juga akan menimbulkan perubahan dari aspek pola pembelajaran, metode, dan guru. Seharusnya setiap siswa harus dapat beradaptasi dengan segala perubahan sehingga tidak menimbulkan masalah yang terjadi baik di rumah ataupun di sekolah.

Sekolah merupakan tempat yang menyenangkan bagi anak untuk bisa mengembangkan potensi yang dimiliki sehingga seluruh tahap perkembangannya dapat dilalui dengan baik tanpa ada masalah yang mengganggu (Santrock, 2020). Sekolah harus dapat menyediakan iklim dan suasana belajar yang bebas dari tekanan sehingga siswa tidak merasa terbebani (Ferdianto & Muhid, 2020). Namun, terdapat indikator perilaku

siswa yang mengalami tekanan dan masalah di sekolah. Dari 191 responden terdapat 65,7 % siswa yang tidak mengikuti pelajaran di dalam kelas dan 73,9% tidak masuk sekolah tanpa memberikan keterangan (Charlina, 2013). Berdasarkan penelitian dari Nursalmin (dalam Oktaviani, 2018) sebanyak 75% siswa menolak untuk sekolah dengan berbagai alasan seperti takut dengan mata pelajaran dan gurunya, takut dengan lingkungannya, malas. Hal tersebut tentu menjadi indikator bahwa siswa SMP masih banyak mengalami berbagai masalah di sekolah sampai tidak ingin pergi kesekolah.

Adanya hal tersebut tentunya dari pihak keluarga, sekolah, dan siswa yang bersangkutan perlu adanya pemecahan masalah. Untuk memecahkan masalah tersebut tentunya harus ada kerja sama antara 3 komponen tersebut. Namun, siswa yang bersangkutan memegang peran yang paling penting dalam memecahkan masalah tersebut. Kemampuan dalam memecahkan masalah sangat erat kaitannya dengan keyakinan siswa tersebut. Keyakinan dalam memecahkan masalah akan memengaruhi hasil belajar. Keyakinan diri seringkali disebut dengan istilah *self-efficacy*. *Self-efficacy* merupakan suatu keyakinan yang dimiliki individu untuk mengatur dan melakukan tindakan yang diperlukan untuk menyelesaikan suatu masalah atau tugas tertentu (Bandura, 1997). Selain itu, menurut Nuzulia (2010) *self-efficacy* merupakan bagian hasil dari proses berpikir seseorang yang dilihat dari suatu keyakinan dan keputusan-keputusan tentang individu dalam memperkirakan kemampuan yang digunakan untuk suatu tindakan tertentu sehingga tercapainya hasil yang diharapkan. Ketika siswa memiliki *self-efficacy* yang baik maka akan berguna untuk membantu dalam menentukan usaha dan pilihan dalam mengatasi masalah, Ketekunan serta kegigihan yang dilakukan dalam menghadapi masalah serta memiliki ketenangan dalam menghadapi tugas atau masalah yang dirasa sulit atau berat (Sunaryo, 2017).

Persepsi individu terkait *self-efficacy* dapat dijadikan acuan untuk melihat seberapa besar usaha dan kegigihan seseorang untuk tetap bertahan dan gigih dalam kondisi sulit untuk menyelesaikan hambatan yang dialaminya. Siswa yang mempunyai *self-efficacy* tinggi cenderung akan berusaha menyelesaikan tugas dan masalahnya walaupun seberapa sulitnya. Sedangkan siswa yang mempunyai *self-efficacy* rendah akan lebih menghindari tugas dan masalah yang dianggap sulit sehingga akan memunculkan perasaan malas untuk sekolah (Sunaryo, 2017). Adanya hal tersebut mengartikan bahwa seorang siswa harus mempunyai *self-efficacy* tinggi sebagai modal untuk bisa menghadapi berbagai masalah di sekolah.

Penelitian ini dilakukan untuk menguji skala *self-efficacy* yang dirancang oleh Bandura (1997) dengan menggunakan analisis *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA). Skala ini dibuat dari hasil modifikasi 3 dimensi yang mencakup 37 item yang masing-masing nilai pada skala Likert 4 poin yang berkisar dari 1 (sangat tidak setuju) hingga 4 (sangat setuju). Pada hal ini memuat 3 dimensi yaitu level, *generality* dan *strength* (Bandura, 1997).

Dimensi pertama yaitu level berkaitan dengan tingkat kesulitan tugas yang mampu dihadapi oleh siswa dengan *self-efficacy* tinggi, meliputi memiliki keyakinan dapat menyelesaikan tugas sulit, dan memiliki perencanaan dalam mengatur diri terhadap pengerjaan tugas. Dimensi kedua yaitu *generality* yang berkaitan dengan luas bidang tingkah laku yang mana siswa dengan efikasi diri tinggi merasa yakin akan kemampuannya yang meliputi memiliki keyakinan penguasaan terhadap tugas yang diberikan, memiliki keyakinan penguasaan materi pembelajaran, dan mampu memecahkan masalah dan menemukan solusi. Dimensi ketiga yaitu *strength* lebih menekankan pada tingkat keyakinan yang dimiliki oleh siswa dengan *self-efficacy* tinggi yaitu meliputi berusaha keras untuk menyelesaikan tugas akademik, memiliki konsistensi dalam mencapai tujuan, dan menjadikan pengalaman sebagai dasar untuk meningkatkan keyakinan (Bandura, 1997).

Belum adanya penelitian hasil modifikasi alat ukur dari teori yang dikembangkan oleh Bandura (1997) menjadikan hal ini menarik untuk dilakukan. Dalam hal ini analisis yang digunakan melalui *Exploratory Factor Analysis* (EFA) untuk mengidentifikasi struktur data yang mendasari sehingga mendapatkan wawasan baru dalam pengembangan teori. Sedangkan, *Confirmatory Factor Analysis* (CFA) digunakan karena untuk mengkonfirmasi atau memvalidasi struktur instrumen pengukuran dan mengonfirmasi teori yang sudah ada sebelumnya. Dari pemaparan tersebut dapat dipahami bahwa pentingnya untuk menggunakan analisis faktor sebagai bentuk untuk mendapatkan struktur data yang baik dan untuk mengkonfirmasi suatu teori yang telah ada.

2. Metode Penelitian

Partisipan dalam penelitian ini adalah siswa Sekolah Menengah Pertama (SMP) dalam tahapan perkembangan remaja awal yang mengalami pertumbuhan dan perkembangan secara pesat dalam aspek fisik maupun psikologis (Azmi dkk, 2023). Instrumen yang digunakan yaitu hasil modifikasi dari teori *self-efficacy* (Bandura, 1997). Dari instrumen tersebut terdapat 37 item. Pertanyaan tertutup di ukur dengan menggunakan skala ordinal berupa Likert dari 1 (sangat tidak setuju) sampai 4 (sangat setuju). Metode penelitian yang digunakan bersifat psikometrik yaitu dengan *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA). Jumlah minimal sampel untuk dapat melakukan analisis faktor (CFA) adalah 50 – 100 dianggap sudah memadai (Sapnas & Zeller, 2002). Sedangkan sampel yang dapat digunakan untuk analisis *Exploratory*

Factor Analysis (EFA) minimal berjumlah 50 (de Winter dkk., 2009). Sampel dalam penelitian ini adalah 151 yang berarti sudah dapat menggunakan *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA). Pengolahan data *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA) menggunakan *software* jamovi 2.3.28.

3. Hasil dan Diskusi

Exploratory Factor Analysis (EFA)

Analisis EFA dilakukan untuk mengidentifikasi faktor yang mendasari untuk sekelompok butir (Pangalila & Budiarto, 2017). Uji kelayakan data pada penelitian ini menghasilkan nilai KMO = 0,787. KMO yang cenderung mendekati angka 1 akan menunjukkan data yang digunakan memadai untuk dilakukan uji analisis faktor (Yong & Pearce, 2013). Hasil *Bartlett's Test of Sphericity* pada penelitian ini menghasilkan signifikansi sebesar 0,001 artinya < 0.05 yang menunjukkan bahwa butir dalam alat ukur *self-efficacy* membentuk dalam faktor. Dari kedua data tersebut dapat disimpulkan bahwa penelitian ini memenuhi syarat untuk dilakukan uji analisis faktor.

Pada analisis faktor, batas minimum faktor *loading* yang digunakan dalam penelitian ini adalah $> 0,4$. Faktor *loading* dianggap baik apabila nilai faktor *loading* $> 0,4$ (Aviesta & Sutedjo, 2023). Hasil dari faktor *loading* yang tidak memenuhi syarat $< 0,4$ maka akan dibuang. Faktor pertama terdiri dari 13 item, setelah diuji EFA maka tersisa 11 item (3 item tidak memenuhi syarat). Faktor kedua terdiri dari 12 item, setelah diuji EFA maka tersisa 8 item (5 item tidak memenuhi syarat). Faktor ketiga terdiri dari 12 item, setelah diuji EFA maka tersisa 6 item (6 item tidak memenuhi syarat).

Tabel. 1. Hasil Uji Analisis Faktor Setelah Rotasi

Item	Factor		
	1	2	3
1. Saya tepat waktu untuk mengumpulkan tugas walaupun dianggap sulit	0.431		
2. Saya dapat mengerjakan soal yang sulit walaupun ada sebagian yang salah	0.490		
4. Saya selalu mengusahakan hasil yang maksimal dalam melewati semua mata pelajaran dikelas	0.483		
5. Saya merasa senang saat mengerjakan PR dari guru	0.424		
6. Saya tidak yakin bisa menyelesaikan PR yang sulit			0.444
8. Saya dapat mengerjakan PR yang dirasa sulit dengan mandiri tanpa bantuan orang lain	0.467		
9. Saya berusaha keras untuk mendapatkan nilai yang tinggi pada mata pelajaran yang dianggap sulit	0.566		
10. Saya merasa senang ketika dihadapkan dengan soal yang sulit	0.568		
12. Saya merasa pusing saat menghadapi soal yang sulit			0.735
13. Orang disekitar saya mengingatkan agar selalu belajar		0.628	
16. Saya tidak yakin menjadi juara kelas		0.618	
17. Saya seringkali tidak mencatat apa yang dijelaskan oleh guru		0.416	
19. Saya merasa bosan mengerjakan PR	0.482		
21. Saya tidak akan melakukan kesalahan yang sama ketika mengerjakan PR			0.465
22. Saya akan mengurangi jam bermain untuk belajar agar mendapatkan nilai yang baik		0.550	
24. Saya tidak akan belajar walaupun mengetahui nilai saya kurang bagus		0.540	
25. Saya seringkali kesal ketika dihadapkan dengan tugas yang sulit			0.664
26. Saya yakin dapat menguasai tugas yang diberikan	0.651		
27. Saya selalu membuat jadwal belajar yang teratur untuk memahami tugas yang diberikan	0.528		
28. Saya merasa takut ketika menghadapi ulangan harian			0.488
29. Saya merasa kurang mampu menguasai beberapa tugas yang diberikan			0.595

30. Saya yakin dapat mengerjakan soal yang diberikan dengan cara belajar terlebih dahulu sebelum ujian	0.486
31. Saya tidak menggunakan waktu belajar secara efektif	0.523
32. Saya hanya menyelesaikan soal latihan jika diawasi guru saja	0.470
36. Saya tidak bekerja keras dalam belajar	0.486

Berdasarkan hasil analisis EFA dapat dipahami bahwa dimensi *level* termasuk di dalam item nomor 1, 2, 4, 5, 8, 9, 10, 19, 26, 27, 30. Dimensi *generality* didalam item nomor 13, 16, 17, 22, 24, 31, 32, 36. Sedangkan dimensi *strength* didalam item nomor 6, 12, 21, 25, 28, 29. Perbedaan faktor *loading* pada penelitian lain dipengaruhi oleh faktor perbedaan jumlah sampel yang digunakan. Jumlah sampel yang besar diizinkan untuk menggunakan faktor *loading* yang kecil (Yong & Pearce, 2013). Sedangkan sampel yang digunakan didalam penelitian ini sebanyak 151 yang tergolong relatif kecil.

Confirmatory Factor Analysis (CFA)

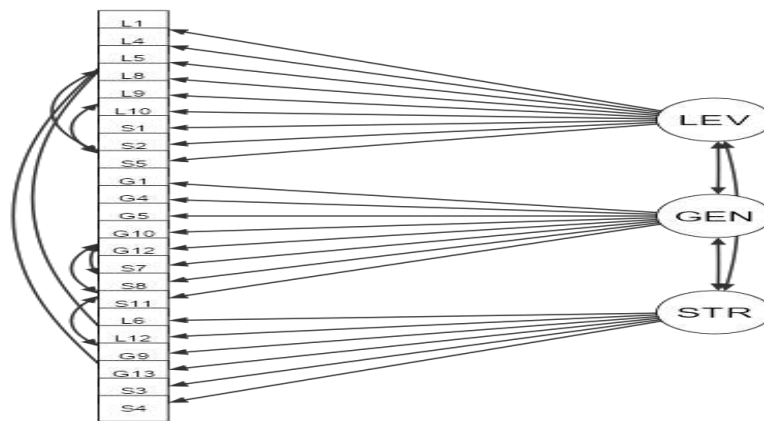
Setelah dilakukan analisis EFA terbentuk sebuah model baru. Langkah selanjutnya dengan melakukan analisis melalui *Confirmatory Factor Analysis* (CFA) berdasarkan hasil EFA yang sudah dilakukan penyesuaian. Namun, setelah dilakukannya analisis menggunakan metode *Confirmatory Factor Analysis* (CFA) didapatkan hasil *Goodness Of Fit* yang kurang diharapkan atau kurang fit. Langkah yang diambil peneliti untuk memperbaiki hasil *Goodness Of Fit* dengan melakukan *modification indices* yang bertujuan untuk memperbaiki model untuk meningkatkan hasil *fit* model (Dima dkk, 2023). Di dalam penelitian ini terdapat model persamaan struktural yang kurang sesuai sehingga diperlukan *modification index* yang dilakukan sebanyak 8 kali untuk menemukan model persamaan struktural yang sesuai atau mencapai kriteria *Goodnes Of Fit* yang baik. Hasil tersebut dijelaskan pada tabel dibawah ini.

Tabel 2. Goodness Of Fit

No	Ukuran Model Fit	Tingkatan <i>Fit</i> Yang Baik	Hasil	Penilaian
1.	Chi-square p-value	$P > 0,05$	<0.001	Tidak Fit
2.	CFI	$> 0,9$	0.921	Fit
3.	TLI	$\geq 0,90$	0.908	Fit
4.	SRMR	$<0,08$	0.0664	Fit
5.	RMSEA	$<0,08$	0.0461	Fit

Chi square seharusnya bernilai lebih besar dari 0.005 sehingga dikatakan fit, namun hal ini sangat sensitif dengan ukuran data. Seharusnya dianjurkan dapat menggunakan sampel yang besar (lebih dari 200) agar model dapat fit (Yamin & Kurniawan, 2009). Sedangkan sampel dalam penelitian ini berjumlah 151 sehingga dugaan besar model *chi square* dikategorikan tidak fit karena sampel yang kurang mencukupi.

Comparative Fit Index (CFI) merupakan pengembangan dari *Normed Fit Index* (NFI) yang menjadi salah satu alternatif model fit (Syafitri, 2017). Ukuran yang ideal untuk *Comparative Fit Index* (CFI) adalah lebih besar dari 0,9 sehingga dikatakan fit (Lu & Zhu, 2010). TLI digunakan untuk mengatasi permasalahan yang timbul akibat kompleksitas model. Ukuran yang dianjurkan agar dapat dikategorikan fit adalah lebih besar sama dengan 0,90 (Lippert, 2007). *Standardized RMR* (SRMR) merupakan cara alternatif untuk menguji model yang lain dengan ukuran idealnya lebih kecil dari 0,08 dan 0 dikategorikan *perfect fit* (Hooper et al, 2008). *Root Mean Square Error Of Approximation* (RMSEA) digunakan untuk mengukur penyimpangan nilai parameter pada suatu model dengan matrik kovarian populasinya (Syafitri, 2017). Ukuran yang ideal untuk dapat dikatakan fit adalah lebih kecil dari 0,08 (Zhao dkk., 2013).



Gambar 1. Path Diagram

Uji Reliabilitas

Setelah dilakukan uji validitas, langkah yang harus dilakukan selanjutnya dengan menguji reliabilitas. Uji reliabilitas digunakan untuk melihat kekonsistenan atau kestabilan jawaban responden terhadap pertanyaan yang disajikan (Rahim & Saputra, 2018). Pengujian reliabilitas dilakukan untuk mengetahui sebesar mana suatu instrumen dapat dipercaya untuk mengukur atribut psikologi dengan cara yang sama setiap alat ukur digunakan. Di dalam penelitian ini didapatkan hasil uji reliabilitas yang dilihat dari *Cronbach Alpha* bernilai 0.506. Guilford (1965) mengatakan jika *Cronbach Alpha* berkisar 0.40 – 0.60 maka dikategorikan sedang.

4. Kesimpulan

Sebelum dianalisis faktor (EFA), skala dimodifikasi dari Bandura. Skala *self-efficacy* terdapat 3 dimensi yaitu level (12 item), *generality* (13 item), dan *strength* (12 item). Setelah dilakukan analisis faktor (EFA) didapatkan 25 item dari 37. Hal ini disebabkan terdapat item yang gugur karena faktor *loading* yang tidak memenuhi batas ($<0,04$) sehingga terdapat 12 item yang gugur. Berdasarkan hasil analisis CFA, didapatkan model yang tidak fit berdasarkan kriteria *chi-square*, TLI, CFI, SRMR, dan RMSEA sehingga dilakukan proses *modification indices* untuk mendapatkan hasil yang fit. *Modification indices* sebanyak 8 kali untuk mencapai model yang fit berdasarkan kriteria yang sudah ditentukan. Namun, untuk kriteria *chi-square* mendapatkan hasil yang tidak fit dikarenakan sampel yang kurang 200. Model fit setelah dilakukan *modification indices* adalah TLI = 0.908, CFI= 0.921, SRMR= 0.0664, RMSEA= 0.0461 sehingga alat ukur ini dapat dikategorikan fit. Uji reliabilitas untuk alat ukur ini adalah $\alpha = 0.506$ yang dikategorikan sedang.

5. Referensi

- Aviesta, K. A., & Sutedjo, B. (2023). Pengaruh Brand Trust dan Brand Image terhadap Customer Satisfaction dan terhadap Minat Beli Ulang Pada Produk Scarllet Whitening. *SEIKO : Journal of Management & Business*, 6(2), 245–253.
- Azmi, D. N., Mahardika, I. K., Mutmainah, N., & Lestari, P. (2023). Pengertian perkembangan dan pertumbuhan anak usia SMP ditinjau dari pemahamannya terhadap pembelajaran IPA. *Jurnal Pendidikan Tambusai*, 7(3), 27171-27176.
- Bandura, A. (1997). *Self-Efficacy. The Exercise of Control*. New York. W.H. Froeman and Company.
- Charlina, T. R. (2013). Pengaruh Behavior Modification Terhadap Perubahan Frekuensi Perilaku Seksual pada Remaja di Kota Batu. Universitas Muhammadiyah Malang.
- de Winter, J. C., Dodou, D., & Wieringa, P. A. (2009). Exploratory Factor Analysis With Small Sample Sizes. *Multivariate behavioral research*, 44(2), 147–181. <https://doi.org/10.1080/00273170902794206>
- Dima, A., Kleden, M. A., & Atti, A. (2023). Analisis Faktor-Faktor yang Mempengaruhi Prestasi Siswa Menggunakan Metode Structural Equation Modeling (SEM). *Statistika*, 23(2), 132-146.
- Ferdianto, F. & Muhid, A. (2020). Stres Akademik Pada Siswa: Menguji Peranan Iklim Kelas Dan School Well-Being. *Jurnal Psikologi Indonesia*, 9(1), 140-156.
- Guilford, J. P. (1965). *Fundamental Statistics in Psychology and Education*: New York: McGraw-Hill.
- Hooper, D., Coughlan, J., & Mullen, M. (2008). *Structural Equation Modelling : Guidelines for*

- Determining Model Fit Structural equation modelling : guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60. <https://doi.org/10.1037/1082989X.12.1.58>.
- Lippert, S. K. (2007). Investigating postadoption utilization: An examination into the role of interorganizational and technology trust. *IEEE Transactions on Engineering Management*, 54(3), 468–483. <https://doi.org/10.1109/TEM.2007.900792>.
- Lu, C. T., & Zhu, D. S. (2010). The study on the determinants of the online consumers' intention to return. *Proceedings - 9th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2010*, 289–294. <https://doi.org/10.1109/ICIS.2010.88>.
- Nuzulia, S. 2010. *Dinamika Stress Kerja, Self-Efficacy dan Strategi Coping*. Semarang, Penerbit UNDIP.
- Oktaviani, T. L. (2018). Penerapan Konseling Kelompok Rasional Emotif Perilaku (Rep) Guna Mengurangi School Refusal (Penolakan Sekolah) Siswa Kelas Viii Smpn 1 Cerme. State University Of Surabaya.
- Pangalila, S. A. C., & Budiarto, Y. (2017). Factor Analysis of Rotter's Interpersonal Trust scale. *Humanitas*, 14(2), 150.
- Rahim, A., & Saputra, H. (2018). Exploratory factor analysis (EFA) pada penyerapan anggaran pendapatan dan belanja negara (APBN) tahun 2017 di Provinsi Sumatera Barat. *Indonesian Treasury Review: Jurnal Perbendaharaan, Keuangan Negara dan Kebijakan Publik*, 3(3), 236-254.
- Santrock, J. W. (2020). *Life-Span Development* McGraw-Hill Higher Education. In Boston, MA (14th ed.). Mc Graw Hill.
- Sapnas, K. G., & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement*, 10, 135–154.
- Sunaryo, Y. (2017). Pengukuran self-efficacy siswa dalam pembelajaran matematika di MTs N 2 Ciamis. *Teorema: Teori dan Riset Matematika*, 1(2), 39-44.
- Syafitri, W. (2017). Confirmatory Factor Analysis (CFA) Model Penerimaan Teknologi Berbasis Keamanan Informasi. *SITEKIN: Jurnal Sains, Teknologi dan Industri*, 14(2), 157-168.
- Yamin, S., & Kurniawan, H. (2009). *SPSS Complete*. Jakarta: Salemba Infotek
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Qualitative Methods for Psychology*, 9(2), 79-94.
- Zhao, W., Liu, J., Ye, D., & Wei, J. (2013). Mining user daily behavior patterns from access logs of massive software and websites. *Proceedings of the 5th Asia-Pacific Symposium on Internetware - Internetware '13*, 1–4. <https://doi.org/10.1145/2532443.2532462>.



ADAPTASI ALAT UKUR *LEARNING AGILITY* PADA MAHASISWA BARU: VERSI BAHASA INDONESIA

Yasa Widayati¹, Zulfa Indira Wahyuni², Nia Tresniasari³, Sitti Evangeline I. Suaidy⁴, Ayunni Santicha⁵, Niken Ayu Lestari⁶, Yasmin Syahidah Al-Bisyri⁷

^{1,2,3,4,5,6,7} Fakultas Psikologi, UIN Syarif Hidayatullah, Jakarta, Indonesia

Email: zulfa.indira@uinjkt.ac.id

Received :	Revised	Accepted :	Published
17 Januari 2025	25 Maret 2025	14 April 2025	30 April 2025

Abstract

Students are required to be flexible and responsive learners in the midst of uncertainty and changes in global situations. Learning agility in college students allows new students to adapt and create new values when they are in ambiguous situations. Therefore, a valid learning agility instrument is needed so that the adaptation of the measuring instrument to the instrument proposed by Gravett & Caldwell (2016) is carried out to test the learning agility measuring instrument for new students by implying four dimensions, namely people agility, change agility, results agility, and mental agility. The participants of this study were 222 new students of the Faculty of Psychology, Syarif Hidayatullah Islamic State University Jakarta. Confirmatory Factor Analysis (CFA) was used as a data processing method with the help of Lisrel 8.80 software. Based on CFA analysis, it can be concluded that there are two items dropped because they are not significant in measuring the dimensions of learning agility.

Keywords: learning agility, CFA, adaptation, learning

Abstrak

Abstrak dalam bahasa Indonesia,

Mahasiswa dituntut untuk menjadi pembelajar yang fleksibel dan cepat tanggap di tengah ketidakpastian dan perubahan situasi global. *Learning agility* pada mahasiswa memungkinkan mahasiswa baru untuk beradaptasi dan menciptakan nilai-nilai baru ketika mereka berada dalam situasi yang ambigu. Maka dari itu, diperlukan instrumen *learning agility* yang valid sehingga adaptasi alat ukur terhadap instrumen yang dikemukakan oleh Gravett dan Caldwell (2016) dilakukan untuk menguji alat ukur *learning agility* pada mahasiswa baru dengan mengimplikasikan empat dimensi, yaitu *people agility*, *change agility*, *results agility*, dan *mental agility*. Partisipan penelitian ini adalah 222 mahasiswa baru Fakultas Psikologi Universitas Islam Negeri Syarif Hidayatullah Jakarta. *Confirmatory Factor Analysis* (CFA) digunakan sebagai metode pengolahan data dengan bantuan software Lisrel 8.80. Berdasarkan analisis CFA dapat disimpulkan bahwa terdapat dua item yang di *drop* karena tidak signifikan dalam mengukur dimensi *learning agility*.

Kata kunci: kelincahan belajar, CFA, adaptasi, belajar

1. Pendahuluan

Perubahan dari siswa menjadi mahasiswa merupakan tantangan tersendiri bagi mahasiswa baru. Tidak hanya statusnya yang berubah, mahasiswa mengalami banyak perubahan seperti perubahan lingkungan, budaya akademik, beban akademik, serta budaya sosial. Perubahan-perubahan ini menuntut mahasiswa untuk dapat beradaptasi dengan baik. Ketidakmampuan dalam beradaptasi dapat menimbulkan berbagai permasalahan seperti tekanan psikologis dan gejala kelelahan akibat stres ekstrem (Lee & Larson, 2000). Stres ekstrem terutama stress akademik, dalam waktu yang lama dan tidak terselesaikan dapat menimbulkan *academic burnout* (Jeon et al., 2022). *Academic burnout* merupakan salah satu permasalahan yang mengganggu pelajar dan memiliki dampak yang signifikan terhadap kesehatan mental maupun prestasi akademik mahasiswa (Rahmatpour et al., 2019). *Academic burnout* menjadi salah satu indikator dari *learning engagement* di mana *learning engagement* juga dikaitkan dengan prestasi akademik yang baik. Selain itu, *learning engagement*

juga berkaitan dengan *learning agility*, di mana orang dengan *learning agility* yang tinggi cenderung memiliki *learning engagement* yang tinggi (Jeon et al., 2022).

Mahasiswa juga dituntut untuk menjadi pembelajar yang fleksibel dan cepat tanggap di tengah-tengah ketidakpastian pengetahuan yang tinggi dan perubahan situasi global, sebagai prasyarat untuk terjun ke dalam masyarakat (Kim et al., 2018). *Learning agility* pada mahasiswa memungkinkan mahasiswa baru untuk beradaptasi dan menciptakan nilai-nilai baru ketika mereka berada dalam situasi yang ambigu. Dalam proses belajar, *agility* merupakan faktor penting yang dapat memprediksi bagaimana kinerja seseorang. Sebagaimana digambarkan bahwa individu dengan *learning agility* yang sangat baik cenderung menikmati belajar dan tumbuh dengan terus mencari tantangan baru, secara aktif meminta umpan balik dari orang lain, dan merefleksikan pengalaman mereka (Jeon et al., 2022). Hal ini juga didukung oleh Korn Ferry (2017) bahwa individu dengan *learning agility* akan terus berkembang; memiliki kemampuan baru, belajar dari pengalaman, menerima tantangan baru, dan memiliki kinerja yang baik dalam situasi dinamis.

Menurut Lombardo & Eichinger (2001), *learning agility* adalah kemauan dan kemampuan untuk mempelajari kompetensi baru dalam kondisi yang baru pertama kali, sulit atau berbeda. Gravett & Caldwell (2016) menyatakan bahwa *learning agility* adalah kemampuan belajar yang berkaitan dengan kemampuan beradaptasi dan kemauan untuk menghadapi hal yang tidak diketahui dan digunakan untuk memprediksi kinerja seseorang dalam mengerjakan tugas-tugas baru. Sementara itu Hallenbeck (2016) menuliskan *learning agility* adalah faktor kunci yang membedakan mereka yang mampu mengekstraksi paling banyak pembelajaran dari pengalaman apapun dan kemudian menerapkannya. *Learning agility* seseorang dapat dilihat melalui kecepatan belajarnya (yaitu, kemampuan untuk memahami sesuatu dengan cepat) dan juga kemampuannya untuk berpindah dari satu ide ke ide yang lain atau dari satu situasi ke situasi yang lain (DeRue et al., 2012). DeRue et al. (2012) kemudian mendefinisikan *learning agility* sebagai kemampuan untuk cepat tanggap dalam memahami suatu situasi dan bergerak secara fleksibel dalam melayani pembelajaran baik di dalam maupun di luar pengalaman.

Dalam dunia akademik, *learning agility* dapat dipahami sebagai kemampuan seseorang dalam memahami pelajaran pada situasi yang baru serta kemampuan seseorang dalam mencari tantangan baru dan aktif dalam belajar (Bayu et al., 2023). Seseorang yang memiliki *learning agility* cenderung merefleksikan diri dan mengevaluasi pengalaman sebagai pembelajaran berikutnya. *Learning agility* sering dimanifestasikan sebagai keinginan untuk mencoba ide-ide baru, belajar dari kesalahan, memperbanyak pengetahuan, dan mencari atau mengembangkan teknik-teknik baru untuk memecahkan masalah yang sulit (DeRue et al., 2012). Selain itu, *learning agility* juga merupakan kemampuan seseorang dalam mengambil keputusan yang cepat, bahkan tanpa adanya data-data yang lengkap. Hal ini menggambarkan kemampuan seseorang untuk merangkul ambiguitas yang akan menjadi ide-ide baru untuk mengatasi masalah dalam situasi yang tidak biasa (DeRue et al., 2012).

Learning agility dikembangkan berdasarkan empat dimensi, yaitu *people agility*, *results agility*, *mental agility*, dan *change agility*.

- 1 *People Agility* : sejauh mana individu mengenal dirinya dengan baik, belajar dari pengalaman, memperlakukan orang lain secara konstruktif, dan bersikap dingin dan tangguh di bawah tekanan perubahan.
- 2 *Change Agility* : tingkat keingintahuan individu, memiliki hasrat untuk ide, suka bereksperimen dengan kasus uji, dan terlibat dalam kegiatan pengembangan keterampilan.
- 3 *Results Agility* : sejauh mana individu mendapatkan hasil dalam kondisi sulit, menginspirasi orang lain untuk tampil di luar normal, dan menunjukkan jenis kehadiran yang membangun kepercayaan pada orang lain.
- 4 *Mental Agility* : tingkat dimana individu memikirkan masalah dari sudut pandang baru dan merasa nyaman dengan kompleksitas, ambiguitas, dan menjelaskan pemikiran mereka kepada orang lain.

Beberapa penelitian terdahulu telah membahas *learning agility* pada mahasiswa. Jeon et al., (2022) menemukan bahwa *learning agility* merupakan faktor penting yang memengaruhi perilaku persiapan karier melalui tantangan akademis dan adaptasi terhadap kehidupan kampus. Namun, masih sedikit penelitian yang menghasilkan instrumen *learning agility* pada kalangan tersebut di Indonesia. Maka dari itu, penulis mengadaptasi instrumen *learning agility* yang dikemukakan oleh Gravett & Caldwell (2016) dan menguji alat ukur *learning agility* pada mahasiswa baru dengan mengimplikasikan empat dimensi, yaitu *people agility*, *change agility*, *results agility*, dan *mental agility*.

2. Metode Penelitian

Penelitian ini bertujuan untuk menguji alat ukur *learning agility* pada mahasiswa dengan melibatkan dimensi *people agility*, *result agility*, *mental agility*, dan *change agility*. Penelitian ini adalah penelitian kuantitatif. Populasi dalam penelitian ini merupakan mahasiswa baru fakultas Psikologi UIN Syarif Hidayatullah Jakarta Angkatan 2023, adapun sampel yang digunakan adalah semua anggota populasi sebanyak 222 responden.

Instrumen

Penelitian ini berfokus pada adaptasi alat ukur untuk *learning agility* (kelincahan belajar) yang mulanya dikembangkan oleh Gravett & Caldwell (2016). Kemudian pada sebuah studi terdahulu yang dilakukan oleh Lombardo & Eichinger (2000), menetapkan kerangka kerja yang terdiri dari empat faktor, dimana dimensi *people agility*, *results agility*, *mental agility*, dan *change agility* berkorelasi laten. Gravett & Caldwell (2016) kemudian menyusun *Learning agility Self-Assessment* (LASA) yang terdiri dari 24 item yang mengukur empat dimensi yang sama seperti yang dikemukakan oleh Lombardo & Eichinger (2000). Menggunakan skala likert 5 poin yaitu 1 = hampir tidak pernah, 2 = jarang, 3 = kadang-kadang, 4 = sering, 5 = selalu.

Semua item yang digunakan dalam penelitian ini diterjemahkan terlebih dahulu ke dalam Bahasa Indonesia oleh peneliti dan kemudian dilakukan uji keterbacaan untuk memastikan item-item tersebut dapat dipahami dan tidak multitafsir.

Tabel 1. Blueprint Learning Agility Self-Assessment (LASA)

Aspek	Indikator	No. Item	Contoh item
<i>Mental Agility</i>	Memikirkan masalah dari sudut pandang baru. Merasa nyaman dengan kompleksitas, ambiguitas. Dapat menjelaskan pemikiran mereka kepada orang lain.	1, 6, 8, 9, 12, 20	Jika saya mendengar kata yang tidak dikenal dalam percakapan, saya mencari artinya.
<i>People Agility</i>	Sejauh mana individu mengenal dirinya. Memperlakukan orang lain secara konstruktif. Bersikap dingin dan tangguh di bawah tekanan perubahan.	2, 14, 19, 22, 23, 25	Saya senang bekerja dengan orang lain dalam memecahkan masalah
<i>Change Agility</i>	Rasa ingin tahu yang tinggi. Memiliki hasrat untuk ide. Suka bereksperimen. Terlibat dalam kegiatan pengembangan keterampilan.	4, 5, 7, 11, 13, 16	Pengalaman baru bukanlah kesempatan belajar hal baru bagi saya
<i>Results Agility</i>	Mendapatkan hasil dalam kondisi sulit. Menginspirasi orang lain untuk tampil. Menunjukkan jenis kehadiran yang membangun kepercayaan pada orang lain.	3, 10, 15, 17, 18, 21, 24	Jika saya tidak mengetahui sesuatu, saya senang untuk bertanya

Adapun instrumen *Learning agility Self-Assessment* (LASA) yang terdiri 25 item disajikan dalam tabel berikut:

Tabel 2. Instrumen *Learning agility Self-Assessment (LASA)*

No Aitem	Aitem Asli	Terjemahan	Penyesuaian
1	<i>If I hear an unfamiliar word in a conversation, I look up the meaning</i>	Jika saya mendengar kata yang tidak dikenal dalam percakapan, saya mencari artinya	Ketika saya mendengar istilah asing dalam percakapan, saya mencari tahu artinya
2	<i>I enjoy working with others to try to solve problems.</i>	Saya menikmati bekerja dengan orang lain untuk mencoba menyelesaikan masalah	Saya senang bekerja dengan orang lain untuk memecahkan masalah
3	<i>If I don't know the answer to something, I am comfortable asking</i>	Jika saya tidak tahu jawaban untuk sesuatu, saya merasa nyaman bertanya	Jika saya tidak mengetahui sesuatu, saya tidak malu untuk bertanya
4	<i>I'm comfortable when circumstances that affect my work change.</i>	Saya merasa nyaman ketika keadaan yang memengaruhi perubahan pekerjaan saya.	Saya merasa nyaman ketika keadaan yang mempengaruhi pekerjaan saya berubah
5	<i>New experiences are not learning opportunities for me</i>	Pengalaman baru bukanlah kesempatan belajar bagi saya	Pengalaman baru bukanlah kesempatan belajar hal baru bagi saya
6	<i>I easily retain new information</i>	Saya dengan mudah menyimpan informasi baru.	Saya mudah dalam menyimpan informasi baru
7	<i>I can deliver results amidst changing circumstances</i>	Saya dapat memberikan hasil di tengah-tengah keadaan yang berubah.	Saya dapat memiliki pencapaian di tengah keadaan yang berubah
8	<i>I'm optimistic that I can learn new information</i>	Saya optimis bahwa saya dapat mempelajari informasi baru	Saya optimis dapat mempelajari informasi baru
9	<i>I don't make an effort to learn from my mistake</i>	Saya tidak berusaha belajar dari kesalahan saya	Saya tidak berusaha belajar dari kesalahan
10	<i>I'm very logical when I solve problems</i>	Saya sangat logis saat menyelesaikan masalah.	Saya berpikir sangat logis ketika memecahkan masalah
11	<i>I enjoy a change of pace in the way I do thing</i>	Saya menikmati perubahan kecepatan dalam cara saya melakukan sesuatu	Saya menyukai perubahan suasana dalam melakukan sesuatu
12	<i>I understand the best approach for me to learn something new</i>	Saya memahami pendekatan terbaik bagi saya untuk mempelajari sesuatu yang baru.	Saya mengetahui cara yang tepat untuk belajar sesuatu yang baru
13	<i>I enjoy researching new information</i>	Saya menikmati meneliti informasi baru	Saya senang mencari informasi baru
14	<i>I seek out feedback on my skills and abilities</i>	Saya mencari umpan balik tentang keterampilan dan kemampuan saya	Saya mencari umpan balik atas keterampilan dan kemampuan saya
15	<i>I don't accept others' information at face value</i>	Saya tidak menerima informasi orang lain dengan nilai nominal	Saya tidak menerima informasi orang lain begitu saja
16	<i>I can put aside a project and switch to another when necessary</i>	Saya bisa mengesampingkan proyek dan beralih ke yang lain saat diperlukan	Saya dapat mengesampingkan sebuah pekerjaan dan beralih ke pekerjaan lain bila diperlukan.
17	<i>I find satisfaction in digging deeply into ways to solve problems</i>	Saya menemukan kepuasan dalam menggali secara	Saya merasa puas ketika menggali secara mendalam

18	<i>If one problem-solving approach doesn't work, I move to another</i>	mendalam cara untuk menyelesaikan masalah Jika satu pendekatan pemecahan masalah tidak berhasil, saya pindah ke yang lain.	cara dalam memecahkan masalah Jika satu cara pemecahan masalah tidak berhasil, saya mengganti cara pemecahannya
19	<i>I seek out people to learn about subjects outside my work field</i>	Saya mencari orang untuk belajar tentang subjek di luar bidang kerja saya	Saya mencari orang lain untuk mempelajari hal-hal di luar lingkup yang saya kuasai
20	<i>I do not enjoy assignments that force me to learn something new</i>	Saya tidak menikmati tugas yang memaksa saya untuk mempelajari sesuatu yang baru	Saya tidak menikmati tugas yang memaksa saya untuk belajar hal baru
21	<i>I can find a way to get things done even when I don't have clear directions</i>	Saya dapat menemukan cara untuk menyelesaikan sesuatu bahkan ketika saya tidak memiliki arah yang jelas.	Saya dapat mencari cara menyelesaikan sesuatu bahkan ketika tidak ada petunjuk yang jelas
22	<i>I seek out people to show me ways to be more efficient and effective</i>	Saya mencari orang untuk menunjukkan kepada saya cara menjadi lebih efisien dan efektif	Saya mencari orang lain untuk menunjukkan cara baru yang lebih efektif dan efisien
23	<i>I enjoy learning from others</i>	Saya menikmati belajar dari orang lain	Saya senang belajar dari orang lain
24	<i>I look for ways to use new knowledge</i>	Saya mencari cara untuk menggunakan pengetahuan baru.	Saya mencari cara untuk menggunakan pengetahuan saya
25	<i>People seem to come to me when they need help for solving problems</i>	Orang-orang tampaknya datang kepada saya ketika mereka membutuhkan bantuan untuk menyelesaikan masalah	Orang cenderung datang ke saya apabila perlu bantuan dalam memecahkan masalah

Pengumpulan Data

Pengumpulan data dilakukan menggunakan kuesioner yang diberikan secara langsung kepada responden. Instrumen dicetak menjadi bentuk buku, kemudian diberikan kepada responden sebagai buku kerja

Analisis Data

Pada penelitian ini terbagi menjadi 2 bagian yaitu bagian adaptasi alat ukur dan bagian pengujian confirmatory factor analysis (CFA) untuk uji validitas. Bagian pertama adalah adaptasi alat ukur. Adaptasi alat ukur dilaksanakan dengan menerjemahkan setiap item ke dalam bahasa Indonesia kemudian menyesuaikannya supaya lebih mudah dipahami. Kemudian dilakukan uji baca untuk memastikan bunyi aitem dapat dipahami sebelum disusun dalam bentuk kuisisioner siap cetak. Bagian kedua yaitu uji validitas konstruk menggunakan *Confirmatory Factor Analysis (CFA)* dengan bantuan software Lisrel 8.80. Adapun langkah-langkah CFA yang dilakukan adalah sebagai berikut (Umar & Nisa, 2020):

1. Dilakukan uji CFA dengan model empat faktor dan dilihat nilai P-value dan RMSEA yang dihasilkan. Model dikatakan fit apabila P-value > 0.05 dan RMSEA < 0.05 yang berarti model dengan empat faktor dapat diterima dan seluruh item benar mengukur suatu faktor.
2. Jika model tidak fit maka perlu dilakukan modifikasi model dengan membiarkan kesalahan pengukuran berkorelasi bebas.

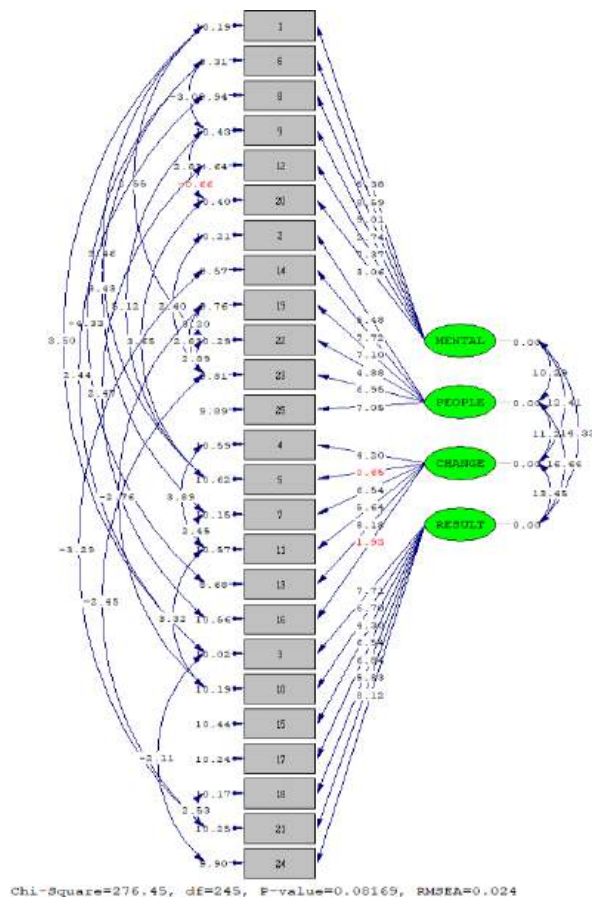
Setelah model fit, maka dilakukan analisis item dengan melihat *t-value* item lebih besardari 1.96 juga melihat apakah muatan faktor masing-masing item memiliki nilai koefisien yang positif. Jika *t-value* > 1.96 dan memiliki nilai koefisien positif, maka item tersebut dinyatakan signifikan dalam mengukur faktor yang hendak diukur dan tidak perlu didrop.

3. Hasil dan Diskusi

Pada bagian pertama, yaitu adaptasi alat ukur, aitem yang berbahasa Inggris diterjemahkan ke dalam bahasa Indonesia. Setelah itu dilakukan uji baca, kemudian dipersiapkan dalam layout buku berukuran a5 untuk pengambilan data. Setelah data terkumpul, kemudian dilakukan uji validitas dan reliabilitas dengan

CFA menggunakan lisrel 8.80. Dari hasil uji validitas, setelah membiarkan residu berkorelasi, model 4 faktor dari learning agility dikatakan fit dengan nilai signifikansi p-value sebesar 0.0816. Nilai ini lebih besar dari kriteria signifikansi 0.05 sehingga dapat dikatakan bahwa perbedaan yang ada antara data dan model tidak signifikan. Selain melihat dari nilai p-value, validitas konstruk dapat melihat nilai RMSEA. Pada model empat faktor ini nilai RMSEA sebesar 0.024. Nilai ini lebih kecil dari kriteria 0.05 dengan signifikansi dari RMSEA sebesar 1.00. Nilai tersebut lebih besar dari kriteria P-value <0.05 sehingga dapat dikatakan nilai RMSEA tersebut tidak signifikan. Artinya, menurut RMSEA, model yang digunakan fit karena tidak terdapat perbedaan yang signifikan antara data dan model.

Gambar 1. Path Diagram hasil CFA



a. Mental Agility

Tabel 3. Muatan Faktor Item Mental Agility

No. Item	Koefisien	Standard Error	T-value	keterangan
Item 1	0.38	0.07	5.38	V
Item 6	0.61	0.07	8.59	V
Item 8	0.61	0.07	9.01	V
Item 9	0.20	0.07	2.74	V
Item 12	0.51	0.07	7.37	V
Item 20	0.22	0.07	3.06	V

Berdasarkan tabel 3 di atas, dapat dilihat bahwa nilai t bagi koefisien muatan faktor 6 item semuanya signifikan. Hal ini ditunjukkan dengan nilai t yang lebih besar dari kriteria 1.96. Selanjutnya melihat muatan faktor dari item, diketahui tidak terdapat item dengan muatan faktor negatif. Sehingga keseluruhan item dari dimensi mental agility tidak perlu ada yang didrop. Selanjutnya akan dilihat apakah kesalahan pengukuran pada level item saling berkorelasi. Jika ternyata suatu item memiliki kesalahan pengukuran yang berkorelasi dengan kesalahan pengukuran pada banyak item lain, maka berarti bahwa item tersebut bersifat

multidimensional, karena mengukur lebih dari satu hal. Dalam dimensi ini, kesalahan pengukuran pada item-item masih bisa diterima dan diikutsertakan dalam perhitungan faktor skor.

b. People Agility

Tabel 4. Muatan Faktor Item People Agility

No Item	Koefisien	Standar Error	T-value	Keterangan
Item 2	0.38	0.07	5.48	V
Item 14	0.53	0.07	7.72	V
Item 19	0.49	0.07	7.10	V
Item 22	0.34	0.07	4.88	V
Item 23	0.49	0.07	6.95	V
Item 25	0.48	0.07	7.05	V

Berdasarkan tabel 4 di atas, dapat dilihat bahwa nilai t bagi koefisien muatan faktor 6 item semuanya signifikan. Hal ini ditunjukkan dengan nilai t yang lebih besar dari kriteria 1.96. Selanjutnya melihat muatan faktor dari item, diketahui tidak terdapat item dengan muatan faktor negatif. Sehingga keseluruhan item dari dimensi people agility tidak perlu ada yang didrop. Selanjutnya akan dilihat apakah kesalahan pengukuran pada level item saling berkorelasi. Jika ternyata suatu item memiliki kesalahan pengukuran yang berkorelasi dengan kesalahan pengukuran pada banyak item lain, maka berarti bahwa item tersebut bersifat multidimensional, karena mengukur lebih dari satu hal. Dalam dimensi ini, kesalahan pengukuran pada item-item masih bisa diterima dan diikutsertakan dalam perhitungan faktor skor.

c. Change Agility

Tabel 5. Muatan Faktor Item Change Agility

No. Item	Koefisien	Standar Error	T-value	keterangan
Item 4	0.27	0.06	4.20	V
Item 5	0.04	0.06	0.65	X
Item 7	0.45	0.07	6.54	V
Item 11	0.37	0.07	5.64	V
Item 13	0.58	0.07	8.18	V
Item 16	0.12	0.06	1.93	X

Berdasarkan tabel 5 di atas, dapat dilihat bahwa muatan faktor pada enam item memiliki nilai koefisien positif, namun nilai t bagi item 5 dan item 16 tidak signifikan. Hal ini ditunjukkan dengan nilai t yang lebih kecil dari kriteria 1.96. Artinya item-item tersebut tidak signifikan dalam mengukur change agility dan dapat dihilangkan. Sedangkan empat item lainnya dapat digunakan. Melihat hal tersebut, item 5 dan item 16 dapat didrop sehingga tersisa 4 dari 6 item yang mengukur *change agility*.

Nilai t-value yang kecil menandakan item tersebut tidak memiliki karakter yang sama dengan item lain. Artinya, sampel memiliki penilaian yang tidak konsisten. Sampel dengan change agility yang tinggi memiliki pandangan yang berbeda-beda dalam menilai item-item tersebut, begitu juga sampel dengan change agility yang rendah. Hal itu terjadi pada item 16 yang berbunyi “Saya dapat mengesampingkan sebuah pekerjaan dan beralih ke pekerjaan lain bila diperlukan.” Selain itu, tata bahasa juga dapat mempengaruhi inkonsistensi dalam jawaban responden. Kalimat yang sulit dipahami akan memicu penilaian yang berbeda dari setiap responden dan menimbulkan pertanyaan. Salah satu item yang menimbulkan pertanyaan dari responden pada saat pengerjaan adalah item 5 yang berbunyi “Pengalaman baru bukanlah kesempatan belajar hal baru bagi saya”.

d. Results Agility

Table 6. Muatan Faktor Item Results Agility

No. Item	Koefisien	Standar Error	T-value	keterangan
Item 3	0.51	0.07	7.71	V
Item 10	0.45	0.07	6.70	V
Item 15	0.30	0.07	4.30	V
Item 17	0.44	0.07	6.54	V

Item 18	0.46	0.07	6.84	V
Item 21	0.40	0.07	5.83	V
Item 24	0.54	0.07	8.12	V

Berdasarkan tabel 6 di atas, dapat dilihat bahwa nilai t bagi koefisien muatan faktor 7 item semuanya signifikan. Hal ini ditunjukkan dengan nilai t yang lebih besar dari kriteria 1.96. Selanjutnya melihat muatan faktor dari item, diketahui tidak terdapat item dengan muatan faktor negatif. Sehingga keseluruhan item dari dimensi *results agility* tidak perlu ada yang didrop. Selanjutnya akan dilihat apakah kesalahan pengukuran pada level item saling berkorelasi. Jika ternyata suatu item memiliki kesalahan pengukuran yang berkorelasi dengan kesalahan pengukuran pada banyak item lain, maka berarti bahwa item tersebut bersifat multidimensional, karena mengukur lebih dari satu hal. Dalam dimensi ini, kesalahan pengukuran pada item-item masih bisa diterima dan diikutsertakan dalam perhitungan faktorskor.

4. Kesimpulan

Berdasarkan hasil analisis *Confirmatory Factor Analysis* (CFA) model empat faktor menunjukkan bahwa setiap dimensi dari *Learning agility*, yaitu *Mental Agility*, *People Agility*, *Change agility*, dan *Results Agility* dapat diterima namun memerlukan modifikasi sebanyak 24 kali untuk mencapai model fit. Meskipun modifikasi dilakukan, karena kesalahan pengukuran pada item-item tetap dapat diterima dan diikutsertakan dalam perhitungan faktor skor.

Semua item terbukti mengukur hanya satu dari empat faktor sehingga model dengan empat faktor dapat diterima. Selanjutnya, melihat koefisien muatan faktor pada setiap item, tidak ada yang bermuatan negatif. Namun, untuk signifikansi T-value, ditemukan dua item yang memiliki nilai T-value < 1.96 yaitu pada item 5 dan item 16, sehingga kedua item tersebut harus di drop dan tidak dimasukkan pada penghitungan faktor. Nilai T-value < 1.96 terjadi bila item tersebut tidak memiliki karakter yang sama dengan item lain sehingga terjadi penilaian yang tidak konsisten dari sampel. Selain itu tata bahasa yang sulit dipahami juga dapat mempengaruhi ketidakkonsistenan. Kedua hal tersebut menjadi faktor penyebab item 5 dan item 16 tidak signifikan dalam mengukur dimensi *change agility*.

Oleh karena itu, dapat disimpulkan bahwa setelah dilakukan uji validitas didapatkan 2 item yang tidak signifikan, sehingga *Learning agility Self-Assessment* (LASA) yang pada awalnya memiliki 25 item menjadi 23 item yang mencakup dimensi *mental agility*, *people agility*, *change agility*, dan *results agility*. *Learning agility Self-Assessment* LASA dengan 23 item dapat dianggap sebagai alat ukur yang valid dan telah teruji dengan CFA sehingga dapat digunakan untuk mengukur *learning agility* pada individu. Adapun penelitian lebih lanjut dapat dilakukan dengan memperluas sampel yang lebih beragam untuk memastikan generalitas hasil. Dengan demikian, diharapkan bahwa pengembangan ini akan meningkatkan validitas, reliabilitas, dan ketepatan LASA sebagai alat yang efektif dalam mengukur *learning agility* pada berbagai populasi.

5. Referensi

- Bayu, I Gede Wira., Widiana, I Wayan., Yudiana, I Kadek. (2023). Learning Science with Numbered Heads Together (NHT) based on Growth Mindset Improving Science Literacy and Learning agility of Elementary School Students. *Pegegog Journal of Education and Instruction*. <https://doi.org/10.47750/pegegog.13.04.21>
- De Meuse, K. P. (2023). A Meta-Analytic Examination of the Relationship between Learning agility and General Cognitive Ability. *Journal of Managerial Issues*, XXXV(1), 18-33.
- De Meuse, K. P., Dai, G., & Hallenbeck, G. S. (2010). Learning agility: A Construct Whose Time Has Come. *Consulting Psychology Journal: Practice and Research*, 22(2), 119-130. 10.1037/a0019988
- De Meuse, K. P., Dai, G., & Swisher, V. V. (2012). Leadership Development: Exploring, Clarifying, and Expanding Our Understanding on Learning agility. *Industrial and Organizational Psychology*, 5, 280-315.
- De Meuse, K. P., & Harvey, V. S. (2022). The Science and Application of Learning agility: Introduction to The Special Issue. *Consulting Psychology Journal*, 74(3), 207-214.
- Derue, D. S., Ashford, S. J., & Myers, C. G. (2012). Learning agility: In Search of Conceptual Clarity and Theoretical Grounding. *Industrial and Organizational Psychology*, 5, 258-279.
- Gravett, Linda S., Caldwell, Sheri A. (2016). *Learning agility: The Impact on Recruitment and Retention*. New York : Palgrave Macmillan.
- Hallenbeck, George. (2016). *Learning agility: Unlock The Lessons of Experience*. USA : Center for Creative Leadership.
- Jeon, Mi-Kyeong., Lee, Insook., Lee, Mi-Young. (2022). The multiple mediating effects of grit and learning agility on academic burnout and learning engagement among Korean university students : a cross-sectional study. *Annals of Medicine*, Vol 54 : 1. doi: 10.1080/07853890.2022.2122551.

- Kim, Hye Jeong., Hong, Ah Jeong., Song Hae-Dok. (2018). The Relationships of Family, Perceived Digital Competence and Attitude, and Learning agility in Sustainable Student Engagement in Higher Education. *Sustainability*, 10(12), <https://doi.org/10.3390/su10124635>
- Korn Ferry. (2017). *The Organisational X-factor: Learning agility*. Korn Ferry Focus. Retrieved August 11, 2023, <https://focus.kornferry.com/the-organisational-x-factor-learning-agility/>
- Lee, Meery., Larson, Reed. (2000). The Korean ‘Examination Hell’’: Long Hours of Studying, Distress, and Depression. *Journal of Youth and Adolescence*, 29: 249-271. <https://link.springer.com/article/10.1023/A:1005160717081>
- Lombardo, Michael.M., Eichinger, Robert.W. (2001). High potentials as high learners. *Human Resource Management*, 39 (4). [https://doi.org/10.1002/1099-050X\(200024\)39:4<321::AID-HRM4>3.0.CO;2-1](https://doi.org/10.1002/1099-050X(200024)39:4<321::AID-HRM4>3.0.CO;2-1)
- Özgenel, M., & Yazıcı, Ş. (2021). Learning agility of School Administrators: An Empirical Investigation. *International Journal of Progressive Education*, 17(1), 247-261. 10.29329/ijpe.2020.329.16
- Rahmatpour, Pardis., Chehrzad, Minoomitra., Ghanbari, Atefeh., Sadat-Ebrahimi, Seyyed-Reza. (2019). Academic burnout as an educational complication and promotion barrier among undergraduate students : A cross-sectional study. *Journal of Education and Health Promotion*. Vol 8(1), doi: 10.4103/jehp.jehp_165_19.
- Saputra, N., Abdinagoro, S. B., & Kuncoro, E. A. (2018). The Mediating Role of Learning agility on the Relationship between Work Engagement and Learning Culture. *Social Sciences & Humanities*, 26, 117-130
- Umar, Jahja., Nisa, Yunita Faela. (2020). Uji Validitas Konstruk dengan CFA dan Pelaporannya. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia*, 9 (2), 1-11. <http://dx.doi.org/10.15408/jp3i.v9i2.XXXXXX>
- Wardhani, N. S., Sulastiana, M., & Ashriyana, R. (2022). Adaptasi Alat Ukur Learning agility pada Karyawan untuk Meningkatkan Organizational Agility: Versi Bahasa Indonesia. *Psikologika*, Volume 27(2). DOI:10.20885
- Yantiningtyas, I. D. (2020, December 9). *Pentingnya Learning agility di Masa Pandemi dan VUCA*. SWA. Retrieved August 10, 2023, from <https://swa.co.id/swa/my-article/pentingnya-learning-agility-di-masa-pandemi-dan-vuca>

ANALISIS PSIKOMETRIS *THE MORAL INJURY SCALE***Yasmin Siti Aulia¹, Aditya Jayadinata², Ally Allena Greene³, Diah Zaleha Wyandini⁴, Ghinaya Ummum Mukminin Hidayat⁵**^{1,2,3,4,5} Fakultas Ilmu Pendidikan, Universitas Pendidikan Indonesia, Bandung, IndonesiaEmail: yasmnsaa@upi.edu

Received :	Revised	Accepted :	Published
10 Januari 2025	08 April 2025	22 April 2025	30 April 2025

Abstract

This study aims to develop a valid and reliable measurement tool for assessing moral injury among university students in higher education settings. Moral injury is the psychological impact resulting from violations of deeply held moral values, potentially leading to guilt, shame, and existential crises. The developed instrument was tested through item analysis, validity, and reliability assessments. Results indicated that all items had item-total correlations above 0.30 and factor loadings ≥ 0.40 , demonstrating psychometric validity. Reliability, measured using Cronbach's α , achieved a score of 0.963, indicating a very high level of consistency. However, model fit analysis revealed that the instrument did not fully meet the model fit criteria (GFI = 0.811; RMSEA = 0.089). These findings highlight the need for further model refinement and the development of psychological interventions to support student well-being. This study makes a significantly contributes educational psychology literature and provides a foundation for improving measurement tools to assess moral injury more effectively

Keywords: Moral Injury, Validity, Reliability, Model Fit, Psychological Measurement Tools.

Abstrak

Penelitian ini bertujuan mengembangkan alat ukur moral injury pada mahasiswa yang valid dan reliabel dalam konteks pendidikan tinggi. *Moral injury* didefinisikan sebagai dampak psikologis akibat pelanggaran nilai moral yang diyakini, yang dapat memicu rasa bersalah, malu, dan krisis eksistensial. Instrumen yang dikembangkan diuji melalui analisis item, validitas, dan reliabilitas. Hasil menunjukkan bahwa seluruh item memiliki korelasi item-total di atas 0,30 dan factor loading $\geq 0,40$, yang menandakan validitas dan kelayakan secara psikometrik. Reliabilitas diukur menggunakan koefisien *Cronbach's α* dengan nilai 0,963, menunjukkan tingkat konsistensi yang sangat tinggi. Namun, analisis model fit menunjukkan bahwa alat ini belum sepenuhnya memenuhi kriteria kecocokan model (GFI = 0,811; RMSEA = 0,089). Temuan ini menyoroti kebutuhan akan perbaikan model serta intervensi psikologis untuk mendukung kesejahteraan mahasiswa. Penelitian ini memberikan kontribusi signifikan dalam literatur psikologi pendidikan dan menyediakan dasar untuk pengembangan instrumen yang lebih efektif.

Kata kunci: Moral Injury, Validitas, Reliabilitas, Model Fit, Alat Ukur Psikologi.

1. Pendahuluan

Moral injury adalah jenis trauma psikologis tertentu yang ditandai dengan rasa bersalah, rasa malu, dan krisis spiritual yang intens, yang dapat berkembang ketika seseorang melanggar keyakinan moralnya, dikhianati, atau menyaksikan orang yang dipercaya melakukan kekejaman (Jinkerson, 2016). Istilah ini awalnya muncul dalam konteks militer untuk menggambarkan dampak psikologis dari peristiwa yang melibatkan tindakan yang bertentangan dengan prinsip moral, seperti kekerasan dalam perang (Litz et al., 2009). Namun, seiring dengan berkembangnya penelitian, *moral injury* juga telah diteliti pada berbagai konteks dan kelompok populasi, seperti dunia medis (Ford, 2019), tokoh agama (Greene, Bloomfield & Billings, 2020), pendidikan (Levinson, 2015; Murray, Krahe & Goodsmann, 2018), orang tua (Haight, Sugrue, Calhoun, dan Black, 2017). Berkaitan dengan konteks pendidikan di Indonesia, *moral injury* dapat berdampak pada peserta didik, yaitu siswa dan atau

mahasiswa. *Moral injury* dalam konteks pendidikan merujuk pada dampak psikologis yang dialami individu, terutama siswa, akibat pelanggaran norma moral yang mereka saksikan atau alami, seperti *bullying* dan kekerasan komunikasi. Beberapa penelitian menyebutkan bahwa akibat pelanggaran norma yang disaksikan ataupun yang dialami siswa dapat berdampak pada prestasi akademik, kesehatan mental, dan juga perubahan perilaku (Sunanih et al., 2025; Putri, 2022). pengukuran, terdapat beberapa pengembangan alat ukur *moral injury*. Pada mahasiswa, dampak kesehatan mental dari *moral injury* salah satunya adalah *self-harm* (Dara, 2024). *Self-harm* sangat terkait dengan pikiran untuk bunuh diri dan sekitar 27-30% remaja dengan *Non-Suicidal Self-Injury* (NSSI) telah melaporkan upaya bunuh diri (Gillies et al., 2018). Mahasiswa sebagai kelompok yang rentan terhadap tekanan akademik, sosial, dan emosional, dapat mengalami *moral injury* melalui berbagai pengalaman traumatis, seperti ketidakadilan dalam sistem pendidikan, konflik nilai, atau pengkhianatan oleh individu yang dipercayai (Shay, 2014). Mengingat dampak dari *moral injury* pada mahasiswa yang dapat berakibat adanya upaya bunuh diri, maka peneliti berasumsi perlu untuk mengidentifikasi *moral injury* secara spesifik pada mahasiswa.

Berdasarkan hasil penelusuran literatur terdapat beberapa pengembangan alat ukur *moral injury* (Bryan et al., 2015; Yeterian et al., 2019; Wingerter, 2023). Bryan et al. (2015 dan Yeterian et al. (2019) mengembangkan alat ukur *moral injury* dengan sampel dari kalangan militer yang masih aktif bekerja (Bryan et al., 2015) dan yang sudah pensiun dari kemiliterannya atau veteran perang (Yeterian et al., 2019). Sedangkan sampel penelitiannya Wingerter (2023) adalah orang dewasa secara umum. Di Indonesia sendiri sejauh penelusuran peneliti belum menemukan alat ukur yang secara khusus mengukur *moral injury*. Untuk itu peneliti bermaksud untuk mengembangkan alat ukur *moral injury* versi Indonesia dengan melakukan adaptasi alat ukur yang dikembangkan oleh Wingerter (2023). Peneliti memilih alat ukur dari Wingerter (2023) karena sampel dalam pengembangan alat ukurnya lebih relevan dengan sampel mahasiswa, yaitu orang dewasa secara umum. Pengembangan alat ukur *moral injury* versi bahasa Indonesia ini diharapkan dapat mengidentifikasi *moral injury* pada mahasiswa yang sesuai dengan konteks dan budaya Indonesia.

2. Metode Penelitian

Partisipan Penelitian

Populasi dalam penelitian ini adalah mahasiswa aktif di Kota Bandung. Jumlah mahasiswa di Kota Bandung tidak ada jumlah pasti maka peneliti mengambil ketentuan minimal sampel dari Hair et al. (2010) yang menyatakan bahwa jumlah minimal sampel adalah 5 kali jumlah item. Teknik *sampling* yang digunakan adalah *non probability sampling* dengan *accidental sampling*. Partisipan dalam penelitian ini sebanyak 222 orang mahasiswa yang berusia 18 sampai dengan 35 tahun. Jumlah ini sudah melebihi kriteria sampel minimal dari Hair et al. (2010), jumlah item *The Adult Moral Injury Scale* adalah 39 dikali 5 yaitu 198 orang.

Instrumen

Instrumen yang digunakan dalam penelitian ini adalah *The Adult Moral Injury Scale* dari Wingerter (2023) yang telah dimodifikasi dan diterjemahkan ke dalam Bahasa Indonesia. Alat ukur ini terdiri dari 39 item dengan kategori respon mulai dari Sangat Tidak Setuju hingga Sangat Setuju.

Tabel 1. Dimensi, Indikator, dan Contoh Item

Dimensi	Indikator	Contoh Item Asli	Contoh Item Adaptasi
<i>Perpetration - Based Moral Injury</i>	<i>Spiritual/existential crisis</i>	<i>Having done bad things in the past, I have less confidence in the existence of God (or other higher power).</i>	Saya tidak terlalu yakin dengan kehadiran Tuhan, karena hal buruk yang pernah saya lakukan di masa lalu
		<i>I no longer trust myself to know right from wrong.</i>	Saya tidak lagi percaya pada diri sendiri untuk membedakan mana yang benar dan mana yang salah.
	<i>Guilt</i>	<i>I feel I must protect others from myself.</i>	Saya merasa harus melindungi orang lain dari diri Saya sendiri.
	<i>Shame</i>	<i>I don't take care of myself.</i> <i>I feel ashamed of things I have done in the past.</i>	Saya tidak merawat diri sendiri. Saya merasa malu atas hal-hal yang telah Saya lakukan di masa lalu.

<i>Betrayal-based Moral Injury</i>	<i>Spiritual/existential crisis</i>	<i>Since being betrayed/double-crossed/backstabbed/hurt in the past, I question my purpose.</i>	Sejak dikhianati/ disakiti di masa lalu, Saya mulai mempertanyakan tujuan Saya.
	<i>Anger</i>	<i>Since being betrayed/double-crossed/backstabbed/hurt in the past, I am less ready to trust others. I fantasize about hurting those who hurt me.</i>	Sejak dikhianati/ disakiti di masa lalu, Saya jadi kurang siap untuk mempercayai orang lain. Saya sering berfantasi untuk menyakiti mereka yang menyakitiku. Saya sering memikirkan orang-orang yang menyakiti saya
	<i>Feelings of betrayal and humiliation</i>	<i>I often think about those who have hurt me.</i>	Sejak dikhianati/disakiti di masa lalu, Saya cenderung lebih cuek pada perasaan atau menyakiti orang lain
		<i>Since being betrayed/double-crossed/backstabbed/hurt in the past, I tend to care less about offending or hurting others.</i>	Hal-hal yang terjadi pada Saya di masa lalu adalah sesuatu yang tak termaafkan.
		<i>Things have happened to me in the past that are unforgivable.</i>	Saya merasa malu atau terhina karena telah mempercayai seseorang/sesuatu yang melanggar kepercayaanku.
		<i>I feel embarrassed or humiliated for having trusted someone/something that violated my trust.</i>	

Desain Penelitian

Pendekatan penelitian yang dilakukan adalah pendekatan kuantitatif dengan metode survei.

Teknik Analisis

Prosedur pengumpulan data diawali dengan tahap persiapan, di mana instrumen diterjemahkan kemudian melakukan uji validitas isi dengan *expert judgement* dari ahli bahasa dan ilmuwan psikologi. Setelah itu, pengumpulan data dilakukan secara daring melalui media sosial serta secara luring di kampus. Selanjutnya, data yang telah terkumpul diverifikasi untuk memastikan kelengkapan dan konsistensinya.

Analisis data dilakukan melalui beberapa tahapan. Pertama, analisis item menggunakan korelasi item-total untuk menilai kelayakan setiap item. Kedua, analisis faktor melibatkan *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA) untuk menguji validitas konstruksi. Reliabilitas diukur menggunakan koefisien *Cronbach's Alpha*.

3. Hasil

Analisis Item

Menurut Azwar (2012), salah satu kriteria untuk menentukan kelayakan item dalam instrumen adalah berdasarkan nilai korelasi item-total. Biasanya, item dengan nilai korelasi lebih dari 0,30 dianggap memenuhi syarat psikometrik dan layak digunakan dalam pengukuran. Sebaliknya, item dengan nilai korelasi kurang dari 0,30 dinilai tidak layak dan sebaiknya tidak digunakan. Berdasarkan analisis yang sudah dilakukan menggunakan JASP, 39 item pada alat ukur ini memiliki nilai korelasi item-total di atas 0,30, dengan rentang nilai antara 0,396 hingga 0,758. Artinya, semua item tersebut, mulai dari item 1 hingga item 39, dinyatakan memenuhi kriteria kelayakan psikometrik dan dapat digunakan dalam pengukuran.

Confirmatory Factor Analysis

Byrne (2016) menjelaskan bahwa *Confirmatory Factor Analysis* (CFA) adalah teknik statistik yang digunakan untuk menguji hipotesis mengenai struktur faktor yang telah ditentukan sebelumnya. Dalam CFA, peneliti mengajukan hipotesis tentang hubungan antara variabel-variabel dan memanfaatkan data yang ada untuk mendukung atau menolak model tersebut. Tahap pertama yang perlu diperhatikan adalah melihat nilai KMO MSA (*Kaiser Mayer-Olkin Measure of Sampling Adequacy* dan *Bartlett's Test*), hal ini digunakan guna mengukur apakah suatu variabel layak untuk dianalisis faktor atau tidak (Hayurika & Arief, 2015)

Nilai KMO secara keseluruhan sebesar 0,938 menunjukkan bahwa data memiliki tingkat sampling adequacy yang sangat baik. Nilai MSA (*Measure of Sampling Adequacy*) untuk setiap indikator harus berada di atas 0,50 agar dapat digunakan dalam analisis faktor (Hair et al., 2010). Berdasarkan hasil yang didapat, semua indikator memiliki nilai MSA lebih dari 0,50, sehingga tidak ada indikator yang perlu dieliminasi. Hasil ini menunjukkan bahwa seluruh variabel layak untuk dilanjutkan ke tahap analisis faktor.

Model Fit

Menurut Wijayanto (2008), analisis *model fit index* membutuhkan evaluasi terhadap beberapa fit indeks. Fit indeks yang digunakan meliputi GFI (*Goodness of Fit Index*), NFI (*Normed Fit Index*), NNFI (*Non-normed Fit Index*), dan CFI (*Comparative Fit Index*). Sebuah model dikatakan fit apabila nilai GFI, NFI, NNFI, dan CFI berada dalam rentang >0,9 hingga <1,00, sehingga dapat disimpulkan bahwa model tersebut memiliki *Good fit*. Selain itu, RMSEA (*Root Mean Square Error of Approximation*) dianggap fit jika nilainya kurang dari 0,05.

Tabel 2. Model Fit Indeks

Indeks	Value	Kriteria	Keterangan
Chi-Square	<0,001	>0,05	Tidak Fit
GFI (Goodness of Fit Index)	0,811	>0,9 dan <1,00	Tidak Fit
NFI (Normed Fit Index)	,689	>0,9 dan <1,00	Tidak Fit
NNFI (Non-normed Fit Index)	0,760	>0,9 dan <1,00	Tidak Fit
CFI (Comparative Fit Index)	0,774	>0,9 dan <1,00	Tidak Fit
RMSEA (Root Mean Square Error of Approximation)	0,089	<0,05	Tidak Fit

Hasil analisis menunjukkan bahwa model yang diajukan memiliki *p-value* pada uji *Chi-Square* sebesar <0,001, yang lebih kecil dari ambang batas >0,05. Hal ini mengindikasikan bahwa model tidak fit. Selain itu, indeks *Goodness of Fit Index* (GFI) menunjukkan nilai sebesar 0,811, yang juga berada di bawah ambang batas >0,9, sehingga model ini tidak memenuhi kriteria fit berdasarkan GFI. Selanjutnya, nilai *Normed Fit Index* (NFI), *Non-normed Fit Index* (NNFI), dan *Comparative Fit Index* (CFI) masing-masing sebesar 0,689, 0,760, dan 0,774, yang semuanya lebih rendah dari ambang batas >0,9. Dengan demikian, model dinyatakan tidak fit berdasarkan ketiga indeks ini. Di sisi lain, nilai RMSEA sebesar 0,089 menunjukkan bahwa model ini tidak fit karena melebihi ambang batas yang ditetapkan, yaitu <0,05. Secara keseluruhan, berdasarkan hasil analisis model fit di atas, dapat disimpulkan bahwa alat ukur *The Adult Moral Injury Scale* yang dirancang tidak fit. Model ini tidak memenuhi kriteria kecocokan yang ditentukan sesuai pedoman yang digunakan.

Factor Loadings

Menurut Sudaryono (2019), sebuah alat ukur dianggap valid jika setiap item memiliki factor loading $\geq 0,4$. Data pada Tabel 3 menunjukkan bahwa seluruh item memiliki nilai *factor loading* yang berada dalam rentang 0,413 hingga 0,802. Oleh karena itu, dapat disimpulkan bahwa semua item, yaitu item 1 hingga item 39, memenuhi syarat validitas dan menjadi bukti bahwa alat ukur ini valid untuk digunakan.

Tabel 3. Muatan dan Penamaan Faktor

Faktor	Item	Muatan Faktor
Faktor 1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14	0,615 – 0,784
Faktor 2	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25	0,413 – 0,802
Faktor 3	26, 27, 28, 29, 30, 31, 32	0,495 – 0,773
Faktor 4	33, 35, 36, 37	0,605 – 0,788
Faktor 5	38, 39	0,695 – 0,741

Tabel 3 menunjukkan bahwa faktor yang sebelumnya terdapat 6 faktor terbagi menjadi lima faktor, secara sistematis tidak ada perpindahan item yang banyak dari satu faktor ke faktor lain, namun terdapat satu faktor yang terdiri dari satu item tergabung ke faktor sebelumnya. Penamaan faktor disesuaikan dengan item-item yang tergabung di dalam faktor 1 dinamakan *spiritual/ existential crisis*, faktor kedua dinamakan *guilt*, faktor ketiga *existential betrayal*, faktor ke empat *anger*, faktor kelima *Feelings of betrayal and humiliation*.

Reliabilitas

Penghitungan reliabilitas dilakukan untuk setiap faktor, dan secara keseluruhan menggunakan software JASP. Dalam analisis ini, reliabilitas diukur menggunakan koefisien *Alpha Cronbach (Cronbach's α)*. Dapat dilihat pada Tabel 4 hasil perhitungan pada Faktor 1, nilai *Cronbach's α* sebesar 0,933, menunjukkan tingkat reliabilitas yang sangat tinggi. Faktor 2 memiliki nilai *Cronbach's α* sebesar 0,889, yang juga termasuk dalam kategori reliabilitas tinggi. Selanjutnya, pada Faktor 3, nilai *Cronbach's α* sebesar 0,863, yang menunjukkan reliabilitas tinggi. Untuk Faktor 4, nilai *Cronbach's α* sebesar 0,799, yang masih masuk dalam kategori cukup baik. Namun, pada Faktor 5, nilai *Cronbach's α* adalah 0,680, yang mengindikasikan tingkat reliabilitas yang lebih rendah dibandingkan dengan faktor lainnya, tetapi tetap cukup memadai. Berdasarkan interpretasi reliabilitas menurut Azwar (2012), semakin mendekati nilai 1,00, maka reliabilitas semakin tinggi. Dengan demikian, dapat disimpulkan bahwa secara keseluruhan, *The Adult Moral Injury Scale* memiliki tingkat reliabilitas yang sangat baik, meskipun terdapat variasi dalam reliabilitas pada masing-masing faktor. Pada penghitungan secara keseluruhan dengan total 39 item didapatkan nilai reliabilitas sebesar 0.963, yang berarti dapat disimpulkan bahwa alat ukur "*The Adult Moral Injury Scale*" yang kami adaptasi memiliki hasil reliabilitas yang tinggi

Tabel 4. Reliabilitas

Faktor	Koefisien α
Faktor 1	0,933
Faktor 2	0,889
Faktor 3	0,863
Faktor 4	0,7990
Faktor 5	0,680
Total	0,963

4. Pembahasan

Analisis statistik yang dilakukan menunjukkan bahwa keseluruhan item pada skala ini memiliki korelasi item-total di atas 0,30, dengan rentang nilai antara 0,396 hingga 0,758. Hasil ini mengindikasikan bahwa semua item memenuhi syarat kelayakan psikometrik. Selanjutnya, analisis faktor menggunakan nilai *Kaiser-Meyer-Olkin (KMO)* sebesar 0,938 memperkuat kesimpulan bahwa data memenuhi persyaratan sampling adequacy untuk analisis faktor lebih lanjut. Berdasarkan analisis faktor eksploratori dan konfirmatori, seluruh item memiliki *factor loading* di atas 0,40, yang membuktikan validitas item tersebut. Reliabilitas skala diukur menggunakan koefisien *Cronbach's α* , yang mencapai nilai total sebesar 0,963. Hal ini menunjukkan bahwa alat ukur ini memiliki reliabilitas yang sangat tinggi dan konsisten dalam mengukur konstruk yang dimaksud. Setiap dimensi juga memiliki nilai reliabilitas yang baik, dengan rentang 0,680 hingga 0,933, menunjukkan bahwa setiap subskala dapat diandalkan.

Meskipun validitas dan reliabilitas alat ukur ini tinggi, analisis model fit menunjukkan bahwa model ini belum sepenuhnya sesuai dengan kriteria fit yang ditentukan. Nilai *Chi-Square* ($p < 0,001$), *GFI* (0,811), *NFI* (0,689), *NNFI* (0,760), dan *CFI* (0,774) tidak memenuhi ambang batas $>0,9$ yang diharapkan. Selain itu, nilai *RMSEA* sebesar 0,089 melebihi batas 0,05, yang menandakan bahwa model ini perlu direvisi untuk meningkatkan kesesuaian. Ketidaksiapan model fit ini menunjukkan perlunya perbaikan lebih lanjut, seperti melakukan modifikasi model atau menambahkan parameter yang lebih sesuai dengan data yang diperoleh. Faktor-faktor seperti kompleksitas item dan struktur faktor yang dihipotesiskan dapat menjadi penyebab model tidak fit, selain itu jumlah sampel dan variasi sampel juga bisa menjadi salah satu faktor penyebab model tidak fit.

Hasil yang menunjukkan model fit kurang optimal menuntut revisi lanjutan, sehingga untuk penelitian selanjutnya sangat disarankan untuk melakukan modifikasi model pada alat ukur ini untuk mencapai kesesuaian indeks fit. Selain itu, program dukungan berbasis *mindfulness* dan konseling individual dapat diintegrasikan untuk membantu mahasiswa mengatasi *moral injury* yang teridentifikasi. Implikasi ini juga memperkuat kebutuhan akan kebijakan yang mendukung kesehatan mental di lingkungan akademik.

5. Kesimpulan

Berdasarkan analisis statistik yang sudah dilakukan, hasil penelitian menunjukkan adanya kompleksitas dalam pengembangan alat ukur *The Adult Moral Injury Scale*, yaitu dalam analisis model fit masih belum sepenuhnya memenuhi kriteria indeks model fit yang optimal. Hal ini mengindikasikan bahwa alat ukur masih perlu penyesuaian kembali. Namun meskipun demikian, temuan lain menunjukkan bahwa alat ukur ini memiliki kualitas psikometrik yang cukup baik, dimana korelasi item total seluruh item $> 0,3$, muatan faktor seluruh item $> 0,4$, selain itu nilai koefisien reliabilitas alat ukur sangat memuaskan yaitu 0,963. Hal ini menunjukkan konsistensi internal yang kuat. Dengan demikian meskipun alat ukur masih membutuhkan perbaikan, *The Adult Moral Injury*

Scale memiliki potensi yang signifikan untuk digunakan dalam penelitian lain, tentunya setelah dilakukan penyempurnaan pada alat ukur. Penelitian selanjutnya disarankan untuk melakukan modifikasi model untuk mencapai kesesuaian indeks fit yang lebih optimal.

6. Referensi

- Azwar, S. (2012). *Reliabilitas dan Validitas*. Sleman. Pustaka Pelajar.
- Bryan, C. J. (2015) Measuring Moral Injury: Psychometric Properties of the Moral Injury Events Scale in Two Military Samples. *Assessment* 1–14 Sage Pub DOI: 10.1177/1073191115590855
- Byrne, B. M. (2016). *Structural Equation Modeling with Amos: Basic Concepts, Applications, and Programming* (3rd ed.). Routledge.
- Dara, R.A (2024) Kebermaknaan Hidup Mahasiswa Psikologi Yang Melakukan Self-Harm. Skripsi. Program Studi Psikologi Jurusan Psikologi Fakultas Kedokteran Universitas Malikussaleh
- Ford, Eric W. PhD. Stress, Burnout, and Moral Injury: The State of the Healthcare Workforce. *Journal of Healthcare Management* 64(3):p 125-127, May-June 2019. | DOI: 10.1097/JHM-D-19-00058
- Gillies, D., Christou, M. A., Dixon, A. C., Featherston, O. J., Rapti, I., GarciaAnguita, A., Villasis-Keever, M., Reebye, P., Christou, E., Al Kabir, N., & Christou, P. A. (2018). Prevalence and Characteristics of Self-Harm in Adolescents: Meta-Analyses of Community-Based Studies 1990–2015. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(10), 733–741
- Greene, T., Bloomfield, M. A. P., & Billings, J. (2020). Psychological trauma and moral injury in religious leaders during COVID-19. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1), S143–S145. <https://doi.org/10.1037/tra0000641>
- Haight, Sugrue, Calhoun, Black (2017) “Basically, I look at it like combat”: Reflections on moral injury by parents involved with child protection services. *Children and Youth Services Review*. Elsevier <https://doi.org/10.1016/j.childyouth.2017.10.009>
- Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010). *Multivariate Data Analysis. 7th Edition*. Pearson, New York.
- Hayurika, Turina L., & Sandy Arief. (2015). Analisis faktor-faktor yang mempengaruhi minat siswa dalam pengambilan keputusan memilih jurusan akuntansi kelas X di SMK N 1 Demak. *Jurnal Pendidikan Ekonomi Dinamika Pendidikan*, 10(1), 88-103.
- Jinkerson, J. D. (2016). Defining and assessing moral injury: A syndrome perspective. *Traumatology*, 22(2), 122–130. <https://doi.org/10.1037/trm0000069>
- Levinson, M. (2015). “Moral Injury and the Ethics of Educational Injustice.” *Harvard Educational Review* 85 (2) (June): 203–228. doi:10.17763/0017-8055.85.2.203.
- Litz, B. T., Stein, N. R., Delaney, E., & Lebowitz, L. (2009). Moral injury and mental health: The role of the military. *Journal of Traumatic Stress*, 22(5), 453–460.
- Murray, E., Krahé, C., & Goodsmann, D. (2018). Are medical students in prehospital care at risk of moral injury?. *Emergency medicine journal : EMJ*, 35(10), 590–594. <https://doi.org/10.1136/emered-2017-207216>
- Putri, E. D. (2022). Kasus Bullying di Lingkungan Sekolah : Dampak Serta Penanganannya. *Keguruan: Jurnal Penelitian, Pemikiran Dan Pengabdian*, 10, 24–30.
- Shay, J. (2014). Moral Injury. *Psychoanalytic Psychology*, 31(2), 182–191. <https://doi.org/10.1037/n0036090>
- Sunanih, S., Nurhaliza, A., Shakila, A., Ulpah, D. N., Rahmaldi, D., Farida, D. N., Maulida, I., Ashilah, M., Rahmawati, N. A., Saputra, R. F., Qurani, S. N., Utami, W., & Santi, D. (2025). Analisis Pengaruh Perilaku Bullying terhadap Prestasi Akademik Siswa di Sekolah Dasar. *Dewantara: Jurnal Pendidikan Sosial Humaniora*, 4(1), 31–45. <https://doi.org/10.30640/dewantara.v4i1.3767>
- Sudaryono. (2019). *Metodologi Penelitian*. Depok. Rajagrafindo Persada.
- Wijayanto, S. H. (2008). *Structural Equation Modeling dengan Lisrel 8.8: Konsep dan Tutorial*. Yogyakarta. Graha Ilmu.
- Wingert, R. (2023). *The Adult Moral Injury Scale: An Initial Validation Study* [George Fox University]. <https://digitalcommons.georgefox.edu/psyd/489>
- Yeterian, J. D., Berke, D. S., Carney, J. R., McIntyre-Smith, A., St Cyr, K., King, L., Kline, N. K., Phelps, A., Litz, B. T., & Members of the Moral Injury Outcomes Project Consortium (2019). Defining and Measuring Moral Injury: Rationale, Design, and Preliminary Findings From the Moral Injury Outcome Scale Consortium. *Journal of traumatic stress*, 32(3), 363–372. <https://doi.org/10.1002/jts.22380>

JPPP

Jurnal Penelitian dan Pengukuran Psikologi
Fakultas Pendidikan Psikologi
Universitas Negeri Jakarta

