

Volume 14 No. 2, Oktober 2025



# Jurnal Penelitian & Pengukuran **Psikologi**

**JPPP**

Volume  
14

Nomor  
02

Hal.  
97 - 196

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](mailto: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](mailto:santiyudhistira@unj.ac.id) atau [jppp@unj.ac.id](mailto: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

<b>Nama</b>	<b>Judul Artikel</b>	<b>Halaman</b>
<i>Dewi Fransiska Simanjuntak, Fauiziah Wijayanti, Nanda Putri Adhiningtyas, Muhammad Fathraka Yusuf, Julia Angel</i>	Pengaruh Rasa Syukur dan Dukungan Sosial terhadap <i>Psychological Well-Being</i> pada Ibu Bekerja	97 – 105
<i>Rahmat S. Bintang, Isabell Nicole Fischer</i>	<i>Body Conciousness Questionnaire</i> (BCQ) Versi Bahasa Indonesia: Pengujian Properti Psikometri Menggunakan CFA dan Rasch Model	106 – 113
<i>Qurrota Aini, Herdiyan Maulana, Mira Ariyani</i>	Efektivitas Intervensi <i>Self-Compassion Writing</i> terhadap Stres dan <i>Well-Being</i> Mahasiswa	114 – 126
<i>Rerepih Julung Merekah Wanodya Sinandang, Ratri Pratiwi</i>	Uji Psikometri Skala <i>Cyberstalking</i> Versi Indonesia: <i>Confirmatory Factor Analysis</i> dan Invariansi Gender	127 – 134
<i>Galih Adriansyah, Arinta Aisyah Putri, Galih Khalisha Qalbyna, Safira Febriani, Diah Zaleha Wyandini, Ghinaya Ummul Mukminin Hidayat</i>	Analisis Psikometri <i>Morbid Curiosity Scale</i> Versi Bahasa Indonesia	135 – 145
<i>Amr Yazid Pikoli, Miftahul Huda, Benidha Roos Feri Nisa</i>	Modifikasi Skala Regulasi Akademik Mahasiswa	146 – 158
<i>Muhammad Ardi, Nilam Widyarini</i>	Konstruksi dan Evaluasi Psikometrik Skala <i>Counterproductive Work Behavior</i>	159 – 168
<i>Rizki Nur Pratama, Muchamad Salamun Abdul Rahman, Nova Mukhlina,, Muthmainnah Imtiyaz, Rifda Amira Mulika, Ricky Budhi Hartanto, Ismi Uswatun Khasanah</i>	Pengembangan Skala Kesiapan Menghadapi Perubahan Pada Mahasiswa Baru di Indonesia dan Uji Invariansi: Pendekatan Mimic Model	169 – 177

<b>Nama</b>	<b>Judul Artikel</b>	<b>Halaman</b>
<i>Abiyu Arib Mahiyuddin</i>	<i>Pengembangan Student's Metacognitive Awareness Scale (SMAS) pada Siswa di Indonesia</i>	178 – 187
<i>Renata V. Gunjaya, Cecilia Ang, Sania Sania, P. Tommy Y. S. Suyasa</i>	<i>Contextual Verbal Creativity Test (Cvct): A New Tool For Understanding Verbal Creativity Among Students Of University X</i>	188 – 196



## PENGARUH RASA SYUKUR DAN DUKUNGAN SOSIAL TERHADAP PSYCHOLOGICAL WELL-BEING PADA IBU BEKERJA

Dewi Fransiska Simanjuntak<sup>1</sup>, Fauziah Wijayanti<sup>2</sup>, Nanda Putri Adhiningtyas<sup>3</sup>, Muhammad Fathraka Yusuf<sup>4</sup>, Julia Angel<sup>5</sup>

<sup>1,2,3,4,5</sup>, Prodi Psikologi, Fakultas Psikologi, Universitas Negeri Jakarta, Jakarta, Indonesia

Email: [dewi.fransiska@unj.ac.id](mailto:dewi.fransiska@unj.ac.id)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
18 Agustus 2025	13 Oktober 2025	14 Oktober 2025	30 Oktober 2025

### Abstract

*This study aims to investigate the contribution of gratitude and social support to the psychological well-being of working mothers in Indonesia. This study used a quantitative approach with 106 working mothers in Indonesia, age range above 20 until 60 years, participating in the research. Data were collected via online questionnaires consisted of, Gratitude Questionnaire-Six Item Form (GQ-6), Multidimensional Scale of Perceived Social Support (MSPSS), and Ryff's Psychological Well-being Scale (Short Form). Data analysis was conducted using a regression technique to test the effect between the variables. The multiple regression analysis indicated that gratitude and social support together accounted for 27.1% of the variance in psychological wellbeing ( $R^2 = 0.271$ ;  $F = 19.099$ ;  $p < 0.001$ ). This finding suggests that the model is statistically significant, although the majority of variance in psychological well-being is explained by other factors beyond the scope of this study. The findings of this study can provide practical input to working mothers on how to increase their gratitude and can inform people involved in supporting working mothers that the support they provide can enhance mothers' psychological well-being.*

*Keywords: gratitude, social support, psychological well-being, working mother*

### Abstrak

Penelitian ini bertujuan untuk menginvestigasi kontribusi rasa syukur (*gratitude*) dan dukungan sosial terhadap kesejahteraan psikologis ibu bekerja di Indonesia. Penelitian ini menggunakan pendekatan kuantitatif dengan melibatkan 106 ibu bekerja di Indonesia berusia antara 20 hingga 60 tahun. Data dikumpulkan melalui kuesioner daring yang terdiri atas *Gratitude Questionnaire-Six Item Form (GQ-6)*, *Multidimensional Scale of Perceived Social Support (MSPSS)*, dan *Ryff's Psychological Well-Being Scale (versi singkat)*. Analisis data dilakukan dengan teknik regresi untuk menguji pengaruh antar variabel. Hasil analisis regresi berganda menunjukkan bahwa rasa syukur dan dukungan sosial secara bersama-sama menjelaskan 27,1% variasi kesejahteraan psikologis ( $R^2 = 0,271$ ;  $F = 19,099$ ;  $p < 0,001$ ). Temuan ini mengindikasikan bahwa model yang digunakan signifikan secara statistik, meskipun sebagian besar variasi kesejahteraan psikologis dijelaskan oleh faktor lain di luar lingkup penelitian ini. Temuan penelitian ini dapat memberikan masukan praktis bagi ibu bekerja untuk meningkatkan rasa syukur, serta bagi pihak-pihak yang terlibat dalam memberikan dukungan kepada ibu bekerja bahwa dukungan tersebut dapat berkontribusi pada peningkatan kesejahteraan psikologis mereka.

Kata kunci: rasa syukur, dukungan sosial, *psychological well-being*, ibu bekerja, peran ganda

### 1. Pendahuluan

Banyak faktor yang menyebabkan semakin banyaknya wanita bekerja, salah satunya akses internet yang luas menunjang pertumbuhan ekonomi digital yang pesat dan membuka peluang baru bagi perempuan di Indonesia untuk berpartisipasi dalam dunia kerja (Halim, Johnson, & Perova, 2022). Data dari Badan Pusat Statistik (BPS, 2023) menunjukkan bahwa partisipasi angkatan kerja perempuan di Indonesia terus meningkat, termasuk di sektor formal yang menuntut profesionalisme dan keterlibatan penuh waktu. Perkembangan industri membuat kebutuhan akan tenaga kerja semakin meningkat termasuk tenaga kerja perempuan, dan sebagian diantaranya adalah perempuan yang juga berperan sebagai ibu, hal ini membuat ibu bekerja memiliki peran ganda yaitu sebagai pekerja dan juga ibu dalam rumah tangga (Wulandari, 2015). Peran yang dijalankan oleh ibu bekerja memiliki tuntutan dan tanggung jawab yang berbeda, untuk berperan sebagai tenaga kerja yang produktif sekaligus

pengelola rumah tangga dan pengasuh utama bagi anak (Wijayanto & Fauziah, 2020). Peran ganda didefinisikan oleh Greenhaus dan Beutell, (1985) sebagai konflik antar peran dimana beberapa pekerjaan dan tanggung jawab dalam keluarga tidak memiliki kinerja yang sesuai. Dengan beban dan tanggungjawab yang ada dalam menjalankan peran ganda, ibu bekerja mudah mengalami stress (Wahyudi & Hadi, 2017). Dalam menjalankan peran ganda ini, ibu bekerja perlu menjaga kesehatan mental yang baik, karena kondisi fisik dan mental merupakan dua hal yang saling berkaitan. Ketika kondisi mental seseorang terganggu, masalah fisik dapat muncul, dan begitu juga sebaliknya (Dewi, 2012). Sementara peran ganda yang dijalankan ibu bekerja dapat memengaruhi *psychological well-being* atau kesejahteraan psikologis. Fava (2012) menyatakan bahwa *psychological well-being* yang rendah merupakan faktor yang dapat menyebabkan depresi, maka menjaga *psychological well-being* merupakan hal yang sangat penting untuk ibu bekerja agar dapat menyeimbangkan perannya sebagai ibu dan juga pekerja profesional.

Ryff (1989) mendefinisikan *psychological well-being* sebagai kondisi optimal yang mencerminkan aktualisasi potensi diri secara penuh. Individu yang memiliki kesejahteraan psikologis ditandai dengan kemampuan untuk menerima kelebihan dan kekurangannya, memiliki tujuan hidup yang jelas, menjalin hubungan positif dengan orang lain, bersikap mandiri, serta terus mengalami pertumbuhan pribadi. Terdapat 6 dimensi dalam *psychological well-being* diantaranya: pertama *Self-Acceptance* (Penerimaan Diri) yaitu sejauh mana individu mampu bersikap positif terhadap dirinya sendiri, penerimaan diri yang baik akan membuat individu cenderung memahami dan menghargai perasaan yang dimilikinya. Kedua, *Positive Relations with Others* (Hubungan Positif dengan Orang Lain), yaitu kemampuan yang dimiliki individu untuk membangun dan menjaga hubungan interpersonal yang hangat, penuh empati, dan bermakna. Dimensi ketiga, *Personal Growth* (Pertumbuhan Pribadi) yang menggambarkan kapasitas individu untuk terus berkembang dan merealisasikan potensi dirinya. Dimensi keempat adalah, *Purpose in Life* (Tujuan Hidup) yang menggambarkan bagaimana individu memiliki arah dan makna di dalam hidupnya. Dimensi kelima adalah, *Environmental Mastery* (Penguasaan Lingkungan) yaitu kemampuan individu dalam menciptakan dan mengelola lingkungan hidupnya secara efektif. Dimensi keenam adalah *Autonomy* (Otonomi) yang menggambarkan kemampuan individu untuk membuat keputusan secara mandiri, tidak mudah terpengaruh oleh tekanan sosial, serta mampu menentukan tujuan hidupnya sendiri tanpa membutuhkan persetujuan dari orang lain.

*Psychological well-being* pada ibu bekerja merupakan topik yang sering diteliti. Terdapat beberapa penelitian terdahulu yang membahas tentang hubungan antara *psychological well-being* ibu bekerja. Meriko & Hadiwirawan (2019) menemukan bahwa ibu bekerja memperoleh *psychological well-being* dalam menjalankan peran ganda mereka. Anggarwati & Thamrin (2019) menemukan bahwa konflik dalam rumah tangga memiliki hubungan negatif signifikan dengan *psychological well-being* pada ibu bekerja, semakin tinggi konflik antara pekerjaan dan keluarga, maka *psychological well-being* ibu bekerja akan menurun. Penelitian lain yang dilakukan oleh Rabbani & Yuniardi (2024) menggali hubungan antara dukungan sosial dengan *psychological well-being* ibu bekerja yang dimediasi oleh *work family conflict*, hasil dari penelitian ini menunjukkan jika dukungan sosial meningkat maka *psychological well-being* pada ibu bekerja juga akan meningkat. Penelitian yang dilakukan oleh Kirwin & Ettinger (2022) menemukan bahwa perubahan situasi eksternal seperti pandemi Covid-19 dapat menurunkan *psychological well-being* ibu bekerja. Status sosial ekonomi (SES) berhubungan dengan *psychological well-being*, individu dengan tingkat SES yang lebih tinggi, cenderung memiliki akses sumber daya dan kecenderungan yang lebih besar untuk memperoleh *psychological well-being* (Navarro-Carrillo et al., 2020).

Penelitian yang sudah disebutkan sebelumnya, menunjukkan bahwa *psychological well-being* dapat dipengaruhi oleh faktor-faktor eksternal. Namun, terdapat beberapa faktor internal yang juga berpengaruh terhadap tingkat *psychological well-being*. Usia, merupakan faktor internal yang berpengaruh terhadap dimensi *psychological well-being*, seperti dimensi *personal growth* dan *purpose in life* yang cenderung lebih tinggi pada usia dewasa muda, sementara dimensi lainnya seperti *autonomy* dan *enviromental mastery* meningkat seiring bertambahnya usia (Suberry et al., 2024). Jenis kelamin juga dapat berpengaruh terhadap *psychological well-being*, dimensi *autonomy* lebih tinggi pada laki-laki, sementara dimensi *positive relations* lebih tinggi pada perempuan (Wan et al., 2025).

Faktor-faktor yang sudah disebutkan menunjukkan bahwa *psychological well-being* ini dipengaruhi oleh faktor eksternal dan juga dan faktor internal. Namun, penelitian tentang faktor internal yang mungkin berpengaruh terhadap *psychological well-being* seperti rasa syukur masih jarang diteliti, sehingga penelitian untuk melihat hubungan pengaruh syukur dan dukungan sosial terhadap *psychological well-being* ibu bekerja, perlu dilakukan. Dalam konteks ibu yang bekerja, faktor-faktor yang berpengaruh terhadap *psychological well-being* tersebut menjadi semakin kompleks karena adanya peran ganda yang harus dijalani, yakni sebagai pekerja profesional dan sekaligus sebagai pengasuh serta pengelola rumah tangga. Rasa syukur (*gratitude*) didefinisikan sebagai kecenderungan individu untuk mengenali dan menghargai kebaikan yang diterima dari orang lain maupun dari kehidupan secara umum (Emmons & McCullough, 2003). Rasa syukur dapat membuat ibu bekerja terbantu dalam melihat sisi positif dari situasi yang sedang dihadapinya, hal ini dapat membantu ibu bekerja untuk mengelola tekanan dalam pekerjaan dan tuntutan rumah tangga dengan lebih baik, karena rasa syukur dapat mengurangi perasaan cemas (Lovell & Wetherell, 2023).

Penelitian menunjukkan bahwa individu dengan tingkat rasa syukur yang tinggi cenderung memiliki pandangan hidup yang lebih positif, lebih resilien terhadap stres, dan mampu mempertahankan emosi positif dalam jangka panjang (Wood et al., 2010; Lovell & Wetherell, 2023). Intervensi berbasis psikologi positif, seperti journaling rasa syukur, juga terbukti meningkatkan kesejahteraan mental dan mengurangi stres serta kecemasan pada ibu (AlHugail, 2025).

Dukungan sosial juga merupakan faktor eksternal yang berpengaruh terhadap *psychological well-being* ibu bekerja. Menurut Zimet et al., (1988) dukungan sosial bukan berarti sekedar ada orang lain di sekitar individu, tetapi lebih dari itu, yaitu persepsi individu terhadap ketersediaan bantuan yang berarti untuk individu. Dukungan sosial dibedakan ke dalam tiga bentuk, antara lain: Pertama, dukungan emosional yaitu perhatian, empati, dan kasih sayang yang diberikan oleh orang lain. Kedua, dukungan instrumental yaitu bantuan nyata dalam menyelesaikan tugas atau menghadapi masalah. Ketiga, dukungan informasional, yaitu saran, arahan, atau informasi yang membantu individu mengambil keputusan atau menghadapi situasi sulit. Dukungan sosial berfungsi sebagai pelindung terhadap stres (Zimet et al., 1998). Beberapa penelitian sebelumnya menunjukkan bahwa terdapat korelasi positif dan signifikan antara dukungan sosial dan *psychological well-being* pada ibu bekerja (Arfianto et al., 2020), dukungan sosial dari orang lain berkorelasi positif dengan adaptabilitas karir (Giffari & Suhariadi, 2017; Atqakum et al., 2022). Kemudian penelitian yang dilakukan Sulaeman dan Kustimah (2024), dukungan sosial memiliki hubungan positif yang kuat dengan *psychological well-being* pada ibu bekerja di Indonesia.

Hubungan antara rasa syukur, dukungan sosial, dan *psychological wellbeing* menjadi penting untuk diteliti lebih lanjut pada konteks ibu bekerja di Indonesia. Meskipun studi di negara lain telah menemukan kaitan positif antara rasa syukur dan *psychological wellbeing* (Emmons & McCullough, 2003; Wood et al., 2010), serta antara dukungan sosial dan *psychological wellbeing* (Taylor, 2011; Cohen & Wills, 1985), penelitian dengan fokus pada ibu bekerja di Indonesia masih relatif terbatas. Mengingat adanya perbedaan konteks budaya, struktur sosial, dan peran keluarga di Indonesia, pemahaman yang lebih mendalam tentang hubungan ketiga variabel ini dapat memberikan kontribusi pada pengembangan strategi intervensi dan kebijakan yang mendukung *psychological well-being* ibu bekerja. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menguji pengaruh rasa syukur dan dukungan sosial terhadap *psychological well-being* pada ibu bekerja.

## 2. Metode Penelitian

### Partisipan

Partisipan dalam penelitian ini adalah 106 ibu bekerja yang berdomisili diberbagai wilayah di Indonesia. Data diperoleh dengan menggunakan teknik *accidental sampling* dengan rentang usia 20 tahun sampai dengan 50 tahun. Sebagian besar partisipan berada pada kelompok usia 31–35 tahun (45,3%), diikuti oleh kelompok usia 36–40 tahun (28,3%). Seluruh partisipan mengikuti penelitian secara sukarela setelah mendapatkan penjelasan mengenai tujuan, prosedur, dan hak mereka sebagai responden. Usia pernikahan di bawah 5 tahun (16%), sebagian besar partisipan ada dalam rentang usia pernikahan 5 sampai 10 tahun (54,7%), usia pernikahan dalam rentang 11–15 tahun (21,7%), usia pernikahan dalam rentang 16–20 tahun (2,8%), dan terdapat (4,7%) usia pernikahan dalam rentang di atas 20 tahun. Lama bekerja partisipan pun berbeda-beda, terdapat (12,3%) partisipan yang bekerja di bawah 5 tahun, sementara partisipan yang bekerja dalam rentang 5–10 tahun sejumlah (40,6%), terdapat (36,8%) partisipan yang bekerja dalam rentang 11–15 tahun. Sementara (6,6%) partisipan sudah bekerja dalam rentang 16–20 tahun, dan (3,8%) partisipan bekerja di atas 20 tahun. Partisipan juga merupakan ibu bekerja yang memiliki 1 anak (40,6%), sementara (45,3%) memiliki 2 anak, sementara ibu bekerja yang memiliki 3 anak sejumlah (13,2%), dan Ibu yang memiliki 5 anak (0,9). Sebelum mengisi kuesioner, partisipan memberikan *informed consent* sebelum melanjutkan pengisian kuesioner secara *online*.

### Desain Penelitian

Penelitian ini menggunakan desain kuantitatif dengan pendekatan korelasional. Tujuan dari desain ini adalah untuk mengetahui pengaruh antara rasa syukur dan dukungan sosial, terhadap *psychological well-being* pada ibu bekerja. Data dikumpulkan melalui penyebaran kuesioner secara daring menggunakan *Google Form*, yang berisi instrumen *The Gratitude Questionnaire-Six Item Form* (GQ-6) dari McCullough et al. (2002), *Multidimensional Scale of Perceived Social Support* (MSPSS) dari Canty-Mitchell & Zimet (2000), dan *Ryff's Psychological Well-Being Scale* versi *short scale* (Ryff & Keyes, 1995).

### Prosedur

Penelitian ini diawali dengan pengajuan proposal penelitian kepada Lembaga Penelitian dan Pengabdian Masyarakat Universitas Negeri Jakarta (LPPM UNJ), yang disertai pengurusan izin kepada komite etik dengan mengisi *form* penilaian mandiri klirens etik (*self assessment ethical clearance*). Setelah mengisi *form* klirens etik, peneliti menyusun kuesioner daring melalui *google form* yang terdiri atas lembar persetujuan partisipasi (*informed consent*), pertanyaan demografis seperti rentang usia partisipan, usia pernikahan, lama bekerja, dan jumlah anak, kemudian partisipan diminta mengisi tiga instrumen utama penelitian. Kuesioner disebarluaskan melalui media sosial (*WhatsApp*, *Instagram*, dan *Facebook*) peneliti melakukan pengambilan data selama tiga minggu. Partisipan yang

bersedia berpartisipasi mengisi kuesioner secara mandiri. Peneliti memastikan kerahasiaan data responden dan menjelaskan bahwa data hanya digunakan untuk kepentingan penelitian.

### **Instrumen**

Penelitian ini menggunakan tiga instrumen psikologis yang telah teruji validitas dan reliabilitasnya untuk mengukur variabel rasa syukur, dukungan sosial, dan *psychological well-being* pada ibu bekerja.

#### **1. *Gratitude Questionnaire-Six Item Form (GQ-6)***

Rasa syukur diukur menggunakan *Gratitude Questionnaire-Six Item Form (GQ-6)* yang dikembangkan oleh McCullough, Emmons, dan Tsang (2002). Instrumen ini terdiri dari enam item yang mengukur kecenderungan individu untuk merasakan dan mengekspresikan rasa syukur dalam kehidupan sehari-hari. Respon diberikan menggunakan skala Likert 1–7, dimana “1” = sangat tidak setuju dan “7” = sangat setuju, di mana skor yang lebih tinggi menunjukkan tingkat rasa syukur yang lebih tinggi. GQ-6 telah banyak digunakan dalam penelitian lintas budaya dan menunjukkan konsistensi internal yang baik ( $\alpha = 0,82$ ) (McCullough et al., 2002). Penelitian di Indonesia yang dilakukan oleh Moningga & Soewastika (2023) menggunakan alat ukur ini memperoleh konsistensi internal sebesar ( $\alpha = 0,73$ ). Dalam penelitian ini, uji reliabilitas skala dengan enam item menunjukkan nilai Cronbach’s Alpha sebesar 0,654 dan meningkat menjadi 0,773 apabila item G6 dihapus. Hal ini mengindikasikan konsistensi internal skala berada pada kategori cukup hingga baik.

#### **2. *Multidimensional Scale of Perceived Social Support (MSPSS)***

Untuk menilai persepsi dukungan sosial atau *social support* dari orang-orang terdekat, yaitu keluarga, teman, dan orang-orang penting, Canty-Mitchell dan Zimet (2000) mengembangkan *Multidimensional Scale of Perceived Social Support (MSPSS)*. Alat ukur ini telah digunakan dalam studi yang melibatkan siswa sekolah menengah pertama (SMP) dan sekolah menengah atas (SMA) di Turki dan Indonesia (Turan et al., 2014; Laksmi et al., 2020), serta mahasiswa sarjana (S1) di Hongkong, Amerika Serikat bagian Midwest, dan Indonesia (Hui et al., 2018; Ghosh & Fouad, 2017; Salsabhilla & Panjaitan, 2019). Model MSPSS yang telah diterjemahkan dalam bentuk bahasa Indonesia memiliki sifat psikometrik yang baik untuk mengukur dukungan sosial, sehingga skala MSPSS Indonesia dapat direkomendasikan untuk mengukur *social support* (Sulistiani et al., 2022). *Multidimensional Scale of Perceived Social Support (MSPSS)* awalnya dikembangkan oleh Zimet et al. (1988) dan kemudian divalidasi ulang pada berbagai populasi, termasuk remaja oleh Canty-Mitchell dan Zimet (2000). Di Indonesia, versi adaptasi MSPSS telah dikembangkan dan divalidasi oleh Sulistiani, Fajrianti, dan Kristiana (2022). Penelitian tersebut menggunakan pendekatan model Rasch untuk menilai validitas dan reliabilitas instrumen, sehingga menghasilkan bukti bahwa versi Bahasa Indonesia dari MSPSS layak digunakan untuk mengukur persepsi dukungan sosial pada konteks budaya Indonesia. Instrumen dukungan sosial terdiri dari 12 item yang terbagi ke dalam tiga dimensi sumber dukungan sosial: keluarga, teman, dan orang signifikan lainnya. Setiap item diukur menggunakan skala Likert 1–4, dimana “1” = sangat tidak setuju dan “4” = sangat setuju. Skor total yang lebih tinggi menunjukkan persepsi dukungan sosial yang lebih tinggi. MSPSS memiliki reliabilitas yang tinggi ( $\alpha > 0,85$ ) dan telah teruji pada berbagai kelompok demografis, termasuk populasi pekerja perempuan.

#### **3. *Ryff’s Psychological Well-Being Scale (Short Form)***

*Psychological well-being* diukur menggunakan *Ryff’s Psychological Well-Being Scale* versi singkat yang dikembangkan oleh Ryff (1989) dan disesuaikan menjadi versi *short form* untuk efisiensi pengukuran. Versi ini terdiri dari 18 item yang mencakup enam dimensi kesejahteraan psikologis: penerimaan diri, hubungan positif dengan orang lain, otonomi, penguasaan lingkungan, tujuan hidup, dan pertumbuhan pribadi. Respon diukur dengan skala Likert 1–4, “1” = sangat tidak setuju dan “4” = sangat setuju. Skor yang lebih tinggi mencerminkan tingkat *psychological well-being* yang lebih baik. Skala ini memiliki reliabilitas yang baik ( $\alpha$  berkisar antara 0,70–0,84 pada tiap dimensi) dan telah banyak digunakan dalam penelitian di berbagai konteks pekerjaan. *Ryff’s Psychological Well-Being Scale (RPWB)* versi pendek dengan 18 item telah diadaptasi ke Bahasa Indonesia dan digunakan dalam berbagai penelitian. Humaidah dan Mulyono (2022) menguji validitas konstruk RPWB-18 pada mahasiswa Indonesia melalui analisis faktor konfirmatori (CFA) dan menemukan bahwa model enam dimensi menunjukkan goodness of fit yang memadai, dengan semua item signifikan.

### **Analisis Statistik**

Teknik regresi ganda dilakukan untuk membuktikan hipotesis mengenai pengaruh *gratitude* dan dukungan sosial terhadap *psychological well-being* pada ibu bekerja. Data yang diperoleh menunjukkan penyebaran residual yang normal serta tidak memenuhi kriteria multikolinearitas. Dengan demikian, uji regresi ganda dapat dilakukan. Pengujian dilakukan menggunakan program *SPSS* versi 23.

### 3. Hasil dan Diskusi

#### Data Demografis

Tabel 1 merangkum data demografis partisipan, beberapa data demografis yang diminta dalam penelitian ini, antara lain rentang usia partisipan, usia pernikahan, jumlah anak dan lama bekerja. Partisipan dalam penelitian ini berjenis kelamin perempuan (100%), berusia 20 sampai 50 tahun (100%), memiliki jumlah anak 1 (43%), 2 anak (48%), lama bekerja direntang 5-10 tahun (43%).

**Tabel 1. Hasil Analisis Data Demografis**

Kategori	<i>n</i>	%
<b>Jenis Kelamin</b>		
Perempuan	106	100%
<b>Usia</b>		
20 -25	1	0,9%
26-30	13	12,3%
31-35	48	45,3%
36-40	30	28,3%
41-45	9	8,5%
46-50	4	3,8%
>50	1	1.2%
<b>Usia Pernikahan</b>		
< 5 tahun	17	16%
5-10 tahun	58	54,7%
11-15 tahun	23	21,7%
16-20 tahun	3	2,8%
>20 tahun	5	4,7%
<b>Jumlah Anak</b>		
1	43	40.6%
2	48	45,3%
3	14	13,2%
5	1	0,9%
<b>Lama Bekerja</b>		
< 5	13	12.3%
5 - 10	43	40.6%
11-15	39	36.8%
16-20	7	6.6%
> 20	4	3.8%

Data demografis partisipan dalam penelitian ini meliputi usia, usia pernikahan, jumlah anak dan lama bekerja, dikumpulkan untuk mendeskripsikan karakteristik sampel penelitian. Analisis utama difokuskan pada hubungan antarvariabel penelitian, sehingga data demografis tidak digunakan sebagai variabel perbandingan dalam analisis.

#### Uji Hipotesis

Sebelum melakukan analisis regresi, terlebih dahulu dilakukan analisis korelasi Pearson untuk menguji hubungan antara variabel-variabel penelitian. Hasil analisis korelasi disajikan pada Tabel 2.

**Tabel 2. Hasil Analisis Korelasi Pearson Antar Variabel**

Variabel	1	2	3
1. <i>Psychological Well-Being</i>	—		
2. Rasa Syukur	.508**	—	
3. Dukungan Sosial	.389**	.590**	—

Catatan. N = 106

\*\*p < .01

Berdasarkan Tabel 2, dapat disimpulkan bahwa rasa syukur memiliki hubungan positif yang signifikan dengan *psychological well-being* ( $r = .508, p < .01$ ). Dukungan sosial juga berhubungan positif secara signifikan dengan

*psychological well-being* ( $r = .389, p < .01$ ). Selain itu, terdapat hubungan positif yang signifikan antara rasa syukur dan dukungan sosial ( $r = .590, p < .01$ ). Analisis regresi linier berganda dilakukan untuk menguji pengaruh rasa syukur dan dukungan sosial terhadap *psychological well-being* pada ibu bekerja. Hasil analisis regresi disajikan dalam Tabel 2.

**Tabel 3. Hasil Analisis Regresi**

Prediktor	B	SE	$\beta$	t	p
Konstanta	32.484	3.642		8.918	< .001
Rasa Syukur	0.370	0.090	.428	4.103	< .001
Dukungan Sosial	0.161	0.122	.137	1.314	.192

Catatan.  $R^2 = .271$ ;  $F(2, 103) = 19.099, p < .001$

Berdasarkan Tabel 3, model regresi secara keseluruhan signifikan ( $F(2, 103) = 19.099, p < .001$ ) dengan nilai  $R^2$  sebesar .271. Hal ini menunjukkan bahwa kedua variabel prediktor secara bersama-sama menjelaskan 27.1% varians dalam *psychological well-being*, sementara 72.9% dijelaskan oleh faktor-faktor lain di luar variabel penelitian tersebut. Rasa syukur terbukti berpengaruh positif dan signifikan terhadap *psychological well-being* ( $\beta = .428, p < .001$ ). Hal ini menunjukkan bahwa semakin tinggi tingkat rasa syukur yang dimiliki ibu bekerja, maka semakin tinggi pula tingkat *psychological well-being* yang dimilikinya. Sebaliknya, variabel dukungan sosial tidak berpengaruh signifikan terhadap *psychological well-being* ( $\beta = .137, p = .192$ ) ketika dinilai bersama-sama dengan rasa syukur dalam model regresi.

Temuan ini mengindikasikan bahwa meskipun secara terpisah kedua variabel berkorelasi signifikan dengan *psychological well-being*, namun ketika dianalisis secara bersama-sama dalam model regresi, hanya rasa syukur yang memberikan pengaruh yang signifikan. Tingginya korelasi antara rasa syukur dan dukungan sosial ( $r = .590$ ) menunjukkan adanya multikolinearitas, dimana variabel rasa syukur tampaknya "menyerap" sebagian besar varians yang dapat dijelaskan oleh dukungan sosial.

### Pembahasan

Hasil penelitian ini menunjukkan bahwa rasa syukur dan dukungan sosial berkontribusi secara signifikan terhadap *psychological well-being* pada ibu bekerja. Temuan ini sejalan dengan berbagai studi sebelumnya yang menemukan bahwa rasa syukur merupakan salah satu prediktor positif *psychological well-being* Rabbani & Yuniardi (2024). Menurut Emmons dan McCullough (2003), rasa syukur memfasilitasi individu untuk fokus pada aspek positif dalam hidup, sehingga mengurangi emosi negatif seperti stres dan kecemasan yang sering dialami oleh ibu bekerja. Dalam konteks ini, rasa syukur membantu ibu bekerja untuk memaknai tantangan pekerjaan dan peran ganda secara lebih positif, sehingga meningkatkan dimensi-dimensi *psychological well-being* seperti *purpose in life* dan *self-acceptance* (Ryff, 1989).

Temuan penelitian ini menunjukkan bahwa rasa syukur merupakan prediktor signifikan terhadap *psychological well-being* pada ibu bekerja. Hasil ini selaras dengan teori Emmons & McCullough (2003) yang mendefinisikan rasa syukur sebagai penggerak kesejahteraan psikologis yang membantu individu menyadari dan menghargai kebaikan dalam hidupnya. Pada konteks ibu bekerja yang menghadapi tuntutan peran ganda (Greenhaus & Beutell, 1985; Wulandari, 2015), rasa syukur berfungsi sebagai mekanisme koping yang adaptif.

Analisis berdasarkan dimensi-dimensi *psychological well-being* Ryff (1989) memberikan gambaran yang lebih mendalam. Pertama, rasa syukur secara langsung memperkuat *self-acceptance* (penerimaan diri), dimana ibu bekerja dapat lebih menerima keterbatasan dan kelebihan mereka di tengah tuntutan peran ganda, mengurangi kecenderungan untuk merasa gagal (Lovell & Wetherell, 2023). Kedua, rasa syukur mendorong *personal growth* (pertumbuhan pribadi) dengan membantu ibu bekerja memandang tantangan sebagai peluang untuk berkembang, bukan sekadar beban. Ketiga, rasa syukur memperkuat *purpose in life* (tujuan hidup) dengan membantu mereka menemukan makna yang lebih dalam dari aktivitas sehari-hari, baik di rumah maupun di tempat kerja (Wood et al., 2010). Temuan ini juga didukung oleh penelitian terbaru AlHugail (2025) yang membuktikan efektivitas intervensi *gratitude journaling* dalam meningkatkan kesejahteraan mental pada ibu.

Meskipun analisis korelasi menunjukkan hubungan positif antara dukungan sosial dan *psychological well-being*, sejalan dengan temuan Sulaeman & Kustimah (2024) dan teori Zimet et al. (1988) tentang fungsi dukungan sosial sebagai *buffer* stres, analisis regresi parsial mengungkapkan bahwa dukungan sosial tidak berpengaruh signifikan secara langsung. Temuan yang tampaknya bertolak belakang ini dapat dijelaskan melalui mekanisme mediasi. Dalam konteks ibu bekerja di Indonesia, dukungan sosial, baik secara emosional, instrumental, maupun informasional, ternyata hanya efektif meningkatkan *positive relations with others* dan *environmental mastery* jika diinternalisasi melalui rasa syukur. Artinya, kehadiran dukungan dari pasangan, keluarga, atau rekan kerja (seperti yang diteliti oleh Arfianto et al., 2020) saja tidak cukup. Seorang ibu bekerja perlu memiliki kapasitas rasa syukur untuk memaknai bantuan tersebut bukan sebagai beban balasan atau bentuk ketidakmampuannya, melainkan

sebagai wujud kepedulian yang tulus. Dengan demikian, rasa syukur memungkinkan dukungan eksternal memperkuat *autonomy* (kemandirian) dan mengurangi konflik peran (Wijayanto & Fauziah, 2020).

Temuan ini memberikan kontribusi penting dengan menunjukkan bahwa dalam konteks budaya kolektivistik Indonesia, faktor internal (rasa syukur) justru lebih menentukan dalam membangun ketahanan psikologis daripada faktor eksternal (dukungan sosial). Implikasi praktisnya, program intervensi untuk meningkatkan psychological well-being ibu bekerja—seperti yang disarankan oleh AlHugail (2025)—harus menitikberatkan pada penguatan fondasi internal melalui pelatihan rasa syukur (*gratitude training*), baru kemudian dikombinasikan dengan penguatan jejaring dukungan sosial di tempat kerja dan keluarga. Penelitian ini memiliki beberapa keterbatasan, termasuk sampel yang terbatas secara geografis, penggunaan *self-report* yang berpotensi bias, dan desain *cross-sectional* yang tidak dapat menetapkan hubungan kausal. Untuk penelitian selanjutnya, disarankan untuk menggunakan desain longitudinal guna menelusuri dinamika hubungan antara variabel dari waktu ke waktu, serta mengeksplorasi variabel mediator lain seperti *coping stress* atau variabel moderator seperti tipe pekerjaan dan dukungan kebijakan perusahaan (Kirwin & Ettinger, 2022). Pendekatan *mixed-methods* juga akan sangat berguna untuk menggali pengalaman subjektif ibu bekerja dalam memaknai rasa syukur dan dukungan sosial secara lebih mendalam.

#### 4. Kesimpulan

Berdasarkan hasil penelitian, disimpulkan bahwa rasa syukur dan dukungan sosial secara bersama-sama memberikan kontribusi signifikan terhadap psychological well-being pada ibu bekerja. Temuan ini mengindikasikan bahwa ibu bekerja yang memiliki tingkat rasa syukur tinggi dan memperoleh dukungan sosial yang memadai cenderung mengalami kondisi psikologis yang lebih positif. Analisis statistik lebih lanjut mengungkapkan temuan penting bahwa secara parsial, kontribusi dukungan sosial tidak signifikan dalam memengaruhi *psychological well-being*. Hal ini menunjukkan bahwa dukungan sosial—baik dari pasangan, keluarga, maupun lingkungan kerja—tidak berpengaruh secara langsung, melainkan memerlukan faktor mediator untuk dapat berkontribusi secara optimal. Dukungan sosial akan efektif meningkatkan *psychological well-being* apabila ibu bekerja telah memiliki kapasitas rasa syukur yang memadai untuk memaknai dukungan tersebut secara positif. Dengan demikian, faktor internal berupa rasa syukur memegang peran yang lebih sentral dibandingkan faktor eksternal seperti dukungan sosial. Implikasi praktis dari penelitian ini menekankan pentingnya pengembangan program intervensi yang tidak hanya berfokus pada penguatan jejaring dukungan sosial, tetapi lebih menitikberatkan pada pengembangan rasa syukur sebagai fondasi utama. Bagi penelitian selanjutnya, disarankan untuk mengeksplorasi faktor-faktor lain yang mungkin berinteraksi dengan temuan ini, seperti beban kerja, gaya pengasuhan, maupun kebijakan perusahaan yang dapat memengaruhi *psychological well-being* ibu bekerja.

#### 5. Referensi

- AlHugail, M. A. (2025). Effects of a Positive Psychology Intervention on Mental Health and Well-Being Among Mothers in Riyadh, Saudi Arabia. *Healthcare*, 13(15), Article 1925. <https://doi.org/10.3390/healthcare13151925>
- Anggarwati, P. I., & Thamrin, W. P. (2019). *Work Family-Conflict dan Psychological Well-Being pada Ibu Bekerja*. *Jurnal Psikologi*, 12(2), 200-212. <https://doi.org/10.35760/psi.2019.v12i2.2444>
- Arfianto, M. A., Mustikasari, Wardani, I. Y. (2020). Apakah dukungan sosial berhubungan dengan kesejahteraan psikologis ibu pekerja?. *Jurnal Keperawatan Jiwa*, 8(4), 505-514.
- Atqakum, L., Daud, M., & Nurdin, M. N. H. (2022). *Pengaruh dukungan sosial terhadap adaptabilitas karier mahasiswa tingkat akhir*. *PESHUM: Jurnal Pendidikan, Sosial dan Humaniora*, 1(6), 576
- Badan Pusat Statistik. (2023). Keadaan ketenagakerjaan Indonesia Agustus 2023. BPS RI. <https://www.bps.go.id>.
- Canty-Mitchell, J., & Zimet, G. D. (2000). Psychometric properties of the Multidimensional Scale of Perceived Social Support in urban adolescents. *American Journal of Community Psychology*, 28(3), 391–400. <https://doi.org/10.1023/A:1005109522457>.
- Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98(2), 310–357. <https://doi.org/10.1037/0033-2909.98.2.310>
- Dewi, K. S. (2012). *Buku ajar kesehatan mental*. Semarang: Lembaga Pengembangan dan Penjaminan Mutu Pendidikan, Universitas Diponegoro.
- Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: An experimental investigation of gratitude and subjective well-being in daily life. *Journal of Personality and Social Psychology*, 84(2), 377–389. <https://doi.org/10.1037/0022-3514.84.2.377>
- Fava, G. A. (2012). The clinical role of psychological well-being. *World Psychiatry* 11(2), 102-103. <https://doi.org/10.1016/j.wpsyc.2012.05.018>

- Kirwin, M. A., & Ettinger, A. K. (2022). Working mothers during COVID-19: A cross-sectional study on mental health status and associations with the receipt of employment benefits. *BMC Public Health*, 22(1), Article 435. <https://doi.org/10.1186/s12889-021-12468-z>
- Ghosh, A., & Fouad, N. A. (2017). *Social support and career adaptability among Midwestern college students*. *Journal of Career Assessment*, 25(4), 705–719. <https://doi.org/10.1177/1069072716679927>
- Greenhaus, J. H., & Beutell, N. J. (1985). Sources of conflict between work and family roles. *Academy of Management Review*, 10(1), 76–88. <https://doi.org/10.5465/amr.1985.4277352>
- Halim, D., Johnson, H., & Perova, E. (2022). Heterogeneous impact of internet availability on female labor market outcomes in Indonesia. *World Development*, 157, 105941. <https://doi.org/10.1016/j.worlddev.2022.105941>
- Hui, E. K. P., Wong, E., & Tsang, K. W. (2018). *Career adaptability and social support among graduating college seniors in Hong Kong*. *Journal of Vocational Behavior*, 108, 124–133. <https://doi.org/10.1016/j.jvb.2018.07.003>
- Humaidah, A., & Mulyono, R. (2025). Adaptasi *Psychological Well-Being Short Scale* pada mahasiswa. *Edukatif: Jurnal Ilmu Pendidikan*, 7(1), 141–149. <https://edukatif.org/index.php/edukatif/index>
- Laksmi, D., Wahyuni, S., & Nugraha, D. (2020). *Validation of the Indonesian version of the Multidimensional Scale of Perceived Social Support (MSPSS) among adolescent disaster victims*. *PLoS ONE*, 15(12), e0229958. <https://doi.org/10.1371/journal.pone.0229958>
- Lovell, B., & Wetherell, M. A. (2023). Social support mediates the relationship between dispositional gratitude and psychological distress in caregivers of autistic children. *Psychology, Health and Medicine*, 28(8), 2084–2094. <https://doi.org/10.1080/13548506.2022.2162939>
- Meriko, C., & Hadiwirawan, O. (2019). Kesejahteraan psikologis perempuan yang berperan ganda. *Seurune: Jurnal Psikologi UNSYIAH*, 2(1), 45–56. <https://doi.org/10.24815/j-p.v2i1.13273>
- McCullough, M. E., Emmons, R. A., & Tsang, J.-A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, 82(1), 112–127. <https://doi.org/10.1037/0022-3514.82.1.112>
- Salsabihilla, A., & Panjaitan, R. U. (2019). DUKUNGAN SOSIAL DAN HUBUNGANNYA DENGAN IDE BUNUH DIRI PADA MAHASISWA RANTAU. *Jurnal Keperawatan Jiwa (JKJ) : Persatuan Perawat Nasional Indonesia*, 7(1), 107-114. <https://doi.org/10.26714/jkj.7.1.2019.107-114>
- Sulaeman, M.R., & Kustimah., K. (2024). Hubungan *Perceived Social Support* dengan *psychological well-being* pada ibu bekerja. *Jurnal Psikologi Sains dan Profesi*, 8 (1), 53-65. <https://doi.org/10.24198/jpsp.v8i1.53033>
- Sulistiani, W., Fajrianti, F., & Kristiana, I. F. (2022). Validation of the Indonesian Version of the Multidimensional Scale of Perceived Social Support (MSPSS): A Rasch Model Approach. *Jurnal Psikologi*, 21(1), 89-103. <https://doi.org/10.14710/jp.21.1.89-103>
- Navarro-Carrillo, G., Torres-Marín, J., Carretero-Dios, H., & Valor-Segura, I. (2020). Dispositional optimism and psychological well-being: The mediating role of perceived social support. *Journal of Happiness Studies*, 21(3), 1165–1183. <https://doi.org/10.1007/s10902-019-00127-9>
- Rabbani, D., & Yuniardi, M. S. (2024). The role of work-family conflict in mediating the influence of social support on the psychological well-being of career women. *Jurnal Ilmiah Psikologi Terapan*, 12(2), 114-120. <https://doi.org/10.22219/jipt.v12i2.33759>
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57(6), 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>
- Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69(4), 719–727. <https://doi.org/10.1037/0022-3514.69.4.719>
- Taylor, S. E. (2011). Social support: A review. In M. S. Friedman (Ed.), *The handbook of health psychology* (pp. 189–214). Oxford University Press.
- Turan, N., Gencoz, T., & Zimet, G. D. (2014). *The psychometric properties of the Turkish version of the Multidimensional Scale of Perceived Social Support (MSPSS) in adolescents*. *Journal of Child and Family Studies*, 23(3), 437–444. <https://doi.org/10.1007/s10826-013-9721-2>
- Wahyudi, A., & Hadi, C. (2017). Hubungan antara konflik peran ganda dengan stres kerja pada perawat wanita di RSUD Kardinah Kota Tegal. *Jurnal Empati*, 6(1), 112–116. Retrieved from <https://ejournal3.undip.ac.id/index.php/empati/article/view/15647>
- Wan, J., Wee, L. H., Siau, C. S., & Wong, Y. H. (2025, February 27). Psychological well-being and its associated factors among university students in Sichuan, China. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2025.1473871>
- Wijayanto, A. Y., & Fauziah, N. (2020). KERJA DI GENGGAMANKU KELUARGA DI HATIKU Interpretative Phenomenological Analysis Tentang Work-Family Balance Pada Ibu Bekerja. *Jurnal EMPATI*, 7(1), 76-83. <https://doi.org/10.14710/empati.2018.20149>

- Wulandari, J. (2015). Tinjauan tentang konflik peran ganda dan dukungan sosial suami terhadap stres kerja (Studi pada dosen perempuan Fakultas Ilmu Sosial dan Politik Universitas Lampung). *Ecodemica*, 3(2), 520–529. <https://ejournal.bsi.ac.id/ejurnal/index.php/ecodemica/article/view/79>
- Wood, A. M., Froh, J. J., & Geraghty, A. W. (2010). Gratitude and well-being: A review and theoretical integration. *Clinical Psychology Review*, 30(7), 890–905. <https://doi.org/10.1016/j.cpr.2010.03.005>
- Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment*, 52(1), 30–41. [https://doi.org/10.1207/s15327752jpa5201\\_2](https://doi.org/10.1207/s15327752jpa5201_2)



## **BODY CONCIOUSNESS QUESTIONNAIRE (BCQ) VERSI BAHASA INDONESIA: PENGUJIAN PROPERTI PSIKOMETRI MENGGUNAKAN CFA DAN RASCH MODEL**

**Rahmat S. Bintang<sup>1</sup>, Isabell Nicole Fischer<sup>2</sup>**

<sup>1,2</sup>, Fakultas Psikologi, Universitas Bosowa, Makassar, Indonesia

*Email:* [rahmatsbintang@universitasbosowa.ac.id](mailto:rahmatsbintang@universitasbosowa.ac.id)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



<i>Received</i>	<i>Revised</i>	<i>Accepted</i>	<i>Published</i>
12 Februari 2025	4 September 2025	30 September 2025	30 Oktober 2025

### **Abstract**

*Body consciousness is a psychological construct that refers to an individual's awareness of their body, encompassing private body consciousness, public body consciousness, and body competence. This study aimed to adapt the Body Consciousness Questionnaire (BCQ) into Indonesian and examine its psychometric properties using a combination of CFA and Rasch model analyses. CFA was employed to test the factor structure, while the Rasch model was used to obtain more robust parameter estimates. The sample consisted of 304 respondents (male = 108; female = 196) aged 18–24 years (M = 20.58, SD = 1.52). A non-probability purposive sampling technique was applied. Results from the CFA indicated that the Indonesian version of the BCQ demonstrated good model fit (unidimensional). These findings were further supported by the Rasch model analysis, which confirmed the assumptions of unidimensionality and local independence. Additional results related to item fit and reliability were also satisfactory. Overall, the Indonesian version of the BCQ was found to be a valid and reliable instrument for measuring body consciousness among adolescents in Indonesia.*

*Keywords:* adaptation, body consciousness, CFA, Rasch model

### **Abstrak**

*Body consciousness merupakan konsep psikologis yang merujuk pada kesadaran individu terhadap tubuhnya, baik dari aspek private body consciousness, public body consciousness, dan body competence. Penelitian ini bertujuan untuk mengadaptasi instrumen Body Consciousness Questionnaire (BCQ) ke dalam Bahasa Indonesia dan kemudian melakukan pengujian properti psikometri dengan menggunakan penggabungan analisis CFA dan Rasch model. Analisis data menggunakan CFA untuk menguji struktur faktor dan analisis Rasch model untuk mendapatkan estimasi parameter yang lebih baik. Sampel dalam penelitian ini sebanyak 304 responden (laki-laki = 108; perempuan = 196) dengan rentang usia 18 - 24 tahun (mean usia = 20.58; SD usia = 1.52). Metode sampling yang digunakan adalah non-probability sampling dengan teknik purposive sampling. Hasil penelitian dengan menggunakan CFA menunjukkan bahwa instrumen BCQ versi Bahasa Indonesia memiliki model fit terhadap data (unidimensional). Selain itu, temuan ini didukung oleh analisis Rasch model di mana asumsi unidimensionalitas terpenuhi, begitupun dengan local independence. Temuan lainnya berkaitan dengan item fit dan reliabilitas menunjukkan hasil yang baik. Item-item dalam instrumen BCQ versi Bahasa Indonesia dinyatakan valid dan reliable untuk mengukur konstruk Body Consciousness remaja di Indonesia.*

*Kata kunci:* adaptasi, body consciousness, CFA, Rasch model

## **1. Pendahuluan**

*Body consciousness adalah kesadaran individu terhadap berbagai aspek tubuh mereka, baik yang bersifat internal maupun yang dapat diamati oleh orang lain. Ini mencakup pemahaman terhadap sensasi internal seperti perasaan, pikiran, dan emosi yang berkaitan dengan diri sendiri (private body consciousness), serta kesadaran terhadap penampilan fisik, perilaku, dan interaksi dengan orang lain (public body consciousness). Body consciousness dibagi menjadi dua dimensi utama, yakni private body consciousness dan public body consciousness (Miller et al., 1981). Private body consciousness menurut Turk et al., (1979) yaitu mengacu pada kesadaran terhadap sensasi dan pengalaman internal tubuh yang tidak dapat diamati oleh orang lain, seperti detak jantung dan napas. Sementara itu, public body consciousness berkaitan dengan kesadaran terhadap aspek eksternal tubuh yang dapat dilihat oleh orang lain, seperti penampilan dan postur.*

Penelitian mengenai *Body consciousness* menunjukkan bahwa aspek *private body consciousness* dan *public body consciousness* merupakan dua hal yang terpisah namun berkorelasi satu sama lain. Hal ini menunjukkan bahwa individu dengan kesadaran tinggi terhadap satu aspek cenderung juga memiliki kesadaran tinggi terhadap aspek lainnya. Kedua aspek ini penting untuk memahami bagaimana individu berinteraksi dan memahami tubuh mereka dalam berbagai konteks (Fenigstein et al., 1975). Terdapat beberapa instrumen pengukuran yang dapat digunakan untuk mengukur konstruk *Body consciousness*, di antaranya *Male Body Image Self-Consciousness Scale*, *Body Consciousness Questionnaire* (BCQ), *Objective Body Consciousness Scale*.

*Male Body Image Self-Consciousness Scale* dikembangkan oleh McDonagh et al., (2008). Alat ukur ini terdiri dari dimensi kesadaran diri terhadap citra tubuh saat dalam intimasi fisik, kekhawatiran terhadap penampilan dan fungsi alat genital, serta kekhawatiran terhadap tingkat/bentuk otot tubuh dengan jumlah item sebanyak 17 item. Selanjutnya ada instrumen *Body Consciousness Questionnaire* (BCQ) yang dikembangkan oleh Miller et al., (1981) dari *University of Texas at Austin*. Instrumen ini dirancang untuk mengukur kesadaran tubuh individu melalui tiga dimensi utama yaitu *private body consciousness*, *public body consciousness*, dan *body competence*. BCQ terdiri dari 15 item yang mengukur kesadaran tubuh individu.

Penelitian ini bertujuan untuk mengadaptasi instrumen BCQ ke dalam Bahasa Indonesia dan kemudian melakukan pengujian properti psikometri. Hal ini dilakukan karena instrumen BCQ belum pernah diadaptasi ke dalam Bahasa Indonesia. Selain itu, instrumen BCQ ingin diuji struktur faktornya dengan sampel Indonesia. Adapun pendekatan analisis data yang akan digunakan adalah *Confirmatory Factor Analysis* (CFA). Dengan pendekatan CFA, instrumen bisa diuji (dikonfirmasi) untuk melihat sejauh mana seluruh item dari sebuah instrumen memang mengukur tentang satu hal saja, yaitu apa yang hendak diukur (Umar & Nisa, 2020). Selain itu, untuk mendapatkan hasil yang lebih baik, penelitian ini juga akan melengkapi analisis dengan menggunakan Rasch model. Penggabungan analisis CFA dengan Rasch model memberikan keunggulan dalam analisis properti psikometri instrumen (Kelly et al., 2007).

Penggunaan Rasch model dalam penelitian ini karena adanya beberapa kelebihan yaitu *specific objectivity* dan *linearity* (Rasch 1966). Dengan menggunakan pendekatan CFA dan Rasch model secara bersamaan, penelitian ini memberikan perspektif yang berbeda dan lebih mendalam dibandingkan metode sebelumnya. Pendekatan yang berbeda ini memungkinkan analisis yang lebih spesifik dalam validasi model. Oleh karena itu, penelitian ini memiliki nilai penting karena menawarkan hasil yang lebih terperinci melalui metode yang berbeda (Mishra, 2016). Selain itu, hasil penelitian ini diharapkan dapat bermanfaat untuk penelitian-penelitian selanjutnya mengenai *body consciousness* dalam konteks Indonesia dan keterkaitannya dengan variabel-variabel psikologi lainnya.

## 2. Metode Penelitian

### Partisipan

Sampel dalam penelitian sebanyak 304 responden (responden laki-laki = 108; responden perempuan = 196) dengan rentang usia antara 18 tahun sampai 24 tahun (mean usia = 20.58; SD usia = 1.52. Metode sampling yang digunakan adalah *non probability sampling* dengan teknik *purposive sampling*. Responden dalam penelitian bersifat sukarela dan informasi mengenai responden bersifat rahasia (*confidential*) hanya untuk keperluan penelitian.

### Instrumen Penelitian

Penelitian ini menggunakan instrumen *Body Consciousness Questionnaire* (BCQ) yang dikembangkan oleh Miller et al., (1981). BCQ merupakan instrumen yang dikembangkan untuk menilai kesadaran tubuh dan persepsi individu terhadap tubuh mereka sendiri. Proses adaptasi ke dalam Bahasa Indonesia dengan mengacu pada tahapan adaptasi yang diusulkan oleh Beaton et al., (2000). Instrumen BCQ menggunakan skala Likert 5 poin. Poin 1 untuk "sangat tidak sesuai", 2 untuk "tidak sesuai", 3 untuk "sedikit sesuai", 4 untuk "sesuai", dan 5 untuk "sangat sesuai". Proses adaptasi yang dilakukan mengikuti pedoman adaptasi oleh Beaton et al., (2000) dan tetap mengikuti skala aslinya yaitu terdiri dari tiga dimensi yang mewakili 15 item pernyataan, yaitu: aspek *private body consciousness*, *public body consciousness*; dan *body competence*. Selain itu, skala likert yang digunakan pada instrumen BCQ versi Bahasa Indonesia tetap menggunakan skala likert 5 poin dan semua item bersifat favorable. Untuk lebih jelasnya, *blueprint* instrumen BCQ versi Bahasa Indonesia dapat dilihat pada Tabel 1 dan pada Tabel 2 menunjukkan bunyi itemnya.

**Tabel 1. Blueprint instrumen BCQ**

No.	Dimensi	Item	Jumlah
1.	<i>Private Body Consciousness</i>	3, 4, 5, 8, 12	5
2.	<i>Public Body Consciousness</i>	1, 7, 10, 11, 13, 15	6
3.	<i>Body Competence</i>	2, 6, 9, 14	4
<b>Jumlah</b>			<b>15</b>

### Pengumpulan Data

Pengumpulan data dilakukan melalui *Google Form* yang disebar melalui berbagai *platform* media sosial.

Proses pengumpulan data dimulai pada tanggal 13 Mei 2024 dan berlangsung hingga 26 Mei 2024. Melalui formulir *Google* ini, responden diundang untuk berpartisipasi dalam penelitian dengan mengisi kuesioner yang telah disiapkan. Media sosial yang digunakan adalah *WhatsApp*, *instagram*, dan *X*. Data yang terkumpul selama periode tersebut akan menjadi dasar untuk analisis dan temuan dalam penelitian ini.

### Teknik Analisis Data

Analisis data dalam penelitian ini menggunakan dua pendekatan yaitu *confirmatory factor analysis* dan *Rasch model*. CFA digunakan untuk menguji apakah struktur faktor dari variabel sesuai dengan yang dihipotesiskan sebelumnya. CFA dapat digunakan untuk memodifikasi model agar kesesuaiannya dengan data semakin baik. Ini mengevaluasi seberapa baik model cocok dengan data. Ada dua hal yang perlu dilihat dalam analisis dengan CFA, yaitu melihat model fit dan menentukan item valid. Pertama, untuk melihat apakah model fit didukung oleh data, terdapat beberapa indeks *goodness-of-fit* di antaranya: *Root Mean Square Error of Approximation (RMSEA)*, *Comparative FIT Indeks (CFI)*, *Tucker-Lewis Indeks (TLI)*, dan *Standardized Root Mean Square Residual (SRMR)*. Untuk dapat dikatakan *fit*, nilai  $RMSEA = < 0.06$ ;  $CFI$  dan  $TLI = > 0.90$ ; dan  $SRMR = < 0.08$  (Hu & Bentler, 1999). Kedua, untuk melihat parameter item dikatakan valid dan tidak valid, digunakan kriteria yaitu nilai koefisien muatan faktor bernilai positif dan nilai  $z$ -value  $> 1.96$  (Umar & Nisa, 2020). Dengan memperhatikan kriteria ini, peneliti dapat menentukan apakah item-item dalam instrumen pengukuran tersebut sesuai dengan model yang dihipotesiskan dan memiliki kontribusi yang valid dalam mengukur konstruk yang diinginkan. Analisis CFA dilakukan dengan menggunakan software JASP dengan metode estimasi *maximum likelihood (ML)*.

Pendekatan kedua setelah dilakukan analisis CFA, maka dilakukan analisis dengan *Rasch model*. *Rasch model* memiliki beberapa keunggulan seperti *specific objectivity*, *sufficiency*, dan *parameter separation* (Wright & Master, 1982; Suryadi et al., 2020; Rahayu et al., 2023). Analisis *Rasch* dalam penelitian ini menggunakan model Partial Credit Model (PCM) yang dikembangkan oleh Masters (1982). Pengujian properti psikometri dalam penelitian ini adalah pengujian asumsi *unidimensionality* dan *local independence*, melaporkan item *fit*, *reliabilitas*, *wright map*. Analisis *Rasch model* dengan menggunakan software WINSTEPS versi 3.73 dengan metode estimasi *Joint (Unconditional) Maximum Likelihood (JMLE)*.

## 3. Hasil dan Diskusi

### Statistik Deskriptif

Sebelum melakukan uji CFA, tahap pertama yang dilakukan adalah menguji asumsi univariat normality. Asumsi univariat normality terpenuhi ketika nilai *skewness* masing-masing item berada pada rentang nilai -2 sampai +2 (Ahmed et al., 2022). Berdasarkan hasil analisis, dapat dilihat bahwa asumsi univariat normality pada instrumen BCQ terpenuhi di mana nilai *skewness* setiap item berada pada range yang ditetapkan (Lihat Tabel 2).

Tabel 2. Statistik deskriptif

Item	Instrumen BCQ versi Bahasa Indonesia	Mean	SD	Skewness
1.	Saat bersama orang lain, saya ingin tangan saya bersih dan terlihat bagus.	3.891	0.939	-0.575
2.	Saya merasa cukup ideal untuk ukuran tubuh saya	4.191	0.724	-0.622
3.	Saya merasa peka jika tubuh saya tegang	3.908	0.922	-0.605
4.	Saya segera menyadari ketika mulut atau tenggorokan saya kering	3.984	0.906	-0.878
5.	Saya merasakan jantung saya berdebar-debar	3.964	0.830	-0.805
6.	Saya merasa lebih gesit berjalan dibandingkan dengan orang lain.	3.928	0.965	-0.674
7.	Saya merasa penting untuk kulit saya terlihat baik, misalnya tidak memiliki bekas luka	4.181	0.802	-0.802
8.	Saya cepat merasakan rasa lapar pada perut saya	4.168	0.853	-0.939
9.	Saya mampu bergerak dengan cepat	4.145	0.795	-0.860
10.	Saya dapat menyadari akan tampilan wajah terbaik dan terburuk saya	4.191	0.706	-0.572
11.	Saya suka memastikan bahwa rambut saya terlihat bagus	3.796	1.036	-0.550
12.	Saya cepat menyadari ketika terjadi perubahan pada suhu tubuh saya	4.076	0.862	-0.799
13.	Saya berpikir banyak tentang bentuk tubuh saya	4.066	0.869	-0.675
14.	Saya merasa lebih terkoordinasi dibandingkan kebanyakan orang.	4.033	0.878	-0.859
15.	Saya merasa khawatir dengan postur tubuh saya	4.016	0.822	-0.891

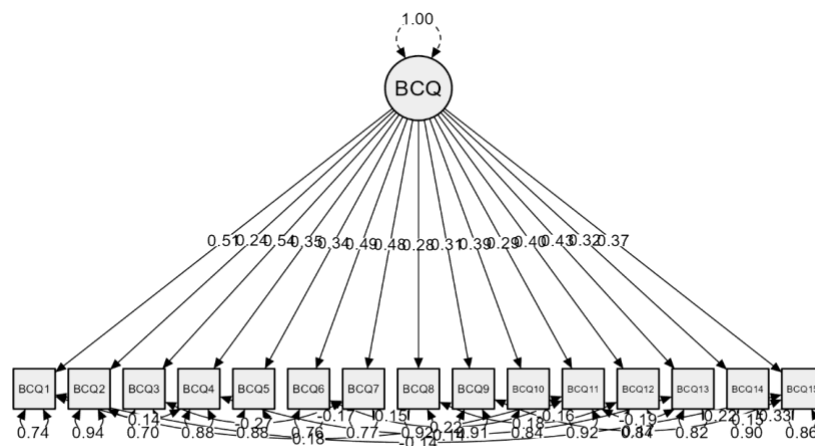
### Model Testing CFA

Setelah asumsi *univariat normality* terpenuhi, maka analisis CFA dapat dilakukan. Penelitian ini menggunakan software JASP versi 0.19.0.0 dengan metode estimasi *Maximum Likelihood (ML)*. Hasil analisis awal diperoleh nilai Chi-square = 249.826 (df = 90) dengan p-value < .001; RMSEA = 0.076 (90% C.I. = 0.065, 0.088); CFI = 0.699; TLI = 0.649; SRMR = 0.068. Berdasarkan nilai-nilai tersebut menunjukkan bahwa model belum fit di mana semua nilai *goodness-of-fit* belum terpenuhi. Untuk membuat model fit, salah satu cara yang dapat dilakukan adalah dengan melakukan modifikasi model. Modifikasi yang dilakukan dalam hal ini adalah membebaskan residual antar item saling berkorelasi (Umar & Nisa, 2020). Residual yang dibebaskan saling berkorelasi adalah residual antar item yang terbesar.

Setelah dilakukan modifikasi model sebanyak 15 kali yaitu item 14 dengan 15; item 6 dengan item 11; item 1 dengan item 11; item 1 dengan item 15; item 4 dengan item 13; item 3 dengan item 7; item 13 dengan item 14; item 11 dengan item 13; item 9 dengan item 15; item 10 dengan item 11; item 8 dengan item 12; item 7 dengan item 8; item 2 dengan item 4; item 6 dengan item 7; dan item 13 dengan item 15, diperoleh model fit dengan nilai Chi-square = 91.880 (df = 75) dengan p-value 0.090; RMSEA = 0.027 (90% C.I. = 0.000, 0.045), CFI 0.968, TLI = 0.956, dan SRMR = 0.045. Setelah model fit, tahapan berikutnya adalah mengidentifikasi item-item dari instrumen BCQ untuk melihat item valid dan tidak valid. Berdasarkan hasil analisis, ditemukan bahwa semua item valid (Lihat Tabel 3).

**Tabel 3. Koefisien Muatan Faktor Instrumen BCQ**

Item	Faktor Loading	Std. Error	z-value	p-value	Keterangan
1.	0.464	0.060	7.752	<.001	valid
2.	0.190	0.047	4.024	<.001	valid
3.	0.487	0.059	8.262	<.001	valid
4.	0.330	0.060	5.533	<.001	valid
5.	0.289	0.053	5.436	<.001	valid
6.	0.427	0.061	6.973	<.001	valid
7.	0.364	0.052	6.961	<.001	valid
8.	0.266	0.056	4.746	<.001	valid
9.	0.248	0.052	4.754	<.001	valid
10.	0.280	0.045	6.193	<.001	valid
11.	0.293	0.073	4.029	<.001	valid
12.	0.352	0.055	6.364	<.001	valid
13.	0.403	0.056	7.224	<.001	valid
14.	0.301	0.058	5.192	<.001	valid
15.	0.341	0.054	2.295	<.001	valid



**Gambar 1. Model Fit Instrumen BCQ versi Bahasa Indonesia**

### Analisis Rasch model

#### *Unidimensionality dan local Independence*

Dua asumsi Rasch model yang harus dipenuhi dalam penggunaan Rasch model yaitu *unidimensionality* dan *local independence*. *Unidimensionalitas* berarti bahwa instrumen yang dianalisis hanya mengukur satu konstruk

saja (Bond & Fox, 2015). Sementara itu, *local independence* menunjukkan bahwa jawaban responden pada satu item tidak boleh dipengaruhi oleh jawaban pada item yang lain (de Ayala, 2009). Pengujian asumsi *unidimensionality* dalam Rasch model bertujuan untuk melihat apakah sekumpulan item yang diteorikan mampu mengukur satu konstruk yang sama dengan menggunakan *Principal Component Analysis of Residual* (PCAR) (Chou dan Wang, 2010). Asumsi *unidimensionality* terpenuhi ketika nilai *raw variance explained by measures* > 20% (Reckase, 1979). Hasil penelitian menunjukkan bahwa asumsi *unidimensionality* terpenuhi di mana didapatkan nilai *raw variance explained by measures* sebesar 22.9% (Lihat Tabel 5).

Setelah asumsi *unidimensionality* terpenuhi, selanjutnya dilakukan pengujian asumsi *local independence*. Untuk menguji asumsi *local independence* digunakan metode  $Q_3$  (Yen, 1984). Kriteria indeks  $Q_3$  yaitu korelasi residual antar pasangan item tidak boleh > 0.30 (Aryadoust et al., 2020; Christensen et al., 2017). Hasil analisis menunjukkan bahwa tidak terjadi pelanggaran asumsi *local independence* pada instrumen BCQ versi Bahasa Indonesia karena tidak ada item yang memiliki nilai korelasi residual > 0.30. Nilai korelasi residual instrumen BCQ versi Bahasa Indonesia yaitu -0.21 sampai 0.25.

**Tabel 5. Asumsi Unidimensionality**

	Eigenvalues	Observed	Expected
<i>Total raw variance</i>	19.4	100%	100%
<i>Raw variance explained by measures</i>	4.4	<b>22.9%</b>	22.5%
<i>Raw variance explained by person</i>	1.6	8.5%	8.3%
<i>Raw variance explained by items</i>	2.8	14.4%	14.2%
<i>Raw variance (total)</i>	15.0	77.1%	77.5%

#### Item Fit Statistik

Untuk melihat item fit dalam Rasch model, dengan melihat nilai *Infit Mean Square* (MNSQ) dan *Outfit Mean Square* (MNSQ) (Linacre, 2018). Nilai *Infit* dan *Outfit* yang dapat diterima antara 0.5 sampai 1.5 (Bond & Fox, 2015). Berdasarkan hasil penelitian, semua item pada instrumen BCQ versi Bahasa Indonesia fit terhadap Rasch model (Lihat Tabel 6). Selain itu, pada tabel 6 juga menunjukkan bahwa item-item pada instrumen BCQ versi Bahasa Indonesia telah diurutkan dari item yang paling sukar disetujui yaitu item 11 sampai item paling mudah disetujui yaitu item 10. Lokasi tingkat kesukaran item adalah berada pada range -0.270 (item10) sampai 0.383 logit (item 11).

**Tabel 6. Item fit**

Item	Measure (Tingkat Kesukaran)	Infit MNSQ	Outfit MNSQ	PTMEA
Item 11	0.383	1.30	1.25	0.42
Item 1	0.241	0.97	0.96	0.54
Item 3	0.215	0.96	0.93	0.53
Item 6	0.185	1.04	1.02	0.54
Item 5	0.127	0.92	0.89	0.40
Item 4	0.095	1.11	1.04	0.41
Item 15	0.041	0.84	0.84	0.46
Item 14	0.014	0.99	0.98	0.46
Item 13	-0.042	0.99	0.97	0.48
Item 12	-0.060	0.96	0.97	0.50
Item 9	-0.183	0.95	0.96	0.39
Item 8	-0.226	1.14	1.19	0.38
Item 7	-0.251	0.95	0.95	0.48
Item 2	-0.270	0.91	0.92	0.34
Item 10	-0.270	0.79	0.85	0.39

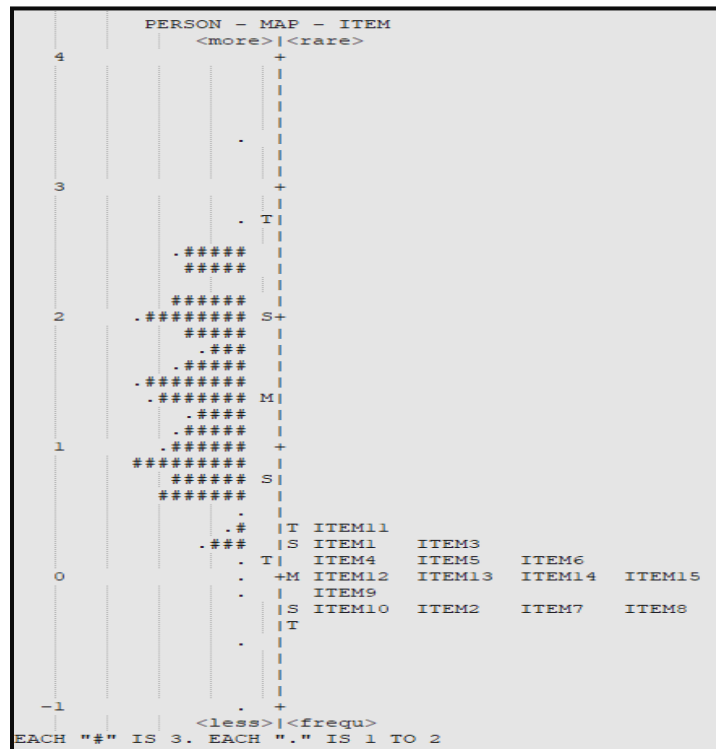
#### Reliability

Rasch model memiliki koefisien reliabilitas yang memiliki konsep yang berbeda dari *Classical Test Theory* (misal Cronbach's alpha) yang dikembangkan oleh Wright & Master (1982). Koefisien reliabilitas dalam Rasch model memiliki estimasi yang lebih baik untuk mengukur orang melalui *person separation reliability* (PSR) dan untuk item melalui *item separation reliability* (ISR). PSR merupakan estimasi seberapa baik instrumen dapat membedakan responden pada variabel yang diukur, sementara itu ISR merupakan seberapa baik item memilah responden yang mengikuti tes (Wright & Master, 1982). Kriteria PSR untuk dapat diterima dikatakan memiliki reliabilitas yang baik adalah > 0.70 (Tennant & Conaghan, 2007). Hasil analisis menunjukkan bahwa nilai PSR

pada instrumen BCQ versi Bahasa Indonesia adalah 0.72. Sementara itu, nilai ISR dari instrumen BCQ versi Bahasa Indonesia sebesar 0.77. Artinya, instrumen BCQ versi Bahasa Indonesia cukup dapat diandalkan, tetapi belum optimal (< 0.90).

**Wright Map**

*Wright map* merupakan peta yang menunjukkan lokasi pengukuran orang dan kalibrasi item pada skala yang sama. *Wright map* memudahkan untuk melakukan identifikasi lokasi antara tingkat kemampuan orang dengan tingkat kesukaran item. Dalam Rasch model, rata-rata tingkat kesukaran item adalah 0.00 (Bond & Fox, 2015). Berdasarkan hasil analisis pada instrumen BCQ versi Bahasa Indonesia (lihat gambar 2), dapat diketahui bahwa rata-rata tingkat kemampuan orang berada pada 1.5 logit di atas rata-rata tingkat kesukaran item.



**Gambar 2. Wright map instrumen BCQ versi Bahasa Indonesia**

Berdasarkan *Wright Map* di atas, terlihat bahwa distribusi rata-rata tingkat kemampuan orang relatif lebih tinggi dibandingkan dengan rata-rata tingkat kesulitan item instrumen. Secara umum responden berada pada rentang 0 hingga 2 logit, yang menunjukkan kemampuan menengah hingga tinggi, sementara sebagian besar item pada instrumen BCQ versi Bahasa Indonesia justru berada pada tingkat kesulitan rendah hingga sedang. Kondisi ini mengindikasikan bahwa beberapa item, seperti item 10, item 2, dan item 7 serta item 8, tergolong terlalu mudah sehingga hampir semua responden mampu memilih sangat setuju pada item-item tersebut, yang berpotensi mengurangi daya diskriminasi item. Sebaliknya, butir dengan tingkat kesulitan lebih tinggi, seperti item 12, item 13, dan item 14, cukup dapat membedakan responden dengan kemampuan tinggi, meskipun jumlah responden yang mampu menjawab sangat setuju sedikit. Secara umum, sebaran item sudah mencakup berbagai level kesulitan, namun belum sepenuhnya seimbang dengan kemampuan responden.

**Pembahasan**

Penelitian ini merupakan studi pertama yang mengadaptasi instrumen *Body Consciousness Questionnaire* (BCQ) versi Bahasa Indonesia dan menguji properti psikometrinya dengan menggunakan penggabungan analisis CFA dan Rasch model. Penelitian ini telah memberikan metode analisis yang baru dibandingkan dengan penelitian sebelumnya. Instrumen BCQ versi Bahasa Indonesia terbukti memiliki properti psikometri yang baik. Hasil analisis CFA menunjukkan bahwa struktur 1-faktor (unidimensi) instrumen BCQ versi Bahasa Indonesia terbukti fit dengan data. Diperoleh model fit dengan nilai Chi-square = 91.880 (df = 75) dengan p-value 0.090; RMSEA = 0.027 (90% C.I. = 0.000, 0.045), CFI 0.968, TLI = 0.956, dan SRMR = 0.045. Selain itu, semua item juga terbukti valid di mana diperoleh nilai koefisien muatan faktor bernilai positif dan nilai z-value > 1.96 dengan signifikansi p-value < 0.01 (Hu & Bentler, 1999; Umar & Nisa, 2020).

Temuan dari hasil analisis CFA didukung oleh hasil analisis rasch model yang menunjukkan bahwa asumsi unidimensionality instrumen BCQ versi Bahasa Indonesia juga terpenuhi. Diperoleh nilai *raw variance explained by measures* sebesar 22.9%. Hasil ini telah memenuhi sesuai kriteria yang dicetuskan oleh Reckase (1979) yaitu sebesar > 20%. Selain itu, tidak ditemukan pelanggaran asumsi *local independence* pada instrumen BCQ versi Bahasa Indonesia. Hasil analisis menunjukkan bahwa tidak ada item yang memiliki nilai korelasi residual > 0.30 (Aryadoust et al., 2020; Christensen et al., 2017). Temuan lainnya berkaitan dengan reliabilitas. *Person Separation Reliability* (PSR) dari instrumen BCQ versi Bahasa Indonesia adalah 0.72. Nilai PSR sebesar 0.72 telah melebihi dari kriteria yang ditetapkan yaitu 0.70 (Tennant & Conaghan, 2007). Hasil ini menunjukkan bahwa instrumen BCQ versi Bahasa Indonesia mampu membedakan individu yang memiliki *body consciousness* yang tinggi dengan *body consciousness* yang rendah.

Selain itu, nilai reliabilitas pada *item separation reliability* (ISR) diperoleh nilai sebesar 0.77. Hasil ini menunjukkan bahwa instrumen ini cukup dapat diandalkan, tetapi belum optimal, akan lebih baik jika nilai ISR > 0.90 (Bond & Fox, 2015). Nilai ISR sebesar 0.77 juga menunjukkan bahwa masih ada ruang untuk perbaikan pada instrumen BCQ versi Bahasa Indonesia, misalnya diperlukan penambahan item dengan tingkat kesulitan yang lebih tinggi agar instrumen mampu mengukur responden dengan kemampuan di atas rata-rata. Temuan ini juga menunjukkan bahwa meskipun instrumen dapat digunakan untuk mengukur konstruk *body consciousness*, optimalisasi dalam penyusunan item masih perlu dilakukan. Dengan demikian, kualitas instrumen akan semakin baik dalam memberikan informasi yang lebih akurat mengenai variasi kemampuan responden.

Hasil analisis lainnya yaitu semua item pada instrumen BCQ versi Bahasa Indonesia terbukti fit dengan Rasch model. Hasil ini mendukung temuan pada analisis CFA sebelumnya. Lokasi tingkat kesukaran item yaitu dari item yang paling mudah disetujui yaitu item nomor 10 (*saya dapat menyadari akan tampilan wajah terbaik dan terburuk saya*) dengan tingkat kesukaran -0.270 logit sampai item yang paling sukar untuk disetujui yaitu item nomor 11 (*saya suka memastikan bahwa rambut saya terlihat bagus*) dengan nilai 0.383 logit. Adapun keterbatasan dalam penelitian ini adalah sampel yang hanya terdiri dari remaja yang berusia 18-24 tahun, sehingga ruang lingkup teoritis dan generalisasi *body consciousness* dalam penelitian ini sangat sempit, padahal *body consciousness* berkaitan dengan berbagai kelompok usia. Berdasarkan keterbatasan ini, peneliti mengakui bahwa temuan mengenai karakteristik psikometri instrumen BCQ versi Bahasa Indonesia hanya berlaku untuk karakteristik sampel yang sama. Oleh karena itu, diperlukan studi lebih lanjut untuk karakteristik usia lainnya. Selain itu, penelitian ini juga belum melakukan uji *measurement invariance*, misal apakah ada item yang berlaku tidak sama antar budaya, suku, maupun jenis kelamin.

#### 4. Kesimpulan

Penelitian ini merupakan studi pertama yang mengadaptasi instrumen *Body Consciousness Questionnaire* (BCQ) versi Bahasa Indonesia dan menguji properti psikometrinya dengan menggunakan penggabungan analisis CFA dan Rasch model. Hasil analisis menunjukkan bahwa instrumen ini memiliki kesesuaian model yang baik (*model fit*) yang didukung oleh data empiris, sehingga dapat dikatakan memiliki properti psikometri yang valid dan reliabel untuk digunakan dalam mengukur *body consciousness* pada remaja Indonesia. Temuan ini memperkuat bukti bahwa BCQ versi Bahasa Indonesia dapat dijadikan alat ukur yang representatif dalam konteks budaya Indonesia, khususnya pada populasi remaja. Meskipun demikian, penelitian lanjutan masih diperlukan untuk mengeksplorasi keterkaitan BCQ dengan konstruk psikologis lainnya, seperti *self-esteem*, *body image*, maupun regulasi emosi. Dengan demikian, BCQ versi Bahasa Indonesia berpotensi tidak hanya sebagai alat asesmen, tetapi juga sebagai instrumen yang mendukung pengembangan intervensi psikologis yang lebih tepat sasaran.

#### 5. Referensi

- Ahmed O, Hossain KN, Hiramoni FA, Siddique RF and Chung S (2022). Psychometric properties of the Bangla version of the stress and anxiety to viral epidemics-6 Items scale among the general population in Bangladesh. *Front. Psychiatry* 13:804162. <https://doi.org/10.3389/fpsy.2022.804162>
- Aryadoust, V., Ng, L. Y., & Sayama, H. (2021). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing*, 38(1), 6–40. <https://doi.org/10.1177/0265532220927487>
- 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., & Fox, C. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Lawrence Erlbaum & Associates. <https://doi.org/10.4324/9781315814698>
- Chou, Y.-T., & Wang, W.-C. (2010). Checking dimensionality in item response models with principal component analysis on standardized residuals. *Educational and Psychological Measurement*, 70(5), 717–731. <https://doi.org/10.1177/0013164410379322>

- 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>
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43(4), 522–527. <https://doi.org/10.1037/h0076760>
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kelly, P. A., Kallen, M. A., & Suarez-Almazor, M. E. (2007). A combined-method psychometric analysis recommended modification of the multidimensional health locus of control scales. *Journal of Clinical Epidemiology*, 60(5), 440–447. <https://doi.org/10.1016/j.jclinepi.2006.08.005>
- Linacre, J. M. (2018). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149–174. <https://doi.org/10.1007/BF02296272>
- McDonagh, L. K., Morrison, T. G., & McGuire, B. E. (2008). The Naked Truth: Development of a Scale Designed to Measure Male Body Image Self-Consciousness during Physical Intimacy. *The Journal of Men's Studies*, 16(3), 253–265. <https://doi.org/10.3149/jms.1603.253>
- Miller, L. C., Murphy, R., & Buss, A. H. (1981). Consciousness of body: Private and public. *Journal of Personality and Social Psychology*, 41(2), 397–406. <https://doi.org/10.1037/0022-3514.41.2.397>
- Mishra, M. (2016). Confirmatory factor analysis (CFA) as an analytical technique to assess measurement error in survey research: A Review. *Paradigm*, 20(2), 97–112. <https://doi.org/10.1177/0971890716672933>
- Rahayu, W., Hayat, B., & Putra, M. D. K. (2023). *Analisis rasch: aplikasi dan interpretasi*. UNJ Press
- Rasch, G. (1966). An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology*, 19(1), 49–57. <https://doi.org/10.1111/j.2044-8317.1966.tb00354.x>
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational and Behavioral Statistics*, 4(3), 207–230. DOI: 10.3102/10769986004003207
- Suryadi, B., Hayat, B., & Putra, M. D. K. (2020). Evaluating psychometric properties of the Muslim Daily Religiosity Assessment Scale (MUDRAS) in Indonesian samples using the Rasch model. *Mental Health, Religion & Culture*, 23(3-4), 331–346. <https://doi.org/10.1080/13674676.2020.1795822>
- Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research*, 57(8), 1358–1362. <https://doi.org/10.1002/art.23108>
- Turk, D. C., Meichenbaum, D. H., & Berman, W. H. (1979). Application of biofeedback for the regulation of pain: A critical review. *Psychological Bulletin*, 86(6), 1322–1338. <https://doi.org/10.1037/0033-2909.86.6.1322>
- Umar, J & Nisa, Y., F. (2020). Uji validitas dengan CFA dan pelaporannya. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)*, 9(2), 1-11. <http://dx.doi.org/10.15408/jp3i.v9i2.XXXXXX>
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125–145. <https://doi.org/10.1177/014662168400800201>



## EFEKTIVITAS INTERVENSI SELF-COMPASSION WRITING TERHADAP STRES DAN WELL-BEING MAHASISWA

Qurrota Aini<sup>1</sup>, Herdiyan Maulana<sup>2</sup>, Mira Ariyani<sup>3</sup>

<sup>1,2,3</sup>, Magister Sains Psikologi, Fakultas Psikologi, Universitas Negeri Jakarta, Indonesia

Email: [hi.qurrota@gmail.com](mailto:hi.qurrota@gmail.com)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
24 Juni 2025	6 Oktober 2025	14 Oktober 2025	30 Oktober 2025

### Abstract

Addressing mental health issues among Indonesian college students is crucial due to the high prevalence of stress. Students face significant stress from academic workloads, challenging assignments, and financial constraints, which can lead to long-term consequences such as academic decline, diminished quality of life, and social dysfunction. Despite this need, universities often lack sufficient mental health support due to social-psychological barriers and resource limitations. While individuals may independently manage early-stage stress symptoms—referred to as coping—students require empirically validated, effective, and accessible coping strategies. One emerging coping model is instilling self-compassion (SC), which involves showing unconditional kindness to oneself when facing pressure. SC can be delivered through writing interventions as a means of self-help. Data were collected from 32 student participants using a within-subject experimental design to measure stress (Perceived Stress Scale/PSS), well-being (Satisfaction with Life/SWLS), and self-compassion (Self-Compassion Scale/SCS). The data analysis technique used was repeated-measure ANOVA. The study results showed that the self-compassion writing intervention was not effective in reducing stress, nor was it effective in improving well-being and self-compassion.

*Keywords:* college student, coping strategy model, self-compassion writing, stress, well-being

### Abstrak

Mengatasi masalah kesehatan mental di kalangan mahasiswa Indonesia sangat penting karena tingginya prevalensi stres. Mahasiswa menghadapi stres dari beban akademik, tugas yang menantang, dan keterbatasan keuangan, yang dapat menyebabkan konsekuensi jangka panjang seperti penurunan prestasi akademik, penurunan kualitas hidup, dan disfungsi sosial. Meski tingginya kebutuhan, universitas sering kali kekurangan dukungan kesehatan mental yang memadai karena hambatan sosial-psikologis dan keterbatasan sumber daya. Meskipun individu dapat secara mandiri mengelola gejala stres tahap awal—yang disebut sebagai *coping*—mahasiswa memerlukan strategi *coping* yang efektif, teruji secara empiris, dan mudah diakses. Salah satu model *coping* yang sedang berkembang adalah menanamkan *self-compassion* (SC), yang melibatkan menunjukkan kebaikan tanpa syarat kepada diri sendiri saat menghadapi tekanan. SC dapat disampaikan melalui intervensi menulis sebagai sarana *self-help*. Data dikumpulkan dari 32 partisipan mahasiswa dengan teknik eksperimen *within-subject design* untuk mengukur stres (*Perceived Stress Scale*/PSS), *well-being* (*Satisfaction with Life*/SWLS), serta *self-compassion* (*Self-Compassion Scale*/SCS). Teknik analisis data yang digunakan adalah *Repeated Measure ANOVA*. Hasil penelitian menunjukkan bahwa intervensi *self-compassion writing* tidak efektif dalam menurunkan stres, serta tidak efektif dalam meningkatkan *well-being* dan *self-compassion*.

Kata kunci: mahasiswa, model strategi coping, *self-compassion writing*, stres, *well-being*

### 1. Pendahuluan

Pendidikan tinggi adalah masa yang menimbulkan stres bagi banyak mahasiswa. Di Amerika Serikat, 87% dari 3.409 orang dewasa mengatakan bahwa pendidikan mereka merupakan sumber stres yang signifikan (American Psychological Association, 2020). Dalam survei nasional lain dengan sampel 33.204 mahasiswa dari 41 perguruan tinggi, hampir tiga perempat mahasiswa melaporkan mengalami stres psikologis sedang atau parah (National College Health Assessment, American College Health Association, 2022). Hal ini senada dengan Riset Kesehatan Dasar (Riskesdas) 2018 yang menunjukkan lebih dari 19 juta penduduk berusia lebih dari 15 tahun mengalami gangguan mental emosional, dan 12 juta di antaranya mengalami depresi (Badan Penelitian dan Pengembangan Kesehatan Kementerian RI, 2018; Kemkes.go.id, 2021).

Mahasiswa pada jenjang sarjana diwajibkan memenuhi berbagai tuntutan akademik dan non-akademik. Dalam masa studi mereka yang relatif lama (3-5 tahun); beban akademik seperti tugas yang menantang (Kizhakkeveettill, dkk., 2017), masalah keuangan, dan hubungan sosial (Karyotaki, dkk., 2020; Seedhom, dkk., 2019) meningkatkan prevalensi stres dan resiko kesehatan mental lanjut bagi mahasiswa. Tingkat stres yang tinggi berimplikasi pada penurunan performa akademik (Sohail, 2013), kesehatan fisik, serta mental, seperti kerentanan terhadap penyakit dan pola makan tidak sehat (Ng, dkk., 2016), kualitas tidur yang buruk (Gardani, dkk., 2022), hingga peningkatan depresi dan kecemasan (Kroshus, dkk., 2021; Conley, dkk., 2020). Stres juga berhubungan dengan perilaku kesehatan yang bermasalah seperti penggunaan narkoba (Karaman, dkk., 2019; Sharp & Theiler, 2018).

Kerentanan mahasiswa terhadap stres tidak hanya dipengaruhi oleh faktor akademik, tetapi juga faktor perkembangan. Masa perkuliahan merupakan periode transisi menuju dewasa awal (*emerging adulthood*) yang ditandai dengan tuntutan kemandirian, pencarian identitas, serta penyesuaian terhadap lingkungan sosial dan akademik yang baru (Arnett, 2000). Dalam konteks Indonesia, tekanan akademik sering diperparah oleh harapan orang tua yang tinggi dan seringkali memicu *fear of failure* serta distress psikologis pada mahasiswa (Setyaningrum, dkk., 2024). Selain itu, keterbatasan dukungan finansial membuat mahasiswa dari latar ekonomi rendah lebih rentan terhadap stres akademik (Zega & Zega, 2025). Mahasiswa yang merantau juga menghadapi tantangan adaptasi lingkungan dan keterbatasan dukungan sosial (Fitri, dkk., 2024). Faktor-faktor ini membuat mahasiswa berada pada posisi yang lebih rentan dibanding kelompok usia lain, sehingga risiko stres kronis dan gangguan mental emosional menjadi semakin besar.

Perawatan kesehatan mental bagi mahasiswa terbukti berdampak positif, seperti dilaporkan oleh Center for Collegiate Mental Health (CCMH, 2022). Berdasarkan data dari 180 pusat konseling universitas di Amerika Serikat tentang 190.907 mahasiswa yang mencari perawatan kesehatan mental, ditemukan bahwa mahasiswa baru dengan tingkat stres akademik yang tinggi dan memiliki riwayat rawat inap psikiatri 48% lebih mungkin untuk mengundurkan diri dari sekolah. Ada beberapa faktor protektif yang mengurangi risiko pengunduran diri dari sekolah, termasuk ketika stres akademik berkurang secara signifikan selama konseling dan mahasiswa secara bersamaan berpartisipasi dalam kegiatan ekstrakurikuler; mereka 51% lebih kecil kemungkinannya untuk mengundurkan diri dari sekolah. Temuan ini menyoroti peran penting pusat konseling perguruan tinggi: ketika kondisi mahasiswa meningkat selama layanan di pusat konseling, mereka lebih cenderung bertahan di perguruan tinggi.

Dalam kondisi ideal, perguruan tinggi dapat memberikan layanan kesehatan mental untuk membantu mahasiswa mengatasi stres. Namun, kendala psikologis seperti rasa enggan atau malu (Kompas.com, 2021), mahasiswa menganggap perawatan yang tersedia tidak nyaman dan meragukan efektivitas perawatan (Lattie, dkk., 2019), ditambah keterbatasan fasilitas dan SDM pada layanan kesehatan jiwa (Xiao, dkk., 2017; Center for Collegiate Mental Health, 2021), mengakibatkan penanganan kesehatan mental sangat terbatas untuk didapatkan.

Berdasarkan laporan tahun 2022 oleh WHO, nyaris di semua negara berkembang, layanan kesehatan mental masih belum optimal. Ketersediaan layanan tersebut sering diremehkan dan tidak dibekali dengan sumber daya yang memadai (World Health Organization, 2022). Di Indonesia, jumlah tenaga profesional kesehatan mental yang minim menjadi salah satu hambatan pencarian akses perawatan pada sektor ini. Hanya terdapat 3.605 psikolog klinis dengan 2.627 psikolog (72,8%) tersebar di Pulau Jawa (Pusat Data Strategis dan Statistik Ikatan Psikolog Klinis Indonesia, 2023). Psikiater sejumlah 1.221 orang saja dan mayoritas (68,49%) berada di Pulau Jawa (Antaraneews.com, 2022). Rasio tenaga kesehatan mental dibandingkan dengan jumlah penduduk Indonesia mencapai 1:200.000, jauh dari jumlah ideal yang direkomendasikan WHO yaitu 1:30.000 (Kumparan.com, 2022).

Meski dengan dukungan yang minim pada pelayanan kesehatan mental, secara teoretik, setiap orang memungkinkan secara mandiri mengatasi stres dengan melakukan upaya preventif dan penanganan gejala stres pada tahap awal—yang disebut sebagai *coping* (Lazarus & Folkman, 1984). Namun, untuk dapat melakukan hal tersebut, mahasiswa membutuhkan model strategi *coping* yang adaptif, efektif, terjangkau, serta mudah diaplikasikan untuk bertahan seumur hidup serta melindungi *well-being* mereka (Graves, dkk., 2021). Hal senada ditegaskan oleh Córdova Olivera, dkk., (2023) bahwa *workshop* mengenai manajemen stres dan teknik *coping* sangat diperlukan mahasiswa. Perguruan tinggi juga memiliki peran penting dalam memaksimalkan potensi dari intervensi untuk mengurangi stres (Karyotaki, dkk., 2020).

Kajian di bidang psikologi dan kesehatan mengindikasikan bahwa perilaku *self-compassion* sangat terkait dengan isu stres. *Self-compassion* berarti memberikan kebaikan tanpa syarat kepada diri sendiri saat menghadapi kesulitan, yang terbukti berkontribusi positif dan memicu strategi *coping* yang adaptif. Individu yang menyayangi diri sendiri tidak terlibat dalam kritik diri yang keras; mereka memiliki lebih banyak sumber daya untuk mengaktifkan perilaku *self-care*, menangani stresor, dan menangani emosi. Individu yang welas asih tidak terlibat dalam strategi yang disfungsi untuk kesehatan mental jangka panjang, seperti bertindak seolah-olah pemicu stres tidak nyata dan menyalahkan diri sendiri (Ewert, dkk., 2021).

Finlay-Jones (2017) menemukan bahwa lebih menjadi lebih *self-compassionate* dalam kehidupan sehari-hari dapat memengaruhi sifat stresor yang dihadapi individu (misalnya, melalui perilaku *self-care* yang protektif), penilaian (*reappraisal*) terhadap stresor yang dihadapi, tanggapan fisiologis dan afektif terhadap stres, dan *coping*

*strategy* yang digunakan individu saat mencoba mengatasi stres. Memiliki sikap *self-compassion* memungkinkan individu untuk membawa kesadaran pada emosi mereka, mendekati kesulitan dengan kebaikan dan pengertian, serta memiliki rasa kemanusiaan bersama. Dengan pendekatan yang lebih adaptif terhadap stresor, individu dapat secara bertahap mengubah emosi negatif mereka menjadi lebih positif, memiliki pemahaman situasi yang lebih baik, juga mengadopsi *coping strategy* yang lebih efektif untuk mengatasi stresor mereka (Neff, 2003). Dari paparan di atas, dapat disimpulkan bahwa *self-compassion* berperan sebagai faktor protektif untuk stres.

Selain berperan sebagai faktor protektif stres, *self-compassion* juga berkaitan dengan *well-being* karena perannya dalam menciptakan *mindset* positif (Zessin, dkk., 2015) dan membantu orang merasa aman serta terlindungi (Neff, 2011). Dalam meta-analisis tersebut (Zessin, dkk., 2015), dibuktikan adanya keterkaitan antara berbagai jenis *well-being* dengan *self-compassion*—dan bahkan ditemukan adanya hubungan sebab-akibat antar keduanya di beberapa subsampel. *Self-compassion* juga termasuk salah satu mediator dalam hubungan antara stres akademik dan *well-being*, yang artinya *self-compassion* dapat mengurangi dampak dari stres akademik terhadap *well-being* mahasiswa (Poots & Cassidy, 2020). Studi lain oleh Stoerber, dkk., (2020) menemukan dampak positif *self-compassion* pada *subjective well-being*. Sementara studi Karakasidou, dkk. (2023) mengkonfirmasi bahwa *self-compassion* dapat memprediksi *emotional well-being* dan mendukung gagasan bahwa *self-compassion* dapat berperan sebagai faktor protektif terhadap dampak negatif pada kesehatan mental. Karakasidou, dkk. (2023) bahkan menyarankan perlunya intervensi yang meningkatkan *self-compassion* untuk mendukung kesehatan mental dan *well-being* secara keseluruhan pada mahasiswa selama pandemi COVID-19.

Berbagai pelatihan untuk meningkatkan *self-compassion* (misalnya, “*Gestalt two-chair*” *exercise*, *compassionate mind training exercise*, dan *mindful self-compassion program*) dapat berfungsi sebagai strategi regulasi emosi yang efektif, yang akhirnya menyebabkan peningkatan kebahagiaan dan pengurangan tekanan emosional, depresi, kecemasan, serta peningkatan kemampuan untuk menenangkan dan meyakinkan diri sendiri. Sebagian besar intervensi *self-compassion* memerlukan bantuan dan pedoman dari profesional untuk diterapkan—yang membuatnya kurang layak untuk digunakan sebagai alat *self-help* (Wong & Mak, 2016), terlebih lagi sebagai strategi *coping* yang dibutuhkan penggunaannya dalam menghadapi stres sehari-hari.

Sebagai alternatif, pendekatan berbasis tulisan seperti *Written Emotional Disclosure Paradigm* (WEDP) menawarkan bentuk intervensi yang lebih sederhana dan mudah diakses. Intervensi ini—dalam bentuk awalnya yang disebut *expressive writing* atau menulis ekspresif—pertama kali diperkenalkan secara sistematis oleh Pennebaker & Beall (1986) dan menjadi dasar bagi lahirnya berbagai model intervensi menulis. Hingga 2021, telah terdapat 1.809 studi intervensi menulis dengan berbagai model (Gao, 2022). Selain model tradisional berupa *expressive writing* (Vukčević Marković, dkk., 2020; Mariën, dkk., 2022; Argudo, 2021; Castillo, dkk., 2019), jenis penulisan lainnya yaitu *positive writing* (Mariën, dkk., 2022), *benefit-finding writing* (Chu, dkk., 2020; Gallagher, dkk., 2020) *self-esteem writing* (Barbeau, dkk., 2022; Dupasquier, dkk., 2020) hingga model penulisan terbaru yaitu *self-compassion writing* (Dupasquier, dkk., 2020; Dreisoerner, dkk., 202; Urken & LeCroy, 2021; Wong & Mak, 2016; Ziemer, dkk., 2019). Intervensi menulis banyak digunakan sebagai alat *self-help* karena terbukti bermanfaat bagi kesehatan mental maupun fisik (Frattaroli, 2006). Selain karena murah dan mudah dilakukan, intervensi ini juga dapat diakses oleh individu yang enggan atau tidak memiliki kesempatan untuk mengikuti psikoterapi formal (Pascoe, 2017).

Pemilihan *expressive writing* sebagai titik awal dalam penelitian ini didasarkan pada relevansinya dengan konteks stres mahasiswa. Melalui kegiatan menulis ekspresif, individu diberi ruang untuk mengungkapkan dan mengolah pikiran serta perasaan yang selama ini ditekan. Sebagaimana dijelaskan oleh Zheng, dkk. (2019), *expressive writing* dapat meningkatkan proses kognitif dan regulasi stres melalui pencarian makna terhadap pengalaman yang menekan. Hasil tinjauan pustaka oleh Savera, dkk. (2022) terhadap 30 studi nasional dan internasional juga memperkuat temuan tersebut: *expressive writing* efektif dalam menurunkan berbagai bentuk stres serta membantu individu menghadapi depresi, trauma, dan tekanan emosional lainnya. Penelitian lain juga mendukung manfaat serupa, di mana *expressive writing* mampu menurunkan stres akademik dan meningkatkan kesejahteraan mahasiswa (Argudo, 2021).

Dalam perkembangannya, pendekatan *expressive writing* diadaptasi untuk menekankan dimensi penerimaan diri dan kebaikan terhadap diri sendiri, melahirkan model *self-compassion writing*. Seperti dijelaskan oleh Wong & Mak (2016), dengan mengadopsi metodologi *expressive writing*, *self-compassion writing* dapat menjadi cara yang efektif dan terjangkau untuk meningkatkan *self-compassion*, regulasi emosi, dan *well-being*.

Intervensi *self-compassion writing* termasuk protokol penulisan terbaru dengan studi pertama dilakukan di Amerika Serikat oleh Leary, dkk. (2007). Dalam satu-satunya *systematic review* yang membahas *self-compassion writing* (Aydoğdu & Dirik, 2022), 31 artikel menunjukkan beragam manfaat pada kelompok yang menerima intervensi *self-compassion writing*, yaitu: peningkatan afek positif, penurunan afek negatif, peningkatan rasa percaya diri, penurunan depresi, kinerja mental yang lebih baik, dan pengurangan gejala fisik seperti nyeri. Aydoğdu & Dirik (2022) menyimpulkan bahwa *self-compassion writing* adalah metode intervensi *self-help* yang tidak memerlukan biaya dan dapat menghasilkan efek dalam waktu singkat, serta dapat diterapkan pada kelompok sampel yang berbeda, dalam peristiwa kehidupan yang berbeda, dan memberikan hasil yang layak serta efektif.

Penelitian mengenai *self-compassion writing* sudah cukup banyak dilakukan di luar negeri dan terbukti memberi manfaat bagi kesehatan mental. Namun, konteks budaya Indonesia memiliki tantangan unik. Mahasiswa Indonesia umumnya tidak memiliki kebiasaan menulis secara rutin (Fahmi, 2018), terlebih lagi menggunakan menulis sebagai sarana ekspresi emosional maupun refleksi diri. Hal ini sejalan dengan temuan lintas budaya bahwa individu dengan latar Asia cenderung menekan ekspresi emosional karena pengaruh nilai budaya seperti rasa malu dan rasa bersalah (Liw, dkk., 2022), sehingga ekspresi emosi secara terbuka—termasuk melalui tulisan—belum menjadi praktik yang lazim.

Penelitian tentang *self-compassion writing* di Indonesia pun masih sangat terbatas. Hingga kini, baru terdapat tiga studi: Windiastuti & Kaloeti (2022) pada mahasiswa dengan depresi ringan, Muchsin & Murdiana (2024) pada perempuan korban *body shaming*, serta Inaya, Bakar, & Hamid (2024) yang meneliti regulasi emosi pada mahasiswa baru. Kebaruan pada penelitian ini adalah penggunaan intervensi *self-compassion writing* sebagai model strategi *coping stres* bagi mahasiswa Indonesia dengan menggunakan pendekatan eksperimen. Berdasarkan hal tersebut, penelitian ini menguji hipotesis bahwa intervensi *self-compassion writing* akan (1) menurunkan tingkat stres, (2) meningkatkan *well-being*, serta (3) meningkatkan *self-compassion* mahasiswa.

## 2. Metode Penelitian

### Desain Penelitian dan Sampel

Penelitian ini menggunakan rancangan eksperimen *within-subjects design*. Partisipan akan diberikan intervensi *self-compassion writing* selama tiga hari berturut-turut setelah diberikan *pre-test*. *Post-test* diukur pada hari ketiga, dan *follow-up* diukur pada 2 minggu setelah intervensi diberikan. Teknik pengambilan sampel menggunakan metode *purposive sampling*, yaitu penentuan sampel dengan pertimbangan tertentu (Sugiyono, 2014). Kriteria inklusi partisipan dalam penelitian ini adalah: 1) mahasiswa aktif S1 Fakultas Pendidikan Psikologi, Universitas Negeri Jakarta, 2) memiliki tingkat stres sedang hingga tinggi (diukur dengan skala stres pada saat rekrutmen *pre-experiment*), 3) memahami penggunaan Google Form, dan 4) bersedia mengikuti pelatihan menulis (diukur dari kesediaan memberikan *inform consent*). Skor batas tingkat stres sedang hingga tinggi dipilih untuk memungkinkan lebih banyak ruang untuk mendeteksi perubahan dan untuk menawarkan intervensi kepada mereka yang berisiko tertinggi.

Ukuran sampel ditentukan dengan *a priori power analysis* menggunakan G\*Power 3.1 untuk desain *within-subjects repeated measures ANOVA* ( $\alpha = 0,05$ ,  $power = 0,80$ ,  $effect\ size\ f = 0,25$ ). Hasil perhitungan menunjukkan minimal 28 partisipan diperlukan. Dengan demikian, jumlah akhir partisipan yang menyelesaikan penelitian ( $N = 32$ ) memenuhi syarat untuk mendeteksi efek sedang.

### Prosedur

Prosedur *self-compassion writing* yang digunakan diadaptasi dari penelitian Leary, dkk. (2007) serta Wong & Mak (2016). *Self-compassion writing* dapat diartikan sebagai bentuk intervensi menulis berbasis *self-compassion*; yaitu menulis dengan memasukkan komponen-komponen *self-compassion* (*mindfulness*, *common humanity*, *self-kindness*). *Self-compassion writing* akan mengarahkan individu untuk menuliskan pengalaman yang sulit, tidak menyenangkan, atau peristiwa yang membuat seseorang merasa buruk tentang diri sendiri—sesuatu yang melibatkan kegagalan, penghinaan, atau penolakan. Kemudian, individu diminta merespons pengalaman tersebut dengan sikap *self-compassion*, yaitu dengan 1) menyadari dan menerima pikiran-perasaan yang muncul, 2) mengingat bahwa peristiwa buruk adalah bagian dari pengalaman hidup dan orang lain pun mengalami hal yang sama, juga 3) menyampaikan kepedulian, kasih sayang, dan kelembutan kepada diri sendiri.

Tiga komponen *self-compassion* akan digunakan dalam intervensi penulisan selama 3 hari berturut-turut. Dalam studi ini, yang membedakan instruksi per hari hanyalah konteks peristiwa negatif yang dialami oleh partisipan. Pada hari pertama, konteks peristiwa negatif yang ditulis terkait dengan tempat kuliah (contoh: isu terkait biaya kuliah, masalah finansial, tugas perkuliahan, beban akademik, relasi dengan dosen, dan sebagainya). Pada hari kedua, konteks peristiwa negatif yang ditulis terkait dengan rumah (contoh: permasalahan di rumah, kejadian atau hal di masa lalu yang hingga kini masih membekas, relasi dengan kerabat atau orangtua, dan sebagainya). Pada hari ketiga, konteks peristiwa negatif yang ditulis terkait dengan relasi sosial (contoh: isu dengan teman atau pasangan). Ketiga konteks peristiwa negatif dipilih berdasarkan pertimbangan bahwa sumber penyebab stres mahasiswa beragam, baik berupa faktor akademik maupun non-akademik, mulai dari isu keuangan, percintaan, hingga hubungan dengan teman dan keluarga (Karyotaki, 2020).

Tiap sesi dalam intervensi *self-compassion writing* akan berlangsung maksimal selama 60 menit, dengan alokasi waktu penulisan selama kurang lebih 20 menit. Psikoedukasi mengenai *self-compassion* akan diberikan pada hari pertama, sebelum sesi menulis pertama dilakukan. Psikoedukasi dilakukan berdasarkan temuan dari studi Williamson dan Blackhart (2021), bahwa tingkat *self-compassion* partisipan yang diberikan penjelasan tambahan semakin meningkat dibandingkan mereka yang tidak. Partisipan yang diberikan penjelasan mengingat lebih banyak detail tentang *self-compassion* dan mereka merasa instruksinya lebih mudah dipahami. Meski

instruksi dalam *self-compassion writing* diberikan dalam bahasa yang jelas, *self-compassion* mungkin merupakan konsep yang baru bagi partisipan (Williamson dan Blackhart, 2021).

Dalam pelaksanaannya, seorang fasilitator akan mendampingi partisipan untuk membacakan instruksi penulisan, memberikan penjelasan, dan menjawab pertanyaan selama intervensi berlangsung. Tiap partisipan diberikan sebuah modul berisikan panduan penulisan dan sebagai wadah untuk menulis. Modul juga dilengkapi dengan instruksi penulisan secara detil yang mengarahkan partisipan untuk menulis tiap komponen *self-compassion*.

### **Kontrol Terhadap Variabel Luar (*Extraneous Variables*)**

Untuk meminimalkan pengaruh variabel luar yang dapat memengaruhi hasil intervensi, prosedur penelitian dirancang dengan sejumlah kontrol standar. Semua sesi intervensi dilaksanakan dalam kondisi yang seragam: instruksi disampaikan oleh fasilitator yang sama menggunakan modul tertulis yang distandarisasi; setiap sesi berlangsung pada ruangan yang tenang dan terjadwal pada rentang waktu yang konsisten (selama 3 hari, pukul 08.00—09.00 WIB) untuk mengurangi variasi karena waktu. Penulisan dilakukan selama  $\pm 20$  menit per sesi dengan total durasi sesi maksimal 60 menit, termasuk alokasi waktu untuk psikoedukasi singkat, tanya jawab, serta memberi kesempatan partisipan menyelesaikan tulisannya. Partisipan diminta menonaktifkan perangkat komunikasi (misalnya, notifikasi ponsel), tidak berbicara dengan sesama partisipan, dan fokus penuh pada aktivitas menulis selama sesi berlangsung.

Sebelum intervensi, partisipan diseleksi melalui kriteria inklusi/eksklusi yang mencakup: mahasiswa aktif S1 Fakultas Pendidikan Psikologi, Universitas Negeri Jakarta, 2) memiliki tingkat stres sedang hingga tinggi (diukur dengan skala stres pada saat rekrutmen *pre-experiment*, 3) memahami penggunaan Google Form, dan 4) bersedia mengikuti pelatihan menulis (diukur dari kesediaan memberikan *inform consent*). Kepatuhan terhadap instruksi diverifikasi melalui *manipulation check* saat partisipan selesai menulis (selama 3 hari), berupa persepsi subjektif terhadap intervensi menulis dengan skala Likert dan 1 *open-ended question*.

### **Instrumen Penelitian dan Teknik Analisis Statistik**

Tingkat stres diukur dengan instrumen *Perceived Stress Scale* (PSS). PSS disusun oleh Cohen (1994) dan diadaptasi dalam Bahasa Indonesia oleh Argalina (2019). PSS adalah instrumen psikologis yang paling banyak digunakan untuk mengukur persepsi terhadap stres. Item dirancang untuk menilai seberapa tidak terduga, tidak terkendali, dan kelebihan beban (*overload*) yang dirasakan responden dalam hidup mereka. Pertanyaan dalam PSS bersifat umum sehingga relatif bebas dari konten spesifik untuk kelompok sub-populasi apapun. Instrumen PSS terdiri dari 10 item dengan contoh item seperti “Selama sebulan terakhir, seberapa sering Anda marah karena sesuatu yang tidak terduga?”. Ke-10 item diukur dengan skala Likert 5 poin (0=tidak pernah hingga 4=sangat sering) dengan cara skoring yaitu seluruh skor harus dijumlahkan. Skor berkisar antara 1-13 berarti stres rendah, skor berkisar antara 14-26 berarti stres sedang, sementara skor berkisar antara 27-40 berarti stres tinggi.

*Well-being* diukur menggunakan *Satisfaction with Life Scale* (SWLS). SWLS telah diadaptasi ke dalam Bahasa Indonesia oleh Muttaqin (2022). SWLS adalah instrumen singkat berisi 5 item yang dirancang untuk mengukur penilaian kognitif global tentang kepuasan seseorang terhadap hidupnya. SWLS biasanya memerlukan waktu sekitar satu menit untuk diisi dengan skala Likert 7 (1=sangat tidak setuju, 7=sangat setuju). Pertanyaan-pertanyaan dalam SWLS terbuka untuk interpretasi, menjadikan SWLS cocok untuk orang dewasa dengan berbagai latar belakang dan paling sesuai digunakan dalam populasi non-klinis.

Tingkat *self-compassion* diukur menggunakan instrumen *Self-Compassion Scale* (SCS) yang disusun oleh Neff (2003) kemudian diadaptasi oleh Sugianto, dkk. (2020). Instrumen SCS terdiri dari 6 subskala dengan total 26 item dan diukur dengan skala Likert 5 (1=hampir tidak pernah hingga 5=hampir selalu). Untuk menghitung skor total *self-compassion*, skor butir-butir subskala negatif (*self-judgement*, *isolation*, *overidentification*) diberi skor secara terbalik. Rata-rata tiap komponen akan dihitung, kemudian dihitung rata-rata totalnya (*grand mean*).

Pada penelitian ini, *manipulation check* diberikan saat partisipan selesai menulis (selama 3 hari berturut-turut), berupa persepsi subjektif terhadap intervensi menulis dengan skala Likert dan 1 *open-ended question* (Jelaskan tugas apa yang sedang diberikan padamu?). Terdapat lima pertanyaan yang diadaptasi dari studi *self-compassion writing* oleh Ziemer, dkk. (2019). Partisipan akan menjawab skala Likert 1 (tidak sama sekali) sampai 7 (sangat banyak) dengan pertanyaan sebagai berikut: 1) Seberapa banyak tulisanmu berisi hal-hal yang privat tentang dirimu? 2) Sejauh mana kamu mengungkapkan emosi dalam tulisanmu? 3) Apakah kamu yakin bahwa apa yang kamu tulis mempengaruhi cara kamu berpikir tentang kejadian tersebut? 4) Apakah kamu yakin bahwa menghadapi sesuatu yang ada dalam tulisanmu mampu mengubah caramu merasa tentang hal tersebut? 5) Sejauh mana menulis tentang hal tersebut membuat kamu merasa dipahami dan lebih menerima dirimu sendiri.

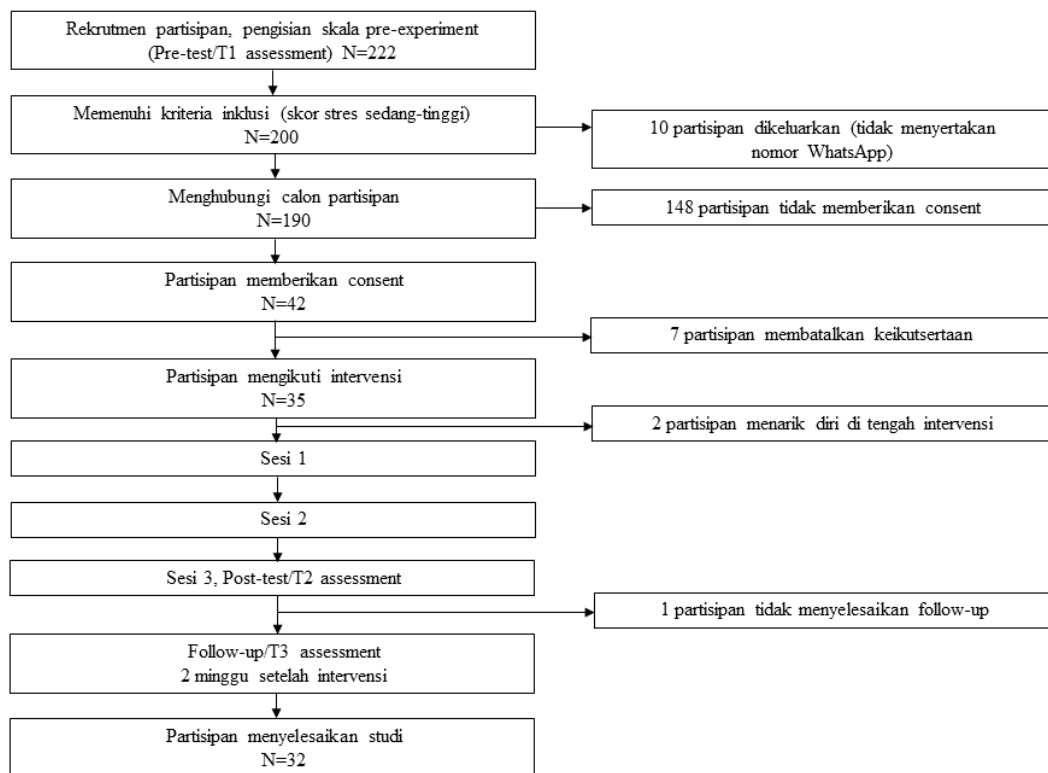
Selain itu, *program's satisfaction & acceptability* ditanyakan kepada partisipan pada saat pengisian *follow-up* untuk menggali kepuasan terkait program menulis, berupa 3 item yang akan dijawab dengan skala Likert 1 (sangat tidak setuju) hingga 5 (sangat setuju). Item pertanyaan sebagai berikut: 1) Apakah kamu puas dengan program

ini? 2) Apakah kamu ingin terus melanjutkan menulis? 3) Apakah kamu ingin merekomendasikan program ini kepada orang lain?

Analisis statistik untuk menguji asumsi dan hipotesis menggunakan bantuan *software* SPSS 25. Uji asumsi yang dilakukan yaitu uji normalitas, sementara uji hipotesis yang digunakan adalah uji *Repeated Measure ANOVA* untuk melihat ada tidaknya perbedaan dari hasil penelitian yang dilakukan berulang pada suatu variabel penelitian dengan minimal 3 sampel berpasangan.

### 3. Hasil dan Diskusi

Dalam proses rekrutmen partisipan, peneliti berhasil menjangkau sebanyak 222 mahasiswa yang kemudian diminta untuk mengisi skala *pre-experiment* sebagai *baseline*. Dari jumlah tersebut, 200 mahasiswa memenuhi kriteria inklusi. Kemudian, 10 mahasiswa dikeluarkan karena tidak memberikan nomor WhatsApp sehingga tidak dapat dihubungi. Selanjutnya, peneliti menghubungi 190 mahasiswa, mengirimkan detail penelitian beserta *informed consent*, dan menawarkan keikutsertaan kepada calon partisipan. Dari 190 mahasiswa, hanya 42 mahasiswa yang bersedia dan mengonfirmasi keikutsertaan. Tujuh mahasiswa kemudian membatalkan keikutsertaan dengan berbagai alasan (sakit, sedang berada dalam kondisi mental tidak baik, jadwal intervensi yang tidak sesuai, hingga mundur tanpa pemberitahuan). Dua partisipan menarik diri di tengah intervensi dan 1 partisipan tidak menyelesaikan asesmen *follow-up*. Tingkat putus partisipasi adalah 23,8%; 32 dari 42 partisipan yang dikonfirmasi menyelesaikan penelitian, yang didefinisikan sebagai mengikuti intervensi penulisan selama 3 hari dan menyelesaikan *follow-up* 2 minggu setelah intervensi (lihat Gambar 1 untuk rincian).



Gambar 1. Proses Rekrutmen Partisipan

**Tabel 1. Deskripsi Karakteristik Partisipan**

Karakteristik	n (%)
<b>Jenis Kelamin</b>	
Perempuan	9 (28%)
Laki-laki	23 (72%)
<b>Usia</b>	
18 Tahun	1 (3.1%)
19 Tahun	16 (50%)
20 Tahun	11 (34.3%)
21 Tahun	3 (9.3%)
22 Tahun	1 (3.1%)
<b>Mahasiswa Tahun ke-</b>	
Tahun kedua	32 (100%)
<b>Senang mengekspresikan perasaan/pengalaman/pikiran tentang kehidupan sehari-hari melalui menulis</b>	
Ya	20 (62.5%)
Tidak	12 (37.5%)
<b>Pernah mengekspresikan perasaan/pengalaman/pikiran tentang kehidupan sehari-hari melalui tulisan?</b> Misalnya melalui journaling, gratitude journaling, bullet journaling, atau semacamnya	
Pernah	23 (72%)
Tidak	9 (28%)
<b>Seberapa sering mengekspresikan perasaan/pengalaman/pikiran tentang kehidupan sehari-hari melalui tulisan</b>	
Tidak pernah	6 (18.7%)
1-2 kali dalam setahun	6 (18.7%)
1-2 kali dalam sebulan	11 (34.3%)
1-3 kali dalam seminggu	7 (21.8%)
5-7 kali dalam seminggu	2 (6.3%)

Hasil dari uji normalitas menunjukkan distribusi normal pada variabel stres, *well-being*, dan *self-compassion* dengan melihat skor *Shapiro-Wilk*. Nilai p signifikan pada ketiga variabel seluruhnya lebih besar dari nilai signifikansi  $\alpha = 0.05$  yang berarti data normal.

**Tabel 2. Hasil Uji Normalitas**

Uji normalitas (N = 32)	Kolmogorov-Smirnov	Shapiro-Wilk
<b>Pre-test</b>		
Stres	0.200	0.322
Well-being (WB)	0.200	0.200
Self-compassion	0.200	0.069
<b>Post-test</b>		
Stres	0.200	0.435
Well-being (WB)	0.200	0.401
Self-compassion	0.197	0.138
<b>Follow-up (2 minggu)</b>		
Stres	0.063	0.247
Well-being (WB)	0.066	0.063
Self-compassion	0,200	0.451

**Tabel 3. Hasil Uji Hipotesis**

Variabel	<i>Repeated Measure ANOVA (N=32)</i>						F (2, 30)	$\eta^2$	Wilk's Lambda	p
	Pre		Post		Follow-up					
	M	SD	M	SD	M	SD				
Stres	18.94	5.65	19.16	6.77	17.34	6.04	2.30	0.13	0.87	0.12
WB	4.07	1.03	3.98	0.88	4.11	1.05	0.40	0.03	0.97	0.40
SC	3.19	0.49	3.22	0.46	3.31	0.46	3.01	0.17	0.83	0.06

Setelah diketahui data terdistribusi normal, selanjutnya dilakukan analisis *Repeated Measure ANOVA*. Berdasarkan hasil perhitungan, ditemukan bahwa *self-compassion writing* tidak efektif untuk menurunkan skor stres pada mahasiswa. Analisis *Repeated Measure ANOVA* menemukan bahwa Wilk's Lambda = 0.867,  $F(2,30) = 2.30$ ,  $p = 0.118$ ,  $\eta^2 = 0.13$ . Selanjutnya, *self-compassion writing* tidak efektif untuk menaikkan skor *well-being* pada mahasiswa. Analisis *Repeated Measure ANOVA* menemukan bahwa Wilk's Lambda = 0.974,  $F(2,30) = 0.399$ ,  $p = 0.675$ ,  $\eta^2 = 0.03$ . Terakhir, *self-compassion writing* tidak efektif untuk menaikkan skor *self-compassion* pada mahasiswa. Analisis *Repeated Measure ANOVA* menemukan bahwa Wilk's Lambda = 0.833,  $F(2,30) = 3.011$ ,  $p = 0.064$ ,  $\eta^2 = 0.17$ .

*Program's satisfaction & acceptability* ditanyakan kepada partisipan untuk menggali kepuasan terkait program menulis, berupa 3 item yang dijawab dengan skala Likert 1 (sangat tidak setuju) hingga 5 (sangat setuju). Sebanyak 93,7% partisipan dilaporkan puas dengan program menulis, dan 90,6% menyarankan intervensi ini untuk orang lain. Sekitar 62,5% responden yang menunjukkan niat untuk terus menulis (lihat Tabel 4).

**Tabel 4. Program's Satisfaction and Acceptability**

Pertanyaan (N=32)	Sangat tidak setuju	Tidak Setuju	Netral	Setuju	Sangat Setuju
Apakah kamu puas dengan program ini?	0	0	2 (6.3%)	18 (56.2%)	12 (37.5%)
Apakah kamu ingin terus melanjutkan menulis?	0	1 (3.1%)	11 (34.4%)	11 (34.4%)	9 (28.1%)
Apakah kamu ingin merekomendasikan program ini kepada orang lain?	0	1 (3.1%)	2 (6.3%)	20 (62.5%)	9 (28.1%)

### Pembahasan

Hasil penelitian menunjukkan bahwa *self-compassion writing* tidak efektif dalam menurunkan stres, meningkatkan *well-being*, maupun meningkatkan *self-compassion* mahasiswa. Meskipun *systematic review* oleh Aydoğdu & Dirik (2022) menunjukkan beragam manfaat pada kelompok yang menerima intervensi *self-compassion writing*, penelitian ini menemukan hasil yang kontraindikatif. Sejumlah faktor kemungkinan berkontribusi terhadap ketidakefektifan tersebut, mulai durasi dan prosedur intervensi, tingkat pemahaman peserta terhadap konsep *self-compassion*, hingga konteks budaya menulis dan ekspresi emosional di Indonesia.

Efektivitas *Written Emotional Disclosure Paradigm* (WEDP) atau pengungkapan emosi melalui tulisan—dengan bentuk paling awal berupa *expressive writing*—telah menjadi topik penelitian sejak 1980-an. Namun, hingga kini hasilnya masih bervariasi (Pennebaker dan Chung, 2007). Beberapa meta-analisis *expressive writing* menunjukkan inkonsistensi efek; Smyth (1998) menemukan bahwa *expressive writing* lebih efektif jika dilakukan dalam waktu yang lebih panjang dan diselingi antar-sesi—bahwa menulis 1 kali setiap minggu selama 1 bulan mungkin lebih efektif daripada menulis 4 kali dalam 1 minggu—sedangkan Frattaroli (2006) menemukan bahwa sekitar 25% studi menunjukkan efek negatif dan 70% studi menunjukkan efek positif. Frattaroli juga mengidentifikasi beberapa karakteristik penelitian yang cenderung menghasilkan efek positif: (1) partisipan memiliki masalah kesehatan atau pengalaman traumatis, (2) kondisi menulis yang nyaman dan privat, (3) adanya insentif partisipasi, (4) jumlah sesi minimal tiga kali, (5) topik yang masih relevan secara emosional, dan (6) evaluasi tindak lanjut (*follow-up*) yang dilakukan dalam waktu relatif singkat (kurang dari 1 bulan).

Temuan-temuan ini menunjukkan bahwa keberhasilan intervensi menulis sangat bergantung pada rancangan dan konteks pelaksanaannya. Jika dikaitkan dengan penelitian ini, durasi 3 hari berturut-turut dapat dikategorikan terlalu singkat untuk menghasilkan perubahan bermakna. Selain itu, tema penulisan yang mencakup tiga ranah berbeda (akademik, rumah, dan relasi sosial) mungkin membuat proses refleksi partisipan menjadi terlalu tersebar.

Akibatnya, emosi yang diungkapkan dalam tulisan tidak berfokus pada sumber stres yang paling dominan bagi masing-masing individu, sehingga efek terhadap *perceived stress* secara umum menjadi lemah.

Dalam konteks penelitian ini, hasil yang tidak signifikan juga dapat dikaitkan dengan sifat intervensinya sendiri yang menuntut pemahaman lebih kompleks dibandingkan *expressive writing*. Berbeda dengan *expressive writing* yang relatif mudah diterapkan tanpa psikoedukasi—yaitu meminta partisipan menuliskan pikiran, perasaan, dan pengalaman terkait peristiwa traumatis & emosional (Pennebaker & Evans, 2014)—*self-compassion writing* menuntut pemahaman terhadap tiga komponen utama: *mindfulness*, *common humanity*, dan *self-kindness*. Tanpa pemahaman yang mendalam, penulisan bisa jadi hanya menjadi pengungkapan emosi, bukan refleksi yang penuh *self-compassion*.

Meskipun telah diberikan psikoedukasi sebelum menulis, dalam penelitian ini terdapat partisipan yang gagal mengikuti instruksi dengan benar—fenomena yang juga ditemukan pada penelitian Williamson (2014) dan Windiasuti & Kaloeti (2022). Dalam penelitian ini, psikoedukasi singkat mungkin belum cukup untuk memunculkan perubahan emosional signifikan. Sebaliknya, program *Mindful Self-Compassion* (MSC) oleh pionir *self-compassion*, Neff & Germer (2013) yang berlangsung selama 8 minggu penuh terbukti menghasilkan peningkatan *self-compassion*, *mindfulness*, dan *well-being* yang bertahan hingga 6 bulan kemudian. Selain menulis, peserta juga memiliki ragam kegiatan lain; contohnya, peserta dalam kelompok kecil berdiskusi tentang bahasa yang mereka gunakan saat mengkritik diri sendiri dan cara menggunakan bahasa yang lebih lembut dan mendukung. Peserta juga diajarkan latihan informal seperti meletakkan tangan di dada saat stres atau mengulangi frasa *self-compassion* untuk digunakan dalam kehidupan sehari-hari.

Meta-analisis Han & Kim (2023) juga menemukan bahwa intervensi *self-compassion* efektif menurunkan stres hanya jika dilakukan rata-rata selama 5 minggu atau lebih dan mencakup latihan tambahan seperti meditasi, visualisasi, hingga praktik harian dengan panduan berupa rekaman audio. *Self-compassion writing* menjadi bagian dari rangkaian intervensi yang diberikan, tetapi bukan satu-satunya intervensi. Dengan demikian, pendeknya durasi intervensi dan keterbatasan pemahaman peserta terhadap konsep *self-compassion* dapat menjadi alasan utama atas tidak signifikannya hasil penelitian ini.

Selain itu, penelitian oleh Swee dkk. (2023) menunjukkan bahwa peningkatan *self-compassion* cenderung merupakan proses yang lebih lambat dibandingkan penurunan emosi negatif seperti rasa malu atau kritik diri. Hal ini menunjukkan bahwa perubahan pada *self-compassion* mungkin memerlukan waktu dan latihan lebih panjang agar efeknya muncul.

Faktor lain yang perlu dipertimbangkan adalah konteks budaya dan kebiasaan menulis mahasiswa Indonesia. Penelitian menunjukkan bahwa mahasiswa di Indonesia umumnya tidak memiliki kebiasaan menulis secara rutin (Fahmi, 2018), terlebih menggunakan tulisan sebagai sarana ekspresi emosional maupun refleksi diri. Hal ini diperkuat oleh temuan lintas budaya bahwa individu dengan latar Asia cenderung menekan ekspresi emosional karena nilai-nilai budaya seperti rasa malu dan rasa bersalah (Liw dkk., 2022). Dalam konteks seperti ini, menulis reflektif dengan nuansa *self-compassion* mungkin terasa asing dan sulit diinternalisasi.

Dengan demikian, hasil non-signifikan dalam penelitian ini kemungkinan tidak hanya disebabkan oleh durasi yang singkat, tetapi juga oleh kesenjangan antara karakteristik intervensi dengan kesiapan budaya dan emosional partisipan.

#### 4. Kesimpulan

Penelitian ini menguji efektivitas *self-compassion writing* terhadap stres, *well-being*, dan *self-compassion* mahasiswa di Fakultas Psikologi Pendidikan, Universitas Negeri Jakarta. Hasil penelitian menunjukkan bahwa *self-compassion writing* tidak efektif dalam menurunkan stres, meningkatkan *well-being*, dan meningkatkan *self-compassion*.

Temuan ini menyoroti bahwa efektivitas *self-compassion writing* sangat dipengaruhi oleh kedalaman psikoedukasi, durasi dan karakteristik intervensi, fokus tema, serta konteks budaya peserta. Untuk penelitian selanjutnya, diperlukan durasi intervensi yang lebih panjang, topik yang lebih terfokus, serta kombinasi dengan latihan *self-compassion* lain seperti meditasi atau *compassionate body scan*. Analisis kualitatif terhadap isi tulisan juga perlu dilakukan untuk memastikan sejauh mana partisipan benar-benar mengintegrasikan ketiga komponen *self-compassion* dalam proses menulisnya.

Selain memberikan kontribusi metodologis, temuan ini juga memiliki implikasi praktis bagi pengembangan intervensi psikologis di lingkungan pendidikan. Meskipun hasilnya tidak signifikan, penelitian ini mengindikasikan bahwa *self-compassion* bukanlah keterampilan yang dapat dipelajari secara instan. Mahasiswa membutuhkan proses pembelajaran yang lebih mendalam dan berkelanjutan agar mampu memahami serta menginternalisasi nilai-nilai *self-compassion*. Oleh karena itu, intervensi berbasis *self-compassion writing* sebaiknya diposisikan sebagai langkah awal menuju program pengembangan diri yang lebih komprehensif, bukan sebagai satu-satunya strategi peningkatan kesejahteraan psikologis.

Selanjutnya, faktor budaya dan kebiasaan menulis mahasiswa Indonesia yang belum terbiasa mengekspresikan emosi melalui tulisan perlu mendapat perhatian khusus. Fasilitator direkomendasikan untuk memberikan

pendampingan intensif, psikoedukasi yang kontekstual, serta contoh penulisan reflektif agar peserta dapat menulis dengan lebih efektif dan penuh *self-compassion*.

Dengan demikian, penelitian ini memberikan kontribusi penting dalam mengembangkan intervensi *self-compassion* yang lebih kontekstual, berjangka panjang, dan relevan dengan karakteristik mahasiswa Indonesia.

## 5. Referensi

- American College Health Association. (2022). *American College Health Association-National College Health Assessment III: Reference Group. Executive Summary Fall 2021*. Silver Spring, MD: American College Health Association.
- American Psychological Association. (2020). *Stress in America™ 2020: A national mental health crisis*.
- Antara News. (2022). PDSKJI: Pemerataan rumah sakit jiwa strata utama rampung 2027. <https://www.antarane.ws.com/berita/3159909/pdskji-pemerataan-rumah-sakit-jiwa-strata-utama-rampung-2027>
- Argalina, B. M. (2019). *Gambaran tingkat stres pada siswa dalam penerapan full day school SMA Negeri di Kota Surakarta* [Skripsi, Sekolah Tinggi Ilmu Kesehatan ‘Aisyiyah Surakarta (AISKA)].
- Argudo, J. (2021). Expressive writing to relieve academic stress at university level. *Profile: Issues in Teachers’ Professional Development*, 23(2), 17–33. <https://doi.org/10.15446/profile.v23n2.90448>
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469–480. <https://doi.org/10.1037/0003-066X.55.5.469>
- Aydođdu, B. E., & Dirik, G. (2022). Self-compassion-focused writing practices: A systematic review. *Current Approaches in Psychiatry/Psikiyatri Guncel Yaklasimlar*, 14(4). <https://doi.org/10.18863/pgy.1004339>
- Badan Penelitian dan Pengembangan Kesehatan Kementerian RI. (2018). *Riset Kesehatan Dasar (Riskesdas)*. Kementerian Kesehatan Republik Indonesia.
- Barbeau, K., Guertin, C., Boileau, K., & Pelletier, L. (2022). The effects of self-compassion and self-esteem writing interventions on women’s valuation of weight management goals, body appreciation, and eating behaviors. *Psychology of Women Quarterly*, 46(1), 82–98. <https://doi.org/10.1177/03616843211012411>
- Castillo, Y. A., Fischer, J., & Davila, S. (2019). Managing emotions: Relationships among expressive writing and emotional intelligence. *Integrated Research Advances*, 6(1). <http://pubs.iscience.in/ira>
- CCMH. (2021). *Center for Collegiate Mental Health 2021 Annual Report*. Pennsylvania State University, University Park, PA.
- CCMH. (2022). *Center for Collegiate Mental Health 2022 Annual Report*. Pennsylvania State University, University Park, PA.
- Chu, Q., Wu, I. H. C., & Lu, Q. (2020). Expressive writing intervention for posttraumatic stress disorder among Chinese American breast cancer survivors: The moderating role of social constraints. *Quality of Life Research*, 29(4), 891–899. <https://doi.org/10.1007/s11136-019-02385-5>
- Cohen, S. (1994). *Perceived Stress Scale*. Mind Garden.
- Conley, C. S., Shapiro, J. B., Huguene, B. M., & Kirsch, A. C. (2020). Navigating the college years: Developmental trajectories and gender differences in psychological functioning, cognitive-affective strategies, and social well-being. *Emerging Adulthood*, 8(2), 103–117. <https://doi.org/10.1177/2167696818811120>
- Córdova Olivera, P., Gasser Gordillo, P., Naranjo Mejía, H., La Fuente Taborga, I., Grajeda Chacón, A., & Sanjinés Unzueta, A. (2023). Academic stress as a predictor of mental health in university students. *Cogent Education*, 10(2), 2232686. <https://doi.org/10.1080/2331186X.2023.2232686>
- Dreisöerner, A., Junker, N. M., & van Dick, R. (2021). The relationship among the components of self-compassion: A pilot study using a compassionate writing intervention to enhance self-kindness, common humanity, and mindfulness. *Journal of Happiness Studies*, 22(1), 21–47. <https://doi.org/10.1007/s10902-019-00217-4>
- Dupasquier, J. R., Kelly, A. C., Moscovitch, D. A., & Vidovic, V. (2020). Cultivating self-compassion promotes disclosure of experiences that threaten self-esteem. *Cognitive Therapy and Research*, 44(1), 108–119. <https://doi.org/10.1007/s10608-019-10050-x>
- Ewert, C., Vater, A., & Schröder-Abé, M. (2021). Self-compassion and coping: A meta-analysis. *Mindfulness*, 12(5), 1121–1136. <https://doi.org/10.1007/s12671-020-01563-8>
- Fahmi, A. (2020). Students’ writing habit of English study program at Faculty of Teacher Training and Education (FKIP) Jambi University. In *ELLit: 2nd Online National Seminar on English Linguistics and Literature*. Jambi University.
- Finlay-Jones, A. L. (2017). The relevance of self-compassion as an intervention target in mood and anxiety disorders: A narrative review based on an emotion regulation framework. *Clinical Psychologist*, 21(2), 90–103. <https://doi.org/10.1111/cp.12131>

- Fitri, N. A., dkk. (2024). Tantangan dan strategi mahasiswa perantauan UNNES dalam menjaga kestabilan mental dan pikiran: Studi kasus pada mahasiswa perantauan UNNES. *Jurnal Majemuk*, 3(4), 673–689.
- Fratraroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin*, 132(6), 823–865. <https://doi.org/10.1037/0033-2909.132.6.823>
- Gallagher, S., O'Sullivan, L., Hughes, Z., & O'Connell, B. H. (2020). Building resources in caregivers: Feasibility of a brief writing intervention to increase benefit finding in caregivers. *Applied Psychology: Health and Well-Being*, 12(2), 513–531. <https://doi.org/10.1111/aphw.12195>
- Gao, X. (2022). Research on expressive writing in psychology: A forty-year bibliometric analysis and visualization of current status and research trends. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.825626>
- Gardani, M., Bradford, D. R. R., Russell, K., Allan, S., Beattie, L., Ellis, J. G., & Akram, U. (2022). A systematic review and meta-analysis of poor sleep, insomnia symptoms and stress in undergraduate students. *Sleep Medicine Reviews*, 61, 101565. <https://doi.org/10.1016/j.smrv.2021.101565>
- Graves, B. S., Hall, M. E., Dias-Karch, C., Haischer, M. H., & Apter, C. (2021). Gender differences in perceived stress and coping among college students. *PLoS ONE*, 16(8), e0255634. <https://doi.org/10.1371/journal.pone.0255634>
- Han, A., & Kim, T. H. (2023). Effects of self-compassion interventions on reducing depressive symptoms, anxiety, and stress: A meta-analysis. *Mindfulness*, 14, 1–29. <https://doi.org/10.1007/s12671-023-02148-x>
- Inaya, A. A., Bakar, R. M., & Hamid, H. (2024). Efektivitas self-compassion writing untuk meningkatkan regulasi emosi pada mahasiswa baru. *Innovative: Journal of Social Science Research*, 4(1), 5460–5469.
- Karakasidou, E., Raftopoulou, G., Papadimitriou, A., & Stalikas, A. (2023). Self-compassion and well-being during the COVID-19 pandemic: A study of Greek college students. *International Journal of Environmental Research and Public Health*, 20(6), 4890. <https://doi.org/10.3390/ijerph20064890>
- Karaman, M. A., Lerma, E., Vela, J. C., & Watson, J. C. (2019). Predictors of academic stress among college students. *Journal of College Counseling*, 22(1), 41–55. <https://doi.org/10.1002/jocc.12113>
- Karyotaki, E., Cuijpers, P., Albor, Y., Alonso, J., Auerbach, R. P., Bantjes, J., ... & Kessler, R. C. (2020). Sources of stress and their associations with mental disorders among college students: Results of the World Health Organization World Mental Health Surveys International College Student Initiative. *Frontiers in Psychology*, 11, 1759. <https://doi.org/10.3389/fpsyg.2020.01759>
- Kemenkes. (2021, December 7). Kemenkes beberkan masalah permasalahan kesehatan jiwa di Indonesia. *Sehat Negeriku*. <https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20211007/1338675/kemenkes-beberkan-masalah-permasalahan-kesehatan-jiwa-di-indonesia/>
- Kizhakkeveetil, A., Vosko, A. M., Brash, M., & Philips, M. A. (2017). Perceived stress and fatigue among students in a doctor of chiropractic training program. *Journal of Chiropractic Education*, 31(1), 8–13. <https://doi.org/10.7899/jce-15-27>
- Kompas.com. (2021). 5 alasan orang enggan pergi ke psikolog meski membutuhkannya. <https://www.kompas.com/sains/read/2021/08/09/201500123/5-alasan-orang-enggan-pergi-ke-psikolog-meski-mebutuhkannya>
- Kroshus, E., Hawrilenko, M., & Browning, A. (2021). Stress, self-compassion, and well-being during the transition to college. *Social Science & Medicine*, 269, 113514. <https://doi.org/10.1016/j.socscimed.2020.113514>
- Kumparan. (2022). Krisis kesehatan mental di Indonesia. <https://kumparan.com/andrew-pierre/krisis-kesehatan-mental-di-indonesia-1zGuYzjk48k>
- Lattie, E. G., Adkins, E. C., Winkquist, N., Stiles-Shields, C., Wafford, Q. E., & Graham, A. K. (2019). Digital mental health interventions for depression, anxiety and enhancement of psychological well-being among college students: Systematic review. *Journal of Medical Internet Research*, 21(7), e12869. <https://doi.org/10.2196/12869>
- Lazarus, R., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Publishing Company.
- Leary, M. R., Tate, E. B., Adams, C. E., Batts Allen, A., & Hancock, J. (2007). Self-compassion and reactions to unpleasant self-relevant events: The implications of treating oneself kindly. *Journal of Personality and Social Psychology*, 92(5), 887–904. <https://doi.org/10.1037/0022-3514.92.5.887>
- Liw, L., Ciftci, A., & Kim, T. (2022). Cultural values, shame and guilt, and expressive suppression as predictors of depression. *International Journal of Intercultural Relations*, 89, 90–99. <https://doi.org/10.1016/j.ijintrel.2022.04.003>
- Mariën, S., Poels, K., & Vandebosch, H. (2022). Think positive, be positive: Expressive writing changes young people's emotional reactions towards the COVID-19 pandemic. *Frontiers in Education*, 6, 755896. <https://doi.org/10.3389/feduc.2021.755896>

- Muchsin, N. R., & Murdiana, S. (2024). Efektivitas self-compassion writing pada perempuan korban body shaming. *Jurnal Psikologi Malahayati*, 6(1).
- Muttaqin, D. (2022). Gender and age invariance of the Indonesian version of Satisfaction with Life Scale. *Jurnal Psikologi Ulayat*, 9(2), 208–215. <https://doi.org/10.24854/jpu438>
- Neff, K. (2003a). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity*, 2(2), 85–101. <https://doi.org/10.1080/15298860309032>
- Neff, K. (2003b). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250. <https://doi.org/10.1080/15298860309027>
- Neff, K. (2011). Self-compassion, self-esteem, and well-being. *Social and Personality Compass*, 5(1), 1–12. <https://doi.org/10.1111/j.1751-9004.2010.00330.x>
- Neff, K. D., & Germer, C. K. (2013). A pilot study and randomized controlled trial of the mindful self-compassion program. *Journal of Clinical Psychology*, 69(1), 28–44. <https://doi.org/10.1002/jclp.21923>
- Ng, K. C., Chiu, W. K., & Fong, B. Y. (2016). A review of academic stress among Hong Kong undergraduate students. *Journal of Modern Education Review*, 6(8), 531–540. [https://doi.org/10.15341/jmer\(2155-7993\)/08.06.2016/001](https://doi.org/10.15341/jmer(2155-7993)/08.06.2016/001)
- Pascoe, P. E. (2017). Using patient writings in psychotherapy: Review of evidence for expressive writing and cognitive-behavioral writing therapy. *American Journal of Psychiatry Residents' Journal*, 11(3), 3–6. <https://doi.org/10.1176/appi.ajp-rj.2017.110303>
- Pavlicic, J. M., Buchanan, E. M., Maxwell, N. P., Hopke, T. G., & Schulenberg, S. E. (2019). A meta-analysis of expressive writing on posttraumatic stress, posttraumatic growth, and quality of life. *Review of General Psychology*, 23(2), 230–250. <https://doi.org/10.1177/1089268019831641>
- Pennebaker, J. W., & Beall, S. W. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology*, 95(3), 274–281. <https://doi.org/10.1037/0021-843X.95.3.274>
- Pennebaker, J. W., & Chung, C. K. (2007). Expressive writing, emotional upheavals, and health. In H. Friedman & R. Silver (Eds.), *Handbook of health psychology* (pp. 263–284). Oxford University Press.
- Pennebaker, J. W., & Evans, J. E. (2014). *Expressive writing: Words that heal*. Enumclaw, WA: Idyll Arbor.
- Pennebaker, J. W., & Smyth, J. M. (2016). *Opening up by writing it down: How expressive writing improves health and eases emotional pain*. New York, NY: The Guilford Press.
- Poots, A., & Cassidy, T. (2020). Academic expectation, self-compassion, psychological capital, social support and student wellbeing. *International Journal of Educational Research*, 99, 101506. <https://doi.org/10.1016/j.ijer.2019.101506>
- Pusat Data Strategis dan Statistik Ikatan Psikolog Klinis (IPK) Indonesia. (2023). *Statistik Ikatan Psikolog Klinis (IPK) Indonesia*. <https://data.ipkIndonesia.or.id/>
- Savera, N. D., Purwaningrum, R., & Hidayat, R. R. (2022). A literature review: Does expressive writing effective to reduce stress?. *ProGCouns: Journal of Professionals in Guidance and Counseling*, 3(2), 87-99.
- Seedhom, A. E., Kamel, E. G., Mohammed, E. S., & Raouf, N. R. (2019). Predictors of perceived stress among medical and nonmedical college students, Minia, Egypt. *International Journal of Preventive Medicine*, 10(1), 6. [https://doi.org/10.4103/ijpvm.IJPVM\\_6\\_18](https://doi.org/10.4103/ijpvm.IJPVM_6_18)
- Setyaningrum, A., Rahman, A. M., & Ngesti, M. (2024). Dampak harapan orang tua yang tidak realistis terhadap akademik remaja: Kajian sistematis. *Jurnal Psikologi*, 1(4), 16.
- Sharp, J., & Theiler, S. (2018). A review of psychological distress among university students: Pervasiveness, implications and potential points of intervention. *International Journal for the Advancement of Counselling*, 40(3), 193–212. <https://doi.org/10.1007/s10447-018-9321-7>
- Smyth, J. M. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *Journal of Consulting and Clinical Psychology*, 66(1), 174–184. <https://doi.org/10.1037/0022-006X.66.1.174>
- Sohail, N. (2013). Stress and academic performance among medical students. *J Coll Physicians Surg Pak*, 23(1), 67-71.
- Stoeber, J., Lalova, A. V., & Lumley, E. J. (2020). Perfectionism, (self-) compassion, and subjective well-being: A mediation model. *Personality and Individual Differences*, 154, 109708. <https://doi.org/10.1016/j.paid.2019.109708>
- Sugianto, D., Suwartono, C., & Sutanto, S. H. (2020). Reliabilitas dan validitas Self-Compassion Scale versi bahasa Indonesia. *Jurnal Psikologi Ulayat*. Advance online publication. <https://doi.org/10.24854/jpu02020-337>
- Sugiyono. (2014). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Swee, M. B., Klein, K., Murray, S., & Heimberg, R. G. (2023). A brief self-compassionate letter-writing intervention for individuals with high shame. *Mindfulness*, 14(4), 854–867. <https://doi.org/10.1007/s12671-022-02043-9>

- Urken, D., & LeCroy, C. W. (2021). A randomized controlled trial of a self-compassion writing intervention for adults with mental illness. *Research on Social Work Practice, 31*(3), 254–266. <https://doi.org/10.1177/1049731520972779>
- Vukčević Marković, M., Bjekić, J., & Priebe, S. (2020). Effectiveness of expressive writing in the reduction of psychological distress during the COVID-19 pandemic: A randomized controlled trial. *Frontiers in Psychology, 11*, 587282. <https://doi.org/10.3389/fpsyg.2020.587282>
- Williamson, J. R. (2014). Addressing self-reported depression, anxiety, and stress in college students via web-based self-compassionate journaling [Disertasi, East Tennessee State University]. *Electronic Theses and Dissertations, East Tennessee State University*. <https://dc.etsu.edu/etd/2402>
- Williamson, J., & Blackhart, G. C. (2021). Efficacy of guided versus self-induced learning of web-based self-compassionate journaling by college students. *The American Journal of Psychology, 134*(1), 45–59. <https://doi.org/10.5406/amerjpsyc.134.1.0045>
- Windiastruti, M. P., & Kaloeti, D. V. S. (2022). Studi pendahuluan pengaruh self-compassion writing pada mahasiswa dengan depresi ringan. *Jurnal Empati, 11*(4), 234–244.
- Wong, C. C. Y., & Mak, W. W. S. (2016). Writing can heal: Effects of self-compassion writing among Hong Kong Chinese college students. *Asian American Journal of Psychology, 7*(1), 74–82. <https://doi.org/10.1037/aap0000041>
- World Health Organization (WHO). (2022). *World mental health report: Transforming mental health for all*.
- Xiao, H., Janis, R. A., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2017). Are we in crisis? National mental health and treatment trends in college counseling centers. *Psychological Services, 14*(4), 407–415. <https://doi.org/10.1037/ser0000130>
- Zega, C. J. P., & Zega, F. P. (2025). Meta-analisis faktor-faktor yang memengaruhi tingkat stres akademik pada mahasiswa. *Jurnal Ilmu Ekonomi, Pendidikan dan Teknik, 2*(2), 55–60.
- Zessin, U., Dickhäuser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: A meta-analysis. *Applied Psychology: Health and Well-Being, 7*(3), 340–364. <https://doi.org/10.1111/aphw.12051>
- Zheng, L., Lu, Q., & Gan, Y. (2019). Effects of expressive writing and use of cognitive words on meaning making and post-traumatic growth. *Journal of Pacific Rim Psychology, 13*, e5. <https://doi.org/10.1017/prp.2018.31>
- Ziemer, K. S., Lamphere, B. R., Raque-Bogdan, T. L., & Schmidt, C. K. (2019). A randomized controlled study of writing interventions on college women's positive body image. *Mindfulness, 10*(1), 66–77. <https://doi.org/10.1007/s12671-018-0947-7>



## UJI PSIKOMETRI SKALA *CYBERSTALKING* VERSI INDONESIA: *CONFIRMATORY FACTOR ANALYSIS* DAN INVARIANSI *GENDER*

Rerepih Julung Merekah Wanodya Sinandang<sup>1</sup>, Ratri Pratiwi<sup>2</sup>

<sup>1, 2</sup> Fakultas Psikologi Universitas Mercubuana Yogyakarta

Email: [ratri.pratiwi@mercubuana-yogya.ac.id](mailto:ratri.pratiwi@mercubuana-yogya.ac.id)

p-ISSN: 2337-4845  
e-ISSN: 2620-7486



<i>Received</i>	<i>Revised</i>	<i>Accepted</i>	<i>Published</i>
16 Juni 2025	01 Oktober 2025	01 Oktober 2025	30 Oktober 2025

### Abstract

*The Cyberstalking Scale has been used in various countries to measure cyberstalking behavior, which is becoming increasingly widespread in society, but there has been no research testing the psychometric properties of the Cyberstalking Scale in Indonesia. This study aims to examine the psychometric properties of construct validity using CFA and gender measurement invariance of the Indonesian version of the Cyberstalking Scale. Participants consisted of 306 male and female respondents aged 18-22 years who actively use social media. Respondents were selected using purposive sampling techniques. The evaluation of the internal structure validity of the Indonesian version of the Cyberstalking Scale included factor structure, composite reliability, and measurement invariance between genders, which were tested using confirmatory factor analysis, composite reliability, and multigroup confirmatory factor analysis. The results show that the cyberstalking scale is valid and reliable ( $\alpha = 0.922$ ). In addition, the cyberstalking scale is also invariant at the strict stage. Thus, the Indonesian version of the cyberstalking scale can be used for research and intervention related to deviant behavior in the digital space. The limitation of this study is the use of a homogeneous sample, which limits the generalizability of the results.*

*Keywords: measurement invariance, cyberstalking, psychometric properties*

### Abstrak

*Cyberstalking Scale* telah digunakan di berbagai negara untuk mengukur perilaku *cyberstalking* yang semakin meluas dalam masyarakat, tetapi belum terdapat penelitian yang menguji properti psikometrik *Cyberstalking Scale* di Indonesia. Penelitian ini bertujuan untuk melihat properti psikometri validitas konstruk dengan CFA dan invariansi pengukuran *gender Cyberstalking Scale* versi bahasa Indonesia. Partisipan terdiri dari 306 responden laki-laki dan perempuan berusia 18-22 tahun yang aktif menggunakan media sosial. Responden dipilih menggunakan teknik *purposive sampling*. Evaluasi validitas struktur internal *Cyberstalking Scale* versi Indonesia meliputi struktur faktor, reliabilitas komposit, dan invariansi pengukuran antar jenis kelamin yang diuji menggunakan analisis konfirmatori faktor, reliabilitas komposit, dan analisis multigroup konfirmatori faktor. Hasil menunjukkan bahwa skala *cyberstalking* valid dan reliabel ( $\alpha = 0.922$ ). Skala *cyberstalking* versi Indonesia ini dapat digunakan untuk penelitian maupun intervensi terkait. Keterbatasan penelitian ini adalah penggunaan sampel yang homogen, sehingga kemampuan generalisasi hasil menjadi terbatas.

Kata kunci: invariansi pengukuran, *cyberstalking*, properti psikometri

### 1. Pendahuluan

Perkembangan teknologi dan informasi yang pesat kini menjadi sarana utama dalam kehidupan manusia, menyebabkan perubahan sosial yang besar dan cepat (Rosana, 2010). Teknologi informasi mempermudah akses komunikasi, belanja online, pencarian literatur, dan edukasi, serta informasi dalam berbagai bidang seperti seni, budaya, sosial, politik, dan ekonomi (Irawan, 2014). Meski diharapkan mempermudah aktivitas, kemajuan teknologi juga menimbulkan dampak negatif, seperti ketergantungan, keteledoran, dan penyalahgunaan komputer (Widyopramono, 2012). Selain itu internet juga merupakan sumber dari berbagai macam kejahatan di dunia *cyber* atau *cybercrime* (Amarini, 2018).

Dalam arti sempit, *cybercrime* adalah kejahatan terhadap sistem komputer, sedangkan *cybercrime* dalam arti luas mencakup kejahatan terhadap sistem atau jaringan komputer dan kejahatan yang menggunakan sarana komputer (Aldriano, 2022). Penelitian sebelumnya di bidang hukum oleh Ketaren (2016) menganalisis dinamika *cybercrime* serta mengklasifikasikan jenis-jenisnya berdasarkan motif dan jenis aktifitasnya diantaranya *cyberbullying*,

cyberstalking, cybersquatting, hacking, dan lain sebagainya. Salah satu jenis cybercrime yang masih asing dan belum diketahui secara luas adalah cyberstalking.

*Cyberstalking* dipahami sebagai bentuk pelecehan yang dilakukan secara berulang melalui media elektronik, seperti media sosial, email, atau sarana digital lainnya. (Paullet et al., 2009; Nobles et al., 2014; Wilson et al., 2021). Perilaku ini melibatkan komunikasi atau kontak yang tidak diinginkan dan terus-menerus, yang ditujukan kepada seseorang dengan tujuan mengganggu, mengendalikan, mengintimidasi, atau mengancam. Akibatnya, korban sering mengalami kecemasan dan ketakutan yang serius, yang pada akhirnya berdampak pada penurunan kualitas hidup mereka (Santos, 2023).

*Cyberstalking* bertujuan untuk mengawasi aktivitas online seseorang, teror panggilan atau chat, bahkan terbentuk adanya pengadaan akun sehingga banyak didapati akun-akun palsu atau anonim, tuduhan palsu, pencurian identitas, dan lain sebagainya. *Cyberstalking* memiliki dampak yang luas terhadap korban tetapi pemahaman masyarakat terhadap fenomena ini masih terbatas (Baum, Catalano, Rand, dan Rose, 2009). Di Indonesia sendiri, aktivitas cyberstalking semakin umum seiring pesatnya perkembangan teknologi informasi dan komunikasi yang mengubah pola interaksi sosial masyarakat (Sugiyono, 2012). Pola komunikasi yang kini banyak dilakukan melalui media sosial menyebabkan informasi pribadi lebih mudah diakses dan disalahgunakan. Fenomena ini juga diperkuat oleh minimnya literasi digital serta belum optimalnya regulasi yang secara spesifik mengatur perlindungan terhadap korban kejahatan siber, termasuk cyberstalking (Ketaren, 2016).

*Cyberstalking* memiliki dampak psikologis dan fisik bagi korban, seperti perubahan pola tidur dan makan, kecemasan, stres, dan ketakutan, yang serupa dengan gejala PTSD. Korban sering merasa terisolasi, mudah marah, dan bersalah (Short et al., 2015), meningkatkan depresi (Wright, 2018), dan berdampak negative pada kesejahteraan psikologis korban (Derbing, et al., 2014). Penelitian menunjukkan prevalensi cyberstalking bahwa seperempat penelitian melaporkan viktimisasi hingga dua tahun, dengan mayoritas korban, 60%-80%, adalah perempuan (Spitzberg & Cupach, 2007). ). Penelitian ini menunjukkan perempuan lebih rentan menjadi korban *cyberstalking*, dengan dampak serius secara psikologis maupun fisik. Namun, pada penelitian Paullet et al., (2009) tidak menunjukkan adanya kaitan antara jenis kelamin dengan *cyberstalking*. Fenomena ini menekankan pentingnya memahami peran gender dan mengembangkan pengukuran *cyberstalking* yang bebas bias gender.

Penelitian ini menguji properti psikometrik yaitu validitas konstruk dan uji kesetaraan invariansi pengukuran gender *Cyberstalking Scale* versi Indonesia. Pengujian invariansi pengukuran bertujuan untuk memastikan tidak adanya potensi bias pengukuran yang disebabkan oleh perbedaan kelompok, sehingga hasil analisis uji beda dapat dipertanggungjawabkan (Chen, 2007). *Cyberstalking Scale* merupakan sebuah skala yang digunakan untuk mengukur perilaku cyberstalking. Alat ukur cyberstalking dikembangkan oleh Santos pada tahun 2021 di Brazil. Hasil penelitian menunjukkan 10 aitem final tersebut valid dan reliabel ( $\alpha = 0,86$ ). Setelah itu dilakukan uji kecocokan model yang menunjukkan bahwa satu faktor cukup memadai (GFI = 0,98; CFI = 0,99; TLI = 0,99; RMSEA = 0,02; SRMR = 0,06). Sehingga dapat disimpulkan bahwa *Cyberstalking Scale* valid dan reliabel pada budaya di Brazil.

Penelitian lain oleh Spitzberg et.al (2001) mengembangkan *The Cyber-Obsessional Pursuit (COP) Scale* yaitu skala untuk mengukur sejauh mana pengalaman individu dalam mengalami pengejaran obsesif di dunia maya termasuk *cyberstalking*. Skala ini banyak digunakan untuk mengukur perilaku cyberstalking. Hasil penelitian ini adalah reliabilitas internal untuk subskala COP berkisar dari cukup hingga baik (Cronbach's  $\alpha$  berkisar dari 0,73 hingga 0,85). Sinandang dan Pratiwi (2024) telah melakukan adaptasi *cyberstalking scale* dari Santos dan kawan-kawan (2023) ke versi Indonesia pada 153 orang sampel yang menunjukkan hasil valid dan reliabel ( $\alpha = 0,89$ ) dengan hasil kecocokan model yang memadai (CFI = 0,952; GFI = 0,991; TLI = 0,933, SRMR = 0,064; RMSEA = 0,074). Penelitian Sun dan Chen (2024) pada skala internet moral literacy menunjukkan hasil invariansi berdasarkan gender pada tingkat yang lemah. Artinya terdapat perbedaan yang signifikan pada tingkat internet moral literacy pada laki-laki dan perempuan. Sehingga diperlukan kehati-hatian dalam penggunaan skala yang berkaitan dengan etika digital berdasarkan gender. Demikian pula, Aprillia (2024) mengadaptasi *Cyber-Aggression Typology Questionnaire (CATQ)* ke dalam bahasa Indonesia, meskipun pengembangan lebih lanjut diperlukan untuk sifat psikometrik yang optimal. Penelitian tentang *cyberstalking* telah mendapatkan perhatian yang signifikan, namun masih terdapat celah yang substansial dalam pengukuran dan validasi lintas budaya. Kaur dkk. (2020) mengidentifikasi fokus penelitian yang terfragmentasi dan menekankan kebutuhan mendesak untuk validasi temporal dan lintas budaya skala pengukuran, mencatat bahwa studi yang ada kurang jelas dalam karakterisasi dan pengukuran prevalensi. Abu-ulbeh dkk. (2021) mengembangkan model viktimisasi cyberstalking menggunakan teori kriminologi, namun fokus pada konteks Yordania. Studi-studi ini secara kolektif menyoroiti ketidakhadiran skala *cyberstalking* yang valid dan disesuaikan secara budaya untuk populasi Indonesia, mewakili celah penelitian yang signifikan yang memerlukan pengujian multigrup di berbagai konteks budaya.

Penelitian yang lebih luas perlu dilakukan dengan tujuan memeriksa properti psikometri *Cyberstalking Scale* versi bahasa Indonesia. Penelitian ini dilakukan untuk menguji struktur factor *Cyberstalking Scale* pada subjek berusia 18-22 tahun dengan menggunakan prosedur statistik *Confirmatory Factor Analysis (CFA)* kemudian dilanjutkan dengan *Measurement Invariance*. Penelitian ini bertujuan untuk melihat properti psikometrik validitas

konstruk dan uji kesetaraan invariansi pengukuran gender *Cyberstalking Scale* versi bahasa Indonesia sehingga menghasilkan *Cyberstalking Scale* versi bahasa Indonesia yang valid dan tidak memiliki bias pengukuran berdasarkan kelompok gender.

## 2. Metode Penelitian

### Instrumen

Alat ukur yang akan dianalisis properti psikometriknya pada penelitian ini adalah alat ukur *Cyberstalking Scale*. Konsep alat ukur ini didasarkan pada aspek-aspek berikut: *romantic relationships (past, current, and desired)*, *persecution of acquaintances*, dan *people that the perpetrator suspects or dislikes*. Aspek *Romantic relationships (past, current, and desired)* melibatkan pengawasan, pelecehan, atau upaya mengontrol pasangan masa lalu, saat ini, atau yang diinginkan melalui teknologi komunikasi. Aspek *Persecution of acquaintances* merujuk pada tindakan pelecehan, intimidasi, atau pengawasan terhadap teman, rekan sekerja, atau kenalan melalui media sosial atau platform komunikasi lainnya. Sedangkan aspek *People that the perpetrator suspects/dislikes* yaitu penyebaran fitnah, ancaman, pengawasan, atau tindakan merusak terhadap individu yang dicurigai atau tidak disukai meskipun tidak ada hubungan pribadi sebelumnya. Aspek tersebut memberikan pemahaman yang mendalam mengenai fenomena internal yang tidak dapat diamati secara langsung. Setiap aspek mencakup sejumlah item, dengan total keseluruhan 10 item yang tersebar di antara ketiga aspek tersebut. Seluruh item dalam tes ini disusun sebagai item favorable, yang berarti setiap item mencerminkan respons atau sikap positif terhadap aspek yang diukur. Oleh karena itu, instrumen ini tidak memuat item yang bersifat unfavorable. *Cyberstalking Scale* menggunakan skala Likert lima poin, dengan pilihan jawaban yang terdiri dari : “1” = Sangat Tidak Sesuai, “2” = Tidak Sesuai, “3” = Netral, “4” = Sesuai, “5” = Sangat Sesuai.

### Partisipan

Penelitian ini menggunakan metode non-probability sampling, di mana sampel tidak perlu representatif atau acak, namun didasarkan pada kriteria tertentu (Sugiyono, 2016). Teknik pengambilan sampel yang digunakan adalah purposive sampling, yang memungkinkan peneliti membuat generalisasi teoritis, analitik, dan logis. Karakteristik responden adalah remaja laki-laki atau perempuan usia 18-22 tahun, pengguna aktif media sosial. Pemilihan karakteristik responden menyesuaikan dengan Santos dan kawan-kawan (2023). Total sampel berjumlah 306 subjek, sampel terdiri dari 185 perempuan dan 121 laki-laki.

### Prosedur Penelitian

Penelitian ini terdiri dari dua tahap utama, yaitu proses penerjemahan item-item *Cyberstalking Scale* dan pengujian properti psikometrisnya. Tahap pertama proses penerjemahan berdasarkan prosedur dari Beaton (2000), yang terdiri dari 1) *Translation*; 2) *Synthesis*; 3) *Back-translation*; 4) *Expert Committee Review*; dan 5) *Test of the Pre-Final Version*. Tahap kedua pengujian psikometris dengan *Confirmatory Factor Analysis (CFA)* dan multigrup analisis berdasarkan gender. Data penelitian dikumpulkan melalui pengisian skala oleh subjek, yang terlebih dahulu mengakses tautan yang disediakan oleh peneliti. Penyebaran skala dilakukan secara daring melalui Google Form.

### Teknik Analisis Data

Teknik analisis data dalam penelitian ini menggunakan program JASP. Peneliti melakukan CFA. Dalam metode CFA, jumlah faktor telah ditentukan sebelumnya berdasarkan teori atau konstruk yang diuji. Analisis ini digunakan untuk menguji sejauh mana indikator-indikator dalam model merepresentasikan faktor-faktor tersebut, serta untuk melihat hubungan antar faktor (Thompson, 2004). Menurut Cudeck & Browne (1993), nilai RMSEA  $\leq 0,08$  menunjukkan kecocokan yang baik, sementara lebih dari 0,1 ditolak. CFI yang baik  $\geq 0,90$  (Bentler, 1990; Hooper et al., 2008), dan TLI  $\geq 0,9$  menunjukkan akurasi model yang baik (Bentler & Bonner, 1980). GFI yang baik  $\geq 0,90$ .

Setelah peneliti melakukan CFA, peneliti melakukan pengujian multigrup analisis untuk mengevaluasi invariansi berdasarkan gender dari *Cyberstalking Scale* versi Indonesia. Pengujian invariansi gender dilakukan melalui invariansi konfigural, metrik, dan skalar. Menurut Chen (2007), selisih koefisien CFI  $\leq -0,010$  dan RMSEA  $\leq 0,015$  dapat menunjukkan adanya invariansi pengukuran.

## 3. Hasil

Subjek penelitian ini melibatkan 320 subjek yang berusia antara 18 sampai 22 tahun, yang aktif dalam bermain media sosial. Dari total subjek, 306 digunakan dalam analisis sementara 14 data lainnya dieliminasi karena subjek tidak sesuai dengan kriteria yang dibutuhkan. Deskripsi mengenai subjek yang peneliti gunakan terlampir pada tabel 1. Berdasarkan data di tabel 1, dari total 306 responden, terdapat 121 responden berjenis kelamin laki-laki dan 185 responden berjenis kelamin perempuan. Dari kategori usia, rentang usia responden dari usia 18-22 tahun ( $M = 20,6$ ,  $SD = 1,26$ ). Dilihat dari pendidikan terakhirnya, terdapat 228 responden yang tamat SMA/K, 19 responden tamat D3, dan 59 responden tamat S1. Sedangkan pada status pekerjaan, terdapat 206 responden berasal

dari pelajar atau mahasiswa, 71 responden bekerja sebagai karyawan swasta, 22 responden berasal dari wirausaha, dan 5 responden yang tidak bekerja.

**Tabel 1. Data Demografi Subjek**

Variabel	Kategori	Frekuensi	Presentase
Jenis Kelamin	Laki-laki	121	39,54%
	Perempuan	185	60,46%
Usia	18 tahun	23	7,30%
	19 tahun	40	12,70%
	20 tahun	69	21,90%
	21 tahun	85	26,98%
	22 tahun	98	31,11%
Pendidikan Terakhir	SMA/K	228	74,51%
	D3	19	6,21%
	S1	59	19,28%
Status Pekerjaan	Pelajar/Mahasiswa	206	67,76%
	Karyawan Swasta	71	23,36%
	Wirausaha	22	7,24%
	Tidak Bekerja	5	1,64%

Hasil statistik deskriptif dari skala *cyberstalking* dengan total item berjumlah 10, skor terendah adalah 1 dan skor tertinggi adalah 5. Perhitungan hipotetik dengan skor minimal (min) yaitu  $1 \times 10 = 10$  dan skor maksimal (max)  $5 \times 10 = 50$ . Nilai rata-rata (mean) hipotetik dengan perhitungan  $(10 + 50) : 2 = 30$ , jarak sebaran hipotetiknya yaitu  $50 - 10 = 40$  dengan standar deviasi (SD) sebesar  $(50 - 10) : 6 = 6,67$ . Sedangkan data empirik skala *cyberstalking* memiliki skor minimal 10 dan skor maksimal 50. Nilai rata-rata (mean) sebesar 32,60 dan standar deviasi sebesar 9,5.

**Tabel 2. Data Deskriptif Cyberstalking**

Variabel	N	Data Hipotetik			Data Empirik			SD	
		Mean	Skor		Mean	Skor			
			Min	Max		Min	Max		
Cyberstalking	306	30	10	50	6,67	32,60	10	50	9,5

Keterangan :

- N = Jumlah subjek
- Mean = Rata-rata
- Min = Skor minimal atau terendah
- Max = Skor maksimal atau tertinggi
- SD = Standar deviasi

Sebelum dilakukan analisis invariansi pengukuran, dilakukan pengujian model. Pengujian model dilakukan dengan melihat nilai  $RMSEA \leq 0,08$  menunjukkan kecocokan yang baik, sementara lebih dari 0,1 ditolak (Cudeck & Browne, 1993); CFI yang baik  $\geq 0,90$  (Bentler, 1990; Hooper et al., 2008), dan TLI  $\geq 0,9$  menunjukkan akurasi model yang baik (Bentler & Bonner, 1980), serta GFI yang baik  $\geq 0,90$ .

Berdasarkan tabel 3 hasil pengujian model menunjukkan bahwa nilai CFI 0.978 (CFI  $\geq 0,90$ ), RMSEA 0.064 (RMSEA  $\leq 0,08$ ), GFI 0.991 (GFI  $\geq 0,90$ ), TLI 0.968 (TLI  $\geq 0,9$ ) sehingga hasil analisis menunjukkan alat ukur memiliki ketetapan model yang memuaskan.

**Tabel 3. Statistik Model Fit**

Indikator	Kriteria	Hasil	Deskripsi
<i>Chi-Square</i>	$p > .05$	$p < .001$	<i>Not fit</i>
<i>RMSEA</i>	$< .08$	0.064	<i>Good fit</i>
<i>GFI</i>	$\geq .90$	0.991	<i>Good fit</i>
<i>SRMR</i>	$< .08$	0.031	<i>Good fit</i>
<i>CFI</i>	$\geq .90$	0.978	<i>Good fit</i>
<i>TLI</i>	$\geq 0.90$	0.968	<i>Good fit</i>

Menurut Kline (2015) dalam CFA, nilai loading factor sebesar  $\geq 0,6$  dianggap menunjukkan korelasi yang kuat antara indikator dan konstruk laten, sedangkan nilai diatas 0,3 berarti cukup tinggi, dan kurang dari 0,3 umumnya dianggap rendah dan dapat diabaikan. Hal ini terutama berlaku jika ukuran sampel yang lebih besar ( $\geq 300$ ) karena ukuran sampel tersebut cukup besar untuk memastikan kestabilan estimasi parameter yang akurat. Hasil loading factor *Cyberstalking Scale* versi Indonesia pada table 4 menunjukkan nilai factor loadings berkisar antara 0,495 hingga 0,864 dengan  $p < 0,001$ . Peneliti mempertahankan indicator R5 dengan nilai loading factor sebesar 0,495 meski nilai loading factor kurang dari 0,6. Hal ini dikarenakan peneliti menilai bahwa indicator R5 menjadi indicator penting pada faktor romatic relationship pada budaya Indonesia.

**Tabel 4. Factor loadings**

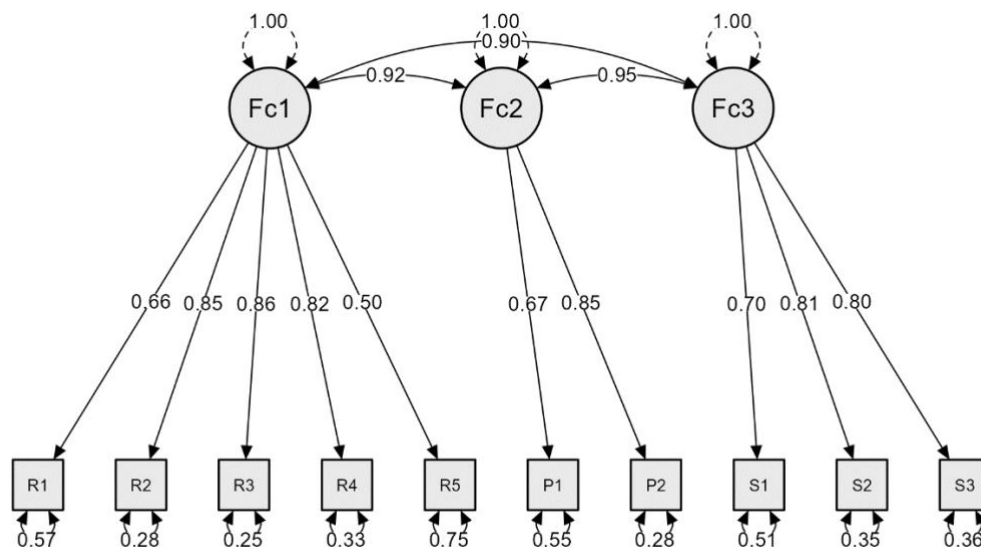
Factor	Indicator	p	Std.Est. (all)
Romantic Relationship	R1	$< .001$	0.658
	R2	$< .001$	0.849
	R3	$< .001$	0.864
	R4	$< .001$	0.821
	R5	$< .001$	0.495
Persecution of Acquaintances	P1	$< .001$	0.673
	P2	$< .001$	0.846
People Perpetrator Suspect	S1	$< .001$	0.699
	S2	$< .001$	0.806
	S3	$< .001$	0.802

Hasil uji validitas converge menunjukkan bahwa pada ketiga faktor nilai AVE  $> 0,5$  (Romantic Relationship = 0,597; Persecution of acquaintances = 0,597; People Perpetrator Suspect = 0,595). Hasil uji reliabilitas internal sebesar  $\alpha = 0,922$  dan reliabilitas konstruk diperoleh melalui nilai Composite Reliability (CR) sebesar 0,989. Sehingga dapat disimpulkan bahwa *Cyberstalking Scale* versi Indonesia validitas dan reliabilitas yang baik.

Menurut Chen (2007) untuk sampel  $\geq 300$  nilai  $\Delta CFI$  direkomendasikan  $\leq .010$  baik pada faktor loading, intercept maupun residual. Yang berarti perbedaan nilai CFI antar model yang diuji harus kecil agar model dapat dikatakan dikatakan invarian. Sementara itu, untuk  $\Delta RMSEA$  selisih antar model adalah  $\leq .015$ , jika selisih RMSEA antar model melebihi nilai ini, maka model dapat dipertimbangkan tidak invarian. Selain itu nilai  $\Delta SRMR$  direkomendasikan sebesar  $\leq .030$  untuk factor loadings dan  $\leq .010$  untuk selisih intercepts dan residuals, jika nilai ini tercapai maka model dianggap fit dan invarian. Berdasarkan hasil analisis pada tabel 4 menunjukkan bahwa terdapat invariansi pengukuran gender dari *Cyberstalking Scale* versi Indonesia dilihat dari invariansi scalar dan strict dengan tiga indikator perbandingan model yaitu CFI, RMSEA, dan SRMR.

**Tabel 5. Invariansi Pengukuran Gender Skala *Cyberstalking***

Invariansi Pengukuran Gender	Indeks Ketetapan Model					Perbandingan Model		
	X	df	CFI	RMSEA	SRMR	$\Delta$ CFI	$\Delta$ RMSEA	$\Delta$ SRMR
1. Invariansi konfigural	111,276	64	0,974	0,069	0,040			
2. Invariansi metrik	129,007	71	0,968	0,073	0,062	-0,006	0,004	0,024
3. Invariansi Scalar	137,048	78	0,967	0,070	0,065	-0,001	-0,003	0,003
4. Invariansi Strict	163,875	88	0,958	0,075	0,064	-0,009	0,005	-0,001



**Gambar 1. Path Diagram Hasil CFA**

#### 4. Diskusi

Hasil pengujian validitas dan reliabilitas pada skala *cyberstalking* versi Indonesia menunjukkan alat ukur memiliki ketetapan model yang memuaskan (CFI = 0,978; RMSEA = 0,064; SRMR = 0,034;  $\alpha$  = 0,922; CR = 0,989; AVE > 0,5). Artinya skala *cyberstalking* versi Indonesia memiliki validitas konstruk dan reliabilitas yang memadai, sehingga skala ini dapat digunakan dalam konteks budaya Indonesia. Skala *cyberstalking* versi Indonesia ini mendukung pengembangan *Cyberstalking Scale* yang telah dilakukan Santos dan kawan-kawan (2023). Hasil menunjukkan kesamaan faktor *Cyberstalking*.

Beberapa penelitian terdahulu pada pengembangan alat ukur perilaku stalking dan pelecehan diruang digital belum dilakukan secara maksimal dan spesifik. Seperti, penelitian Smoker dan March (2017) *Intimate Partner Cyberstalking Scale* (IPCS-scale) pada 21 aitem dengan nilai internal konsistensi  $\alpha$  = 0,92, tidak melakukan pengujian pada validitas konstruk. Penelitian Ayub dan Malik (2024) tentang pengembangan *Cyber Harassment Experience Scale* (CHES) di Paskitan, hanya berdasarkan pengalaman pelecehan diruang digital pada perempuan saja sehingga tidak mewakili perilaku *stalking* ataupun pelecehan pada laki-laki.

Pada hasil peneliti peneliti mempertahankan indikator R5 (Ketika Anda tertarik pada seseorang, tidak salah untuk melihat media sosial kenalan mereka, untuk lebih mengenal orang tersebut) sebagai salah satu indikator penting pada faktor *romantic relationship* pada budaya Indonesia. Berdasarkan Faizun, Angga, Dasi dan Helmi (2020) *cyberstalking* di Indonesia dimanfaatkan dalam membantu individu membangun hubungan romantis. Individu melakukan *cyberstalking* pada target *cyberstalking* memiliki tujuan untuk menciptakan impresi yang baik mengenai dirinya berdasarkan informasi yang telah didapatkan melalui social media target sebelumnya. Selain itu, *cyberstalking* tidak hanya berfokus pada hubungan romantis, tetapi juga pada hubungannya lain dan aitem pada *cyberstalking scales* untuk memperoleh informasi yang lebih spesifik (Santos, et. al., 2023).

Hasil uji invariansi pengukuran berdasarkan gender memperlihatkan bahwa model mencapai invariansi hingga tingkat scalar dan strict, dengan nilai  $\Delta CFI$ ,  $\Delta RMSEA$ , dan  $\Delta SRMR$  masih berada dalam batas rekomendasi (Chen, 2007). Temuan ini menunjukkan bahwa instrumen Cyberstalking Scale versi Indonesia tidak hanya valid dan reliabel, tetapi juga menunjukkan bahwa struktur faktor skala *cyberstalking* konsisten pada kelompok gender, sehingga secara konseptual dipahami dengan cara yang sebanding antara laki-laki dan perempuan. Temuan ini sejalan dengan penelitian Duran-Guerrero et. al., (2025) pada skala *cyber-sexual harassment* yang didalamnya termasuk perilaku *non-consensual sharing* dan seksual *cyber-perpetration* menunjukkan hasil invariansi berdasarkan gender pada tingkat configural hingga strict ( $\Delta CFI \leq 0,003$ ;  $\Delta RMSEA \leq 0,003$ ). Namun, hal berbeda dengan hasil penelitian Santos dan kawan-kawan (2023) yang menunjukkan terdapat perbedaan gender pada *cyberstalking*, meski tidak melalui pengujian CFA multigrup.

Di Indonesia baik laki-laki maupun perempuan memiliki pemahaman yang sama bahwa *cyberstalking* merupakan keingintahuan individu terhadap orang lain yang dikenal dengan tujuan untuk mengenal lebih dekat, dan mayoritas pelaku *cyberstalking* adalah orang yang dikenal oleh target (Faizun, Angga, Dasi & Helmi, 2020). Alkatiri dan Sopacua (2025) menambahkan motif perilaku *cyberstalking* di Indonesia hanya sebatas motif informatif, yakni untuk mencari tahu informasi terhadap target dan motif hiburan. perilaku *cyberstalking* tersebut tidak masuk dalam kategori sebagai kejahatan.

Keterbatasan penelitian ini adalah pada penggunaan sampel yang terbatas pada kelompok usia tertentu, sehingga dapat membatasi generalisasi hasil. Selain itu, pengujian invariansi masih terbatas gender belum mempertimbangkan variasi jenis media sosial yang digunakan, perbedaan status hubungan romantis ataupun usia. Hasil penelitian ini dapat digunakan dalam konteks asesmen seperti identifikasi perilaku menyimpang pada ruang digital. Selain itu, program intervensi psikologi dapat dirancang lintas gender tanpa adanya risiko bias pengukuran.

## 5. Kesimpulan

Berdasarkan hasil penelitian ini dapat disimpulkan bahwa skala *cyberstalking* versi Indonesia memenuhi syarat invariansi pengukuran hingga tingkat strict berdasarkan gender. Artinya skala *cyberstalking* versi Indonesia ini dapat mengukur perilaku *cyberstalking* secara setara antara laki-laki dan perempuan. Dengan demikian, skala ini dapat digunakan secara sah untuk perbandingan skor, analisis hubungan dengan variabel lain, serta asesmen lintas gender tanpa risiko bias. Skala ini juga terbukti memiliki kekuatan psikometrik dan relevan digunakan pada usia dewasa awal di Indonesia.

## 6. Daftar Pustaka

- Abu-ulbeh, W., Altalhi, M., Abualigah, L., Almazroi, A., Sumari, P., & Gandomi, A. H. (2021). Cyberstalking victimization model using criminological theory: A systematic literature review, taxonomies, applications, tools, and validations. *Electronics*, 10(14), 1670. <https://doi.org/10.3390/electronics10141670>
- Aldriano, M. A., & Priyambodo, M. A. (2022). Cyber crime dalam sudut pandang hukum pidana. *Jurnal Kewarganegaraan*, 6(1), 2169–2175.
- Alkatiri, M. A., & Sopacua, Y. (2025). Fenomena perilaku stalking pada second account Instagram di kalangan anak muda Ambon. *Jurnal Ilmu Komunikasi Pattimura*, 4(1), 32–44. <https://doi.org/10.30598/JIKPvol4iss1pp32-44>
- Amarini, I. (2018). Pencegahan dampak negatif perkembangan teknologi informasi terhadap pengguna internet. *Kosmik Hukum*, 18(1).
- Aprillia, M. D. (2024). Adaptasi cyberaggression typology questionnaire (CATQ) ke dalam bahasa Indonesia. <https://doi.org/10.21831/ap.v6i1.72160>
- Ayub, S., & Malik, F. (2024). Construction and validation of Cyber Harassment Experience Scale (CHES). *Pakistan Journal of Psychological Research*, 39(4), 875–894. <https://doi.org/10.33824/PJPR.2024.39.4.47>
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Durán-Guerrero, E., Nocentini, A., Menesini, E., & Sánchez-Jiménez, V. (2025). Online sexual harassment perpetration among peer adolescents: A cross-national and cross-gender study. *Behavioral Sciences*, 15(7), 969. <https://doi.org/10.3390/bs15070969>
- Drebing, H., Bailer, J., Anders, A., Wagner, H., & Gallas, C. (2014). Cyberstalking in a large sample of social network users: Prevalence, characteristics, and impact upon victims. *Cyberpsychology, Behavior, and Social Networking*, 17(2), 61–67. <https://doi.org/10.1089/cyber.2012.0231>
- Faizun, A., Aufa, A. W., Dasi, I., & Helmi, A. F. (2020). Pintu hubungan relasi: Cyberstalking di Indonesia. Dalam *Riset-Riset Cyberpsychology* (hlm. 148–163). Yogyakarta: Zahir Publishing.

- Irawan, P. (2014). Internet dan fenomena bangkitnya peran aktor non negara. *Jurnal Komunikasi*, 6(3), 59–73.
- Kaur, P., Dhir, A., Tandon, A., Alzeiby, E. A., & Abohassan, A. A. (2020). A systematic literature review on cyberstalking: An analysis of past achievements and future promises. *Technological Forecasting and Social Change*, 161, 120426. <https://doi.org/10.1016/j.techfore.2020.120426>
- Ketaren, E. (2016). Cybercrime, cyber space, dan cyber law. *Jurnal Times*, 5(2), 35–42.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Paullet, K. L., Rota, D. R., & Swan, T. T. (2009). Cyberstalking: An exploratory study of students at a Mid-Atlantic university. *Issues in Information Systems*, 10(2), 640–649.
- Rosana, A. S. (2010). Kemajuan teknologi informasi dan komunikasi dalam industri media di Indonesia. *Gema Eksos*, 5(2), 146–148.
- Short, E., Guppy, A., Hart, J. A., & Barnes, J. (2015). The impact of cyberstalking. *Studies in Media and Communication*, 3(2), 23–37. <https://doi.org/10.11114/smc.v3i2.1091>
- Santos, S.I. L., Pimentel, C. E., & Mariano, T. E. (2023). Cyberstalking scale: Development and relations with gender, FOMO and social media engagement. *Current Psychology*, 42(6), 4802–4810. <https://doi.org/10.1007/s12144-021-02603-4>
- Sinandang, R. J. M. W., & Pratiwi, R. (2024). The Indonesian Cyberstalking Scale: Adaptation and psychometric properties. *International Conference on Psychology UMBY*. E-ISSN: 3063-9506.
- Smoker, M., & March, E. (2017). Predicting perpetration of intimate partner cyberstalking: Gender and the dark tetrad. *Computers in Human Behavior*. Advance online publication. <https://doi.org/10.1016/j.chb.2017.03.012>
- Sugiyono. (2012). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Sun, Y.-X., & Chen, I.-H. (2024). Psychometric properties and measurement invariance of the Chinese version of the Internet Moral Literacy Scale. *Asian Journal of Social Health and Behavior*, 7(2), 67–76. [https://doi.org/10.4103/shb.shb\\_374\\_23](https://doi.org/10.4103/shb.shb_374_23)
- Spitzberg, B. H., & Cupach, W. R. (2007). The state of the art of stalking: Taking stock of the emerging literature. *Aggression and Violent Behavior*, 12(1), 64–86. <https://doi.org/10.1016/j.avb.2006.05.001>
- Spitzberg, B. H., Marshall, L., & Cupach, W. R. (2001). Obsessive relational intrusion, coping, and sexual coercion victimization. *Communication Reports*, 14(1), 19–30. <https://doi.org/10.1080/08934210109367733>
- Widyopramono. (2012). *Kejahatan di bidang komputer*. Pustaka Sinar Harapan.
- Wilson, C., Sheridan, L., & Garratt-Reed, D. (2022). What is cyberstalking? A review of measurements. *Journal of Interpersonal Violence*, 37(11–12), NP9763–NP9783. <https://doi.org/10.1177/0886260521997923>
- Wright, M. F. (2018). Cyberstalking victimization, depression, and academic performance: The role of perceived social support from parents. *Cyberpsychology, Behavior, and Social Networking*, 21(2), 110–116. <https://doi.org/10.1089/cyber.2016.0742>



## ANALISIS PSIKOMETRI *MORBID CURIOSITY SCALE* VERSI BAHASA INDONESIA

Galih Adriansyah<sup>1</sup>, Arinta Aisyah Putri<sup>2</sup>, Galih Khalisha Qalbina<sup>3</sup>, Safira Febriani<sup>4</sup>, Diah Zaleha Wyandini<sup>5</sup>, Ghinaya Ummul Mukminin Hidayat<sup>6</sup>

<sup>1,2,3,4,5,6</sup> Fakultas Ilmu Pendidikan, Universitas Pendidikan Indonesia, Bandung, Indonesia

Email: galih.adri10@gmail.com

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
19 Juni 2025	17 September 2025	8 Oktober 2025	30 Oktober 2025

### Abstract

Indonesians have a high interest in the mystical, horror and crime, which is reflected in the popularity of themed content that can lead to risky activities such as urban exploration. This phenomenon is related to morbid curiosity, a curiosity about horrific things such as death and violence, which can influence dangerous explorative behavior. In order to measure this interest and prevent risky behavior, researchers adapted the Morbid Curiosity Scale into an Indonesian version. During the analysis, this measuring instrument was given to 258 participants with male or female criteria, minimum age of 18 years and had a special interest in horror, thriller, supernatural, and crime themes. Based on item analysis using item-rest correlation and Confirmatory Factor Analysis, the reliability of the Indonesian version of the Morbid Curiosity Scale measuring instrument totaling 24 items shows a Cronbach's alpha value of 0.894 and chi-square test p-value of less than 0.001, CFI of 0.840, GFI of 0.954, and RMSEA of 0.079 which means that this scale cannot yet be used because the model does not fit even though the reliability is good. This study has limitations in terms of cultural differences that affect the understanding of items, especially items with a Western cultural context, so further research is needed so that the instrument can be used in the Indonesian cultural context. This instrument is expected to be useful for understanding how curiosity about frightening, mystical, and dangerous phenomena, both in digital media and in everyday life, plays a role in individual behavior.

Keywords: morbid curiosity, adaptation, scale, confirmatory factor analysis, Indonesian

### Abstrak

Masyarakat Indonesia memiliki ketertarikan tinggi terhadap hal-hal mistis, horor, dan kriminal, yang tercermin dari popularitas konten bertema tersebut yang dapat berisiko pada aktivitas membahayakan seperti *urban exploration*. Fenomena ini berkaitan dengan *morbid curiosity* atau keingintahuan terhadap hal mengerikan seperti kematian dan kekerasan yang dapat memengaruhi perilaku eksploratif berbahaya. Guna mengukur ketertarikan tersebut dan mencegah perilaku yang berisiko, maka peneliti melakukan adaptasi alat ukur terhadap *Morbid Curiosity Scale* ke dalam versi bahasa Indonesia. Pengujian alat ukur ini diberikan kepada 258 partisipan dengan kriteria laki-laki atau perempuan dengan minimal berusia 18 tahun dan memiliki ketertarikan khusus terhadap hal-hal bertema horor, *thriller*, supranatural, dan kriminal. Berdasarkan analisis item menggunakan *item-rest correlation* serta analisis faktor menggunakan *Confirmatory Factor Analysis*, reliabilitas alat ukur *Morbid Curiosity Scale* versi bahasa Indonesia yang berjumlah 24 item menunjukkan hasil nilai Cronbach's alpha sebesar 0.894 dan nilai p-value *Chi-square test* kurang dari 0.001, CFI sebesar 0.840, GFI sebesar 0.954, dan RMSEA sebesar 0.079 yang berarti skala ini belum dapat digunakan karena model tidak fit meskipun reliabilitasnya baik. Penelitian ini memiliki keterbatasan dalam hal perbedaan budaya yang memengaruhi pemahaman item, khususnya item dengan konteks budaya Barat sehingga diperlukan penelitian lanjutan agar alat ukur dapat digunakan pada konteks budaya Indonesia. Alat ukur ini diharapkan bermanfaat untuk memahami bagaimana keingintahuan terhadap fenomena menakutkan, mistis, dan berbahaya baik dalam media digital maupun kehidupan sehari-hari berperan dalam perilaku individu.

Kata kunci: morbid curiosity, adaptasi, alat ukur, analisis faktor konfirmatori, Indonesia

## 1. Pendahuluan

Indonesia merupakan negara dengan budaya masyarakat yang beragam. Salah satu budaya masyarakat Indonesia tersebut adalah kepercayaan terhadap hal mistis seperti mitos, legenda, dan lainnya yang mengakar sehingga turut mempengaruhi perilaku individu. Kepercayaan terhadap hal-hal tersebut dapat terlihat dari beragamnya *urban*

*legend* dan ritual tertentu yang sudah ada sejak zaman dahulu dan sifatnya sudah turun temurun, seperti *urban legend* populer Nyi Roro Kidul, daerah Kalimantan dengan sihir, dan Badui dengan ilmu kebal (Justine et al., 2021). Seiring dengan perkembangan teknologi dan dunia digital, konten-konten bertemakan mistis atau supranatural, horor, *thriller*, dan kriminal kerap kali menjadi konten yang populer di Indonesia, terlihat dari banyaknya *views*, *like*, *repost*, dan *comment (engagement)* pada konten-konten tersebut. Hal ini dibuktikan dengan beberapa karya, seperti film KKN di Desa Penari dengan 10 juta penonton (Taqiyya, 2023), hingga banyaknya penonton konten-konten horor di kanal YouTube seperti Jurnal Risa dan Nessie Judge. Begitu pun dengan konten mistis dan horor berbentuk cerita pada utas dalam platform X yang mengundang banyak pembaca. Fenomena ini diperkuat dengan studi oleh Yasim dan Yahya (2024) yang meneliti mengenai berbagai unsur mistis dalam film horor Indonesia yang menarik minat penonton dan mencerminkan kepercayaan masyarakat terhadap kekuatan supranatural. Begitupun juga dengan penelitian yang dilakukan Andianto dan Delliana (2022) yang menunjukkan bahwa salah satu unsur yang menarik dalam sebuah tayangan bagi masyarakat Indonesia adalah unsur mistisme.

Berdasarkan paparan data tersebut, dapat diketahui bahwa konten bertema mistis atau supranatural, horor, *thriller*, dan kriminal memiliki daya tarik yang signifikan dan menjadi konten populer di kalangan masyarakat Indonesia. Ketertarikan masyarakat Indonesia yang tinggi terhadap hal yang tidak menyenangkan dan tidak lazim tersebut dapat mengarah kepada aktivitas yang membahayakan keselamatan dan kesehatan individu yang bersangkutan, seperti salah satunya adalah aktivitas *urban exploration* yang kini banyak dilakukan terutama oleh kalangan anak muda dan menjadi salah satu konten populer di media sosial. *Urban exploration (urbex)* adalah hobi menjelajahi dan mendokumentasikan (melalui foto/video berkualitas tinggi) bangunan terbengkalai seperti pabrik, saluran pembuangan, dan sebagainya (Fulton, 2021). Contohnya eksplorasi bangunan rumah sakit terbengkalai, dapat menyebabkan paparan radiasi hasil rontgen. Adapun paparan patologi dan *blackmold* juga dapat mengancam kesehatan individu apabila individu tidak menggunakan masker dan perlindungan yang cukup (Saputra, 2023). Di Indonesia, aktivitas *urban exploration* menjadi tren dan ketertarikan tersendiri, hal ini dapat terlihat dari munculnya komunitas-komunitas *urbex* di berbagai platform media sosial seperti Facebook, Instagram, dan YouTube dengan jumlah pengikut dan *viewers* yang mencapai ribuan pengguna media sosial di Indonesia, salah satunya seperti komunitas @urbexpeoplelombok. Konten-konten *urban exploration* yang diunggah pada kanal YouTube Joe Kal mencapai 2,1 juta *viewers*. Adapun konten yang diunggah di kanal YouTube Sara Wijayanto mencapai 12 juta *viewers*. Konten yang disajikan dalam "Diary Misteri Sara" berhasil memberikan stimulus yang berdampak pada perubahan sikap audiens terhadap kepercayaan mistis, sehingga tayangan ini dapat dianggap berperan dalam membentuk pengetahuan, memperkuat keyakinan, atau memengaruhi pemikiran responden tentang hal-hal mistis (Nurdianti & Basarah, 2024).

Ketertarikan atau keingintahuan tentang hal-hal yang tidak menyenangkan, khususnya kematian dikaitkan dengan istilah *morbid curiosity* (Scrivner, 2021b). Sifat '*morbid*' secara esensial menempatkan kematian sebagai inti dari objek rasa ingin tahu. Ketertarikan ini secara khusus terarah pada berbagai aspek kematian itu sendiri, termasuk faktor-faktor penyebabnya. Scrivner (2021b) kemudian mengembangkan *Morbid Curiosity Scale* untuk mengukur keingintahuan seseorang tentang hal-hal yang tidak menyenangkan. Lebih lanjut dapat disimpulkan bahwa *morbid curiosity* tidak hanya merujuk pada rasa ingin tahu terhadap kematian itu sendiri, tetapi juga pada berbagai kondisi yang secara potensial dapat menimbulkan kematian. Dalam hal ini, Scrivner merujuk pada Zuckerman dan Litle (1986) yang mengembangkan skala *Curiosity About Morbid Events (CAME)*. CAME terdiri dari item-item yang mencerminkan ketertarikan atau kenikmatan dalam menonton kekerasan dan kematian. Mereka beralasan bahwa faktor pendorong dibalik keingintahuan tentang kejadian-kejadian yang tidak wajar adalah kebutuhan individu akan stimulasi dan gairah baru.

Penelitian tentang preferensi media dan *morbid curiosity* menyebutkan bahwa pada konteks ini, sebagian individu memilih untuk tidak melakukan penghindaran ketika melihat suatu lingkungan yang menjijikan dan menakutkan, tetapi justru merangsang keingintahuan pada area tersebut. Temuan ini menunjukkan bahwa pengalaman terhadap lingkungan semacam itu dapat memicu keingintahuan yang kuat, khususnya untuk memahami atau mengeksplorasi hal-hal yang dianggap mengancam. Hal ini diperkuat dalam penelitian oleh Scrivner dan Stubbersfield (2023) yang menjelaskan bahwa rasa ingin tahu tersebut muncul sebagai solusi untuk mencari dan mengetahui informasi terkait ancaman pada situasi berbahaya karena penghindaran total justru dapat menjadikan individu tidak tahu cara mengidentifikasi dan merespons situasi berbahaya dengan tepat jika situasi tersebut terjadi. Hasil studi ini menunjukkan bahwa rasa ingin tahu tersebut berfungsi sebagai mekanisme adaptif pada individu untuk memperoleh informasi tentang ancaman dalam situasi berbahaya. Berdasarkan pernyataan tersebut dapat dipahami bahwa *morbid curiosity* mampu mendorong perilaku eksploratif terhadap hal-hal berisiko, termasuk kecanduan terhadap obat-obatan terlarang. Studi menunjukkan bahwa individu dengan kecenderungan ini seringkali termotivasi oleh kebutuhan psikologis untuk menghilangkan ketidakpastian, meskipun konsekuensinya negatif. Dorongan ini dapat membuat individu mengeksplorasi fenomena yang ambigu atau berbahaya, seperti kecelakaan atau narkoba, sehingga meningkatkan rasa ingin tahunya dan membuat mereka lebih rentan terhadap kecanduan (Ghaemi et al., 2024a).

Merujuk pada berbagai temuan dan fakta lapangan yang telah dipaparkan di atas, peneliti tertarik untuk mengadaptasi alat ukur *Morbid Curiosity Scale* yang sebelumnya telah dikembangkan oleh Scrivner (2021b) ke dalam versi bahasa Indonesia. Pemilihan alat ukur ini didasarkan pada fakta bahwa sampai saat ini hanya tersedia

dua alat ukur untuk mengukur *morbid curiosity* yaitu *Morbid Curiosity Scale* (Scrivner, 2021b) itu sendiri dan skala *Curiosity About Morbid Events* (CAME) (Zuckerman & Little, 1986). Namun, skala CAME bersifat unidimensional serta reliabilitas dan validitasnya belum dievaluasi secara luas. Selain itu, item-item dalam skala tersebut hanya berfokus pada aspek *violence* dan tidak membahas aspek lain yang berkaitan dengan *morbid curiosity*, sehingga kurang komprehensif untuk menyelidiki *morbid curiosity*. Di sisi lain, *Morbid Curiosity Scale* (Scrivner, 2021b) menunjukkan konsistensi internal yang sangat baik, retest reliability yang kuat, validitas konvergen dan diskriminan yang baik, struktur faktor yang stabil dan kemampuan untuk mengukur berbagai aspek *morbid curiosity* mencakup *interpersonal violence*, *body violations*, *minds of dangerous people*, dan *paranormal danger* (Wang et al., 2023).

Studi pendahuluan yang dilakukan oleh peneliti menemukan sampai saat ini *Morbid Curiosity Scale* telah diadaptasi ke dalam bahasa Persia dengan dimensi *Body*, *Supernatural*, *Violence*, dan *Mind* oleh Ghaemi et al. (2024b). Sampel pada penelitian *Morbid Curiosity Scale* dalam bahasa Persia ini adalah mahasiswa Zanzan University dengan total 500 orang. Adaptasi *Morbid Curiosity Scale* oleh Ghaemi et al. (2024b) menunjukkan nilai *Cronbach's alpha* secara keseluruhan 0.85 dan per dimensi 0.79, 0.79, 0.80, dan 0.81. Hasil penelitian ini menunjukkan bahwa *Morbid Curiosity Scale* versi bahasa Persia ini menunjukkan model yang fit dan dapat digunakan untuk mengukur *morbid curiosity* pada mahasiswa.

Adaptasi *Morbid Curiosity Scale* ke dalam bahasa Indonesia dilakukan karena dua pertimbangan utama, yaitu belum tersedianya versi alat ukur ini dalam bahasa Indonesia dan instrumen asli telah relevan untuk mengukur fenomena yang marak terjadi di Indonesia serta memiliki properti psikometrik yang teruji. Dengan demikian, penelitian ini bertujuan menyediakan *Morbid Curiosity Scale* versi bahasa Indonesia yang memenuhi standar psikometrik untuk kebutuhan pengukuran di populasi Indonesia. Sehingga diharapkan alat ukur ini dapat mengukur tingkat ketertarikan atau keingintahuan masyarakat Indonesia terhadap berbagai aspek atau aktivitas yang berkaitan dengan hal-hal tidak menyenangkan, seperti peristiwa kematian, tidak lazim, atau situasi menyeramkan yang terjadi baik di kehidupan sehari-hari maupun dari film, utas (*thread*), video, cerita pendek, dan konten lainnya yang ada dan populer di berbagai platform.

## 2. Metode Penelitian

### Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain penelitian non-eksperimental menggunakan survei yang selaras dengan tujuan penelitian yaitu mengadaptasi dan menguji properti psikometris dari *Morbid Curiosity Scale* versi bahasa Indonesia. Langkah-langkah pengadaptasian alat ukur merujuk pada International Test Commission for Translating and Adapting Tests (International Test Commission, 2017) yaitu *pre-condition*, *test development*, *confirmation*, *score and scale interpretation*, dan *documentation*.

Tahap pertama mengadaptasi alat ukur ini adalah *pre-conditioning*, yakni peneliti meminta izin kepada pemilik hak cipta untuk mengadaptasi tes melalui *e-mail* dan sudah mendapatkan izin langsung dari pemilik untuk dilakukan adaptasi. Selanjutnya, peneliti melakukan *review* dalam bentuk diskusi bersama ahli di bidang psikologi guna memastikan kesesuaian topik *morbid curiosity* dengan fenomena yang berkembang di masyarakat Indonesia, menilai potensi sensitivitas budaya dalam alat ukur, dan pertimbangan praktis terkait penggunaan alat ukur dalam penggunaan alat ukur dalam penelitian di Indonesia.

Tahap kedua adalah *test development* dengan melakukan penerjemahan langsung (*forward translation*) dari bahasa Inggris ke bahasa Indonesia sekaligus mengubah unsur kata atau konteks bahasa yang tidak relevan agar lebih sesuai dengan aspek kebudayaan Indonesia dan sesuai dengan target populasi yang dituju. Tahap alih bahasa ini diterjemahkan oleh ahli dengan latar belakang psikologi.

Tahap ketiga adalah *confirmation*, yakni peneliti menentukan sampel responden yang sesuai dengan tujuan dari uji alat ukur ini yaitu perempuan dan laki-laki berusia minimal 18 tahun dan pernah atau sering menonton film horor/*thriller*/dokumenter kriminal. Jumlah responden yang dihimpun merujuk pada pernyataan dari Hair et al. (2019) bahwa jumlah sampel minimal lima kali lebih banyak dari jumlah item yang dianalisis yaitu 25 item. Kemudian peneliti menentukan item-item final yang telah disesuaikan dengan budaya dari sampel yang dituju dan melakukan *expert judgement* sebagai uji validitas isi bersama ahli psikometri yang memiliki kompetensi di bidang konstruksi alat ukur dan memahami konstruk *morbid curiosity*. Proses ini berbeda dengan tahapan *review* awal karena pada tahap ini instrumen telah diterjemahkan dan disesuaikan secara budaya. Hasil penilaian digunakan untuk menentukan item-item final yang akan disebarakan kepada partisipan penelitian.

Tahap keempat adalah *administration*. Pada tahap ini, peneliti menentukan instruksi pengerjaan untuk responden yaitu, "*Anda diminta untuk menilai seberapa sesuai pernyataan tersebut dengan diri Anda. Tidak ada jawaban yang benar atau salah. Silakan jawab dengan jujur dan pilih opsi yang paling menggambarkan perasaan Anda yang sebenarnya*". Jangka waktu pengisian kuesioner adalah 15 menit. Dalam penyajiannya, diberikan empat opsi jawaban, yakni (1) Sangat Tidak Sesuai, (2) Tidak Sesuai, (3) Sesuai, (4) Sangat Sesuai. Prosedur pengambilan data dilakukan secara *offline* menggunakan *paper and pencil format*, serta secara *online* menggunakan media Google Form yang tautannya dibagikan ke responden yang sesuai dengan kriteria melalui media sosial Whatsapp, Twitter, dan lainnya. Peneliti menjamin hak responden dengan menyertakan bagian

*informed consent* sebelum pengisian kuesioner. Hal ini bertujuan untuk memastikan partisipasi responden bersifat sukarela dan informasi yang diberikan akurat. peneliti juga memastikan bahwa data penelitian bersifat rahasia dan hanya akan digunakan untuk kepentingan penelitian.

Tahap kelima adalah *score and scale interpretation*. Pada tahap ini peneliti melakukan analisis data dengan mengolah seluruh respons yang diberikan oleh partisipan melalui kuesioner yang disebarluaskan pada tahap *administration*. Selanjutnya, reliabilitas alat ukur diestimasi menggunakan *Cronbach's alpha*. Jenis koefisien reliabilitas ini paling sering digunakan dalam literatur yang mengukur reliabilitas konsistensi internal atau sejauh mana respons konsisten di seluruh item pengukuran (Kline, 2016). Validitas internal pada penelitian ini dilakukan dengan mempertimbangkan nilai *item-rest correlation* dan uji validitas konstruk melalui *Confirmatory Factor Analysis*. CFA memberikan kerangka kerja analitik yang sangat kuat untuk menguji kesetaraan model pengukuran di berbagai kelompok yang berbeda termasuk kelompok yang berbeda budaya (Brown, 2006).

Kemudian pada tahap terakhir, yaitu *documentation*. Peneliti mendokumentasikan seluruh proses adaptasi alat ukur, mencakup langkah-langkah yang digunakan, proses pengadministrasian alat ukur, hasil pengumpulan data, dan data hasil evaluasi data psikometrik. Tahap ini penting untuk dilakukan sebagai bentuk akuntabilitas dan transparansi selama proses adaptasi.

### Partisipan Penelitian

Partisipan dalam penelitian ini dijamin menggunakan teknik *convenience sampling* yaitu penyaringan dengan memilih partisipan yang mudah untuk diakses atau dijangkau (Fraenkel et al., 2012). Kriteria partisipan adalah perempuan dan laki-laki berusia minimal 18 tahun dan pernah atau sering menonton film horor/*thriller*/dokumenter kriminal atau pembunuhan. Jumlah partisipan minimal dalam penelitian ini mengacu pada ketentuan yang dikemukakan oleh Hair et al. (2019) bahwa jumlah sampel minimal lima kali lebih banyak dari jumlah item yang dianalisis. Total partisipan yang terlibat dalam pengisian skala ini berjumlah 258 orang yang secara rinci dijelaskan dalam tabel 1 berikut:

**Tabel 1. Data Demografi Partisipan Penelitian**

Karakteristik	n	%
Jenis Kelamin		
Perempuan	176	68.22%
Laki-laki	82	31.78%
Suku Bangsa		
Sunda	142	55.04%
Jawa	47	18.22%
Betawi	11	4.26%
Lainnya	58	22.48%
Usia		
18-23	208	80.62%
24-29	29	11.24%
30-35	14	5.43%
36 tahun ke atas	7	2.71%

Sebanyak 104 orang partisipan berasal dari pengambilan data secara *offline* dengan *paper and pencil format* di sekitar Kota Bandung dan 154 orang dari pengambilan data secara *online* dengan kuesioner Google Form. Total partisipan berjumlah 258 orang tersebut yang terbagi atas partisipan berjenis kelamin perempuan berjumlah 176 orang (68.22%) dan partisipan laki-laki berjumlah 82 orang (31.78%). Berdasarkan jumlah dan persentase tersebut diketahui bahwa partisipan mayoritas berjenis kelamin perempuan. Seluruh partisipan ini berasal dari berbagai wilayah serta suku bangsa yang berbeda. Usia partisipan paling banyak berada di rentang 18-23 tahun yaitu 208 orang. Berdasarkan suku bangsa, partisipan yang berasal dari suku Sunda mendominasi dengan jumlah 142 orang, dilanjutkan dengan suku Jawa dengan jumlah 47 orang, suku Betawi dan Batak masing-masing berjumlah 11 orang dan 8 orang. Selain itu, terdapat 50 partisipan yang berasal dari berbagai suku lain di Indonesia yang menunjukkan beragamnya etnis partisipan dalam penelitian ini.

### Instrumen Penelitian

Instrumen penelitian yang diadaptasi dalam penelitian ini adalah *Morbid Curiosity Scale* yang dikembangkan oleh Scrivner (2021b) yang terdiri dari item-item yang mencerminkan ketertarikan atau kenikmatan dalam menonton kekerasan dan kematian. Pengembangan alat ukur ini dilakukan Scrivner melalui uji *Exploratory Factor Analysis* dan kemudian dilanjutkan dengan uji *Confirmatory Factor Analysis*. Alat ukur ini terdiri atas 24 item yang terbagi menjadi empat dimensi, yaitu *minds* (mengandung item yang berfokus pada pemahaman akan motivasi individu-individu berbahaya), *paranormal* (mengandung item yang menunjukkan minat terhadap fenomena yang tampak bertentangan dengan ilmu pengetahuan atau hukum alam), *body* (mengandung item yang menunjukkan minat dalam memahami batas-batas tubuh dan apa yang terjadi ketika tubuh mengalami kerusakan),

dan *violence* (mengandung item yang menunjukkan ketertarikan untuk melihat tindakan kekerasan tanpa harus memahami motifnya). Reliabilitas alat ukur secara keseluruhan sebesar 0.94 dengan reliabilitas keempat dimensinya yaitu: *minds* 0.92, *paranormal* 0.90, *body* 0.87, *violence* 0.89. Kemudian dari empat dimensi tersebut peneliti mengembangkan sepuluh indikator. Total item yang dicantumkan dalam *blue print* sejumlah 25 item karena terdapat satu item yang mengandung dua pernyataan (*double-barreled question*) sehingga peneliti memisahkannya menjadi dua item terpisah karena dalam *double-barreled question* responden cenderung hanya memperhatikan salah satu stimulus saja (Menold, 2020). Semua butir item yang dicantumkan merupakan item favorable dengan menggunakan skala penilaian 1) Sangat Sesuai, 2) Sesuai, 3) Tidak Sesuai, 4) Sangat Tidak Sesuai.

**Tabel 2. *Blue Print Morbid Curiosity Scale* Versi Bahasa Indonesia**

Dimensi	Indikator	No. Item	Contoh Item Adaptasi
<i>Minds</i>	Individu memiliki ketertarikan pada kasus kriminal	3, 4	Saya senang membaca berita detail tentang pembunuhan dan kekerasan lainnya
	Individu memiliki ketertarikan pada tayangan mengenai kasus pembunuhan	8, 12, 16	Saya tertarik menonton wawancara narapidana kasus pembunuh berantai yang membahas kejahatannya
	Individu memiliki ketertarikan untuk memahami motif pelaku tindak kekerasan	20, 24	Saya tertarik menjadi orang yang mempelajari kepribadian pembunuh untuk mengidentifikasi pelaku kriminal
<i>Paranormal</i>	Individu tertarik pada topik-topik dan/atau kegiatan supranatural	5, 9, 13, 25	Menurut saya, topik supranatural merupakan topik yang menarik
	Individu tertarik untuk mengetahui cara kerja hal-hal yang bersifat supranatural	17, 21	Sebuah dokumenter mengenai boneka santet ( <i>voodoo doll</i> ) akan memiliki makna yang menarik
<i>Body</i>	Individu tertarik untuk mengetahui apa yang terjadi pada tubuh yang rusak	2, 7, 11, 15	Seandainya operasi transplantasi kepala dapat dilakukan, saya ingin melihat prosedurnya
	Individu tertarik untuk mengetahui batas (limit) dari tubuh manusia	19, 23	Saya penasaran efek yang akan ditimbulkan oleh racun paling mematikan di dunia ini terhadap tubuh
<i>Violence</i>	Individu tertarik merasakan peristiwa kekerasan yang terjadi di masa lalu	1, 6, 14, 22	Jika saya hidup di era Abad Pertengahan, saya tertarik menghadiri eksekusi di publik
	Individu tertarik menonton kekerasan tanpa sensor	18	Saya lebih menyukai film dan acara TV yang berbau kekerasan ditayangkan tanpa sensor
	Individu tertarik dengan kekerasan yang terjadi di muka umum	10	Jika saya melihat tawuran di jalan dan tahu saya tidak bisa melerainya, saya akan mencoba untuk menontonnya

### Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan dengan metode statistik dengan menggunakan bantuan *software* JASP versi 0.19.1 (JASP Team, 2024). Analisis pertama yang dilakukan adalah uji *item-rest correlation*. Menurut (DeVellis, 2017) koefisien *item-rest correlation* dapat digunakan sebagai barometer untuk menentukan item mana yang dapat dibuang. Melalui uji ini, dapat dilihat sejauh mana item tersebut konsisten dengan item-item lainnya secara keseluruhan (Furr, 2011; Zijlmans et al., 2018). Semakin tinggi koefisien *item-rest correlation* dalam sebuah tes akan menghasilkan koefisien  $\alpha$  yang tinggi pula (Lord & Novick dalam Zijlmans et al., 2018). Oleh karena itu, item dengan koefisien *item-rest correlation* terendah harus dieliminasi terlebih dahulu karena tidak berkontribusi secara signifikan. Kriteria pemilihan item berdasarkan *item-rest correlation* umumnya menggunakan batas koefisien  $\geq 0.30$  (Azwar, 2016). Selanjutnya, peneliti melakukan uji validitas konstruk menggunakan *Confirmatory Factor Analysis* (CFA). Uji CFA bertujuan untuk mengevaluasi sejauh mana model tertentu sesuai dengan data dan jika model tersebut cocok, untuk memperkirakan muatan faktor, varian, dan kovarian faktor, dan varian kesalahan residual dari variabel yang diamati (Hox, 2021). indikator yang digunakan untuk menentukan apakah model yang dibuat baik atau fit adalah nilai indeks nilai rasio  $X^2/df < 2$  dengan *p-value*

> 0.05, nilai *Comparative Fit Index* (CFI) > 0.90, nilai *Goodness Fit Index* (GFI) > 0.90, nilai *Root Mean Square Error of Approximation* (RMSEA) antara 0.03-0.08 (Hair et al., 2019).

Kriteria dalam menafsirkan nilai muatan faktor (*factor loadings*) merujuk pada Tabachnik dan Fidell (dalam Harrington, 2009) yaitu, nilai di atas 0.71 adalah *excellent*; 0.63, *very good*; 0.55, *good*; 0.45, *fair*; dan 0.32, *poor*. Aturan ini diperkuat dalam Hair et al. (2019) bahwa standar yang baik adalah nilai *standardized loading estimates* harus besarnya 0.5 atau lebih tinggi, dan idealnya 0.7 atau lebih tinggi. Uji reliabilitas dalam penelitian ini menggunakan kriteria dari DeVellis (2017) yaitu < 0.60 *unacceptable*; antara 0.60-0.65, *undesirable*; antara 0.65-0.70, *minimally acceptable*; antara 0.70-0.80, *respectable*; antara 0.80-0.90, *very good*; jauh di atas 0.90, harus mempertimbangkan untuk memperpendek skala. Kesetaraan pengukuran pada prosedur pengambilan data *offline* dan *online* dilakukan melalui uji invariansi pengukuran menggunakan *Multi-group CFA*. Menurut (Chen, 2007) apabila ukuran sampel kecil ( $n \leq 300$ ) maka invariansi pengukuran tidak terpenuhi jika selisih CFI  $\leq -0.005$ , selisih RMSEA  $\geq 0.010$  atau selisih SRMR  $\geq 0.025$  pada uji *loading invariance* dan selisih CFI  $\leq -0.005$ , selisih RMSEA  $\geq 0.010$  atau selisih SRMR  $\geq 0.005$  pada uji *intercept* dan *residual invariance*.

### 3. Hasil dan Diskusi

#### Uji *Item-rest Correlation*

Analisis item mempertimbangkan besaran skor *item-rest correlation* masing-masing item dalam alat ukur menggunakan bantuan software JASP versi 0.19.1 (JASP Team, 2024). Hasil analisis item disajikan dalam tabel 3 berikut:

**Tabel 3. Hasil Uji *Item-rest Correlation Morbid Curiosity Scale* Versi Bahasa Indonesia**

No. Item	Item	<i>Item-Rest Correlation</i>	Keterangan
1	Jika saya hidup di era Abad Pertengahan, saya tertarik menghadiri eksekusi di publik	0.435	Layak
2	Seandainya operasi transplantasi kepala dapat dilakukan, saya ingin melihat prosedurnya	0.402	Layak
3	Saya senang membaca berita detail tentang pembunuhan dan kekerasan lainnya	0.503	Layak
4	Saya penasaran mengenai tindak kejahatan	0.637	Layak
5	Menurut saya, topik supranatural merupakan topik yang menarik	0.396	Layak
6	Jika saya hidup di era Romawi Kuno, saya tertarik menghadiri pertarungan gladiator	0.434	Layak
7	Saya ingin tahu bagaimana proses otopsi jenazah dilakukan	0.523	Layak
8	Saya tertarik menonton film dokumenter tentang motif suatu pembunuhan	0.487	Layak
9	Saya akan tertarik untuk menonton video tentang <i>exorcism</i> (ruqyah jin, pengusiran setan/roh jahat)	0.509	Layak
10	Jika saya melihat tawuran di jalan dan tahu saya tidak bisa melerainya, saya akan mencoba untuk menontonnya	0.271	Tidak Layak
11	Saya tertarik untuk melihat bagaimana proses amputasi anggota tubuh dilakukan	0.488	Layak
12	Bagian favorit saya dalam tontonan kriminal adalah saat mempelajari bagaimana pembunuh melakukan aksinya	0.570	Layak
13	Saya merasa hal-hal gaib itu menarik	0.401	Layak
14	Saya cukup tertarik melihat pertarungan ala koboi jika saya hidup di era koboi Wild West	0.369	Layak
15	Saya ingin melihat bagaimana proses jenazah dipersiapkan untuk dimakamkan	0.450	Layak
16	Saya tertarik menonton wawancara narapidana kasus pembunuh berantai yang membahas kejahatannya	0.515	Layak
17	Sebuah dokumenter mengenai boneka santet ( <i>voodoo doll</i> ) akan memiliki makna yang menarik	0.484	Layak
18	Saya lebih menyukai film dan acara TV yang berbau kekerasan ditayangkan tanpa sensor	0.574	Layak
19	Menurut saya, pengawetan jenazah seperti taksidermi atau mumifikasi adalah sesuatu yang sangat menarik	0.532	Layak
20	Saya tertarik menjadi orang yang mempelajari kepribadian pembunuh untuk mengidentifikasi pelaku kriminal	0.507	Layak

21	Saya penasaran bagaimana jailangkung bekerja	0.424	Layak
22	Saya sangat penasaran tentang bagaimana pertempuran terjadi di Abad Pertengahan	0.511	Layak
23	Saya penasaran efek yang akan ditimbulkan oleh racun paling mematikan di dunia ini terhadap tubuh	0.537	Layak
24	Saya tertarik dengan pola pikir dari orang-orang yang melakukan kekerasan	0.519	Layak
25	Menurut saya ilmu sihir akan menjadi topik yang menarik untuk dipelajari	0.397	Layak

Berdasarkan analisis item yang sudah dilakukan didapatkan hasil bahwa seluruh item memiliki nilai koefisien > 0.30, kecuali item 10 nilai koefisiennya < 0.30 yaitu 0.271 sehingga item 10 dieliminasi dan tidak disertakan dalam analisis lanjutan.

### Uji *Confirmatory Factor Analysis*

Selanjutnya dilakukan uji validitas konstruk yaitu pengujian untuk menunjukkan bahwa konstruk pengukuran teoretis yang telah dikembangkan di awal perancangan tes adalah benar dalam arti didukung oleh data empiris (Azwar, 2021). Uji validitas konstruk dimulai dengan melakukan analisis koefisien *Measure of Sampling Adequacy* (MSA) dari *Kaiser-Meyer-Olkin* (KMO) test serta koefisien *Bartlett test of sphericity*. Adapun menurut Hair et al. (2019) peneliti perlu memastikan bahwa nilai MSA keseluruhan berada di atas 0.50 dan *Bartlett test* menghasilkan nilai di bawah 0.05 sebelum melakukan analisis faktor. Koefisien MSA menunjukkan hasil 0.883 sedangkan *Bartlett test* menghasilkan *p-value* < 0.001 sehingga analisis faktor dapat dilakukan.

Analisis faktor menggunakan *Confirmatory Factor Analysis* (CFA) memungkinkan peneliti untuk mengevaluasi sejauh mana hipotesis pengukuran yang dibuat sebelumnya konsisten dengan data aktual yang dihasilkan oleh responden (Furr, 2011; Hair et al., 2019). Pengujian reliabilitas dilakukan secara multidimensi dikarenakan konstruk *morbid curiosity* terdiri atas empat dimensi (Scrivner, 2021b). Hasil uji reliabilitas menghasilkan *Cronbach's alpha Morbid Curiosity Scale* versi bahasa Indonesia secara keseluruhan yaitu 0.894 serta reliabilitas masing-masing dimensi yaitu *minds* 0.863, *body* 0.793, *paranormal* 0.861, dan *violence* 0.749.

**Tabel 4. Reliabilitas *Morbid Curiosity Scale* Versi Bahasa Indonesia**

Dimensi	<i>Cronbach's Alpha</i>
<i>Minds</i>	0.863
<i>Paranormal</i>	0.861
<i>Body</i>	0.793
<i>Violence</i>	0.749
<b>Total</b>	0.894

Kemudian untuk mengetahui kesesuaian antara teori dan model yang telah disusun, peneliti menggunakan beberapa jenis indeks kecocokan. Menurut (Hair et al., (2019), peneliti sebaiknya mencantumkan setidaknya satu *incremental indices* dan satu *absolute indices* karena jika hanya menggunakan hanya satu indeks GOF, meskipun nilainya relatif tinggi, tidak secara kuat mendukung hasil *Chi-Square* itu sendiri. Oleh karena itu, peneliti menguji indeks kecocokan menggunakan *Chi-square*, *Standardized Root Mean Residual* (SRMR), *Comparative Fit Index* (CFI), *Root Mean Square Error of Approximation* (RMSEA) yang dapat dilihat pada tabel 5 berikut:

**Tabel 5. Uji Model Fit CFA**

Indeks Fit	Kriteria	Value	Keterangan
<i>Chi-square</i>	>0.05	<0.001	Tidak fit
GFI	>0.90	0.954	Fit
CFI	>0.90	0.840	Tidak fit
RMSEA	0.03-0.08	0.079	Fit

Hasil uji CFA juga menghasilkan nilai *factor loading* dari seluruh item dalam *Morbid Curiosity Scale* versi bahasa Indonesia. Nilai yang dipergunakan merupakan nilai *factor loadings* terstandarisasi (*standardized loadings estimates*) yang dapat dilihat dalam tabel 6 berikut:

**Tabel 6. Factor Loadings Morbid Curiosity Scale Versi Bahasa Indonesia**

Item	Estimate	p
Faktor 1: <i>Minds</i>		
Item3	0.667	< 0.001
Item4	0.739	< 0.001
Item8	0.707	< 0.001
Item12	0.644	< 0.001
Item16	0.735	< 0.001
Item20	0.714	< 0.001
Item24	0.637	< 0.001
Faktor 2: <i>Paranormal</i>		
Item5	0.712	< 0.001
Item9	0.766	< 0.001
Item13	0.745	< 0.001
Item17	0.780	< 0.001
Item21	0.698	< 0.001
Item25	0.539	< 0.001
Faktor 3: <i>Body</i>		
Item2	0.576	< 0.001
Item7	0.694	< 0.001
Item11	0.673	< 0.001
Item15	0.553	< 0.001
Item19	0.687	< 0.001
Item23	0.585	< 0.001
Faktor 4: <i>Violence</i>		
Item1	0.606	< 0.001
Item6	0.687	< 0.001
Item14	0.594	< 0.001
Item18	0.642	< 0.001
Item22	0.554	< 0.001

Berdasarkan tabel 6 dapat diketahui bahwa seluruh item memiliki nilai *factor loadings* di atas kriteria yang ditetapkan karena nilai terkecil berada di angka 0.539 dan nilai tertinggi berada di angka 0.780 serta signifikansi seluruh item adalah < 0.001. Selanjutnya, peneliti melakukan uji invariansi pengukuran melalui *Multi-group CFA* untuk mengetahui sejauh mana pengukuran pada prosedur pengambilan data *offline* dan *online* setara. Hasil uji invariansi pengukuran disajikan dalam tabel 7 berikut:

**Tabel 7. Uji Invariansi Pengukuran Pengambilan Data *Offline* dan *Online***

Model	Indeks Model Fit				Perbandingan Model		
	X <sup>2</sup>	df	p-value	CFI	RMSEA	ΔCFI	ΔRMSEA
Invariansi konfigurasi	959.625	496	< 0.001	0.815	0.085		
Invariansi metrik	993.673	519	< 0.001	0.811	0.084	-0.004	-0.001
Invariansi skalar	1,065.265	538	< 0.001	0.790	0.087	-0.021	0.003
Invariansi kovarians	1,099.020	562	< 0.001	0.786	0.086	-0.004	-0.001

Berdasarkan hasil uji invariansi pengukuran pada tabel 7 dan merujuk pada kriteria yang ditetapkan oleh Chen (2007), model memenuhi sebagian besar tahapan invariansi antara pengambilan data *online* dan *offline*. Pengujian invariansi konfigural, metrik, skalar, dan kovarians, *p-value* menunjukkan hasil < 0.001. Selanjutnya, pengujian invariansi metrik menunjukkan perubahan ΔCFI yaitu -0.004 dan ΔRMSEA yaitu -0.001 yang masih dalam batas toleransi (Chen, 2007). Namun, pada uji invariansi skalar, ditemukan penurunan CFI yang melebihi batas yaitu -0.021 meskipun perubahan RMSEA masih memenuhi yaitu 0.003. Di sisi lain, invariansi kovarians terpenuhi dengan perubahan CFI dan RMSEA masing-masing -0.004 dan -0.001. Hasil uji invariansi pengukuran menunjukkan model yang direkomendasikan adalah model dengan invariansi metrik.

#### 4. Pembahasan

Penelitian ini dilakukan dengan tujuan untuk mengadaptasi *Morbid Curiosity Scale* yang dikembangkan oleh Scrivner (2021b) ke dalam bahasa Indonesia guna tersedianya alat ukur yang dapat mengukur ketertarikan yang tinggi pada masyarakat Indonesia terhadap konten-konten tidak menyenangkan bertema horor, *thriller*, *supernatural*, dan kriminal yang kerap kali menjadi konten populer di Indonesia. Dalam rangka mendorong

pengukuran terhadap hal-hal tersebut di kajian ilmiah Indonesia, perlu tersedianya alat ukur yang sudah teruji secara psikometris.

Hasil analisis statistik menggunakan metode *item-rest correlation* menunjukkan bahwa seluruh item dinilai layak karena memiliki nilai koefisien  $> 0.30$ , kecuali item 10 nilai koefisiennya  $< 0.30$  yaitu 0.271. Oleh karena itu, item 10 yang berisi pernyataan “*Jika saya melihat tawuran di jalan dan tahu saya tidak bisa melerainya, saya akan mencoba untuk menontonnya,*” dieliminasi sehingga item final berjumlah 24 item. Selanjutnya, nilai koefisien MSA dari *KMO test* sebesar 0.883 dan koefisien *Bartlett test of sphericity* menunjukkan  $p\text{-value} < 0.001$  mengartikan bahwa data memenuhi syarat untuk dilakukan analisis faktor. Hasil analisis faktor dengan CFA menunjukkan *Morbid Curiosity Scale* versi bahasa Indonesia yang berjumlah 24 memiliki reliabilitas *Cronbach's alpha* keseluruhan sebesar 0.894 serta reliabilitas masing-masing dimensi berada di rentang 0.749-0.863. Merujuk pada DeVellis et al. (2017) nilai reliabilitas keseluruhan dapat dikategorikan *very good*. Sedangkan pada reliabilitas per dimensi, terdapat dua dimensi yang berada pada kategori *very good* dan dua dimensi lainnya berada pada kategori *respectable*. Berdasarkan hasil uji indeks kecocokan model dalam Tabel 5 yang merujuk pada kriteria yang dibuat oleh Hair et al., (2019), terdapat dua indeks yang menunjukkan hasil tidak fit dan dua indeks yang menunjukkan hasil fit. *Chi-square test* digunakan untuk menguji seberapa dekat kesesuaian antara matriks kovarian sampel penelitian dengan matriks kovarian model (Khairi et al., 2021). Nilai  $p\text{-value Chi-square test}$  yang tidak fit menunjukkan adanya perbedaan signifikan antara struktur faktorial yang diperoleh dari data partisipan dengan struktur faktorial yang diestimasi oleh model. Namun demikian, tes ini sensitif terhadap jumlah sampel, di mana ketika sampel semakin besar maka nilai *Chi-square* juga akan semakin besar ( $p\text{-value}$  akan semakin signifikan). Oleh karena itu, *Chi-square test* tidak dapat dijadikan sebagai satu-satunya indeks dalam menguji kecocokan model (Khairi et al., 2021). Selanjutnya, indeks *Comparative Fit Index* (CFI) membandingkan antara fit model yang diajukan peneliti dengan fit model dasar (*null model*) yang mengasumsikan seluruh item dalam skala tidak saling berhubungan (Roos & Bauldry, 2022). Nilai CFI yang tidak fit mengindikasikan bahwa terdapat perbedaan cara pandang dari partisipan terhadap konsep *morbid curiosity* yang digambarkan melalui beberapa item dalam skala.

Meskipun demikian, *Goodness of Fit Index* (GFI) menunjukkan angka fit yang artinya struktur model adaptasi alat ukur mampu merepresentasikan data partisipan dengan baik. Terakhir, *Root Mean Square Residual* (RMSEA) adalah indeks “kesalahan perkiraan” karena mengukur sejauh mana suatu model menunjukkan kecocokan yang cukup baik dalam populasi (Brown, 2006). Nilai RMSEA yang fit menunjukkan adaptasi *Morbid Curiosity Scale* versi bahasa Indonesia memiliki kecocokan yang cukup baik untuk sampel penelitian ini maupun untuk populasi masyarakat Indonesia yang lebih luas. Selain itu, hasil uji CFA menunjukkan bahwa seluruh item memiliki nilai *factor loading* yang baik karena memenuhi kriteria yaitu berada di rentang 0.539-0.780. Seluruh item menunjukkan koefisien *factor loading*  $> 0.5$  dan nilai signifikansi  $p < 0.001$  yang mengindikasikan bahwa indikator tersebut berhubungan kuat dengan konstruk laten yang diukur dan merupakan salah satu tanda validitas konstruk (Hair et al., 2019). Artinya, pertanyaan-pertanyaan dalam skala ini berhasil mengukur konsep *morbid curiosity* yang sama dengan versi aslinya. Sebagai tambahan, uji invariansi pengukuran dilakukan menggunakan *Multigroup-CFA* untuk memastikan kesetaraan pengukuran antara pengambilan data yang dilakukan secara *offline* dan *online* dalam penelitian.

Hasilnya menunjukkan bahwa pada seluruh tingkat pengujian, nilai  $p\text{-value}$  adalah  $< 0.001$  yang berarti terdapat perbedaan distribusi data antar kelompok, tetapi mengingat *Chi-square test* sensitif terhadap besaran sampel maka perlu dilihat juga perubahan pada nilai CFI dan RMSEA. Hasil akhir menunjukkan bahwa *Morbid Curiosity Scale* versi bahasa Indonesia memenuhi invariansi metrik antara data yang dikumpulkan secara *offline* dan *online* yang berarti bahwa setiap item berkontribusi pada konstruk laten dengan derajat yang serupa di dua kelompok berbeda (Putnick & Bornstein, 2016).

Berdasarkan kriteria yang dibuat oleh Hair et al. (2019) hasil uji kecocokan model menunjukkan bahwa alat ukur yang diadaptasi oleh peneliti tidak masuk ke dalam kategori fit. Dalam model CFA, sumber potensial utama terjadinya ketidaksesuaian model adalah jumlah faktor yang terlalu sedikit atau terlalu banyak, pemilihan indikator, pola indikator - *factor loading*, dan kesalahan teori (misalnya, kesalahan pengukuran yang tidak berkorelasi vs berkorelasi) (Brown, 2006). Alat ukur yang diadaptasi peneliti tidak menghasilkan model yang fit dapat disebabkan adanya perbedaan kultur dan budaya antara target pada alat ukur asli dengan target responden peneliti. Scrivner (2021b) mengembangkan *Morbid Curiosity Scale* berdasarkan konteks budaya Barat, di mana konsumsi media yang bertemakan horor dan kriminal disebabkan oleh ketertarikan individu terhadap fenomena berbahaya atau menakutkan yang dilakukan sebagai ekspresi dari rasa ingin tahu yang berfungsi sebagai mekanisme adaptif pada individu untuk memperoleh informasi tentang ancaman dalam situasi berbahaya. Akan tetapi, dalam konteks budaya di Indonesia, bentuk rasa ingin tahu ini memiliki karakteristik yang berbeda, misalnya genre horor yang sangat populer di Indonesia sering kali berkaitan erat dengan kepercayaan lokal dan mitos tradisional (Tiwahyupriadi & Ayuningtyas, 2020). Keingintahuan tersebut tidak terlalu berfokus pada kejadian atau kronologis itu sendiri secara apa adanya, tetapi bercampur dengan unsur nilai-nilai budaya dan norma sosial seperti pantangan, tabu, kepercayaan supranatural, bahkan ketakutan kolektif terhadap hal-hal mistis dan kekerasan. Hal ini menunjukkan bahwa ketertarikan terhadap konten menakutkan di Indonesia lebih terkait

dengan elemen budaya dan spiritual dibandingkan dengan sekadar rasa ingin tahu terhadap fenomena berbahaya.

Oleh karena itu, ketidaksesuaian model juga dapat dijelaskan oleh tidak relevannya beberapa item yang berakar pada budaya Barat, misalnya item ke-6, “*Jika saya hidup di era Romawi Kuno, saya tertarik menghadiri pertarungan gladiator,*” dan item ke-14 “*Saya cukup tertarik melihat pertarungan ala koboi jika saya hidup di era koboi Wild West*”. Item-item ini masih belum relevan atau sesuai dengan responden penelitian. Maka dari itu, partisipan pada penelitian ini hanya mampu melakukan pengandaian ketimbang mengalami secara langsung atau sesuai dengan situasi tersebut. Hal ini dapat mengakibatkan *item bias* pada responden. Selaras dengan pernyataan van de Vijver dan Tanzer (2004) bahwa *item bias* sering terjadi karena spesifikasi budaya dalam konten dan/atau konotasi item. Hal ini juga selaras dengan hasil observasi peneliti selama melakukan pengambilan data baik secara *offline* maupun *online* bahwa terdapat cukup banyak responden yang kebingungan dan bertanya langsung kepada peneliti terkait maksud kedua item tersebut (item 6 & item 14). Berdasarkan hasil observasi tersebut peneliti menyimpulkan bahwa perbedaan budaya pada item tersebut berpengaruh terhadap pemahaman partisipan selama mengisi kuesioner.

Oleh karena itu, alat ukur yang diadaptasi oleh peneliti belum layak untuk digunakan meskipun reliabilitasnya baik. Apabila model alat ukur yang diajukan tidak menemukan kecocokan model yang ideal maka perlu dilakukan modifikasi. Menurut Brown (2006) modifikasi model dilakukan ketika hasil CFA awal menunjukkan bahwa satu atau lebih dari tiga kriteria utama yang digunakan untuk mengevaluasi penerimaan model tidak terpenuhi yaitu: model tidak cocok secara keseluruhan, tidak mereproduksi beberapa hubungan indikator dengan baik, atau tidak menghasilkan estimasi parameter yang dapat ditafsirkan secara seragam. Penelitian selanjutnya disarankan untuk melakukan modifikasi model sehingga didapatkan hasil kesesuaian indeks fit yang lebih memadai.

## 5. Kesimpulan

Penelitian ini telah mengadaptasi *Morbid Curiosity Scale* ke versi bahasa Indonesia dengan item final berjumlah 24 item yang terbagi ke dalam 4 dimensi yaitu *Mind, Paranormal, Body, Violence*. Hasil analisis menunjukkan bahwa model tidak fit karena jika ditinjau dari nilai *Chi-square test*, CFI, GFI, dan RMSEA hanya nilai GFI dan RMSEA saja yang mempunyai nilai kecocokan yang baik. Namun, uji reliabilitas *Morbid Curiosity Scale* versi bahasa Indonesia menghasilkan koefisien *Cronbach's alpha* sebesar 0.894 sehingga termasuk ke dalam kategori sangat baik serta seluruh item memiliki *factor loadings* yang baik yaitu di atas 0.4. Berdasarkan hasil tersebut, alat ukur yang diadaptasi oleh peneliti belum layak untuk digunakan meskipun memiliki konsistensi internal yang baik. Hasil ini menunjukkan bahwa model adaptasi alat ukur belum sepenuhnya mampu merepresentasikan *morbid curiosity* dalam konteks budaya Indonesia. Penelitian selanjutnya diharapkan perlu memodifikasi model agar dapat mencapai nilai indeks kecocokan model fit yang diharapkan, sehingga alat ukur dapat digunakan dan divariasikan dengan variabel lainnya atau bahkan dapat mengembangkan alat ukur baru untuk mengeksplorasi fenomena-fenomena yang berkaitan dengan *morbid curiosity* pada populasi Indonesia. Keberadaan alat ukur *morbid curiosity* yang valid, reliabel, dan relevan secara budaya dalam jangka panjang diharapkan dapat digunakan bagi praktisi psikologi dan bidang lainnya untuk memahami bagaimana keingintahuan terhadap fenomena menakutkan, mistis, dan berbahaya baik dalam media digital maupun kehidupan sehari-hari berperan dalam perilaku individu. Alat ukur ini juga berpotensi digunakan dalam bidang klinis untuk mengidentifikasi bagaimana *morbid curiosity* berlebihan berpotensi menyebabkan perilaku berisiko serta mencari tahu bagaimana solusi penanganannya.

## 6. Referensi

- Andianto, A., & Delliana, S. (2022). Analisis resepsi viewers terhadap tayangan mistis Sableng TV. *KALBISIANA : Jurnal Mahasiswa Institut Teknologi Dan Bisnis Kalbis*, 8(4), 5348–5358.
- Azwar, S. (2016). *Validitas dan reliabilitas*. Pustaka Pelajar.
- Azwar, S. (2021). *Penyusunan skala psikologi* (3rd ed.). Pustaka Pelajar.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press. <https://doi.org/10.5860/choice.44-2769>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). SAGE Publications.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.
- Fulton, C. (2021). Urban exploration: Traces of the secretly documented, decayed, and disused. *LIBRARY TRENDS*, 69(3), 556–572. <https://doi.org/10.1353/lib.2021.0003>
- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. SAGE Publications Ltd. <https://doi.org/10.4135/9781446287866>
- Ghaemi, A., Vaghef, L., & Shalchi, B. (2024a). Structural relationship between dark personality traits, morbid curiosity, and cognitive abilities with addiction vulnerability. *International Journal of High Risk Behaviors*

- and *Addiction*, 13(3), e145652. <https://doi.org/10.5812/ijhrba-145652>
- Ghaemi, A., Vaghef, L., & Shalchi, B. (2024b). Validity and reliability of Persian version of morbid curiosity scale in students. *Medical Journal of Tabriz University of Medical Sciences*, 45(6), 485–494. <https://doi.org/10.34172/mj.2024.001>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.
- Harrington, D. (2009). *Confirmatory factor analysis*. Oxford University Press.
- Hox, J. J. (2021). Confirmatory factor analysis. *The Encyclopedia of Research Methods in Criminology and Criminal Justice*, 2, 830–832. <https://doi.org/10.1002/9781119111931.ch158>
- International Test Commission. (2017). *The ITC guidelines for translating and adapting tests*. [www.InTestCom.org](http://www.InTestCom.org). <https://doi.org/10.1016/j.prps.2020.06.005>
- JASP Team. (2024). *JASP* (0.19.1). <https://jasp-stats.org/>
- Justine, F., Jodie, K., Alfajri, M. R., Dilo, M. S. A. R. U., & Al Kautsar, Z. H. (2021). Budaya mengaitkan berbagai peristiwa dengan hal mistis oleh masyarakat Indonesia. *Jurnal Kewarganegaraan*, 5(2), 602–611.
- Khairi, M. I., Susanti, D., & Sukono. (2021). Study on structural equation modeling for analyzing data. *International Journal of Ethno-Sciences and Education Research*, 1(3), 52–60. <https://doi.org/10.46336/ijeer.v1i3.295>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. The Guilford Press. <https://doi.org/10.15353/cgjsc-rcessc.v1i1.25>
- Menold, N. (2020). Double barreled questions: An analysis of the similarity of elements and effects on measurement quality. *Journal of Official Statistics*, 36(4), 855–886. <https://doi.org/10.2478/jos-2020-0041>
- Nurdianti, K., & Basarah, F. F. (2024). The influence of horror content “Diary Mistry Sara” on the Sara Wijayanto Youtube channel towards the mystical beliefs. *PROPAGANDA: Journal of Communication Studies*, 4(1), 1–17.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Roos, J. M., & Bauldry, S. (2022). *Confirmatory factor analysis*. SAGE Publications Inc. <https://doi.org/10.4135/9781071938959.n4>
- Saputra, C. A. (2023). *YouTuber ini nekat masuk salah satu bangunan berhantu di dunia, lihat simbol setan di rumah sakit terbengkalai*. Liputan 6. <https://www.liputan6.com/global/read/5324094/youtuber-ini-nekat-masuk-salah-satu-bangunan-berhantu-di-dunia-lihat-simbol-setan-di-rumah-sakit-terbengkalai?page=2>
- Scrivner, C. (2021a). An infectious curiosity: morbid curiosity and media preferences during a pandemic. *Evolutionary Studies in Imaginative Culture*, 5(1), 1–12. <https://doi.org/10.26613/ESIC.5.1.206>
- Scrivner, C. (2021b). The psychology of morbid curiosity: Development and initial validation of the morbid curiosity scale. *Personality and Individual Differences*, 183, 1–10. <https://doi.org/10.1016/j.paid.2021.111139>
- Scrivner, C., & Stubbersfield, J. M. (2023). Curious about threats: Morbid curiosity and interest in conspiracy theories in US adults. *British Journal of Psychology*, 115(1), 129–147. <https://doi.org/10.1111/bjop.12682>
- Taqiyya, A. (2023). *10 film Indonesia terlaris sepanjang masa, ada yang capai 10 juta penonton!* GoodStats Data. <https://data.goodstats.id/statistic/10-film-indonesia-terlaris-sepanjang-masa-ada-yang-capai-10-juta-penonton-5OmVy>
- Tiwahyupriadi, D., & Ayuningtyas, Y. (2020). Indonesian horror film: Deconstruction of repetitive elements of Indonesian urban legend for cultural revitalization, creativity, and critical thinking. *KnE Social Sciences*, 4(12), 115–125. <https://doi.org/10.18502/kss.v4i12.7589>
- van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Europeenne de Psychologie Appliquee*, 54(2), 119–135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Wang, X., Wang, Q., Cai, Y., & Tu, D. (2023). Measurement invariance and latent mean differences of the morbid curiosity scale (MCS) across the United States and China. *Heliyon*, 9(9), e19973. <https://doi.org/10.1016/j.heliyon.2023.e19973>
- Yasim, M. N. R., & Yahya, M. (2024). Film horor agama di Indonesia, kesalehan, dan kesakralan yang terpinggirkan? *Endogami: Jurnal Ilmiah Kajian Antropologi*, 8(1), 111–126.
- Zijlmans, E. A. O., Tijmstra, J., van der Ark, L. A., & Sijtsma, K. (2018). Item-score reliability in empirical-data sets and its relationship with other item indices. *Educational and Psychological Measurement*, 78(6), 998–1020. <https://doi.org/10.1177/0013164417728358>
- Zuckerman, M., & Litle, P. (1986). Personality and curiosity about morbid and sexual events. *Personality and Individual Differences*, 7(1), 49–56. [https://doi.org/10.1016/0191-8869\(86\)90107-8](https://doi.org/10.1016/0191-8869(86)90107-8)



## MODIFIKASI SKALA REGULASI AKADEMIK MAHASISWA

Amr Yazid Pikoli<sup>1</sup>, Miftahul Huda<sup>2</sup>, Benidha Roos Feri Nisa<sup>3</sup>

<sup>1</sup>Fakultas Psikologi, Universitas Gajah Mada, Yogyakarta

<sup>2,3</sup>Fakultas Dakwah, Universitas Islam Cordoba, Banyuwangi

Email: [amryazidpikoli@gmail.com](mailto:amryazidpikoli@gmail.com)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
7 July 2025	4 September 2025	1 Oktober 2025	30 Oktober 2025

### Abstract

The numerous academic demands with high intensity can cause students to experience academic stress. Therefore, in meeting these demands, students must have strategies that must be developed to facilitate the course of life in college so that they can avoid academic stress or other mental illnesses that can hinder studies. These strategies are combined into academic self-regulation that can guide students to navigate college life well. This study used a quantitative cross-sectional approach with academic self-regulation as the variable to be studied. The results showed that all items were normally distributed, but the initial CFA analysis of 66 items did not meet the model fit criteria (RMSEA = 0.094; SRMR = 0.159) so that reduction was carried out until the remaining 17 items with good fit indices (CFI = 0.950; TLI = 0.941; RMSEA = 0.058; SRMR = 0.044). Two indicators were dropped due to low factor loading and overlap between items. The MIMIC test found demographic bias in gender on the forethought aspect ( $p < 0.001$ ) and age on the self-reflection aspect ( $p = 0.002$ ), while other variables did not show significant bias. The Rasch Model test supported construct validity by fulfilling the assumptions of unidimensionality, item homogeneity, local independence, and monotonicity, and the DIF test showed no bias based on semester. Thus, the developed academic self-regulation scale can be declared valid, reliable, and suitable for use, although interpretation of the learning planning and self-reflection aspects needs to consider certain demographic Differences.

Keywords: strategy, self-regulation, academics

### Abstrak

Banyaknya tuntutan akademik dengan intensitas yang tinggi dapat membuat mahasiswa mengalami stres akademik. Oleh karena itu dalam memenuhi hal-hal tersebut mahasiswa harus memiliki strategi-strategi yang wajib disusun untuk melancarkan jalannya kehidupan di perkuliahan sehingga dapat terhindar dari stress akademik ataupun penyakit mental lain yang dapat menghambat studi. Strategi-strategi ini dirangkai menjadi regulasi akademik yang dapat menjadi pemandu bagi mahasiswa untuk mengatur kehidupan perkuliahan dengan baik. Penelitian ini menggunakan pendekatan kuantitatif *cross sectional* dengan variabel regulasi diri akademik sebagai variabel yang akan diteliti. Hasil penelitian menunjukkan bahwa seluruh aitem berdistribusi normal, namun analisis CFA awal terhadap 66 aitem belum memenuhi kriteria *model fit* (RMSEA = 0,094; SRMR = 0,159) sehingga dilakukan reduksi hingga tersisa 17 aitem dengan indeks kesesuaian yang baik (CFI = 0,950; TLI = 0,941; RMSEA = 0,058; SRMR = 0,044). Dua indikator gugur akibat *loading factor* rendah dan tumpang tindih antar-aitem. Uji MIMIC menemukan adanya bias demografis pada jenis kelamin terhadap aspek *forethought* ( $p < 0,001$ ) dan usia terhadap aspek *self-reflection* ( $p = 0,002$ ), sementara variabel lain tidak menunjukkan bias signifikan. Uji Rasch Model mendukung validitas konstruk dengan terpenuhinya asumsi unidimensionality, item homogeneity, local independence, dan monotonicity, serta uji DIF menunjukkan tidak adanya bias berdasarkan semester. Dengan demikian, skala regulasi diri akademik yang dikembangkan dapat dinyatakan valid, reliabel, dan layak digunakan, meskipun interpretasi pada aspek perencanaan belajar dan refleksi diri perlu mempertimbangkan perbedaan demografis tertentu.

Kata kunci: strategi, regulasi diri, akademik

### 1. Pendahuluan

Mahasiswa merupakan pelajar tingkat tinggi yang dituntut untuk memiliki kemampuan yang lebih tinggi dari rata-rata siswa di bawahnya. Dorongan tersebut diperoleh dari anggapan masyarakat bahwa mahasiswa diharapkan dapat menjadi *agent of change* (agen perubahan) dalam tatanan masyarakat (Jannah & Sulianti, 2021). Hal ini tentunya menjadi

tantangan tersendiri bagi mahasiswa. Mahasiswa membutuhkan regulasi diri yang harus didukung oleh lingkungan belajar yang masif dalam rangka meningkatkan performa akademik. Penelitian yang dilakukan oleh (Radović et al., 2024) menyatakan bahwa lingkungan belajar memengaruhi regulasi diri dengan menggunakan beberapa media seperti *website* pembelajaran, dukungan dan refleksi dari orang lain, serta penetapan tujuan. Media-media tersebut dirancang untuk meningkatkan kemampuan siswa dalam mengelola dan mengevaluasi proses pembelajaran siswa sehingga mendukung regulasi diri yang mandiri.

Tuntutan akademik dengan intensitas yang tinggi ditambah dengan harus memenuhi ekspektasi dari banyak orang dapat membuat mahasiswa mengalami stres akademik. Sumber stres yang umum terjadi pada mahasiswa yaitu tekanan terkait penilaian, tugas yang berlebihan, dinamika konflik yang terjadi dalam kehidupan perkuliahan, hubungan pribadi, hubungan teman sebaya, masalah kesehatan, lingkungan belajar, dan kegagalan administratif. Penyebab stres ini dapat berdampak negatif terhadap performa akademik, sehingga menyebabkan persepsi stres yang lebih besar, penurunan kesejahteraan secara keseluruhan, peningkatan depresi, kecemasan, dan kelelahan (Eames et al., 2024). Oleh karena itu dalam memenuhi hal-hal tersebut mahasiswa harus memiliki strategi-strategi yang wajib disusun untuk melancarkan jalannya kehidupan di perkuliahan sehingga dapat terhindar dari stress akademik ataupun penyakit mental lain yang dapat menghambat studi. Strategi-strategi ini dirangkai menjadi regulasi diri akademik yang dapat menjadi pemandu bagi mahasiswa untuk mengarahi kehidupan perkuliahan dengan baik (Huda, et.al. 2024)

Regulasi diri akademik merupakan suatu proses yang bersumber dari konfigurasi pikiran, perasaan, dan tindakan yang dihasilkan untuk mencapai tujuan akademik. Hal tersebut mencakup kemampuan seorang mahasiswa dalam menjalankan proses seperti motivasi diri, kemauan, dan pengaturan perilaku untuk memenuhi kebutuhan otonomi dalam konteks akademik (Zimmerman, 1989). Adapun aspek-aspek regulasi akademik mahasiswa menurut (Zimmerman, 2008) yaitu *forethought*, *performance control*, dan *self reflection*. Pada tahap *forethought*, individu yang belajar menetapkan rencana pembelajaran mereka, mengidentifikasi tugas-tugas yang harus diselesaikan, dan membantu membentuk pola pikir yang positif terkait pembelajaran. Selain itu, tahap ini mencakup aktivitas seperti mengambil inisiatif, merencanakan, dan menetapkan tujuan belajar. Selanjutnya pada tahap *performance control* individu (pelajar) menggunakan berbagai strategi yang mereka miliki untuk meningkatkan pembelajaran mereka. Ini melibatkan upaya untuk mengurangi gangguan atau distraksi selama proses belajar, mengelola waktu dengan efektif, serta menciptakan kondisi yang mendukung untuk memfasilitasi pembelajaran. Terakhir adalah *self reflection* dimana pelajar melakukan evaluasi terhadap proses pembelajaran mereka dengan membandingkannya dengan tujuan yang telah ditetapkan. Evaluasi ini mencakup penilaian terhadap pencapaian tujuan, keberhasilan atau kegagalan dalam merencanakan, manajemen waktu yang kurang efektif, dan kondisi di mana mereka mencapai hasil terbaik.

Penelitian terdahulu yang dilakukan oleh (Mahmood, 2013) dengan menggunakan skala ASRS (*Academic self regulation scale*) menyatakan bahwa ASRS, yang terdiri dari 30 item dalam lima sub-faktor yaitu perencanaan mandiri, pemantauan mandiri, instruksi mandiri, evaluasi diri, dan reaksi diri menunjukkan pengaruh nilai yang kuat (dengan nilai alfa 0,83) dan validitas melalui analisis faktor. Hal ini menunjukkan bahwa skala tersebut merupakan alat yang ampuh untuk mengukur sejauh mana mahasiswa terlibat dalam strategi pembelajaran mandiri. Penelitian ini menggarisbawahi pentingnya regulasi diri dalam pembelajaran dan menyediakan sarana bagi para pendidik untuk mengidentifikasi dan mengembangkan keterampilan-keterampilan pada mahasiswa, selaras dengan perkembangan dunia pendidikan yang menuntut pembelajar yang lebih mandiri. Mahasiswa mengelola regulasi diri akademik mereka melalui kombinasi dari motivasi diri dan pengontrolan kemauan. Motivasi diri melibatkan penetapan tujuan dan belajar dari kesalahan, sedangkan pengontrolan kemauan mencakup pengambilan keputusan strategis dan mengubah motivasi menjadi niat (Valenzuela et al., 2020). Namun pada penerapannya skala regulasi diri akademik yang ada saat ini hanya merupakan skala yang menggunakan subjek penelitian siswa luar negeri sehingga peneliti ingin memodifikasi skala regulasi diri akademik yang disesuaikan dengan budaya akademik yang ada di Indonesia.

Namun demikian, skala regulasi diri akademik yang telah banyak digunakan masih dominan dikembangkan dan diuji pada konteks mahasiswa luar negeri. Kondisi ini menimbulkan kesenjangan penelitian (*research gap*), yakni keterbatasan instrumen yang sesuai dengan konteks budaya akademik Indonesia. Oleh karena itu, penelitian ini penting dilakukan untuk memodifikasi dan mengadaptasi skala regulasi diri akademik agar lebih relevan, valid, dan reliabel digunakan dalam lingkungan pendidikan tinggi di Indonesia.

## 2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan variabel regulasi diri akademik sebagai variabel yang akan diteliti. Metode yang digunakan dalam penelitian ini yakni *cross sectional*. Metode *cross sectional* merupakan suatu jenis penelitian yang menekankan waktu pengukuran hanya satu kali pada satu saat (Azwar, 2010). Tujuan dari penelitian ini adalah memodifikasi alat ukur regulasi diri akademik yang didasarkan pada teori *academic self-regulation* dari Zimmerman (2008) yang disesuaikan dengan budaya yang ada di Indonesia. Pengujian aspek psikometri dilakukan dalam penelitian ini guna memperoleh data penelitian berupa angka yang akan diukur secara statistik, untuk menghasilkan suatu kesimpulan.

### 3. Hasil dan Diskusi

#### Subjek

Mahasiswa Warga Negara Indonesia aktif di Perguruan Tinggi di seluruh Indonesia, dan berusia 18 sampai 25 tahun. Teknik pengambilan sampel yang digunakan pada penelitian ini adalah purposive sampling dan insidental sampling. Purposive sampling merupakan teknik pengambilan sampel dengan kriteria tertentu, sementara insidental sampling yaitu teknik pengambilan sampel dengan unsur kebetulan dimana individu yang secara kebetulan bertemu dengan peneliti dan cocok sebagai sumber data maka dapat digunakan sebagai sampel (Sugiyono, 2013). Kedua teknik tersebut termasuk dalam *non-probability sampling*. Teknik *non-probability sampling* adalah cara pengambilan sampel dimana tidak semua anggota dalam populasi memiliki peluang yang sama untuk dipilih menjadi sampel (Sugiyono, 2013).

#### Instrumen Penelitian

Instrumen dalam penelitian ini menggunakan alat ukur modifikasi yang diambil dari teori Zimmerman (2008). Skala Regulasi Diri Akademik ini terdiri dari 3 aspek, yaitu perencanaan (*forethought*), performa akademik (*performance control*), dan evaluasi diri (*self-reflection*). Aitem dalam skala ini berjumlah 66 aitem dengan 11 indikator serta dibagi menjadi dua kategori, yaitu aitem favorable dan aitem unfavorable. Format respon yang disajikan dalam skala ini menggunakan Skala Likert dengan 4 alternatif jawaban (1 = Sangat Tidak Sesuai/STS, 2 = Tidak Sesuai/TS, 3 = Sesuai/S, 4 = Sangat Sesuai = SS).

**Tabel 1. Blueprint Skala Regulasi diri akademik**

Aspek	Indikator	Bobot	Total Aitem
<i>Forethought</i>	Memiliki kemampuan dalam merancang target belajar	36%	24
	Memiliki rencana dan strategi dalam belajar		
	Memiliki kepercayaan diri dan kegigihan dalam proses belajar		
	Memiliki ketertarikan dan dapat mengantisipasi hasil dalam proses belajar		
<i>Performance Control</i>	Memiliki kemampuan menggunakan strategi untuk mengembangkan proses belajar	36%	24
	Memiliki kemampuan mengidentifikasi prioritas dan alokasi waktu sesuai kebutuhan		
	Mempunyai strategi untuk memonitoring proses pembelajaran dan menghindari distraksi		
	Mampu beradaptasi dengan perubahan dan lingkungan akademik		
<i>Self-Reflection</i>	Mampu mengevaluasi hasil belajar	27%	18
	Merasa puas dengan hasil belajar		
	Mampu mengevaluasi manajemen waktu		

#### Prosedur Penelitian

Penelitian dilakukan dengan mengumpulkan data terlebih dahulu setelah melakukan proses penyusunan skala dan proses *ethical clearance* karena penelitian ini melibatkan manusia sebagai subjek dalam penelitian. Pengumpulan data diperoleh dengan menggunakan kuesioner yang berisi instrumen penelitian dan dibuat dalam bentuk *Google Form*. Kuesioner dilakukan dengan memberikan seperangkat pertanyaan atau pernyataan secara tertulis kepada responden untuk dijawab. Dalam kuesioner terdiri dari 4 bagian, yaitu penjelasan penelitian, *informed consent*, identitas dan data demografis, serta skala penelitian. Kuesioner tersebut disebarluaskan secara daring melalui media sosial seperti WhatsApp, Instagram, dan Twitter agar dapat menjangkau partisipan secara luas. Selanjutnya penelitian ini akan melakukan pengujian validitas faktorial (analisis faktor) menggunakan metode *Confirmatory Factor Analysis* (CFA). Setelah itu akan dilakukan pengujian Validitas dan Reliabilitas. Pengujian validitas menggunakan pendekatan Validitas Faktorial (Analisis Faktor). Validitas faktorial yang

digunakan pada penelitian ini adalah Confirmatory Factor Analysis (CFA). Selanjutnya dilakukan pengujian asumsi dengan MIMIC Model, *Rasch Model*, dan analisis DIF untuk mengetahui ada tidaknya efek bias dalam pengukuran pada salah satu kelompok pada data demografi. Reliabilitas alat ukur akan dianalisis dengan metode Cronbach's alpha. Data akan diolah menggunakan bantuan software Jamovi 2.3.28. Analisis aitem ini dilakukan dengan melihat hasil skor *item-rest correlation* tiap itemnya. Aitem dengan skor kurang dari 0,30 akan digugurkan karena diinterpretasikan memiliki daya diskriminasi yang rendah (Azwar, 2019).

### **Descriptif Statistic Item**

Berdasarkan hasil pengambilan data, berikut tampilan data demografis yang akan dijelaskan pada tabel 2:

**Tabel 2. Data Demografis Subjek Penelitian (N = 584)**

<b>Demografis</b>		<b>N</b>	<b>Persentase (100%)</b>
<b>Jenis Kelamin</b>	Perempuan	433	74%
	Laki-laki	151	26%
<b>Semester</b>	Awal	279	48%
	Akhir	305	52%
<b>Usia</b>	18-21 Tahun	362	62%
	22-25 Tahun	222	38%
<b>Rumpun Ilmu</b>	Saintek	220	38%
	Soshum	364	62%
<b>Status Mahasiswa</b>	Rantau	340	58%
	Non Rantau	244	42%
<b>Durasi Belajar</b>	0-2 Jam/Hari	361	62%
	>2 jam/Hari	223	38%
<b>Kegiatan diluar Kuliah</b>	Ada Kegiatan	411	70%
	Tanpa Kegiatan	173	30%

Hasil pengumpulan data menunjukkan bahwa berdasarkan persentase jenis kelamin, Perempuan lebih unggul dengan persentase 74%, berdasarkan tingkat semester, mahasiswa semester akhir sedikit lebih unggul dengan persentase 52%, berdasarkan tingkat usia, usia 18-21 tahun unggul dengan persentase 62%, berdasarkan rumpun ilmu, soshum unggul dengan persentase 62%, berdasarkan status mahasiswa, mahasiswa rantau unggul dengan persentase 58%, berdasarkan durasi belajar, mahasiswa yang belajar 0-2 jam/hari unggul dengan persentase 62%, dan berdasarkan kegiatan diluar kuliah, mahasiswa yang mengikuti kegiatan diluar perkuliahan unggul dengan persentase 70%.

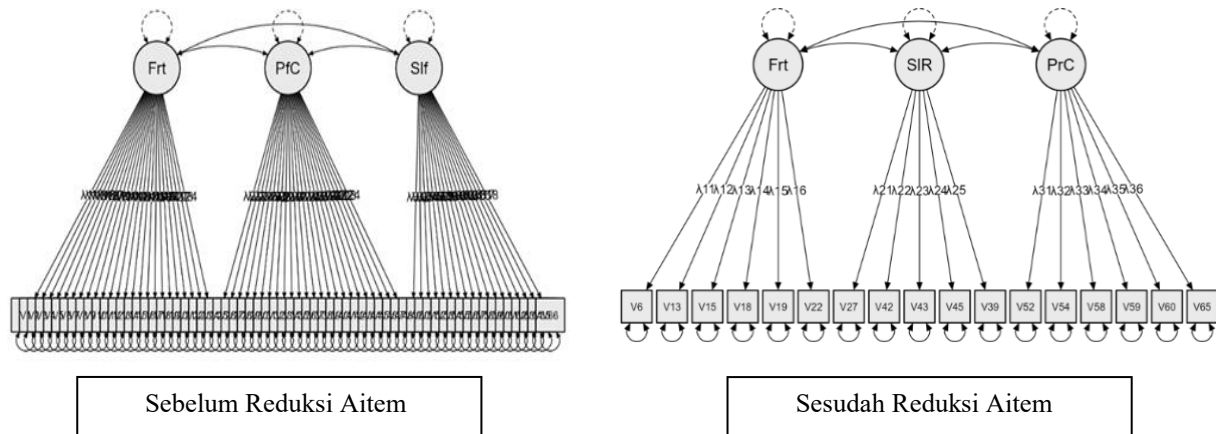
### **Confirmatory Factor Analysis**

*Confirmatory factor analysis* (CFA) adalah metode yang digunakan untuk menguji hipotesis tentang hubungan antara variabel laten (faktor yang tidak bisa diukur langsung) dan variabel yang diamati (yang bisa diukur langsung). CFA membantu peneliti untuk melihat apakah data yang dikumpulkan sesuai dengan model yang diprediksi sebelumnya (Mulaik, 1988). Interpretasi CFA melibatkan pemeriksaan beberapa indeks kecocokan model, seperti *Comparative Fit Index* (CFI) > 0.9, *Tucker-Lewis Index* (TLI) > 0.9, *Root Mean Square Error of Approximation* (RMSEA) < 0.6, dan *Standardized Root Mean Square Residual* (SRMR) < 0.6 (Hair dkk, 2019). Indeks kecocokan model diperlukan untuk mengetahui apakah model yang menjadi hipotesis sesuai dengan data yang didapatkan atau biasanya disebut sebagai model fit.

**Tabel 4. Hasil Uji Confirmatory Factor Analysis**

<b>Model</b>	<b>X<sup>2</sup></b>	<b>df</b>	<b>CFI</b>	<b>TLI</b>	<b>RMSEA</b>	<b>SRMR</b>
CFA versi full item	12893.916	2076	0.451	0.433	0.094	0.159
CFA versi reduksi item	341.936	116	0.950	0.941	0.058	0.044

*Note.* CFI > 0.9; TLI > 0.9; RMSEA < 0.6; SRMR < 0.6.



Gambar 1. Model *Confirmatory Factor Analysis*

Hasil uji CFA disajikan dengan dua versi yakni versi dengan seluruh aitem dan dengan versi setelah aitem direduksi. Fungsi reduksi berguna untuk menurunkan nilai RMSEA dan SRMR serta menaikkan nilai CFI dan TLI agar sesuai dengan asumsi ( $CFI > 0.9$ ;  $TLI > 0.9$ ;  $RMSEA < 0.6$ ;  $SRMR < 0.6$ ) untuk memperoleh model yang fit. Berdasarkan hasil yang diperoleh sebelum reduksi aitem (Tabel 4), baik nilai RMSEA, SRMR, CFI dan TLI masih belum memenuhi syarat model fit uji asumsi, oleh karena itu maka dilakukan reduksi aitem dan akhirnya memperoleh RMSEA 0,058, SRMR 0,044, CFI 0,950, dan TLI 0,941. Hal ini menyatakan bahwa asumsi model CFA telah *fit* dan terpenuhi (Hair dkk, 2019). Reduksi aitem yang dilakukan menyebabkan total aitem setelah CFA berubah menjadi 17 aitem dari total 66 aitem.

**Multiple Indicators Multiple Causes (MIMIC)**

MIMIC (*Multiple Indicators, Multiple Causes*) adalah model yang digunakan untuk mengoperasionalkan variabel format dalam aplikasi empiris. Model ini mengukur variabel laten fokus ( $\eta$ ), dengan beberapa indikator reflektif ( $\gamma_s$ ). Variabel laten terkadang tidak dapat diamati langsung melalui beberapa indikator sehingga MIMIC berfungsi untuk mengukur variabel laten melalui indikator (Lee dkk, 2013). Acuan interpretasi MIMIC model mencakup identifikasi hubungan antara variabel laten dan dimensi merujuk pada nilai *Comparative Fit Index* ( $CFI > 0.9$ ), *Tucker-Lewis Index* ( $TLI > 0.9$ ), *Root Mean Square Error of Approximation* ( $RMSEA < 0.6$ ), dan *Standardized Root Mean Square Residual* ( $SRMR < 0.6$ ), serta hubungan antara dimensi dengan kovariat (demografi) merujuk pada nilai  $< 0.05$  (Schumaker & Lomax, 2016).

Pengujian MIMIC dilakukan dengan dua langkah, yang mana langkah pertama digunakan untuk mengetahui model fit dari seluruh data demografis untuk mengetahui pemenuhan asumsi dari tiap-tiap data demografis. Sedangkan, langkah kedua digunakan untuk mengetahui pengaruh atau ada tidaknya bias dari tiap-tiap data demografis terhadap tiap-tiap aspek. Kedua model ini diperlukan untuk mengetahui apakah skala yang dibuat dapat digunakan dan valid ke tiap-tiap data demografis atau bias kepada tiap-tiap data demografis.

Tabel 5. Hasil uji FIT MIMIC MODEL

Model	X <sup>2</sup>	df	CFI	TLI	RMSEA	SRMR
Mimic Jenis Kelamin	353,812	130	0,951	0,942	0,054	0,044
Mimic Semester	353,145	130	0,950	0,942	0,054	0,043
Mimic Usia	356,046	130	0,950	0,941	0,055	0,043
Mimic Rumpun Ilmu	364,130	130	0,948	0,939	0,056	0,048
Mimic Status Mahasiswa	363,973	130	0,948	0,939	0,056	0,044
Mimic Durasi Belajar Per Hari	380,030	130	0,945	0,935	0,057	0,044
Mimic Kegiatan diluar Kuliah	397,868	130	0,939	0,928	0,059	0,046

Catatan.  $CFI > 0.9$ ;  $TLI > 0.9$ ;  $RMSEA < 0.6$ ;  $SRMR < 0.6$ .

Berdasarkan hasil uji Fit MIMIC model , seluruh data demografis memenuhi asumsi  $CFI > 0.9$ ;  $TLI > 0.9$ ;  $RMSEA < 0.6$ ;  $SRMR < 0.6$  (Schumacker dan Lomax, 2016). Hal ini menunjukkan bahwa seluruh data demografis

sudah *fit* yang berarti seluruh data demografis sudah memenuhi asumsi dengan model data jenis kelamin yang paling *fit* dari seluruh model data demografis yakni dengan RMSEA sebesar 0,054, SRMR sebesar 0,044, CFI sebesar 0,951, dan TLI sebesar 0,942.

**Tabel 6. Model MIMIC**

<b>Demografis</b>	<b>Forethought</b>	<b>Performance</b>	<b>Self Reflection</b>
Jenis Kelamin	< 0.001	0,005	0,216
Semester	0.839	0.128	0.062
Usia	0.163	0.076	0.002
Rumpun Ilmu	0,202	0.412	0.240
Status Mahasiswa	0.809	0.932	0.064
Durasi Belajar	0.092	0.361	0.175
Kegiatan Diluar	0.054	0.683	0.336

Berdasarkan hasil uji MIMIC (Tabel 6), data menunjukkan bahwa beberapa data demografis significant yang berarti memiliki bias terhadap aspek. Data demografis yang dimaksud adalah data jenis kelamin dan juga usia. Data jenis kelamin memiliki bias terhadap aspek *forethought* dengan signifikansi < 0.001. Hal yang sama juga terjadi pada data usia yang memiliki bias pada aspek *self reflection* dengan signifikansi 0.002. Namun, meskipun demikian rata-rata data demografis tidak memiliki bias terhadap tiap-tiap aspek yang menunjukkan model pengukuran dapat digunakan oleh setiap variabel pada data demografis.

### **Rasch Model**

Rasch model adalah sebuah model pengukuran dalam teori respons butir (item response theory) IRT yang digunakan untuk mengukur kemampuan atau karakteristik seseorang berdasarkan respon yang diberikan terhadap serangkaian pertanyaan atau butir. Model ini mengasumsikan bahwa kemampuan individu dan tingkat kesulitan butir dapat diukur dalam satu dimensi. Dengan kata lain, Rasch model memungkinkan untuk membandingkan kemampuan individu secara obyektif tanpa terpengaruh oleh variasi butir yang digunakan dalam pengukuran tersebut (Ludlow dan Haley, 1995).

Lebih lanjut, Ludlow dan Haley (1995) membagi asumsi rasch menjadi 4 asumsi yaitu:

- Unidimensionality*: Asumsi bahwa semua butir atau pertanyaan dalam tes mengukur hanya satu dimensi atau karakteristik yang sama. Artinya, tidak ada dimensi tersembunyi atau tambahan yang diukur oleh aitem tersebut.
- Item Homogeneity*: Asumsi bahwa semua butir atau pertanyaan (aitem) dalam tes setara dalam hal kesulitan dan karakteristik pengukuran. Butir-butir tersebut seharusnya memiliki tingkat kesulitan yang sebanding untuk mengukur kemampuan individu dengan baik.
- Local Independence*: Asumsi bahwa respon individu terhadap setiap butir tidak dipengaruhi oleh respon individu terhadap butir lain dalam tes. Dengan kata lain, respon terhadap setiap butir harus independen satu sama lain.
- Monotonicity*: Asumsi bahwa probabilitas individu untuk menjawab benar suatu butir harus selalu meningkat atau setidaknya tetap konstan dengan meningkatnya kemampuan individu. Artinya, jika seseorang memiliki kemampuan yang lebih tinggi, maka kemungkinan jawaban benar juga harus lebih tinggi atau setidaknya sama.

Asumsi *unidimensionality* dalam Model Rasch terpenuhi karena seluruh aitem terbukti mengukur satu konstruk yang sama, yaitu regulasi diri akademik. Hal ini tercermin dari hasil analisis CFA versi item reduksi yang menunjukkan nilai *fit indices* berada dalam kategori baik (CFI = 0,950 > 0,90; TLI = 0,941 > 0,90; RMSEA = 0,058 < 0,06; SRMR = 0,044 < 0,06). Parameter Rasch memperkuat temuan ini dengan menunjukkan konsistensi aitem dalam mengukur satu dimensi laten tanpa adanya indikasi dimensi tambahan yang tersembunyi. Dengan demikian, model yang digunakan hanya mengukur satu variabel laten yang mencakup tiga aspek utama regulasi diri akademik, yaitu *Forethought*, *Performance Control*, dan *Self Reflection*.

Selanjutnya adalah *item homogeneity* yang memvalidasi bahwa setiap aitem dalam pengukuran memiliki tingkat kesulitan yang sama dalam mengukur kemampuan subjek. Hal tersebut dilihat dari hasil nilai infit dan outfit. Jika hasil nilai menunjukkan angka pada rentang 0,5-1,5 maka asumsi homogeneity terpenuhi (Bond dan Fox, 2015).

**Tabel 7. Item Homogeneity***Item statistics of the rating scale model*

<i>Item</i>	<i>Measure</i>	<i>S.E.Measure</i>	<i>Infit</i>	<i>Outfit</i>
6	-2.24	0.0463	0.895	0.873
13	-1.95	0.0447	0.918	0.940
15	-1.63	0.0440	1.229	1.238
18	-1.49	0.0439	1.065	1.071
19	-1.71	0.0441	1.011	1.042
22	-1.48	0.0439	1.121	1.225
27	-2.03	0.0451	1.339	1.426
39	-1.64	0.0440	0.866	0.864
42	-1.78	0.0442	0.821	0.880
43	-1.95	0.0448	0.826	0.858
45	-2.01	0.0450	1.066	1.032
52	-2.21	0.0461	0.772	0.795
54	-2.49	0.0481	0.775	0.776
58	-2.33	0.0468	0.942	0.937
59	-2.53	0.0485	1.149	1.166
60	-1.98	0.0448	1.051	1.078
65	-2.37	0.0472	1.284	1.364

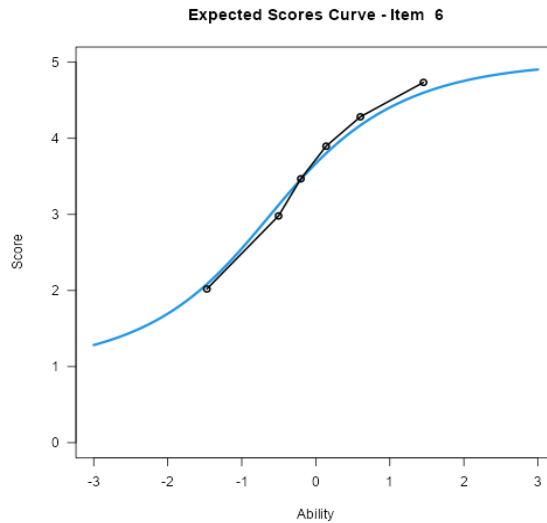
Catatan. *Infit*= Information-weighted mean square statistic; *Outfit*= Outlier-sensitive means square statistic.

Berdasarkan nilai dari tabel *item homogeneity* menunjukkan bahwa nilai setiap aitem berdasarkan nilai *infit* dan *outfit* sudah berada pada rentang 0,5-1,5. Hal ini menunjukkan bahwa asumsi *item homogeneity* sudah terpenuhi (Bond dan Fox, 2015).

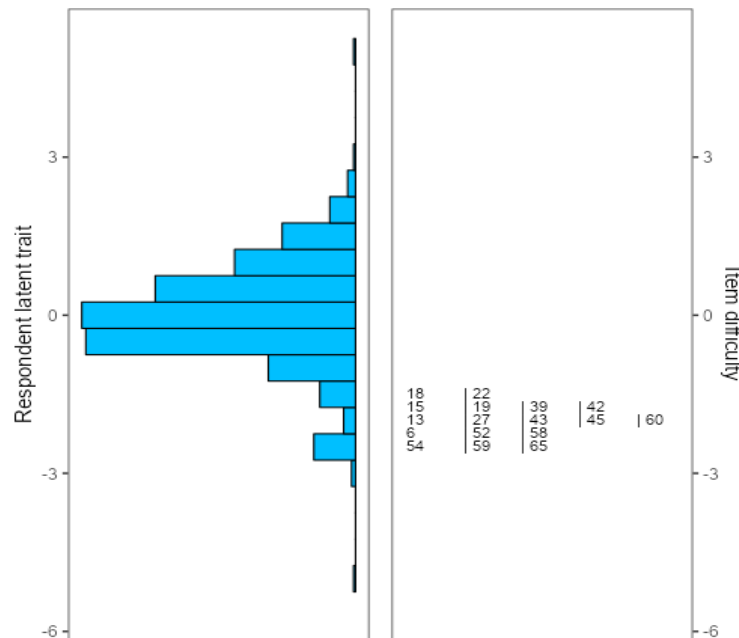
Selanjutnya adalah uji asumsi *local independence*. Uji asumsi tersebut bertujuan untuk mengetahui ada tidaknya pengaruh respon individu terhadap setiap butir dengan pengaruh respon individu dengan respon yang lain. Hal ini dapat terlihat dari nilai hasil uji pada tabel uji asumsi *local individu* (Lampiran 3). Jika aitem memiliki korelasi residual >0.30 maka berarti terdapat pengaruh pada respon individu terhadap setiap butir dengan pengaruh respon individu pada butir yang lain (Yen, 1984).

Berdasarkan hasil pada Q3 Correlation Matrix (Lampiran 3), terlihat bahwa seluruh korelasi setiap butir aitem berada di bawah 0,3. Hal ini menunjukkan bahwa asumsi *local independence* sudah terpenuhi yang berarti setiap aitem dalam pengukuran tidak terdapat pengaruh pada respon individu terhadap setiap butir dengan pengaruh respon individu pada butir yang lain. sehingga tidak ada aitem yang gugur dalam asumsi ini.

Uji asumsi *Rasch Model* yang terakhir yaitu uji asumsi *Monotonicity*. Uji asumsi ini menguji probabilitas individu untuk menjawab dengan mudah suatu butir harus selalu meningkat atau setidaknya tetap konstan dengan meningkatnya kemampuan individu. Artinya, jika seseorang memiliki kemampuan yang lebih tinggi, maka kemungkinan jawaban benar juga harus lebih tinggi atau setidaknya sama. Uji asumsi *Monotonicity* akan di contohkan oleh aitem 6 yang akan menunjukkan konstannya hasil peningkatan probabilitas individu menjawab benar seiring meningkatnya kemampuan individu.



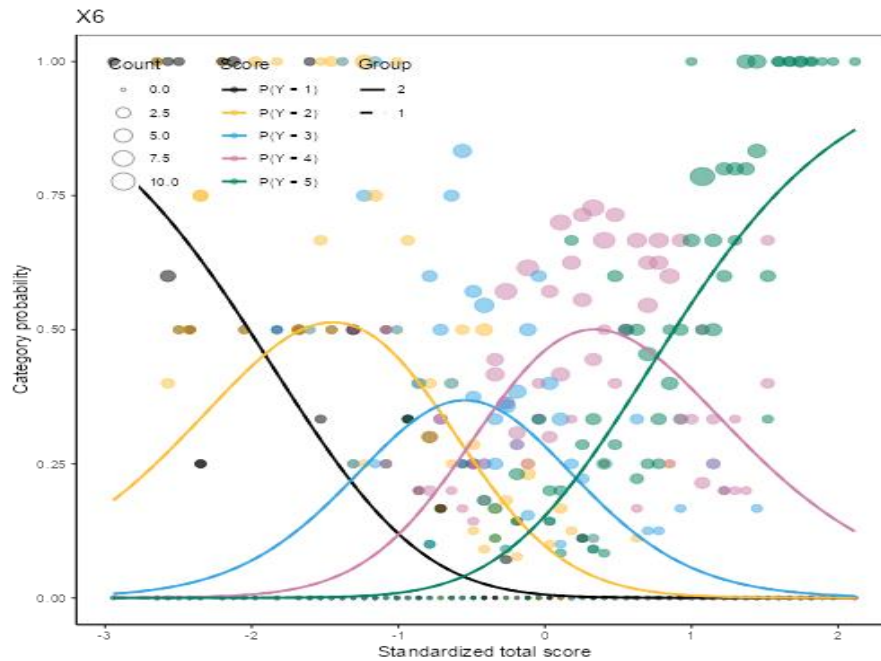
Gambar 2. *Item Characteristic Curve*



Gambar 3. *Wright Map*

### ***DIFferential Item Functioning (DIF)***

DIF (*DIFferential Item Functioning*) adalah teknik psikometrik yang digunakan untuk menilai apakah suatu aitem dalam tes menunjukkan perbedaan fungsi antara kelompok yang berbeda. DIF dinilai dengan membandingkan *Item Characteristic Curves* (ICCs) dari kelompok yang berbeda pada suatu item. Jika ICCs identik atau hampir identik untuk setiap kelompok, maka item tersebut tidak menunjukkan DIF. Namun, jika ICCs berbeda secara signifikan antara kelompok, maka item tersebut dikatakan menunjukkan DIF (Zumbo, 1999). Uji DIF dilakukan pada data demografi semester untuk mengetahui ada tidaknya perbedaan fungsi suatu aitem dalam mengukur regulasi diri akademik antar kelompok pada data demografi berdasarkan semester.



**Gambar 4. Hasil Uji *DIF*ferential Item Functioning**

Berdasarkan gambar tersebut maka data menunjukkan bahwa ICCs hampir identik untuk setiap kelompok, maka item tersebut tidak menunjukkan DIF yang berarti tidak ada bias item terhadap variabel demografi Semester. Hal tersebut berarti tidak ada perbedaan fungsi aitem dalam mengukur regulasi diri akademik antar kelompok pada data demografi berdasarkan semester.

Hal tersebut juga didukung oleh nilai yang terdapat pada ratio Chi-Square (Tabel 8) yang memberikan gambaran bahwa seluruh aitem berada pada rentang  $>0.05$  yang berarti seluruh aitem tidak bias terhadap salah satu kelompok pada data demografi berdasarkan semester.

**Tabel 8. Hasil Uji DIF**

*Likelihood ratio Chi-square statistics*

	Statistic	p	Adj.p	Cronbach's $\alpha$
6	4.008	0.135	0.524	0.913
13	4.302	0.116	0.524	0.915
15	1.385	0.500	0.718	0.918
18	5.380	0.068	0.524	0.915
19	1.960	0.375	0.709	0.917
22	2.096	0.351	0.709	0.920
27	1.450	0.484	0.718	0.920
39	0.668	0.716	0.899	0.914
42	0.213	0.899	0.899	0.914
43	2.906	0.234	0.568	0.914
45	1.360	0.507	0.718	0.915
52	0.226	0.893	0.899	0.913
54	5.241	0.073	0.524	0.914
58	0.246	0.884	0.899	0.914
60	0.262	0.877	0.899	0.916
59	3.445	0.179	0.524	0.915
65	3.374	0.185	0.524	0.918

Note. Adj.p = The adjusted p-values by likelihood ratio test using multiple comparison.

Setelah dilakukan pengujian DIF, uji reliabilitas dilakukan dengan melihat hasil pada Cronbach's  $\alpha$  yang menunjukkan seluruh nilai aitem berada diatas 0,30 yang berarti seluruh aitem sudah reliabel dalam mengukur regulasi diri akademik memiliki daya diskriminasi yang tinggi (Azwar, 2019).

## Pembahasan

Beberapa uji telah dilakukan dalam memverifikasi pengembangan alat ukur ini. Pertama adalah uji normalitas. Uji normalitas yang dilakukan sebanyak dua kali memperlihatkan dua hasil uji berbeda, yang mana fungsi uji yang pertama adalah memperlihatkan hasil uji normalitas dengan hasil uji seluruh aitem berdistribusi normal sehingga tidak ada aitem yang terdiskualifikasi. Selanjutnya uji normalitas yang kedua setelah CFA untuk memverifikasi kenormalan aitem pasca CFA menemukan hasil dengan nilai seluruh aitem berada pada rentang  $\pm 1$  yang menunjukkan seluruh aitem masih berdistribusi normal (Brown, 2015).

Hasil (CFA) *Confirmatory factor analysis* dalam pengukuran ini menunjukkan bahwa ada beberapa aitem yang harus gugur karena dinilai tidak dapat mengukur konstruk yang dimaksudkan. Dari 66 aitem pada skala yang disebar, hanya 17 aitem yang mampu benar-benar mengukur konstruk dari *Academic Self Regulation* yang sudah coba dikembangkan. Hal ini berdasarkan hasil uji CFA versi full aitem yang menunjukkan nilai RMSEA 0.094, dan SRMR 0.159. Hal ini masih sangat jauh dari batas ambang fit nilai RMSEA dan RSMR yang dikemukakan oleh (Hair dkk, 2019), yakni  $< 0,06$ . Hal tersebut juga didukung oleh faktor yang ditunjukkan pada model CFA yang memperlihatkan banyaknya aitem yang membuat skala ini tidak memenuhi syarat model fit berdasarkan nilai RMSEA dan SRMR sehingga beberapa aitem harus tereliminasi dan didapatkan 17 aitem fit dengan nilai fit CFA (RMSEA 0,058, SRMR 0,044, CFI 0,950, dan TLI 0,941) yang dapat dengan benar-benar mengukur konstruk dari skala regulasi diri akademik.

Adapun indikator dari aitem yang tereliminasi atau tidak ada sama sekali aitem yang mewakili setelah pengujian CFA adalah indikator "Memiliki Rencana dan Strategi dalam Belajar" dan indikator "Mengidentifikasi Prioritas dan Alokasi Waktu Sesuai Kebutuhan". Hal tersebut terjadi karena aitem-aitem pada indikator tersebut tidak memiliki faktor loadings yang cukup tinggi atau tidak menunjukkan korelasi dengan model teori yang dihipotesiskan (Byrne, 2013). Namun, meskipun demikian seluruh aspek dari skala pengukuran ini masih terdapat perwakilan aitem dari tiap-tiap aspek yang membuat skala pengukuran ini tetap valid karena masih mencakup semua aspek dari pengukuran (Messick, 1995).

Gugurnya indikator "Memiliki Rencana dan Strategi dalam Belajar" dan indikator "Mengidentifikasi Prioritas dan Alokasi Waktu Sesuai Kebutuhan" dipengaruhi oleh hasil data yang ternyata dalam aitem pertanyaan-pertanyaan yang diberikan pada subject masih bersinggungan dengan pernyataan-pernyataan pada indikator lain yang diberikan. Hal tersebut membuat aitem-aitem yang mewakili indikator-indikator tersebut gugur dan tidak tersisa sama sekali (DeVellis, 2017). Hal ini ditunjukkan oleh pertanyaan tentang distraksi pada indikator "Memiliki Rencana dan Strategi dalam Belajar" yang bersinggungan dengan indikator "Mampu memonitoring proses pembelajaran dan menghindari distraksi" menjadi salah satu bukti adanya singgungan aitem-aitem pertanyaan yang membuat salah satu indikator harus gugur pada proses CFA.

Langkah selanjutnya adalah uji dengan model MIMIC (*Multiple Indicators, Multiple Causes*). Perlu diketahui bahwa sebelum dilakukan uji model MIMIC, seluruh data demografis telah melewati uji fit model MIMIC dengan seluruh hasil dinyatakan fit berdasarkan asumsi dalam penelitian Schumacker dan Lomax, (2016), CFI  $> 0.9$ ; TLI  $> 0.9$ ; RMSEA  $< 0.6$ ; SRMR  $< 0.6$ . Hasil uji dengan model MIMIC menunjukkan bahwa terdapat beberapa data demografis yang bias pada kelompok-kelompok tertentu ketika dilakukan pengukuran dengan menggunakan konstruk alat ukur regulasi diri akademik. Data demografis yang dimaksud adalah data jenis kelamin dan juga usia. Data jenis kelamin memiliki bias terhadap aspek *forethought* dengan signifikansi  $< 0.001$ . Hal yang sama juga terjadi pada data usia yang memiliki bias pada aspek *self reflection* dengan signifikansi 0.002. Namun, meskipun demikian rata-rata data demografis tidak memiliki bias terhadap tiap-tiap aspek yang menunjukkan model pengukuran dapat digunakan oleh setiap variabel pada data demografis.

Aspek *forethought* mencakup aktivitas perencanaan belajar seperti mengambil inisiatif, merencanakan, dan menetapkan tujuan belajar (Zimmerman 1989). Hasil data demografis jenis kelamin yang signifikan terhadap aspek *forethought* memberikan gambaran bahwa jenis kelamin laki-laki dan perempuan masing-masing memiliki cara perencanaan belajar yang berbeda-beda yang tidak dapat disamakan dengan pengukuran yang dilakukan. Berdasarkan hal tersebut, alat ukur yang dikonstruksi setelah dilakukan uji MIMIC berarti menguntungkan salah satu kelompok.

Perbedaan cara perencanaan belajar antara laki-laki dan perempuan disebabkan oleh beberapa faktor yang akan coba dirumuskan dalam penelitian Savariens dan Dam (1994). Laki-laki cenderung memiliki motivasi ekstrinsik yang lebih tinggi, lebih berorientasi pada nilai, dan lebih kompetitif dibandingkan perempuan. Mereka juga cenderung memiliki sikap negatif terhadap belajar dan memperoleh skor lebih tinggi pada skala neurotisme dan ketergantungan. Sebaliknya, perempuan cenderung memiliki motivasi intrinsik yang lebih tinggi. laki-laki lebih cenderung memilih mode pembelajaran konseptualisasi abstrak, sedangkan perempuan lebih cenderung pada

komponen afektif pada pendekatan belajar.

Sementara signifikansi hasil uji MIMIC pada data demografis tingkat usia dengan aspek *self reflection* menunjukkan adanya salah satu kelompok yang diuntungkan dari hasil pengukuran. Hal ini dikarenakan setiap usia pasti memiliki *self reflection* yang berbeda-beda yang mana semakin tinggi usia maka *self reflection* akan meningkat hal tersebut merupakan proses pendewasaan dimana subjek dengan usia lebih tua akan lebih sadar tentang tanggung jawab yang diemban sehingga mereka perlu melakukan *self reflection* dibandingkan dengan subjek dengan usia yang lebih muda. Hal tersebut didukung oleh penelitian yang dilakukan oleh Khan dkk, (2024) yang menemukan bahwa peserta yang lebih senior (di atas 23 tahun) memiliki skor RLS yang lebih tinggi dibandingkan dengan peserta yang lebih muda (di bawah 20 tahun). Hal ini mungkin disebabkan oleh pengalaman yang lebih banyak dalam rotasi klinis, yang memberikan mereka lebih banyak kesempatan untuk mendiskusikan situasi yang tidak terdefinisi dengan baik dan menyelesaikan kasus-kasus sulit bersama dengan orang lain.

Selain aspek *forethought* dengan jenis kelamin dan aspek *self reflection* dengan usia, hasil data menunjukkan bahwa seluruh aspek dapat diterapkan kepada setiap data demografis dan tidak dipengaruhi oleh faktor lain yang berarti tidak bias kepada salah satu kelompok saja. Hal ini menunjukkan pengukuran yang dilakukan sudah sangat baik dan layak untuk digunakan oleh setiap data demografis berdasarkan analisis MIMIC (Finney dan Distefano, 2006).

Selanjutnya adalah uji *Rasch Model*. Uji *Rasch Model* dengan empat asumsi menemukan bahwa pada uji asumsi unidimensionality, semua butir atau pertanyaan dalam tes mengukur hanya satu dimensi atau karakteristik yang sama. Hal tersebut tercerminkan pada hasil uji CFA versi item reduksi yang memiliki nilai CFI  $0.950 > 0.90$ , TLI  $0.941 > 0.90$ , RMSEA  $0.058 < 0.60$ , SRMR  $0.044 < 0.60$ . Artinya, tidak ada dimensi tersembunyi atau tambahan yang diukur oleh aitem tersebut (Ludlow dan Haley, 1995). Setelah itu pada uji asumsi selanjutnya yakni item homogeneity yang menunjukkan hasil bahwa asumsi *item homogeneity* sudah terpenuhi Berdasarkan nilai dari tabel *item homogeneity* menunjukkan bahwa nilai setiap aitem berdasarkan nilai infit dan outfit sudah berada pada rentang 0,5-1,5 (Bond dan Fox, 2015). Hal ini menunjukkan bahwa semua butir atau pertanyaan (aitem) dalam tes setara dalam hal kesulitan dan karakteristik pengukuran.

Selain uji asumsi *unidimensionality* dan *item homogeneity*, uji asumsi pada model rasch juga terdapat uji *local independence* dan *monotonicity*. Berdasarkan hasil uji *local independence*, respon individu terhadap setiap butir tidak dipengaruhi oleh respon individu terhadap butir lain dalam tes. Hal tersebut tercerminkan dari hasil uji *local independence*, seluruh korelasi setiap butir aitem berada dibawah 0,3. Hal ini menunjukkan bahwa asumsi *local independence* sudah terpenuhi yang berarti setiap aitem dalam pengukuran tidak terdapat pengaruh pada respon individu terhadap setiap butir dengan pengaruh respon individu pada butir yang lain (Yen, 1984). Selain itu, asumsi model *rasch* yang terakhir yakni uji *monotonicity* juga terpenuhi oleh Gambaran pada diagram *Wright Map* dan Gambaran pada kurva ICC yang menjelaskan tingkat kesulitan aitem. Terlihat berdasarkan tingkat kesulitan aitem, aitem nomor 18 dan 22 memiliki tingkat kesulitan paling tinggi dari semua aitem. Namun kategori dari aitem tersebut masih masuk dalam range sedang karena masih berada dalam rentang 0-1,37. Hal ini menunjukkan bahwa asumsi *monotonicity* telah terpenuhi yang berarti probabilitas individu untuk menjawab benar suatu butir selalu meningkat atau tetap konstan dengan meningkatnya kemampuan individu (Van Der Linden dan Hambleton, 1997).

Selanjutnya adalah uji DIF. Berdasarkan hasil uji DIF, berarti tidak ada perbedaan fungsi aitem dalam mengukur regulasi diri akademik antar kelompok pada data demografi berdasarkan semester. Hal tersebut ditunjukkan pada gambar pada ICCs hampir identik untuk setiap kelompok, maka item tersebut tidak menunjukkan DIF yang berarti tidak ada bias item terhadap variabel demografi Semester. Hal tersebut juga didukung oleh nilai yang terdapat pada ratio Chi-Square yang memberikan gambaran bahwa seluruh aitem berada pada rentang  $>0.05$  yang berarti seluruh aitem tidak bias terhadap salah satu kelompok pada data demografi berdasarkan semester (Zumbo, 1999). Dengan terpenuhinya seluruh asumsi normalitas hingga validitas aitem, membuktikan bahwa alat ukur yang digunakan telah tervalidasi dan telah dipastikan bahwa alat ukur regulasi diri akademik tersebut benar-benar mengukur apa yang seharusnya diukur (Messick, 1995).

#### 4. Kesimpulan

Penelitian ini menunjukkan bahwa dari 66 aitem skala regulasi diri akademik yang disebar, hanya 17 aitem yang valid dan mampu mengukur konstruk *Academic Self Regulation* setelah melalui proses uji CFA. Beberapa indikator gugur karena adanya tumpang tindih antar-aitem sehingga tidak memiliki perwakilan yang memadai. Uji MIMIC mengungkap adanya bias pada aspek *forethought* berdasarkan jenis kelamin dan pada aspek *self-reflection* berdasarkan usia, yang mencerminkan adanya perbedaan cara perencanaan belajar antara laki-laki dan perempuan serta peningkatan refleksi diri seiring bertambahnya usia. Sementara itu, uji Rasch Model dan DIF menunjukkan bahwa seluruh aitem yang tersisa valid, unidimensional, dan bebas bias berdasarkan semester mahasiswa.

Kontribusi utama penelitian ini adalah menghasilkan skala regulasi diri akademik versi Indonesia dengan 17 aitem valid yang sesuai dengan konteks budaya akademik lokal. Skala ini dapat digunakan sebagai instrumen praktis dalam penelitian psikologi pendidikan maupun intervensi di perguruan tinggi. Implikasi teoretisnya, penelitian ini memperkaya literatur mengenai regulasi diri akademik dengan menekankan pentingnya faktor

demografis dalam pengukuran. Implikasi praktisnya, skala ini dapat membantu pendidik dan konselor akademik dalam mengidentifikasi, memantau, serta mengembangkan strategi regulasi diri mahasiswa untuk meningkatkan performa akademik dan kesejahteraan psikologis mereka.

Saran untuk penelitian selanjutnya adalah memperluas pengujian pada populasi mahasiswa dari berbagai wilayah dan jenis perguruan tinggi di Indonesia guna memastikan generalisasi hasil, serta mengintegrasikan aspek lingkungan belajar yang lebih kompleks. Selain itu, perlu dilakukan pengembangan lanjutan untuk menambah indikator yang hilang agar skala regulasi diri akademik dapat mencakup seluruh aspek yang lebih komprehensif.

## 5. Referensi

- Azwar, Saifuddin. (2010). *Metode Penelitian*. Yogyakarta: Pustaka Pelajar
- Azwar, S. (2019). *Penyusunan Skala Psikologi (Edisi 2)*. Yogyakarta: Pustaka Pelajar
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (3rd ed.). Routledge.
- Brown, T.A. (2015). *Confirmatory Factor Analysis for Applied Research*. The Guilford Press
- Byrne, B. M. (2013). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming* (2nd ed.). Routledge.
- DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). SAGE Publications.
- Eames, D., Thomas, S., Norman, K., Simanton, E., & Weisman, A. (2024). Sociodemographic disadvantage in the burden of stress and academic performance in medical school: implications for diversity in medicine. *BMC Medical Education*, 24(1), 348. <https://doi.org/10.1186/s12909-024-05263-y>
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 269-314). IAP.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
- Huda, M., Aziz, R., & Muallifah, M. (2024). EFEKTIVITAS STRATEGI SIX THINK HATS DALAM MENINGKATKAN BERPIKIR KREATIF SISWA. *JIP (Jurnal Intervensi Psikologi)*, 16(2). <https://doi.org/10.20885/intervensipsikologi.vol16.iss2.art4>
- Jannah, F., & Sulianti, A. (2021). *Perspektif Mahasiswa sebagai Agen of Change melalui Pendidikan Kewarganegaraan*. <https://doi.org/10.47080/propatria.v2i2.593.this>
- Khan, N. F., Saeed, M., Yasmin, R., Butt, A. K., & Khan, A. A. (2018). Age and Gender Based Differences in Self-Assessed Reflectionin-Learning Scale. *Journal of The Pakistan Dental Association*, 27(03), 133–139. <https://doi.org/10.25301/jpda.273.133>
- Lee, N., Cadogan, J. W., & Chamberlain, L. (2013). The MIMIC model and formative variables: problems and solutions. In *AMS Review* (Vol. 3, Issue 1, pp. 3–17). Springer. <https://doi.org/10.1007/s13162-013-0033-1>
- Ludlow, L. H., & Haley, M. S. (1995). Rasch Model Logits: Interpretation, Use, and Transformation. *Boston Colledge*, 55, 967–975.
- Mahmood, N. (2013). *Development and Validation of an Academic Self-Regulation Scale (ASRS) for University Students*. <https://www.researchgate.net/publication/334626448>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749.
- Mulaik, S. A. (1988). Confirmatory Factor Analysis. In *Handbook of Multivariate Experimental Psychology* (pp. 259–285).
- Radović, S., Seidel, N., Menze, D., & Kasakowskij, R. (2024). Investigating the effects of Different levels of students' regulation support on learning process and outcome: In search of the optimal level of support for self-regulated learning. *Computers & Education*, 215, 105041. <https://doi.org/10.1016/j.compedu.2024.105041>
- Severiens, S. E., & Ten Dam, G. T. M. (1994). Gender Differences in learning styles: A narrative review and quantitative meta-analysis. In *Higher Education* (Vol. 27, Issue 4, pp. 487–501). Kluwer Academic Publishers. <https://doi.org/10.1007/BF01384906>
- Schumacker, R. E., & Lomax, R. G. (2016). *A Beginner's Guide to Structural Equation Modeling* (4th ed.). Routledge
- Sugiyono. (2013). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Valenzuela, R., Codina, N., Castillo, I., & Pestana, J. V. (2020). Young University Students' Academic Self-Regulation Profiles and Their Associated Procrastination: Autonomous Functioning Requires Self-Regulated Operations. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00354>
- Van der Linden, W. J., & Hambleton, R. K. (Eds.). (1997). *Handbook of Modern Item Response Theory*. Springer
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125-145.

- Zimmerman. (2008). *Investigating Self-Regulation and Motivation: Historical Background, Methodological Development, and Future Prospect*.
- Zimmerman, B. J. (1989). A Social Cognitive View of Self-Regulated Academic Learning. In *Journal of Educational Psychology* (Vol. 81, Issue 3).
- Zumbo, B. D., & Wild, W. R. (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, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense



## KONSTRUKSI DAN EVALUASI PSIKOMETRIK SKALA COUNTERPRODUCTIVE WORK BEHAVIOR

Muhammad Ardi<sup>1</sup>, Nilam Widyarini<sup>2</sup>

<sup>1,2</sup> Fakultas Psikologi, Universitas Gunadarma, Jakarta, Indonesia

Email: [ardi@nsd.id](mailto:ardi@nsd.id)

p-ISSN: 2337-4845  
e-ISSN: 2620-7486



Received	Revised	Accepted	Published
23 Juli 2025	4 Oktober 2025	27 Oktober 2025	30 Oktober 2025

### Abstract

Platform-based work environments give rise to distinctive forms of counterproductive work behavior (CWB)—ranging from transaction manipulation and risky operational practices to aggression toward others—that are not fully captured by existing CWB instruments. This study aims to construct and psychometrically evaluate a CWB scale tailored to ride-hailing (online motorcycle-taxi) drivers in Indonesia. Participants were 1,416 drivers. Exploratory factor analysis (EFA) mapped items onto two dimensions—CWB toward the organization (CWB-O) and CWB toward individuals (CWB-I); one item was eliminated, yielding 19 items (CWB-O = 14; CWB-I = 5). The two-factor model was confirmed via confirmatory factor analysis (CFA) with good fit (CFI = 0.948; TLI = 0.941; RMSEA = 0.067; SRMR = 0.035; GFI = 0.914). Reliability indicated strong internal consistency: CWB-O ( $\omega = 0.956$ ;  $\alpha = 0.955$ ), CWB-I ( $\omega = 0.759$ ;  $\alpha = 0.781$ ), and the total score ( $\omega = 0.954$ ;  $\alpha = 0.950$ ). These findings suggest that the developed CWB scale is valid and reliable for detecting counterproductive work tendencies among ride-hailing drivers and is suitable for research and practical monitoring.

**Keywords:** counterproductive work behavior, CWB-O, CWB-I, exploratory factor analysis, confirmatory factor analysis, reliability, ride-hailing drivers

### Abstrak

Lingkungan kerja berbasis *platform* memunculkan bentuk *counterproductive work behavior* (CWB) yang khas, mulai dari manipulasi transaksi, perilaku operasional berisiko, hingga agresi terhadap orang lain yang belum sepenuhnya tercakup dalam instrumen *counterproductive work behavior* yang ada. Penelitian ini bertujuan mengonstruksi dan mengevaluasi psikometrik skala *counterproductive work behavior* yang relevan bagi pengemudi ojek *online* di Indonesia. Partisipan berjumlah 1.416 pengemudi ojek *online*. Hasil *exploratory factor analysis* (EFA) digunakan untuk memetakan item ke dua dimensi, yaitu CWB terhadap organisasi (CWB-O) dan CWB terhadap individu (CWB-I), dengan satu butir dieliminasi sehingga tersisa 19 butir (CWB-O = 14; CWB-I = 5). Model dua faktor terkonfirmasi melalui *confirmatory factor analysis* (CFA) dengan kecocokan model yang baik (CFI = 0,948; TLI = 0,941; RMSEA = 0,067; SRMR = 0,035; GFI = 0,914). Reliabilitas menunjukkan konsistensi internal kuat: CWB-O ( $\omega = 0,956$ ;  $\alpha = 0,955$ ), CWB-I ( $\omega = 0,759$ ;  $\alpha = 0,781$ ), dan skor total ( $\omega = 0,954$ ;  $\alpha = 0,950$ ). Temuan ini menunjukkan bahwa skala *counterproductive work behavior* yang dikembangkan valid dan reliabel untuk mendeteksi kecenderungan perilaku kerja kontraproduktif pada pengemudi ojek *online* serta layak digunakan untuk riset dan pemantauan praktis.

**Kata kunci:** *counterproductive work behavior*, CWB-O, CWB-I, *exploratory factor analysis*, *confirmatory factor analysis*, reliabilitas, pengemudi ojek online

### 1. Pendahuluan

Perubahan struktur kerja akibat digitalisasi mendorong tumbuhnya ekosistem kemitraan berbasis *platform*. Dalam konteks ini, relasi antara perusahaan, aplikasi dan pengemudi ojek *online* menuntut tata kelola perilaku kerja yang mengutamakan integritas layanan, keselamatan, dan kepatuhan pada kebijakan sistem. *Counterproductive work behavior* merujuk pada perilaku karyawan yang dilakukan secara sengaja dan berdampak merugikan karyawan, baik dengan menurunkan efektivitas kerja maupun merugikan organisasi melalui perusakan aset atau mengganggu keberfungsian organisasi. Karena hal tersebut, *counterproductive work behavior* kerap diperlakukan setara dengan penyimpangan di tempat kerja (Bennett, Shelly & Lauren, 2018). Rangkaian perilaku *counterproductive work behavior* memiliki dua karakteristik utama, yaitu bersifat volisional (bukan kecelakaan atau karena perintah) dan menimbulkan atau memang dimaksudkan untuk menimbulkan kerugian bagi organisasi

atau para pemangku kepentingan organisasi seperti klien, rekan kerja, pelanggan, dan atasan (Spector & Fox, 2005).

Dalam praktiknya, bentuk *counterproductive work behavior* yang paling sering dijumpai dalam organisasi antara lain penarikan diri (misalnya datang terlambat atau tidak masuk kerja) dan bersikap tidak sopan terhadap orang lain. Meski tidak diharapkan, perilaku-perilaku ini umumnya tidak menimbulkan dampak yang terlalu berat. Sebaliknya, bentuk *counterproductive work behavior* yang lebih jarang seperti sengaja membuang-buang waktu, sabotase, pencurian, tindak kekerasan, penggunaan narkoba atau alkohol saat bekerja, dan pelecehan seksual. Walau jarang terjadi, perilaku ini berpotensi sangat merusak dan pada akhirnya menimbulkan kerugian besar bagi organisasi (Jex & Britt, 2014). Ketika perilaku menyimpang ditujukan kepada rekan kerja, korban dapat mengalami kerugian finansial seperti mengakibatkan cedera atau pencurian barang pribadi, dan juga terbebani secara psikologis seperti merasa menderita, frustrasi, kebingungan, stres, atau rasa takut. Dari sisi sosial, bila perusahaan kehilangan uang akibat perilaku menyimpang contohnya pencurian, kerugian itu sering dialihkan kepada masyarakat dengan menaikkan harga, sehingga pelanggan turut menanggung biaya akibat kenaikan harga. Pelanggan yang mengalami atau menyaksikan perilaku menyimpang dapat mengalami tekanan emosional dan mental serupa dengan karyawan (frustrasi, marah, atau stres) yang dapat berujung pada boikot terhadap perusahaan atau bahkan menempuh jalur hukum (Bennett, Shelly & Lauren, 2018).

Meskipun *counterproductive work behavior* terus mendapat perhatian karena dampaknya terhadap efektivitas organisasi, keselamatan layanan, dan keadilan dalam relasi kerja. Sebagian besar instrumen *counterproductive work behavior* yang tersedia dikembangkan untuk konteks pekerjaan formal, ditandai dengan hubungan kerja tradisional, pengawasan langsung dan indikator yang fokus pada aset fisik organisasi, seperti fasilitas kantor dan peralatan pabrik. Sementara itu, lanskap kerja kontemporer juga mencakup hubungan kerja informal dan juga hubungan kemitraan, seperti hubungan pada ojek *online* (ojol) dengan perusahaan aplikasi yang merupakan hubungan kemitraan. Dalam skema ini, pengemudi yang telah melakukan registrasi dan menyetujui ketentuan layanan dan kebijakan bertindak sebagai mitra pengemudi untuk menyediakan layanan transportasi sesuai permintaan. Dengan demikian, meskipun status pengemudi ojek *online* secara hukum adalah mitra bukan karyawan, aktivitas yang dijalankan tetap merupakan pekerjaan yang diatur oleh seperangkat norma, kebijakan, dan standar keselamatan atau layanan berbasis aplikasi.

Dalam konteks kemitraan berbasis *platform* juga memperlihatkan bentuk-bentuk *counterproductive work behavior* yang khas. Sebuah kasus yang terdokumentasi dalam sistem PT X menunjukkan seorang pengemudi telah melakukan pemesanan atau order fiktif dengan menggunakan 41 akun milik orang lain. Sejak Agustus 2019 hingga 7 Februari 2020, pelaku tercatat melakukan order fiktif sebanyak 30.848 dengan rata-rata sekitar 200 orderan fiktif setiap harinya dan juga menggunakan akun resto untuk melakukan 30 pemesanan. Perbuatan tersebut menyebabkan PT X mengalami defisit yang sangat besar dikarenakan sistem memberikan ongkos kirim dan *voucher* kepada pelaku melalui akun pelanggan sebesar Rp.4000,- per orderan dan *voucher* sebesar Rp.10.000,- setiap orderan akan masuk kedalam saldo *merchant* yang dapat dicairkan keesokan harinya dengan cara ditransfer ke rekening yang di daftarkan para *merchant* (Pangku & Qomaruddin, 2022).

Dari sisi keselamatan pelanggan, studi di Kota Bogor mengidentifikasi perilaku berkendara pengemudi ojek *online* yang dapat menyebabkan kecelakaan lalu lintas, yaitu mendahului dari sisi kiri, menurunkan kecepatan secara tiba-tiba, langsung mengambil jalur kiri ketika berbalik arah dan memacu kendaraan dengan kecepatan tinggi (Axiery, Ruby & Rahma, 2019). Selain itu, tercatat kasus agresi yang melibatkan pengemudi ojek *online*, salah satu kasus yang terjadi pada 4 September 2018 di Kota Jakarta, terjadi pemukulan yang dilakukan oleh pengemudi ojek *online* kepada seseorang berinisial B. Kejadian berawal saat korban sedang mengendarai sepeda motor dengan istrinya dalam perjalanan pulang ke rumah, motor korban mendahului pelaku. Pelaku yang tidak terima didahului akhirnya menabrak ban belakang motor korban dan berujung pada pemukulan korban. Korban juga sempat dikelilingi para pengemudi ojek *online* lainnya (Putri, Maharyantari & Firda, 2022).

Ketiga kasus tersebut menegaskan bahwa *counterproductive work behavior* pada lingkungan kerja berbasis *platform* dan operasional (berkendara tidak aman) dan agresi kepada orang lain yang dapat berdampak pada perusahaan belum sepenuhnya tercakup dalam indikator skala yang sudah ada, sehingga skala perlu disesuaikan dengan konteks pekerjaan pengemudi ojek *online*.

Dengan demikian, meskipun *counterproductive work behavior* dirancang untuk konteks pekerjaan formal, fokus utamanya yakni perilaku kerja yang disengaja dan merugikan organisasi atau pemangku kepentingan tetap relevan untuk kemitraan berbasis *platform*. Agar pengukuran akurat dan dapat diandalkan pada populasi mitra ojek *online*, instrumen *counterproductive work behavior* perlu direkonstruksi. Penelitian ini dilakukan untuk mengkonstruksi dan mengeksplorasi pengembangan awal skala *counterproductive work behavior* yang ditujukan untuk populasi ojek *online*. Langkah ini diharapkan dapat menghasilkan skala *counterproductive work behavior* yang lebih representatif terhadap realitas kerja ojek *online*, sekaligus bermanfaat bagi riset dan kebijakan peningkatan keselamatan serta integritas layanan.

## 2. Metode Penelitian

### Partisipan

Partisipan dalam penelitian ini adalah pengemudi ojek *online*, baik pengemudi roda dua maupun roda empat. Pemilihan sampel menggunakan *nonprobability purposive sampling*, Pendekatan yang banyak digunakan dalam penelitian psikologi ketika tujuan utama adalah mengkaji kelompok tertentu yang memenuhi karakteristik yang telah ditentukan sebelumnya (Campbell et al., 2020). Kriteria mencakup pengemudi ojek *online* yang telah bekerja minimal selama enam bulan, hal ini memastikan bahwa partisipan memiliki pengalaman dan pemahaman yang memadai terhadap aktivitas kerja. Kriteria inklusi mencakup, yakni memiliki akses terhadap perangkat digital dan koneksi internet yang stabil, bersedia mengikuti penelitian secara sukarela, dan mengisi seluruh item dalam instrumen serta melengkapi data demografi. Kriteria eksklusi mencakup, yaitu partisipan yang tidak menyelesaikan seluruh item dalam tes, tidak melengkapi data demografi, atau menunjukkan pola respon yang tidak wajar yang berpotensi mengganggu validitas data.

Pengumpulan data dilakukan secara daring pada Juni hingga Juli 2025. Undangan partisipasi dikirim melalui email yang memuat tautan menuju kuesioner (Google *forms*). Untuk menjamin kualitas data, sebelum analisis dilakukan pembersihan data berdasarkan kriteria eksklusi dan inklusi. Sebanyak 1.416 responden memenuhi kriteria dan menyelesaikan kuesioner, sehingga data dapat digunakan dalam analisis.

### Instrumen

Instrumen dalam penelitian ini dikonstruksi dengan merujuk pada tipologi penyimpangan ditempat kerja yang dikemukakan oleh Bennett dan Sandra (2000). Sejalan dengan tipologi tersebut, *counterproductive work behavior* didimensionalisasi berdasarkan target perilaku, yakni perilaku yang menargetkan organisasi (*organizational deviance*) dan perilaku yang menargetkan individu (*interpersonal deviance*), dengan rentang tingkat keparahan dari ringan hingga serius. Dalam literatur, *counterproductive work behavior* kerap diperlakukan setara dengan penyimpangan di tempat kerja. Selanjutnya *counterproductive work behavior* terdiri dari dua dimensi, yaitu yang menargetkan organisasi (CWB-O) dan yang menargetkan individu (CWB-P atau CWB-I) (Bennett, Shelly & Lauren, 2018). *Counterproductive work behavior organizational* merepresentasikan perilaku yang diarahkan pada organisasi secara keseluruhan, bukan pada individu tertentu dalam lingkungan organisasi. Contohnya membuang waktu, sabotase, datang terlambat, dan berbicara buruk mengenai organisasi kepada orang lain. Sedangkan, *counterproductive work behavior interpersonal* merepresentasikan perilaku yang diarahkan kepada rekan kerja secara perorangan, bukan pada organisasi secara keseluruhan. Contoh yang mungkin paling umum meliputi membuat komentar yang kasar dan menghina orang lain, menyebarkan rumor mengenai orang lain, serta mengecualikan rekan kerja dari aktivitas seperti makan siang atau minum bersama setelah jam kerja (Jex & Britt, 2014).

Berdasarkan kerangka tersebut, peneliti menyusun 20 item yang dipetakan ke dalam dua model dimensi, yaitu *counterproductive work behavior-organizational* (CWB-O) dan *counterproductive work behavior-interpersonal* (CWB-I), masing-masing dimensi terdiri dari 10 item. Seluruh item disusun sebagai favorabel untuk meminimalkan ambiguitas linguistik dalam Bahasa Indonesia dan menghindari miskonsepsi yang kerap muncul pada item dengan format kalimat negatif. Pemilihan 20 item dimaksudkan untuk menyeimbangkan cakupan konten pada kedua dimensi dan beban respon (efisiensi waktu dan keterbacaan) pada populasi ojek *online*. Seluruh item dirancang agar sesuai dengan konteks pekerjaan berbasis platform untuk ojek *online*, mencakup dua ranah utama, yaitu interaksi layanan dengan pelanggan serta kepatuhan terhadap sistem operasional platform. Daftar item *counterproductive work behavior* disajikan pada Tabel 1.

Respon diukur dengan skala Likert lima poin dari 1 (sangat tidak sesuai) hingga 5 (sangat sesuai). Pemilihan lima poin didasarkan pada pertimbangan metodologis bahwa keberadaan kategori tengah (*middle category*) membantu mencegah pergeseran sistematis respon ke arah kategori positif ketika pilihan tengah diadakan. Bukti eksperimental menunjukkan bahwa ketiadaan kategori tengah dapat mengubah distribusi respon, responden terdorong untuk memilih sisi sehingga proporsi jawaban yang lebih positif meningkat sementara jawaban negatif menurun (Wetzelhutter, 2020).

**Tabel 1. Item pernyataan**

No	Item Pernyataan	Dimensi
1	Saya terkadang sengaja menunda perjalanan agar mendapatkan tarif lebih tinggi.	CWB-O
2	Saya pernah menggunakan akun saya untuk mengakses pesanan tanpa izin dari pihak lain.	CWB-O
3	Saya sering mengambil waktu istirahat lebih lama dari yang diperbolehkan oleh aplikasi.	CWB-O
4	Saya pernah menggunakan insentif dari aplikasi untuk keperluan pribadi tanpa izin.	CWB-O
5	Saya pernah memalsukan struk atau bukti perjalanan untuk mendapatkan penggantian lebih besar.	CWB-O
6	Saya terkadang mematikan fitur aplikasi untuk menghindari pesanan yang tidak saya sukai.	CWB-O
7	Saya pernah menyelesaikan pesanan secara asal tanpa memastikan pelanggan mendapatkan pengalaman terbaik.	CWB-O
8	Saya pernah menyimpan barang pelanggan yang tertinggal tanpa melaporkannya ke pihak aplikasi.	CWB-O
9	Saya pernah mengakhiri perjalanan lebih awal tanpa sepengetahuan pelanggan.	CWB-O
10	Saya terkadang mengabaikan instruksi aplikasi karena menurut saya tidak relevan.	CWB-O
11	Saya pernah membicarakan keburukan pelanggan kepada rekan sesama driver.	CWB-I
12	Saya pernah berbicara kasar kepada pelanggan ketika saya merasa kesal.	CWB-I
13	Saya pernah mengolok-olok pelanggan karena permintaan mereka yang menurut saya tidak masuk akal.	CWB-I
14	Saya terkadang menyebarkan rumor negatif tentang rekan driver kepada orang lain.	CWB-I
15	Saya pernah membalas komentar buruk pelanggan dengan nada yang tidak sopan.	CWB-I
16	Saya pernah berbicara kepada pelanggan dengan nada yang tidak profesional.	CWB-I
17	Saya pernah meninggalkan pelanggan menunggu terlalu lama tanpa pemberitahuan.	CWB-I
18	Saya pernah menolak permintaan pelanggan tanpa memberikan alasan yang jelas.	CWB-I
19	Saya pernah mempermalukan pelanggan karena kesalahan yang mereka lakukan.	CWB-I
20	Saya pernah membuat komentar yang tidak pantas tentang pelanggan kepada rekan sesama driver.	CWB-I

### Analisis Data

Pada penelitian ini dilakukan uji validitas dan reliabilitas guna memastikan skala yang dikembangkan akurat dan dapat dipercaya. Bukti validitas konstruk dikaji melalui dua tahap, yaitu *exploratory factor analysis* (EFA) dan *confirmatory factor analysis* (CFA). Teknik analisis *factor analysis* yang awalnya diusulkan oleh Spearman (1904) kini disebut *exploratory factor analysis* (Thompson, 1951). *Exploratory factor analysis* merupakan teknik interdependensi yang tujuan utamanya mendefinisikan struktur dasar yang mendasari sekumpulan variabel dalam analisis. *Exploratory factor analysis* digunakan untuk mengeksplorasi struktur laten dan menentukan jumlah faktor yang paling merepresentasikan data. *Confirmatory factor analysis* diterapkan untuk menguji sejauh mana model pengukuran yang telah ditentukan sebelumnya merepresentasikan data aktual (Hair, William, Barry & Rolph, 2010).

Analisis data dilakukan dengan bantuan aplikasi pengolahan data SPSS untuk *exploratory factor analysis* dan JASP untuk *confirmatory factor analysis*. Pada tahap *exploratory factor analysis*, jumlah faktor ditentukan, item dikelompokkan menurut faktor, dan setiap faktor diberi label sesuai kontennya. Tahap berikutnya, *confirmatory factor analysis* menguji kesesuaian (*fit*) model pengukuran yang diturunkan dari *exploratory factor analysis* dan teori. *Fit model* dinilai menggunakan sejumlah indeks, yaitu *Comparative Fit Index* (CFI), *Tucker-Lewis Index* (TLI), *Root Mean Square Error of Approximation* (RMSEA), *Goodness of Fit Index* (GFI) dan *Standardized Root Mean Square Residual* (SRMR). Model dinyatakan memadai apabila memenuhi kriteria *cutoff* berikut: CFI dan TLI  $\geq 0,95$ , RMSEA  $\leq 0,06$ , SRMR  $< 0,08$  (Hu & Peter, 1999), dan GFI  $\geq 0,90$  (Hooper, Joseph & Michael, 2008). Estimasi parameter yang digunakan adalah *Maximum Likelihood* (ML).

Kemudian dilakukan uji reliabilitas digunakan untuk menilai sejauh mana instrumen pengukuran

menghasilkan hasil yang konsisten. Reliabilitas diukur menggunakan koefisien McDonald's Omega ( $\omega$ ) dan Cronbach's Alpha ( $\alpha$ ). Dalam penelitian psikometri, nilai reliabilitas dianggap baik apabila  $\geq 0.70$  (Hair et al., 2010).

### 3. Hasil dan Diskusi

#### Data Demografi

Responden pada penelitian ini sebanyak 1.416 orang. Berdasarkan tabel data demografi, mayoritas responden berada pada rentang usia produktif 26 hingga 45 tahun (74%). Proporsi laki-laki sebanyak 95% dan perempuan sebanyak 5%. Berdasarkan latar belakang pendidikan terakhir sebagian besar responden merupakan lulusan SMA (64%), diikuti lulusan S1 (16%) dan diploma (8%). Dari sisi masa kerja sebagai pengemudi ojek *online*, sebanyak 45% telah bekerja lebih dari 5 tahun, sementara sisanya tersebar pada rentang masa kerja 1 hingga 4 tahun. Gambaran ini menunjukkan bahwa sampel merepresentasikan mitra pengemudi ojek *online*.

**Tabel 2. Data Demografi**

Demografi	Kriteria	N	Persentase
Usia	< 20	5	0%
	20-25	137	10%
	26-30	293	21%
	31-35	269	19%
	36-40	254	18%
	41-45	222	16%
	46-50	117	8%
	51-55	84	6%
	56-60	31	2%
	> 60	4	0%
Pendidikan	SD	32	2%
	SMP	119	8%
	SMA	909	64%
	DIPLOMA	118	8%
	S1	225	16%
	S2	12	1%
Jenis Kelamin	Laki-laki	1346	95%
	Perempuan	70	5%
Masa kerja berprofesi sebagai pengemudi ojek <i>online</i>	1 Tahun	303	21%
	2 Tahun	139	10%
	3 Tahun	209	15%
	4-5 Tahun	122	9%

#### Analisis Struktur Faktor

Skala *counterproductive work behavior* dikembangkan berdasarkan dua dimensi, yaitu *counterproductive work behavior organizational* (CWB - O) dan *counterproductive work behavior interpersonal* (CWB - I). Item disusun berlandaskan teori tersebut, kemudian dievaluasi menggunakan *exploratory factor analysis* untuk menilai kesesuaian item dengan faktor yang diukur. Mengacu pada tabel 5, suatu item dinyatakan berkontribusi signifikan terhadap suatu faktor apabila memiliki nilai *factor loading* sebesar  $\geq 0,40$ . Nilai ini digunakan sebagai kriteria minimal bahwa item secara substansial memberikan kontribusi dalam terhadap varians faktor laten yang diukur. Sebaliknya, item dengan *factor loading* di bawah 0,40 dianggap lemah dan tidak merepresentasikan konstruk secara memadai, sehingga dipertimbangkan untuk direvisi atau dieliminasi dari model dengan tetap memperhatikan kesesuaian konten dan rasional teoretis.

Berikut tabel 3 yang menunjukkan distribusi nilai *factor loading* pada dua faktor utama beserta nilai *uniqueness* dari masing-masing item:

**Tabel 3. Factor Loadings**

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Uniqueness</i>
Item 14	0.924		0.259
Item 19	0.918		0.281
Item 5	0.890		0.280
Item 16	0.799		0.328
Item 13	0.792		0.371
Item 8	0.790		0.387
Item 15	0.778		0.330
Item 12	0.776		0.361
Item 2	0.745		0.438
Item 9	0.714		0.409
Item 20	0.711		0.407
Item 17	0.646		0.430
Item 7	0.537		0.500
Item 1	0.405		0.552
Item 6		0.808	0.554
Item 10		0.635	0.451
Item 3		0.592	0.672
Item 4		0.482	0.613
Item 11		0.444	0.465
Item 18			0.551

Tabel di atas menunjukkan hasil analisis *exploratory factor analysis* yang memetakan 20 item ke dalam dua faktor utama. Sebagian besar item memuat *loading* tinggi pada Faktor 1 yang mencerminkan konsistensi terhadap dimensi pertama, terlihat pada Item 14 (0.924), Item 19 (0.918), dan Item 5 (0.890) yang masing-masing menunjukkan kontribusi kuat terhadap konstruk tersebut. Sementara itu, Faktor 2 ditandai oleh Item 6 (0.808), Item 10 (0.635), dan Item 3 (0.592) yang menunjukkan keterkaitan dengan dimensi kedua. Nilai *uniqueness* menunjukkan sebagian besar item memiliki proporsi varians yang dijelaskan oleh kedua faktor, meskipun beberapa item seperti item 3 (0.672) dan item 1 (0.552) memiliki nilai residual yang tinggi sehingga kontribusi terhadap struktur faktor lebih rendah.

Namun, perlu dicatat bahwa Item 18 tidak menunjukkan *loading* yang signifikan pada Faktor 1 maupun Faktor 2, serta memiliki nilai *uniqueness* sebesar 0.551. Temuan ini mengindikasikan bahwa item 18 cenderung mengukur dimensi yang berbeda dari dua faktor utama yang teridentifikasi, sehingga item tersebut digugurkan dan tidak digunakan.

Berdasarkan hasil analisis *exploratory factor analysis*, total item pada skala *counterproductive work behavior* berjumlah 19 item. Dimensi *counterproductive work behavior organizational* terdiri dari 14 item, yaitu item 1, 2, 5, 7, 8, 9, 12, 13, 14, 15, 16, 17, 19 dan 20. Dimensi *counterproductive work behavior interpersonal* terdiri dari 5 item, yaitu item 3, 4, 6, 10 dan 11.

Setelah dilakukan *exploratory factor analysis*, analisis dilanjutkan dengan *confirmatory factor analysis* untuk menguji kesesuaian (*fit*) model pengukuran yang diturunkan dari *exploratory factor analysis* dan landasan teoritis. Hasil menunjukkan kecocokan model yang baik dengan nilai CFI (0,948) dan TLI (0,941) berada diatas ambang model *fit* yang ideal. Hair et all (dalam Goey, 2018) menyatakan nilai CFI dikategorikan sebagai *good fit* jika nilainya  $\geq 0,9$ . Adapun nilai TLI berkisar antara 0 sampai 1, dengan nilai mendekati 1 menunjukkan kecocokan model yang baik. Nilai TLI  $\geq 0,9$  dikategorikan sebagai *good fit* (Gunawan, 2016).

**Tabel 4. Indeks Fit Model**

Index	Value
Comparative Fit Index (CFI)	0.948
Tucker-Lewis Index (TLI)	0.941
Bentler-Bonett Non-normed Fit Index (NNFI)	0.941
Bentler-Bonett Normed Fit Index (NFI)	0.940
Parsimony Normed Fit Index (PNFI)	0.830
Bollen's Relative Fit Index (RFI)	0.932
Bollen's Incremental Fit Index (IFI)	0.948
Relative Noncentrality Index (RNI)	0.948

Tabel 4 menyajikan berbagai indeks *fit* tambahan. Nilai RMSEA sebesar 0,067 menunjukkan *fit model* yang tergolong cukup baik, masih dalam rentang *acceptable fit* (0,05 hingga 0,08), dengan interval kepercayaan 90% antara 0,064 hingga 0,071. *Goodness of Fit Index* (GFI) sebesar 0,914. Kombinasi indeks *fit* tersebut menunjukkan kecocokan model yang baik terhadap data.

**Tabel 5. Indeks Fit Tambahan**

Metric	Value
Root mean square error of approximation (RMSEA)	0.067
RMSEA 90% CI lower bound	0.064
RMSEA 90% CI upper bound	0.071
RMSEA p-value	0.000
Standardized root mean square residual (SRMR)	0.035
Hoelter's critical N ( $\alpha = .05$ )	229.184
Hoelter's critical N ( $\alpha = .01$ )	246.444
Goodness of fit index (GFI)	0.914
McDonald fit index (MFI)	0.710
Expected cross validation index (ECVI)	0.847

Tabel 5 menunjukkan *factor loading* item terhadap faktor *counterproductive work behavior organizational* dan *counterproductive work behavior interpersonal*, berdasarkan hasil *exploratory factor analysis* sebelumnya item 18 dianggap gugur sehingga tidak digunakan. *Factor loadings* adalah korelasi antara setiap variabel dan faktor, nilai *loading* menunjukkan derajat kesesuaian antara variabel dan faktor, semakin tinggi nilai *loading*, semakin representatif variabel tersebut terhadap faktor (Hair et al., 2010). *Factor loading* diatas 0,4 dianggap memadai, seluruh item memenuhi kriteria tersebut. Pada CWB - O, *factor loading* berkisar dari 0,645 hingga 0,854, pada CWB - I berkisar dari 0,536 hingga 0,759. Temuan ini menunjukkan bahwa model ini mengkonfirmasi bahwa setiap dimensi berdiri secara konseptual dan empiris sebagai dimensi kunci *counterproductive work behavior*.

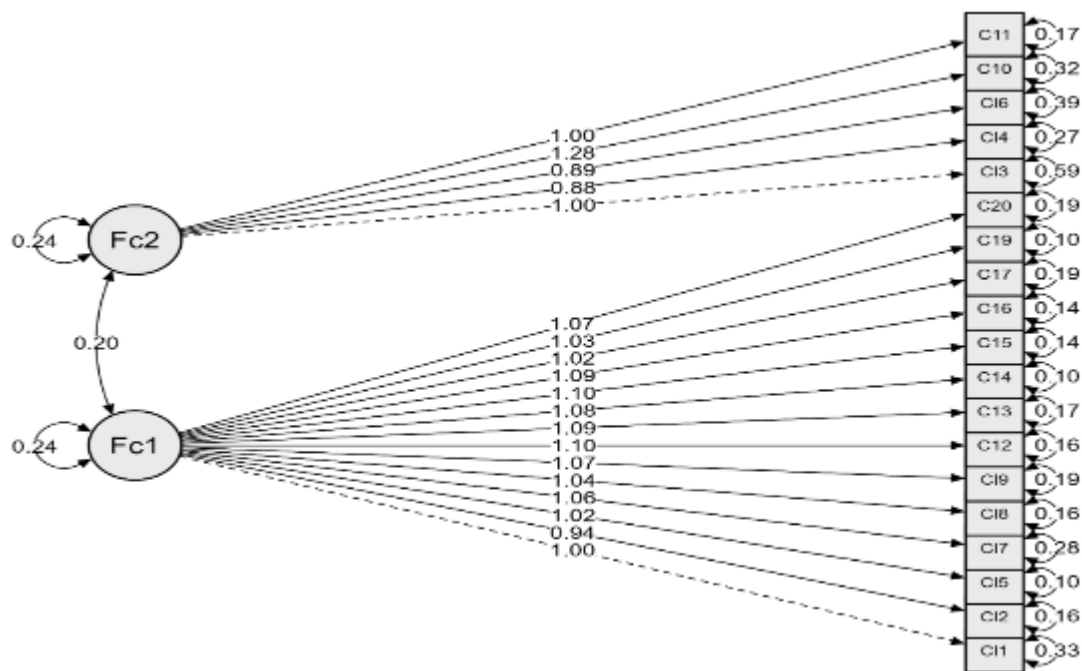
**Tabel 6. Factor Loadings CWB - O dan CWB - I**

Factor	Indicator	Std. estimate	Std. Error	z-value	p	95% Confidence Interval	
						Lower	Upper
CWB-O	Item 1	0.645	0.016	40.166	< .001	0.614	0.677
	Item 2	0.753	0.012	62.366	< .001	0.729	0.777
	Item 5	0.845	0.008	103.255	< .001	0.829	0.861
	Item 7	0.697	0.014	49.106	< .001	0.670	0.725
	Item 8	0.783	0.011	72.299	< .001	0.762	0.805
	Item 9	0.768	0.011	67.070	< .001	0.746	0.791
	Item 12	0.801	0.010	79.294	< .001	0.781	0.821
	Item 13	0.794	0.010	76.200	< .001	0.773	0.814
	Item 14	0.854	0.008	109.600	< .001	0.839	0.869
	Item 15	0.819	0.009	87.857	< .001	0.801	0.837
	Item 16	0.820	0.009	88.438	< .001	0.802	0.838
	Item 17	0.750	0.012	61.470	< .001	0.726	0.774
	Item 19	0.840	0.008	99.577	< .001	0.823	0.856
	Item 20	0.770	0.011	67.627	< .001	0.748	0.792
	CWB-I	Item 3	0.536	0.021	25.290	< .001	0.495
Item 4		0.634	0.018	34.530	< .001	0.598	0.670
Item 6		0.568	0.020	27.913	< .001	0.528	0.608
Item 10		0.738	0.015	49.485	< .001	0.709	0.768
Item 11		0.759	0.014	53.359	< .001	0.731	0.787

Tabel 6 menyajikan hasil analisis kovarians faktor yang menggambarkan hubungan antara faktor-faktor model yang diuji. Terdapat korelasi yang sangat kuat dan signifikan ( $r = 0.836$ ) antara Faktor 1 (CWB - O) dan Faktor 2 (CWB - I). Temuan ini mengindikasikan bahwa kedua faktor tersebut tidak sepenuhnya independen satu sama lain dan mengukur aspek-aspek yang saling berkaitan dari suatu konstruk yang lebih besar. Namun, korelasi ini masih di bawah 0.90, sehingga model dua faktor tetap dapat diterima.

**Tabel 7. Factor Covariances**

						95% Confidence Interval		
			Std. estimate	Std. Error	z-value	p	Lower	Upper
Factor 1	↔	Factor 2	0.836	0.012	67.861	< .001	0.812	0.860



**Gambar 1. Path diagrams model confirmatory factor analysis yang menggambarkan hubungan antara counterproductive work behavior organizational dan counterproductive work behavior interpersonal**

#### Reliabilitas

Uji reliabilitas dilakukan dengan menggunakan koefisien omega ( $\omega$ ) dan koefisien alpha ( $\alpha$ ). Pemilihan keduanya didasarkan pada bukti bahwa ketika asumsi yang mendasari koefisien alpha ( $\alpha$ ) tidak terpenuhi, akurasi dalam mengestimasi konsistensi internal menurun secara bermakna. Koefisien alpha ( $\alpha$ ) masih dapat digunakan secara hati-hati apabila skala terbukti unidimensional dan asumsi model *essentially tau-equivalent* terpenuhi. Namun, kondisi ini jarang dijumpai pada skala psikologis, dan pelanggaran yang tampak ringan dapat secara drastis membiarkan estimasi alpha ( $\alpha$ ). Sejalan dengan konsensus literatur yang menyatakan bahwa penggunaan koefisien alpha ( $\alpha$ ) jarang tepat serta temuan bahwa omega ( $\omega$ ) berkinerja lebih baik ketika asumsi alpha ( $\alpha$ ) tidak terpenuhi, disarankan beralih ke omega ( $\omega$ ) sebagai pengganti alpha ( $\alpha$ ) (Dunn, Thom & Vivienne, 2013). Meski demikian, hasil analisis reliabilitas dengan menggunakan koefisien alpha ( $\alpha$ ) tetap dilaporkan karena masih sering digunakan pada banyak penelitian.

Hasil analisis menunjukkan bahwa dimensi *counterproductive work behavior organizational* memiliki reliabilitas yang tinggi dengan nilai omega ( $\omega$ ) 0,956 dan alpha ( $\alpha$ ) 0,955 menandakan item-item bekerja seragam dalam mengukur konstruk yang sama. Pada dimensi *counterproductive work behavior interpersonal* nilai omega ( $\omega$ ) 0,759 dan alpha ( $\alpha$ ) 0,781 menunjukkan reliabilitas yang memadai untuk keperluan penelitian. Secara keseluruhan, skor total *counterproductive work behavior* mendapatkan nilai omega ( $\omega$ ) 0,954 dan alpha ( $\alpha$ ) 0,950, yang menunjukkan bahwa reliabilitas sangat tinggi. Dengan demikian, skala dapat dinyatakan reliabel.

**Tabel 8. Reliabilitas**

	Coefficient $\omega$	Coefficient $\alpha$
CWB-O	0.956	0.955
CWB-I	0.759	0.781
Total	0.954	0.950

#### 4. Kesimpulan

Penelitian ini berhasil mengonstruksi dan mengevaluasi skala *counterproductive work behavior* yang relevan untuk konteks pengemudi ojek online. *Counterproductive work behavior*, yang terdiri dari dua dimensi utama yaitu *counterproductive work behavior* terhadap organisasi (CWB-O) dan *counterproductive work behavior* terhadap individu (CWB-I), merupakan bentuk perilaku menyimpang yang dapat merugikan sistem kerja dan citra layanan. Dengan 1.416 partisipan, penelitian ini menunjukkan bahwa bentuk-bentuk *counterproductive work behavior* seperti manipulasi aplikasi, perlakuan tidak sopan terhadap pelanggan, dan pelanggaran norma kerja dapat teridentifikasi secara terstruktur menggunakan instrumen yang telah dikembangkan.

Hasil *exploratory factor analysis* memetakan 19 item ke dalam dua faktor, yaitu *counterproductive work behavior organizational* terdiri dari 14 item, yaitu item 1, 2, 5, 7, 8, 9, 12, 13, 14, 15, 16, 17, 19 dan 20. *Counterproductive work behavior interpersonal* terdiri dari 5 item, yaitu item 3, 4, 6, 10 dan 11. Satu item digugurkan yaitu item 18. Model dua faktor tersebut terkonfirmasi melalui *confirmatory factor analysis* dengan kecocokan model yang baik (CFI = 0,948; TLI = 0,941; RMSEA = 0,067; SRMR = 0,035; GFI = 0,914). Reliabilitas menunjukkan konsistensi internal yang kuat, *counterproductive work behavior organizational* ( $\omega = 0,956$ ;  $\alpha = 0,955$ ), *counterproductive work behavior interpersonal* ( $\omega = 0,759$ ;  $\alpha = 0,781$ ), dan skor total ( $\omega = 0,954$ ;  $\alpha = 0,950$ ). Korelasi antarfaktor ( $r = 0,836$ ) menandakan keduanya saling berkaitan namun tetap dapat dipisahkan sebagai dua konstruk yang berbeda. Secara keseluruhan, instrumen yang dikembangkan valid dan reliabel untuk mengidentifikasi kecenderungan perilaku kerja kontraproduktif pada pengemudi ojek online serta layak digunakan dalam penelitian maupun pemantauan praktis di lapangan.

Evaluasi reliabilitas pada penelitian ini berfokus pada konsistensi internal, sehingga temuan reliabilitas perlu dipahami sebagai estimasi awal presisi internal dan belum merefleksikan keandalan instrumen lintas waktu. Oleh karena itu, penelitian selanjutnya disarankan untuk melakukan uji *test-retest*. Selain itu, studi ini belum melaksanakan uji *multigroup confirmatory factor analysis* untuk menilai invariansi pengukuran antargender. Karena itu, disarankan menguji invariansi serta memantau perubahan indeks kecocokan ( $\Delta$ CFI,  $\Delta$ TLI,  $\Delta$ RMSEA). Langkah ini penting agar perbandingan skor antar gender sah dan bebas bias pengukuran.

## 5. Referensi

- Bennett, R. J., & Sandra, L. R. (2000). Development of a Measure of Workplace Deviance. *Journal of Applied Psychology*, 85(3), 349-360.
- Bennett, R. J., Shelly, M., & Lauren, L. (2018). Workplace Deviance. *Oxford Research Encyclopedia of Business and Management*. <https://doi.org/10.1093/acrefore/9780190224851.013.111>
- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive Sampling: Complex or Simple? Research Case Examples. *Journal of Research in Nursing: JRN*, 25 (8): 652–661. <https://doi.org/10.1177/1744987120927206>
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86(3), 425–445.
- Detik. (2023). Viral Video Driver Ojol Aniaya Karyawati di Tangerang Selatan. Diakses dari: <https://www.detik.com>
- Dunn, T. J., Thom, B., Vivienne, B. (2013). From Alpha to Omega: A Practical Solution to the Pervasive Problem of Internal Consistency Estimation. *British Journal of Psychology*, 105(3), 399-412.
- Goey, D. H. H. (2018). Pengaruh employee satisfaction terhadap customer satisfaction dan customer loyalty restoran keluarga di surabaya. *Jurnal ilmiah mahasiswa universitas surabaya*. 7(1), 1752-1771.
- Gunawan, A. (2016). Pengaruh faktor intrinsik dan ekstrinsik terhadap usage intention melalui peran customer satisfaction pada layanan ponsel di surabaya. *Jurnal ilmiah mahasiswa universitas surabaya*. 5(2), 505-523.
- Hair, J. F., William, C. B., Barry, J. B., & Rolph, E. A. (2010). *Multivariate Data Analysis*. USA: Prentice hall.
- Hooper, D., Joseph, C., & Michael, R. M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6 (1): 53-60.
- Jex, S. M., & Britt, T. W. (2014). *Organizational psychology: A scientist-practitioner approach* (3rd ed.). Wiley
- Kompas. (2023). Pengemudi Ojek Online Diduga Tonjok Pelanggan di Pondok Indah. Diakses dari: <https://www.kompas.com>
- Pangku, F., & Qomaruddin, H. (2022). Tinjauan Hukum terhadap Manipulasi Informasi Elektronik oleh Pengemudi Ojek Online. *Bhirawa Law Journal*, 3(2), 102-109.
- Putri, M. A., Dewi, M. P., & Fatimah, F. F. (2022). Kematangan emosi dan perilaku agresi pada pengemudi ojek online. *Arjwa: Jurnal Psikologi*, 1(1), 12-24.
- Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal*, 38(2), 555–572.
- Searle, R. H. (2022). Counterproductive Work Behavior. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of Employee Selection* (2nd ed., pp. 589–606). Routledge.
- Spector, P. E., & Fox, S. (2005). The Stressor–Emotion Model of Counterproductive Work Behavior. In S. Fox & P. E. Spector (Eds.), *Counterproductive Work Behavior: Investigations of Actors and Targets* (pp. 151–174). American Psychological Association.
- Thompson, B. (1951). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. Washington DC: American Psychological Association.
- Wetzelhutter, D. (2020). Scale-Sensitive Response Behavior? Consequences of Offering Versus Omitting a “Don’t Know” Option and/or a Middle Category. *Survey Practice*, 13(1), 1-13.



## PENGEMBANGAN SKALA KESIAPAN MENGHADAPI PERUBAHAN PADA MAHASISWA BARU DI INDONESIA DAN UJI INVARIANSI: PENDEKATAN MIMIC MODEL

Rizki Nur Pratama<sup>1</sup>, Muchamad Salamun Abdul Rahman<sup>2</sup>, Nova Mukhlina<sup>3</sup>, Muthmainnah Imtiyaz<sup>4</sup>, Rifda Amira Mulika<sup>5</sup>, Ricky Budhi Hartanto<sup>6</sup>, Ismi Uswatun Khasanah<sup>7</sup>

<sup>1,2,3,4,5,6,7</sup> Fakultas Psikologi, Universitas Gajah Mada, Yogyakarta, Indonesia

Email: [rizkinurpratama@mail.ugm.ac.id](mailto:rizkinurpratama@mail.ugm.ac.id)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
2 September 2025	8 Oktober 2025	18 Oktober 2025	30 Oktober 2025

### Abstract

*This study aims to modify the Readiness for Change scale, which was previously developed for organizational settings, so that it can be applied to first-year university students in the higher education context. Three dimensions of readiness for change were developed based on the results of a preliminary study involving first-year students as sources for the indicators. The participants consisted of 400 first-year students aged 17–19 years from D1, D2, D3, D4, and S1 educational levels. The instrument testing process involved several stages, including normality testing, model testing using the Confirmatory Factor Analysis (CFA) and Multiple Indicators Multiple Causes (MIMIC) approaches, and reliability analysis to assess internal consistency. The findings indicate that the Readiness for Change Scale for first-year students in Indonesia meets the model fit criteria based on the values of CFI, TLI, RMSEA, and SRMR. The reliability test also shows that the scale demonstrates high internal consistency and is reliable. These results suggest that the developed scale fulfills the requirements for validity and reliability. This study contributes to the field of educational and developmental psychology by providing a valid and reliable instrument to assess students' readiness for change during the transition to university life and serves as a basis for educational institutions to design interventions that effectively support new students' adaptation processes.*

*Keywords: readiness for change, first-year university students, developing scale, CFA, MIMIC Model.*

### Abstrak

Penelitian ini bertujuan untuk memodifikasi skala *Readiness for Change* atau kesiapan menghadapi perubahan yang sebelumnya hanya digunakan dalam konteks organisasi, agar sesuai dengan konteks mahasiswa baru di perguruan tinggi. Tiga dimensi kesiapan menghadapi perubahan dikembangkan berdasarkan hasil studi pendahuluan yang melibatkan mahasiswa baru sebagai sumber indikator. Partisipan penelitian berjumlah 400 mahasiswa baru berusia 17–19 tahun dari jenjang pendidikan D1, D2, D3, D4, dan S1. Pengujian alat ukur dilakukan melalui beberapa tahap, yaitu uji normalitas, analisis model menggunakan pendekatan *Confirmatory Factor Analysis* (CFA) dan *Multiple Indicators Multiple Causes* (MIMIC), serta uji reliabilitas untuk menilai konsistensi internal skala. Hasil analisis menunjukkan bahwa Skala Kesiapan Menghadapi Perubahan pada mahasiswa baru di Indonesia memenuhi kriteria *model fit* berdasarkan nilai CFI, TLI, RMSEA, dan SRMR. Uji reliabilitas juga memperlihatkan bahwa skala ini memiliki konsistensi yang tinggi dan dapat diandalkan (*reliable*). Temuan ini menunjukkan bahwa skala yang dikembangkan telah memenuhi kriteria validitas dan reliabilitas yang baik. Penelitian ini memberikan kontribusi pada bidang psikologi pendidikan dan perkembangan dengan menyediakan alat ukur yang valid dan reliabel untuk menilai kesiapan mahasiswa baru dalam menghadapi perubahan selama masa transisi ke perguruan tinggi, serta menjadi dasar bagi lembaga pendidikan dalam merancang intervensi yang mendukung adaptasi mahasiswa baru secara lebih efektif.

*Kata kunci: kesiapan menghadapi perubahan, mahasiswa baru, pengembangan skala, CFA, MIMIC Model*

### 1. Pendahuluan

Mahasiswa baru berada dalam fase remaja akhir yang memiliki tugas perkembangan untuk menyesuaikan diri dalam berperan di lingkungan sekitarnya (Hurlock, 1990). Pada fase ini, tugas perkembangan remaja melibatkan adaptasi sosial, salah satunya yaitu pergeseran posisi dari senior di Sekolah Menengah Atas (SMA) menjadi junior di

Perguruan Tinggi, yang dikenal sebagai fenomena *top-dog* (Santrock, 2007). Mahasiswa baru tergolong ke dalam *emerging adulthood* yang mana pada fase ini mereka akan dihadapkan dengan pengalaman baru di masa perkuliahan sehingga eksplorasi diri dan kebebasan menjadi tugas perkembangan utama disini (Arnett 2000).

Tugas-tugas baru yang dihadapi mahasiswa menjadi krusial bagi perkembangan psikologisnya. Keberhasilan dalam melalui fase ini akan mengantarkan mahasiswa pada proses penyesuaian (*adjustment*) yang positif dan pada akhirnya berdampak pada performa akademik, kesehatan mental, serta retensi studi yang lebih baik dibandingkan dengan mereka yang mengalami hambatan dalam pemenuhan tugas perkembangannya (*maladjustment*) (Daniels et al., 2019). Sebaliknya, apabila proses transisi ini tidak dikelola dengan baik, mahasiswa baru berisiko mengalami gangguan psikologis seperti kecemasan, depresi, hingga munculnya pemikiran bunuh diri (Duffy et al., 2020).

Untuk mencegah terjadinya *maladjustment*, diperlukan sejumlah faktor protektif yang dapat membantu mahasiswa baru beradaptasi secara optimal. Faktor-faktor tersebut antara lain persepsi dukungan dari teman sebaya melalui hubungan persahabatan yang bermakna, dukungan akademik, motivasi, serta perasaan kesamaan nasib dalam menghadapi tantangan di awal perkuliahan (Cameron & Rideout, 2022). Selain itu, perasaan memiliki dan menjadi bagian dari institusi universitas juga berperan penting dalam mendorong mahasiswa baru menyesuaikan diri dengan budaya dan peran barunya di lingkungan perguruan tinggi.

Studi pendahuluan juga dilakukan oleh peneliti pada September 2024 kepada 10 orang partisipan, dan hasilnya menunjukkan bahwa pertemanan merupakan faktor penting yang mendukung kesiapan menghadapi perubahan, selain dorongan untuk mengeksplorasi diri dan merasakan kebebasan selama menjalani perkuliahan. Temuan ini menguatkan pandangan bahwa dukungan sosial dari teman sebaya memiliki peran sentral dalam menumbuhkan rasa percaya diri mahasiswa baru untuk beradaptasi dengan berbagai tuntutan akademik dan sosial di lingkungan kampus. Dengan demikian, faktor-faktor protektif seperti pertemanan, eksplorasi diri, dan kebutuhan akan kebebasan dapat menjadi pondasi psikologis yang membantu mahasiswa baru menghadapi masa transisi ke dunia perkuliahan secara lebih positif dan adaptif.

Dalam menghadapi perubahan peran dan situasi baru, mahasiswa memerlukan kesiapan diri. Kesiapan menghadapi perubahan (*readiness for change*) merupakan sikap yang dipengaruhi oleh beberapa faktor, yaitu faktor konten yang berkaitan dengan jenis perubahan yang terjadi; faktor proses yang berhubungan dengan cara perubahan tersebut dilaksanakan; faktor konteks yang mengacu pada situasi atau kondisi tempat perubahan berlangsung; serta faktor individu yang mencakup karakteristik personal dari orang yang diminta melakukan perubahan. Kesiapan ini memengaruhi individu secara kognitif dan emosional, sehingga menentukan sejauh mana seseorang siap atau tidak dalam menghadapi perubahan (Holt et al., 2007).

Sejauh ini, topik kesiapan menghadapi perubahan lebih banyak diteliti dalam konteks organisasi. Salah satu alat ukur yang paling dikenal adalah *Readiness for Change* yang dikembangkan oleh Holt et al. (2007). Skala ini terdiri atas empat dimensi, yaitu *appropriateness* atau persepsi individu bahwa perubahan yang terjadi memang dibutuhkan oleh organisasi; *management support*, yakni dukungan dari pihak manajemen terhadap perubahan yang dilakukan; *change efficacy* atau keyakinan diri individu bahwa mereka mampu menjalani perubahan; serta *personal valence*, yaitu sejauh mana individu merasa bahwa perubahan tersebut memberikan manfaat bagi dirinya. Skala *Readiness for Change* merupakan skala multidimensi yang terdiri atas 25 butir dengan lima pilihan jawaban, yaitu Sangat Tidak Relevan (STR), Tidak Relevan (TR), Netral (N), Relevan (R), dan Sangat Relevan (SR). Koefisien reliabilitas alpha skala ini masih berada pada kategori dapat diterima, masing-masing sebesar 0,80 untuk dimensi *appropriateness*, 0,79 untuk *management support*, 0,79 untuk *change efficacy*, dan 0,65 untuk *personal valence*.

Skala *Readiness for Change* (RFC) yang dikembangkan oleh Holt et al. (2007) awalnya dirancang untuk menilai kesiapan individu dalam menghadapi perubahan di lingkungan organisasi. Di tingkat internasional, penelitian seperti yang dilakukan oleh Kondakci dan Zayim (2013) menunjukkan bahwa kesiapan terhadap perubahan dipengaruhi oleh dimensi kognitif dan emosional dalam konteks pendidikan, dan instrumen mereka terbukti valid pada tenaga pendidik. Mengacu pada temuan tersebut, penelitian di Indonesia mulai mengadaptasi RFC ke berbagai konteks. Erlyani, Ardi, dan Suhariadi (2024) berhasil memvalidasi skala ini di pendidikan tinggi dan menemukan empat dimensi utama, yaitu *appropriateness*, *management support*, *change efficacy*, dan *personal valence*. Selain itu, Munawaroh dan Meiyanto (2017) serta Yani dan Soehardi (2017) menunjukkan fleksibilitas RFC dalam konteks organisasi di Indonesia. Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada lingkungan kerja, sehingga pengembangan skala yang relevan dengan konteks mahasiswa baru menjadi penting untuk memahami kesiapan mereka menghadapi perubahan di masa transisi menuju perguruan tinggi.

Di Indonesia, pengembangan skala kesiapan menghadapi perubahan dalam konteks pendidikan masih sangat terbatas, terutama pada kelompok mahasiswa baru yang tengah mengalami masa transisi dari sekolah menengah ke perguruan tinggi. Fase ini menuntut kesiapan psikologis yang tinggi untuk beradaptasi dengan perubahan akademik, sosial, dan peran diri. Oleh karena itu, pengembangan *Skala Kesiapan Menghadapi Perubahan pada Mahasiswa Baru* menjadi penting untuk menghasilkan alat ukur yang sesuai dengan konteks budaya dan perkembangan mahasiswa di Indonesia. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk memodifikasi skala *Readiness for Change* dari Holt et al., (2007) ke dalam konteks mahasiswa baru dengan menyesuaikan dimensi dan

indikatornya terhadap karakteristik responden Indonesia. Hipotesis penelitian ini adalah bahwa skala tersebut terbukti valid dan bebas dari bias pengukuran apabila melibatkan kovariat dalam model analisis faktor.

Penelitian ini memiliki kebaruan dan kontribusi pada kajian psikologi pendidikan dan perkembangan, karena belum banyak penelitian yang menggunakan skala untuk mengukur kesiapan mahasiswa baru dalam menghadapi berbagai perubahan yang terjadi selama masa transisi ke perguruan tinggi, baik di Perguruan Tinggi Negeri maupun Swasta. Selain itu, penelitian serupa yang melibatkan analisis variabel kovariat masih sangat terbatas. Oleh karena itu, peneliti menggunakan *Multiple Indicators Multiple Causes* (MIMIC) model sebagai pendekatan analisis. Model MIMIC dipilih karena mampu menguji pengaruh variabel kovariat (seperti usia, jenis kelamin, dan jenis perguruan tinggi) terhadap konstruk laten, sekaligus mendeteksi adanya potensi bias pengukuran atau ketidaksetaraan (*measurement invariance*) pada skala yang akan dikembangkan (Lee et al., 2013). Dengan demikian, penggunaan MIMIC memungkinkan peneliti untuk memastikan bahwa skala kesiapan menghadapi perubahan yang dikembangkan bersifat adil dan konsisten dalam mengukur konstruk pada berbagai kelompok partisipan.

## 1. Metode Penelitian

### Partisipan dan Prosedur Pengambilan Data

Partisipan penelitian ini adalah mahasiswa baru pada usia 17-19 tahun pada level pendidikan D1, D2, D3, D4, dan S1. Kriteria usia diperoleh berdasarkan definisi dewasa awal menurut Arnett (2000). Sebelum melakukan pengambilan data, peneliti mengajukan persetujuan *ethical clearance* dan peneliti sudah mendapatkan izin *ethical clearance* dari Komite Etika Penelitian Fakultas Psikologi Universitas Gadjah Mada yang diterbitkan pada tanggal 4 Oktober 2024, dengan nomor surat: 8874/UN1/FPSi.1.3/SD/PT.01.04/2024.

Sejumlah 400 partisipan telah berpartisipasi dalam penelitian ini. Pengambilan data dilakukan secara daring menggunakan *google forms* ditujukan pada individu dewasa awal sesuai kriteia yang ada di Indonesia, disebarkan melalui media sosial. Pada lembar awal formulir pengisian diberikan penjelasan penelitian dan lembar *informed consent*, responden mengisikikan kesediaan berpartisipasi, dan mengisikikan data demografi. Setelah itu, responden mengisi skala penelitian sejumlah 45 aitem. Persebaran partisipan berdasarkan data demografi dalam dilihat pada Tabel 1.

**Tabel 1. Data Demografi Partisipan (N=400)**

Demografi	Jumlah	Persentase (%)	Kumulatif (%)
<b>1. Usia</b>			
a. 17 tahun	37	9.3%	9.3%
b. 18 tahun	260	65%	74.3%
c. 19 tahun	103	25.7%	100%
<b>Jumlah Partisipan</b>	<b>400</b>		
<b>2. Jenis Kelamin</b>			
a. Laki-laki	98	24.5%	24.5%
b. Perempuan	302	75.5%	100%
<b>Jumlah Partisipan</b>	<b>400</b>		
<b>3. Jenis Perguruan Tinggi</b>			
a. Perguruan Tinggi Negeri (PTN)	347	86.8%	86.8%
b. Perguruan Tinggi Swasta (PTS)	53	13.2%	100%
<b>Jumlah Partisipan</b>	<b>400</b>		

### Analisis Data

Analisis data dilakukan dengan melakukan uji normalitas sebagai uji asumsi, kemudian untuk menguji validitas konstruk dilakukan analisis *Confirmatory Factor Analysis* (CFA) guna menguji seberapa baik variabel yang diukur dapat mewakili faktor atau konstruk yang terbentuk sebelumnya (Hair, Black, & Babin, 2010) atau dapat dikatakan mengkonfirmasi sejauh mana butir aitem mengukur apa yang hendak diukur (Umar & Nisa, 2020).

Reliabilitas atau keajegan skala ini diukur melalui analisis *Cronbach's Alpha*. Skala ini juga akan dianalisis menggunakan uji *Multiple Indicators Multiple Causes* (MIMIC) yang bertujuan untuk melihat bagaimana variabel kovariat (demografi) berpengaruh terhadap struktur model pengukuran yang dianalisis (Widhiarso & Wijatmiko, 2013) Kemudian, seluruh proses analisis dilakukan melalui *software* Mplus 8.3 dan IBM SPSS Statistics 25.

### Tahap Penetapan Indikator dan Aitem

Setelah melakukan tinjauan literatur dan wawancara studi pendahuluan mengenai kesiapan menghadapi perubahan kepada 10 orang responden. Peneliti memutuskan untuk tidak menggunakan dimensi *appropriateness* karena menurut Holt (2007), dimensi ini mengacu pada sejauh mana karyawan merasakan perubahan yang terjadi penting bagi organisasi, sehingga dimensi ini tidak relevan dengan konteks perubahan yang dialami mahasiswa baru. Peneliti juga mengganti dimensi *management support* menjadi *social support* karena *social support* adalah faktor protektif dominan pada kesiapan menghadapi perubahan mahasiswa baru. Sehingga dimensi pada skala modifikasi ini adalah *change efficacy*, *social support*, dan *personal valence* yang selanjutnya diturunkan ke dalam beberapa indikator yang tertera pada Tabel 2.

Setelah menentukan dimensi dan indikator skala, peneliti memulai untuk menulis aitem. Pada tahap uji coba, peneliti menetapkan tiga dimensi dalam alat ukur ini antara lain *change efficacy*, *social support*, dan *personal valence*. Dimensi *social support* terdiri dari dua indikator, *change efficacy* terdiri dari dua indikator, dan *personal valence* terdiri dari tiga indikator. Total aitem di dalam Skala Kesiapan Menghadapi Perubahan untuk Mahasiswa Baru yang peneliti rancang adalah sebanyak 45 butir.

Sebelum melakukan pengambilan data, dilakukan tinjauan ulang skala oleh dua dosen ahli psikometri Fakultas Psikologi UGM. Ditetapkan penilaian pada aitem *favorable* dari rentang angka 1 sampai 5 yang artinya 1 = sangat tidak sesuai dan 5 = sangat sesuai. Sebaliknya pada aitem *unfavorable* rentang dari angka 5 sampai 1 yang artinya 5 = sangat tidak sesuai dan 1 = sesuai. Semakin tinggi skor yang diperoleh maka artinya semakin tinggi menghadapi perubahan pada mahasiswa baru

**Tabel 2. Blueprint Alat Ukur Kesiapan Menghadapi Perubahan pada Mahasiswa Baru**

No	Dimensi	Indikator	Aitem		Total
			Fav	Unfav	
1.	<i>Social Support</i>	Mahasiswa baru mendapatkan dan bisa memanfaatkan fasilitas dan layanan kampus yang tersedia untuk mendukung keberhasilan akademik	4	2	14
		Mahasiswa baru mendapat dukungan dari lingkungan terdekat, seperti keluarga dan teman	4	4	
2.	<i>Change Efficacy</i>	Mahasiswa baru yakin pada kemampuannya dalam menjalankan tugas sesuai peran dalam bidang akademik dan non-akademik.	3	3	12
		Mahasiswa baru yakin pada kemampuannya dalam membangun hubungan yang baik dengan teman sebaya, dosen, dan staf kampus.	3	3	
3.	<i>Personal Valence</i>	Mahasiswa baru yakin bahwa peran barunya dapat memberi kesempatan untuk mengeksplorasi preferensi, potensi, dan citra diri	5	2	25
		Mahasiswa baru yakin bahwa perubahan peran menjadi mahasiswa akan meningkatkan keterampilan dan kemampuan berpikir	3	3	
		Mahasiswa baru yakin dapat mencari peluang untuk perkembangan karier di masa depan	4	2	
<b>Total</b>			<b>26</b>	<b>19</b>	<b>45</b>

## 2. Hasil dan Diskusi

### Uji Normalitas Data

Peneliti melakukan uji normalitas *skewness* dan *kurtosis* sebelum dan sesudah analisis *Confirmatory Factor Analysis* (CFA) untuk memastikan distribusi data bersifat normal. Menurut Kim (2013), nilai ideal *skewness* dan *kurtosis* agar data dikatakan berdistribusi normal berada dalam rentang  $-2$  hingga  $+2$ . Pada tahap awal, uji normalitas dilakukan terhadap 45 aitem hasil penyusunan awal skala, dan seluruh aitem menunjukkan distribusi normal dengan nilai *skewness* dan *kurtosis* dalam rentang tersebut. Selanjutnya, melalui proses analisis CFA, aitem-aitem yang memiliki nilai *loading factor* rendah ( $<0,50$ ) dan tidak sesuai dengan model teoritis dieliminasi, sehingga diperoleh 26 aitem yang memenuhi kriteria validitas konstruk. Uji normalitas kembali dilakukan terhadap 26 aitem akhir

tersebut, dan hasilnya menunjukkan bahwa data tetap terdistribusi normal dengan nilai *skewness* dan *kurtosis* berada pada rentang -2 hingga +2. Hasil uji normalitas pada model akhir dengan 26 aitem dapat dilihat pada Tabel 3.

**Tabel 3. Uji Normalitas *Skewness* dan *Kurtosis* Skala Akhir**

Aitem	N	Mean	Variance	Skewness	Kurtosis
A1	400	4.06	.789	-.647	-.170
A3	400	3.91	.858	-.647	.080
A4	400	3.50	.952	-.269	-.049
A5	400	3.88	.746	-.428	-.007
A7	400	4.55	.474	-1.538	2.112
A8	400	4.07	1.048	-1.144	.941
A19	400	3.91	1.045	-.663	-.382
A11	400	4.38	.591	-1.288	1.887
A12	400	4.05	.862	-.917	.691
A13	400	4.29	.707	-1.245	1.656
A14	400	3.86	1.310	-.988	.266
A16	400	3.85	.841	-.473	-.298
A22	400	3.95	.822	-.728	.332
A24	400	3.59	1.005	-.340	-.370
A27	400	4.46	.410	-1.070	1.692
A28	400	4.23	.621	-1.039	1.341
A29	400	4.24	.627	-1.136	1.914
A30	400	4.37	.415	-.590	-.349
A32	400	3.70	1.020	-1.063	1.313
A33	400	4.22	.673	-1.083	1.749
A34	400	4.43	.421	-.872	.428
A35	400	4.32	.494	-.713	-.033
A40	400	4.36	.451	-.717	.054
A41	400	4.23	.526	-.945	1.903
A42	400	4.44	.457	-1.042	1.148
A43	400	4.18	.587	-.764	.699

### Hasil Pengujian Model

**Tabel 4. Hasil Pengujian Model**

Model	X <sup>2</sup>	CFI	TLI	RMSEA	SRMR	Diagram	Keterangan
I	9457.530	0.875	0.869	0.053	0.070	-	CFA 45 aitem
II	840.498*	0.914	0.905	0.068	0.063	Gambar 1.	CFA 26 aitem
III	911.575*	0.918	0.911	0.060	0.129	Gambar 2.	MIMIC

Keterangan: *chi-square* ( $\chi^2$ ); *Comparative Fit Index* (CFI); *Tucker-Lewis Index* (TLI); *Root Mean Square Error of Approximation* (RMSEA); *Standardized Root Mean Square Residual* (SRMR).

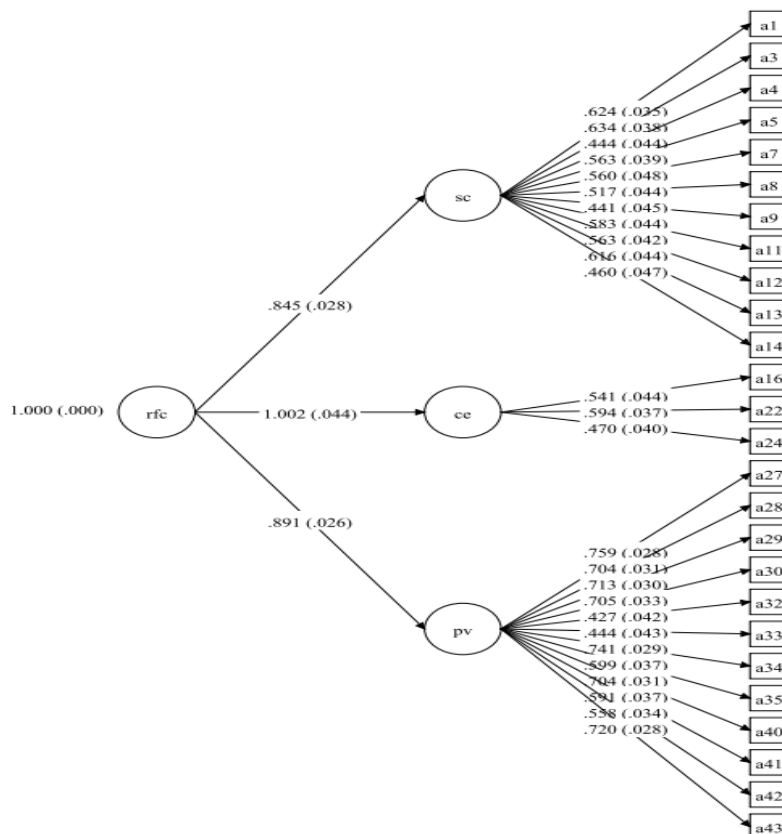
Peneliti melakukan beberapa pengujian model, dan hasilnya disajikan pada Tabel 4. Pengujian model dilakukan sebanyak tiga kali. Model I dan Model II digunakan untuk menguji *model fit* terhadap aitem awal sebanyak 45 aitem dan aitem hasil reduksi sebanyak 26 aitem melalui analisis *CFA*. Proses reduksi dilakukan berdasarkan hasil *CFA* tahap pertama, di mana aitem dengan nilai *factor loading* di bawah 0,40 serta aitem yang menimbulkan *modification indices* tinggi dieliminasi agar model lebih sesuai dengan struktur teoretis. Dengan demikian, 26 aitem yang digunakan dalam Model II merupakan aitem yang memenuhi kriteria validitas konstruk dan tetap merepresentasikan setiap dimensi skala secara proporsional. Selanjutnya, Model III digunakan untuk menguji invariansi skala terhadap variabel demografis menggunakan *Multiple Indicators Multiple Causes* (MIMIC) model, dengan kovariat usia, jenis kelamin, dan jenis perguruan tinggi. Penjelasan mengenai alur analisis masing-masing model disajikan pada subbagian berikut.

### Model I dan II: *Confirmatory Factor Analysis* (CFA)

Pada analisis *CFA*, kriteria *goodness of fit* yang digunakan mencakup nilai *Chi-Square* ( $\chi^2$ ), *Comparative Fit Index* (CFI), *Tucker-Lewis Index* (TLI), *Root Mean Square Error of Approximation* (RMSEA), dan *Standardized Root Mean Square Residual* (SRMR), dengan batasan nilai  $\chi^2$  ( $p > 0,05$ ), CFI dan TLI > 0,90, SRMR < 0,08 (Hu & Bentler, 1999), serta RMSEA < 0,08 (Wang & Wang, 2020). Nilai *Chi-Square model fit* yang diperoleh dari *output*

Mplus dilaporkan bersama *p-value* untuk memastikan kesesuaian dengan kriteria standar tersebut. Pada Model I (lihat Tabel 4), hasil pengujian belum memenuhi nilai yang disyaratkan pada  $\chi^2$ , *CFI*, dan *TLI*, sehingga dilakukan modifikasi dengan menggugurkan aitem yang memiliki *factor loading* < 0,40 dan menyeleksi aitem berdasarkan *modification indices* (*MI*) (McNeish & Wolf, 2023). Proses ini juga mempertimbangkan keseimbangan representasi antar dimensi (*social support*, *change efficacy*, dan *personal valence*).

Berdasarkan ketiga pertimbangan tersebut, tersisa 26 aitem dari 45 aitem awal. Sebanyak 19 aitem yang digugurkan antara lain: aitem 2, 6, dan 20 pada dimensi *social support*; aitem 15, 17, 18, 19, 20, 21, 23, 25, dan 26 pada dimensi *change efficacy*; serta aitem 31, 36, 37, 38, 39, 44, dan 45 pada dimensi *personal valence*. Model hasil modifikasi ini kemudian dianalisis kembali dan menunjukkan *fit indices* yang memenuhi kriteria (lihat Tabel 4, Model II). Meskipun nilai  $\chi^2$  masih signifikan ( $p < 0,001$ ), hal ini dapat diabaikan karena *Chi-Square* sangat sensitif terhadap ukuran sampel dan kompleksitas model (Hair et al., 2019). Oleh karena itu, interpretasi kesesuaian model dalam penelitian ini lebih difokuskan pada indeks relatif seperti *CFI*, *TLI*, *RMSEA*, dan *SRMR* yang lebih stabil

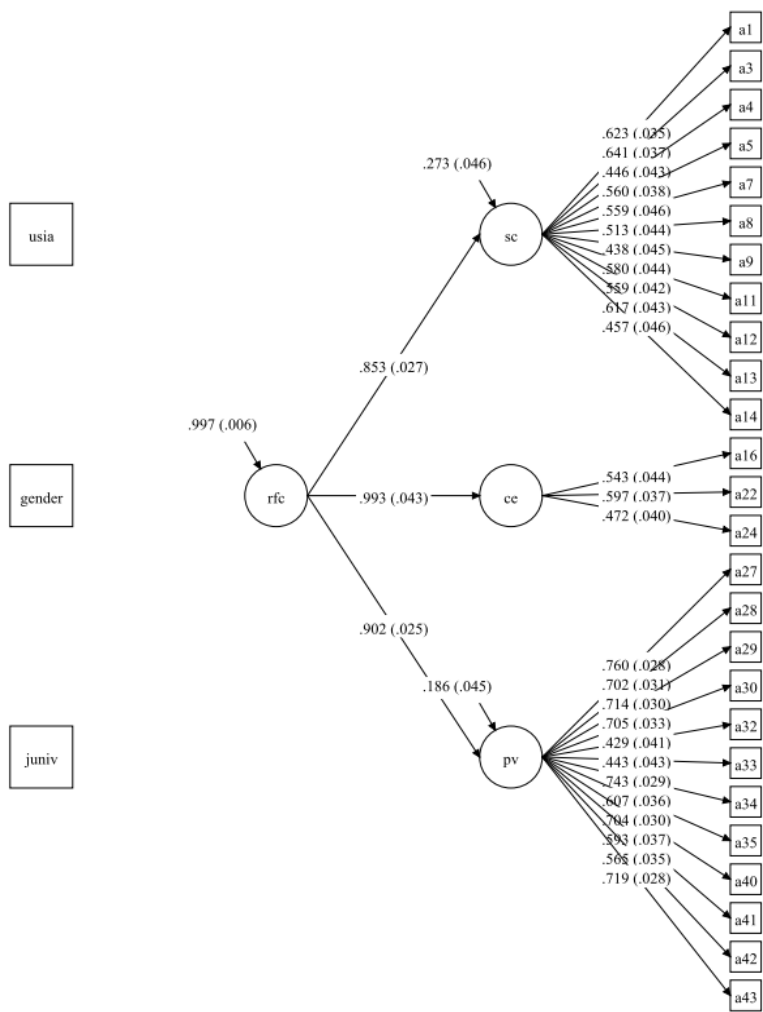


Keterangan: Analisis CFA 26 aitem yang memenuhi kriteria validitas konstruk terhadap variasi jumlah sampel

**Gambar 1. Model II Analisis CFA 26 Aitem**

### Model III: Multiple Indicators Multiple Causes (MIMIC)

Setelah analisis *CFA* menghasilkan model yang fit, peneliti melanjutkan dengan analisis *MIMIC* untuk menguji invariansi alat ukur terhadap kovariat demografis. Kovariat yang melibatkan dalam model adalah usia, jenis kelamin, dan jenis perguruan tinggi. Analisis dilakukan dengan menambahkan ketiga kovariat tersebut ke dalam model *CFA* untuk melihat pengaruh langsungnya terhadap konstruk laten kesiapan menghadapi perubahan. Hasil analisis menunjukkan bahwa setelah kovariat dimasukkan, nilai indeks *fit* model mengalami sedikit penurunan tetapi tetap berada dalam kategori yang dapat diterima (lihat Tabel 4, Model III). Berdasarkan hasil *path coefficient*, ketiga kovariat yaitu usia, jenis kelamin, dan jenis perguruan tinggi tidak menunjukkan hubungan yang signifikan dengan konstruk laten ( $p > 0,05$ ) (lihat Gambar 2). Dengan demikian, tidak ditemukan perbedaan skor kesiapan menghadapi perubahan berdasarkan ketiga kovariat demografis yang diuji.



Gambar 2. Model III Signifikansi Ketiga Kovariats

### Hasil Uji Reliabilitas

Peneliti juga melakukan pengujian reliabilitas data. Berdasarkan 26 aitem yang telah memenuhi model *fit* diperoleh nilai koefisien reliabilitas alpha ( $\alpha$ ) sebesar 0.872 (Lihat pada Tabel 5). Menurut Nunnally dan Bernstein (1994) nilai standar *Cronbach's Alpha* agar dikatakan reliabel harus sebesar  $\alpha > 0.7$ . Sehingga berdasarkan pada hasil uji reliabilitas data yang dilakukan, skala menghadapi perubahan pada mahasiswa baru di Indonesia ini dapat dikatakan reliabel.

Tabel 5. Hasil Reliabilitas *Alpha* ( $\alpha$ )

<i>Reliability Statistics</i>	
<i>Cronbach's Alpha</i>	<i>N of Items</i>
.872	26

### 4. Diskusi

Sesuai dengan tujuan penelitian ini, yaitu mengembangkan *Skala Readiness for Change (Skala Kesiapan Menghadapi Perubahan)* yang dikembangkan oleh Holt et al. (2007), data dianalisis melalui dua tahap, yaitu *Confirmatory Factor Analysis (CFA)* dan model *Multiple Indicators Multiple Causes (MIMIC)*. Hasil penelitian menunjukkan bahwa *Skala Kesiapan Menghadapi Perubahan pada Mahasiswa Baru* di Indonesia, dengan hasil akhir sebanyak 26 aitem yang telah diverifikasi melalui *CFA*, memiliki indeks kecocokan model yang baik sesuai dengan kriteria yang ditentukan.

Pada pengujian validitas konstruk, peneliti mengevaluasi sejauh mana alat ukur ini menilai konstruk yang seharusnya diukur menggunakan *CFA* terhadap beberapa model. Hasil analisis menunjukkan model yang fit dengan

nilai  $CFI = 0,914$ ,  $TLI = 0,905$ ,  $RMSEA = 0,068$ , dan  $SRMR = 0,063$ , dengan mempertimbangkan standar nilai yang telah ditetapkan (Hu & Bentler, 1999; Wang & Wang, 2020). Proses seleksi aitem dilakukan dengan mempertimbangkan nilai *factor loading* dan *modification indices*. Peneliti menggugurkan aitem dengan *factor loading* di bawah 0,40 (McNeish & Wolf, 2023) serta meninjau aitem yang terindikasi melalui *modification indices*, karena perubahan pada aitem tersebut berpotensi meningkatkan kesesuaian model. Dengan memperhatikan keseimbangan representasi aitem dari tiap dimensi, diperoleh 11 aitem pada dimensi *social support* (a1, a3, a4, a5, a7, a8, a9, a11, a12, a13, a14), 3 aitem pada dimensi *change efficacy* (a16, a22, a24), dan 12 aitem pada dimensi *personal valence* (a27, a28, a29, a30, a32, a33, a34, a35, a40, a41, a42, a43).

Model *CFA* yang telah fit kemudian digunakan untuk analisis *MIMIC* dengan menambahkan kovariat demografis, yaitu usia, jenis kelamin, dan jenis perguruan tinggi. Hasil analisis menunjukkan bahwa skala bersifat invarian terhadap ketiga kovariat, yang berarti tidak terdapat perbedaan signifikan pada pengukuran kesiapan menghadapi perubahan berdasarkan faktor demografis tersebut ( $p > 0,05$ ). Menurut Widhiarso (2012), sebuah skala dikatakan invarian apabila menghasilkan informasi yang konsisten dan tidak terpengaruh oleh karakteristik partisipan yang diukur. Hasil ini sejalan dengan penelitian Haque (2008) pada konteks organisasi yang menemukan bahwa *readiness for change* dalam proses belajar tidak menunjukkan perbedaan berdasarkan usia, tingkat pendidikan, maupun level pekerjaan.

Selanjutnya, skala kesiapan menghadapi perubahan pada mahasiswa baru di Indonesia menunjukkan reliabilitas yang baik dengan koefisien *Cronbach's Alpha* sebesar 0,872 ( $\alpha > 0,70$ ) (Nunnally & Bernstein, 1994). Reliabilitas merupakan salah satu kriteria utama kualitas instrumen yang mencerminkan stabilitas, konsistensi internal, dan kesetaraan ukuran (Souza, Alexandre, & Guirardello, 2017). Dengan demikian, dapat disimpulkan bahwa skala ini memiliki stabilitas dan konsistensi yang memadai serta secara empiris valid dalam mengukur kesiapan menghadapi perubahan pada konteks budaya Indonesia.

Namun demikian, penelitian ini memiliki beberapa keterbatasan. Pertama, sampel yang digunakan terbatas pada mahasiswa usia dewasa awal sehingga hasilnya mungkin tidak dapat digeneralisasikan pada kelompok usia perkembangan lain. Kedua, ketimpangan jumlah partisipan antara Perguruan Tinggi Negeri (PTN) dan Perguruan Tinggi Swasta (PTS) cukup besar, di mana jumlah partisipan PTN tiga kali lebih banyak dibandingkan PTS. Penelitian selanjutnya diharapkan dapat menggunakan proporsi sampel yang lebih seimbang serta melibatkan rentang usia partisipan yang lebih beragam agar hasil penelitian lebih representatif bagi seluruh kelompok perkembangan.

## 5. Kesimpulan

Penelitian ini menunjukkan bahwa Skala *Kesiapan Menghadapi Perubahan Pada Mahasiswa Baru* di Indonesia memiliki struktur psikometrik yang baik, yang dibuktikan melalui model *fit* yang memenuhi kriteria serta nilai reliabilitas yang memuaskan, sehingga dapat dinyatakan valid dan reliabel dalam mengukur kesiapan mahasiswa baru menghadapi perubahan pada konteks budaya Indonesia. Secara teoretis, hasil penelitian ini mengindikasikan bahwa kovariat demografis berupa usia, jenis kelamin, dan jenis perguruan tinggi bersifat invarian terhadap konstruk kesiapan menghadapi perubahan, sehingga perbedaan karakteristik demografis tidak memengaruhi hasil pengukuran. Secara praktis, temuan ini diharapkan dapat menjadi dasar bagi Perguruan Tinggi Negeri maupun Swasta dalam mengkaji dan mengembangkan kebijakan, fasilitas, serta program pendukung yang relevan untuk membantu mahasiswa baru beradaptasi dan meningkatkan kesiapan mereka dalam menghadapi dunia perkuliahan.

## 6. Referensi

- Arnett, J. J. (2000). *Emerging adulthood: A theory of development from the late teens through the twenties*. *American Psychologist*, 55(5), 469–480. <https://doi.org/10.1037/0003-066X.55.5.469>
- Cameron, R. B., & Rideout, C. A. (2022). 'It's been a challenge finding new ways to learn': first-year students' perceptions of adapting to learning in a university environment. *Studies in Higher Education*, 47(3), 668-682.
- Daniels, V. Z., Davids, E. L., & Roman, N. V. (2019). The role of family structure and parenting in first-year university adjustment. *South African Journal of Psychology*, 49(3), 446–459. <https://doi.org/10.1177/0081246319839695>
- Duffy, A., Keown-Stoneman, C., Goodday, S., Horrocks, J., Lowe, M., King, N., Pickett, W., McNevin, S. H., Cunningham, S., Rivera, D., Bisdounis, L., Bowie, C. R., Harkness, K., & Saunders, K. E. A. (2020). Predictors of mental health and academic outcomes in first-year university students: Identifying prevention and early-intervention targets. *BJPsych Open*, 6(3), e46. <https://doi.org/10.1192/bjo.2020.2>
- Eryani, N., Ardi, R., & Suhariadi, F. (2024). Readiness for change scale in higher education: Adaptation and validity of the indonesia version. *IJORER: International Journal of Recent Educational Research*, 5(1), 140-156.
- Hair, J. F., Black, W. C., & Babin, B. J. (2010). *Multivariate data analysis: Global edition* (7th ed.). Pearson Education.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage. <https://www.cengage.com/highered>

- Haque, M. M. (2008). *A study of the relationship between the learning organization and organizational readiness for change* [Doctoral dissertation, Pepperdine University]. ProQuest Dissertations Publishing.
- Holt, D. T., Armenakis, A. A., Feild, H. S., & Harris, S. G. (2007). Readiness for organizational change: The systematic development of a scale. *The Journal of Applied Behavioral Science*, 43(2), 232–255. <https://doi.org/10.1177/0021886306295295>
- Hu, L. T., & 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>
- Hurlock, E. B. (1990). *Psikologi perkembangan: Suatu pendekatan sepanjang rentang kehidupan* (D. Meitasari Tjandrasa, Trans.). Erlangga. (Karya asli diterbitkan 1980).
- Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, 38(1), 52–54. <https://doi.org/10.5395/rde.2013.38.1.52>
- Kondakçı, Y., & Zayim K., M. (2013). Development and validation of readiness for change scale. *İlköğretim Online*, pp. 23–35.
- Lee, N., Cadogan, J. W., & Chamberlain, L. (2013). The MIMIC model and formative variables: Problems and solutions. *AMS Review*, 3(1), 3–17. <https://doi.org/10.1007/s13162-013-0033-1>
- McNeish, D., & Wolf, M. G. (2023). Dynamic fit index cutoffs for confirmatory factor analysis models. *Psychological Methods*, 28(1), 61–88. <https://doi.org/10.1037/met000059>
- Munawaroh, L., & Meiyanto, I. S. (2017). Peranan *psychological capital* terhadap kesiapan individu untuk berubah yang dimoderatori oleh persepsi dukungan organisasi. *Jurnal Psikologi*, 44(3), 198-210.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Santrock, J. W. (2007). *Psikologi pendidikan* (Edisi ke-2). Kencana.
- Souza, A. C. D., Alexandre, N. M. C., & Guirardello, E. D. B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, 26(3), 649–659. <https://doi.org/10.5123/S1679-49742017000300022>
- Umar, J., & Nisa, Y. F. (2020). Uji validitas konstruk dengan CFA dan pelaporannya. *JP3I: Jurnal Pengukuran Psikologi dan Pendidikan Indonesia*, 9(2), 1–11. <https://doi.org/10.15408/jp3i.v9i2.17912>
- Wang, J., & Wang, X. (2019). *Structural equation modeling: Applications using Mplus*. John Wiley & Sons.
- Widhiarso, W. (2012). Penerapan model MIMIC untuk menguji konsistensi hasil pengukuran melalui skala. *Journal of Education and Learning (EduLearn)*, 6(2), 109–118. <https://doi.org/10.11591/edulearn.v6i2.166>
- Yani, E. I., & Soehardi, (2017). Pengaruh *transformational leadership* dan *employee engagement* terhadap *readiness for change* pada kelembagaan persandian di lingkungan pemerintah daerah. *Jurnal Ilmiah Manajemen Ubhara*, 4, 90-122.



## PENGEMBANGAN *STUDENT'S METACOGNITIVE AWARENESS SCALE* (SMAS) PADA SISWA DI INDONESIA

**Abiyyu Arib Mahiyuddin**

Magister Psikologi, Program Pascasarjana, Universitas Negeri Makassar, Makassar, Indonesia

Email: [abiyyuarib29@gmail.com](mailto:abiyyuarib29@gmail.com)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
5 Juli 2025	23 Oktober 2025	27 Oktober 2025	30 Oktober 2025

### Abstract

*This study developed the Student's Metacognitive Awareness Scale (SMAS) to assess Indonesian high-school students' metacognitive awareness. Data were collected over 13 days (March 24 - April 6, 2025) using non-probability purposive sampling of students aged 15-18 who met predefined inclusion criteria. Of 212 responses, 208 were analyzed after data cleaning to remove careless responses and incomplete entries. From 52 initial items, 24 items remained following content validation (Aiken's V) and item discrimination screening. Psychometric analyses included EFA and CFA. Reliability was estimated via McDonald's Omega. Findings supported SMAS as a valid and reliable instrument for measuring students' metacognitive awareness*

*Keywords: metacognitive, students, validity*

### Abstrak

Studi ini mengembangkan *Student's Metacognitive Awareness Scale* (SMAS) untuk menilai kesadaran metakognitif siswa SMK/SMK di Indonesia. Data dikumpulkan selama 13 hari (24 Maret - 6 April 2025) menggunakan *sampling purposif non-probabilitas* dari siswa berusia 15-18 tahun yang memenuhi kriteria inklusi yang telah ditentukan. Dari 212 respons 208 dianalisis setelah pembersihan data untuk menghilangkan *careless response* dan respons yang tidak lengkap. Dari 52 aitem awal, 24 aitem tersisa setelah validasi isi (Aiken's V) dan penyaringan daya diskriminasi item. Analisis psikometrik meliputi EFA dan CFA. Reliabilitas diperkirakan melalui McDonald's Omega. Temuan mendukung SMAS sebagai instrumen yang valid dan reliabel untuk mengukur kesadaran metakognitif siswa.

Kata kunci: metakognitif, siswa, validitas

### 1. Pendahuluan

Kesadaran metakognitif merupakan seperangkat proses yang memungkinkan pembelajar untuk merencanakan, memantau, dan mengevaluasi aktivitas kognitifnya secara sadar (Schraw & Dennison, 1994). Konsep ini mencakup tidak hanya pemahaman tentang bagaimana pikiran bekerja, tetapi juga kemampuan untuk mengatur proses tersebut agar lebih efektif dalam menghadapi tugas belajar. Dalam konteks pendidikan, kesadaran metakognitif menjadi fondasi utama bagi pembelajaran yang adaptif di mana siswa dapat mengidentifikasi kekuatan dan kelemahan mereka sendiri, sehingga mendorong otonomi belajar yang lebih tinggi (Flavell, 1979). Namun, di tingkat SMA/SMK di Indonesia, observasi lapangan menunjukkan bahwa strategi belajar siswa sering kali bersifat prosedural, menekankan penyelesaian tugas secara mekanis namun kurang pada pemantauan pemahaman dan refleksi diri (Nederhand et al., 2020).

Fenomena ini tidak hanya terbatas pada kurangnya kesadaran diri, tetapi juga mencerminkan pengaruh faktor budaya dan lingkungan kelas di Indonesia, di mana pendekatan pengajaran tradisional masih mendominasi dengan penekanan pada transmisi pengetahuan daripada pengembangan keterampilan metakognitif (Al-Balushi et al., 2022). Studi observasional di sekolah-sekolah negeri dan swasta menunjukkan bahwa siswa SMA/SMK sering mengalami kesulitan dalam memantau kemajuan belajar mereka, yang menyebabkan ketergantungan berlebih pada bimbingan guru atau teman sebaya. Hal ini sejalan dengan temuan internasional pada populasi remaja, di mana *metacognitive awareness* yang rendah berkorelasi dengan performa akademik yang fluktuatif, terutama pada tugas-tugas kompleks seperti pemecahan masalah sains atau membaca kritis (Köksal et al., 2021).

Konsekuensinya, kemampuan belajar mandiri kurang berkembang secara optimal yang tercermin pada fluktuasi capaian akademik dan kebutuhan intervensi diferensial di kelas (Al-Balushi et al., 2022). Hal ini tidak hanya memengaruhi hasil belajar jangka pendek, tetapi juga kesiapan siswa untuk pendidikan tinggi atau dunia kerja, di mana keterampilan metakognitif menjadi prediktor utama kesuksesan. Oleh karena itu, kesenjangan ini mengindikasikan

perlu alat ukur metakognisi yang sensitif terhadap konteks budaya-bahasa lokal agar guru dan konselor dapat menargetkan intervensi secara presisi (Byrne, 2016; Kline, 2016; Pallant, 2020; Gutiérrez de Blume et al., 2024).

Secara teoritis, terdapat dua komponen metakognisi yaitu *knowledge of cognition* (pengetahuan deklaratif tentang apa yang diketahui, prosedural tentang bagaimana melakukannya, dan kondisional tentang kapan serta mengapa menggunakannya) dan *regulation of cognition* (perencanaan, manajemen informasi, pemantauan, penelusuran kesalahan, evaluasi) menjadi kerangka yang banyak digunakan sebagai landasan operasional indikator (Schraw & Dennison, 1994; González-Cabañes et al., 2022). Komponen *knowledge of cognition* membantu siswa memahami sifat tugas belajar, sementara *regulation of cognition* memungkinkan mereka untuk mengatur proses secara dinamis, seperti memantau pemahaman selama membaca atau mengevaluasi efektivitas strategi setelah tes. Kerangka ini telah divalidasi secara luas dalam berbagai konteks pendidikan, termasuk pada siswa remaja, di mana regulasi sering kali lebih menantang daripada pengetahuan dasar (Tuononen et al., 2023). Namun, penerapannya di Indonesia memerlukan adaptasi untuk mengakomodasi nuansa budaya, seperti nilai kolektivisme yang memengaruhi refleksi diri.

Meskipun demikian, literatur mutakhir menekankan bahwa adaptasi instrumen asing ke populasi baru menuntut perhatian serius pada isu kesetaraan pengukuran (*measurement invariance*), bias terjemahan, dan kesesuaian indikator dengan praktik kelas setempat (Putnick & Bornstein, 2016; Marsh et al., 2020; Hair et al., 2019; Li et al., 2024). Contohnya, studi validasi *Metacognitive Awareness Scale* (MAS) di Cina menunjukkan bahwa terjemahan dan penyesuaian budaya esensial untuk menjaga invariansi struktural, di mana faktor regulasi sering kali lebih sensitif terhadap konteks lokal (Li et al., 2024). Tanpa penanganan yang tepat, struktur laten yang diperoleh dapat bergeser dan mengurangi validitas konstruk, seperti yang terlihat dalam adaptasi instrumen di negara berkembang di mana bahasa dan norma sosial memengaruhi respons self-report (Wang et al., 2023). Di Indonesia, tantangan serupa muncul karena campuran bahasa Indonesia-Inggris dalam kurikulum dan pengaruh nilai gotong royong yang mungkin mengurangi penekanan pada refleksi individu.

Instrumen klasik *Metacognitive Awareness Inventory* (MAI) oleh Schraw dan Dennison (1994) memberi fondasi penting dengan struktur dua faktor yang telah dikonfirmasi melalui CFA di berbagai populasi, termasuk mahasiswa Spanyol (González-Cabañes et al., 2022). Namun, temuan pada populasi remaja Indonesia menunjukkan perlunya rekonstruksi indikator sehingga sesuai dengan perkembangan remaja dan nuansa budaya bahasa (Song et al., 2021). Kritik utama terhadap MAI adalah kurangnya spesifisitas domain yang menyebabkan bias pada siswa sekolah menengah yang lebih familiar dengan tugas konkret daripada abstrak (Ondé et al., 2022). Oleh karena itu, penelitian ini mengembangkan *Student's Metacognitive Awareness Scale* (SMAS) berbasis indikator yang relevan secara kultural, dengan sasaran penggunaan praktis dalam asesmen kelas dan evaluasi intervensi berbasis data, seperti pemantauan kemajuan siswa melalui aplikasi digital.

Secara metodologis, pengembangan SMAS mengacu pada praktik psikometri modern yang komprehensif. Proses dimulai dengan validasi isi oleh pakar pendidikan dan psikologi untuk memastikan kesesuaian indikator dengan konteks SMA/SMK Indonesia, diikuti oleh pembersihan data untuk mengurangi *careless response* melalui deteksi pola respons acak (Curran, 2016). Selanjutnya, eksplorasi struktur laten dilakukan melalui *Exploratory Factor Analysis* (EFA) untuk mengidentifikasi faktor utama, diikuti oleh *Confirmatory Factor Analysis* (CFA) dengan pelaporan indeks fit utama seperti  $\chi^2$ , CFI, TLI, dan RMSEA untuk mengonfirmasi model dua faktor (McNeish, 2018; Flora, 2020; Marsh et al., 2020; Ondé et al., 2022). Reliabilitas dievaluasi menggunakan McDonald's Omega untuk akurasi yang lebih baik daripada Cronbach's alpha, sementara penormaan kategori skor memungkinkan interpretasi praktis bagi guru, seperti klasifikasi rendah-sedang-tinggi metacognitive awareness.

Penelitian ini berkontribusi pada penguatan basis bukti lokal mengenai pengukuran metakognisi siswa di Indonesia, di mana studi sebelumnya masih terbatas pada adaptasi instrumen asing tanpa penyesuaian mendalam (Çini et al., 2023). Secara teoretis, SMAS memperkaya diskursus validitas konstruk dengan menunjukkan bagaimana faktor budaya memengaruhi struktur metakognisi, selaras dengan temuan invariansi di adaptasi internasional (Gutiérrez de Blume et al., 2024). Secara praktis, instrumen ini diharapkan memfasilitasi pemetaan kebutuhan strategi belajar siswa, seperti intervensi diferensial untuk siswa berprestasi rendah, dan perancangan program reflektif di kurikulum Merdeka Belajar.

Lebih lanjut, implikasi SMAS meluas ke evaluasi intervensi sekolah, di mana guru dapat menggunakan skor untuk memantau perubahan *metacognitive awareness* pasca-pelatihan, mirip dengan studi longitudinal pada mahasiswa pendidikan (Alt & Raichel, 2020). Di tingkat kebijakan, temuan ini mendukung integrasi asesmen metakognitif dalam standar nasional pendidikan, membantu mengurangi kesenjangan akses di daerah terpencil. Secara keseluruhan, penelitian ini memberikan panduan pengembangan instrumen dengan standar pelaporan yang transparan dan replikabel, memperkaya *best practice* analisis faktor pada instrumen pendidikan (Byrne, 2016; Hair et al., 2019; Flora, 2020).

## 2. Metode Penelitian

### Desain

Desain penelitian merupakan pengembangan instrumen psikometrik berurutan: (a) konstruksi aitem berbasis dua komponen metakognisi dan delapan indikator operasional, (b) validasi isi oleh validator, (c) uji keterbacaan dan uji bahasa, (d) pembersihan data (deteksi dan eksklusi *careless response*), (e) EFA untuk memeriksa struktur laten, (f) CFA untuk menguji kesesuaian model, (g) estimasi reliabilitas (McDonald's Omega), dan (h) penormaan

skor. Pelaporan mengikuti standar indeks fit dan keputusan model yang transparan (Byrne, 2016; Kline, 2016; Marsh et al., 2020; Flora, 2020).

### Sampel dan Teknik Sampling

Partisipan adalah siswa SMA/SMK berusia 15-18 tahun. Teknik sampling non-probabilitas (*purposive*) digunakan dengan kriteria: status siswa aktif, persetujuan sekolah/orangtua, dan pengisian kuesioner lengkap. Dari 212 respons yang terkumpul, 208 dianalisis setelah pembersihan data. Penentuan ukuran sampel mengacu pada praktik EFA/CFA ( $\geq 200$  partisipan atau 5–10 partisipan per aitem) untuk memastikan stabilitas estimasi dan replikabilitas struktur (Kline, 2016; Hair et al., 2019).

Deteksi *careless response* dilakukan menggunakan indikator majemuk: durasi respon sangat singkat, pola jawaban tetap, dan korelasi antara aitem yang semestinya berhubungan. Respons yang memenuhi kriteria eksklusi dihapus sebelum analisis (Curran, 2016).

### Instrumen

SMAS dikonstruksi dari awal. Indikator *knowledge of cognition* meliputi pengetahuan deklaratif, prosedural, dan kondisional; indikator *regulation of cognition* meliputi perencanaan, manajemen informasi, pemantauan, penelusuran kesalahan, dan evaluasi. Aitem dievaluasi oleh 26 pakar (psikologi pendidikan dan psikologi perkembangan). Skala respons menggunakan Likert 5-poin, keseimbangan aitem favorable/unfavorable dipertimbangkan untuk mengurangi bias gaya menjawab. Validasi isi menggunakan Aiken's V, aitem di bawah ambang ketat direvisi atau dieliminasi (Aiken, 1985).

### Prosedur

Prosedur mencakup perolehan izin/*consent*, *briefing*, pengisian kuesioner secara berkelompok, uji keterbacaan yang melibatkan tiga siswa, revidi sebanyak sebanyak lima *reviewer* dari mahasiswa magister psikologi pendidikan dan perkembangan Universitas Negeri Makassar, dan revidi bahasa dari dosen sekaligus mahasiswa doktoral Pendidikan Bahasa Universitas Negeri Makassar. Pendekatan ini memaksimalkan kejelasan aitem dan kesepadanan makna (Clark & Watson, 2019).

### Analisis Data

EFA dilakukan pada matriks korelasi yang sesuai menggunakan ekstraksi *principal axis factoring* dan rotasi *oblique (promax/oblimin)* mengingat faktor diasumsikan berkorelasi. Penentuan jumlah faktor mempertimbangkan aitem dengan *loading*  $< 0,40-0,50$  atau *cross-loading* tinggi dipertimbangkan untuk dieliminasi. *Heywood case* (*loading*  $> 1$ ) diselidiki melalui komunalitas, properness solusi, dan alternatif rotasi (Marsh et al., 2020; Hair et al., 2019).

CFA menggunakan estimator yang sesuai. Pelaporan fit meliputi  $\chi^2(df, p)$ , CFI, TLI, RMSEA (CI 90%), dan SRMR. Kriteria keputusan: CFI  $\geq 0,90-0,95$ ; TLI  $\geq 0,90$ ; RMSEA  $\leq 0,08$ ; SRMR  $\leq 0,08$ . Aitem valid bila *loading* terstandar  $\geq 0,50$  dan signifikan. Penyesuaian berdasarkan indeks modifikasi dilakukan secara hemat dan harus memiliki rasional teoretis (Byrne, 2016; Kline, 2016).

Reliabilitas dievaluasi menggunakan McDonald's Omega mengingat keunggulannya atas alpha pada kondisi ketidaksamaan muatan (McNeish, 2018; Flora, 2020). Penormaan skor dilakukan dengan menetapkan kategori interpretatif (rendah–sedang–tinggi) berdasarkan distribusi sampel uji (Pallant, 2020).

## 3. Hasil dan Diskusi

### Hasil

Mengembangkan instrument SMAS diawali dengan validasi isi oleh 26 pakar psikologi pendidikan dan psikologi perkembangan Universitas Negeri Makassar menunjukkan tingkat kesepakatan yang tinggi. Indikator pada kedua komponen dinyatakan valid berdasarkan Aiken's V.

**Tabel 1. Hasil Validasi Indikator Keperilakuan**

Aspek	Indikator	Aiken's V	Keterangan
<i>Knowledge of Cognition</i>	Pengetahuan tentang diri sendiri dalam proses pembelajaran	0,90	Valid
<i>Knowledge of Cognition</i>	Pengetahuan tentang cara menggunakan strategi belajar	0,90	Valid
<i>Knowledge of Cognition</i>	Pengetahuan tentang kondisi menggunakan strategi	0,90	Valid
<i>Regulation of Cognition</i>	Perencanaan tujuan pembelajaran	0,80	Valid
<i>Regulation of Cognition</i>	Strategi manajemen informasi	0,90	Valid
<i>Regulation of Cognition</i>	Menilai efektivitas strategi belajar	0,90	Valid
<i>Regulation of Cognition</i>	Strategi penelusuran kesalahan	0,80	Valid
<i>Regulation of Cognition</i>	Evaluasi hasil belajar	0,90	Valid

Validasi indikator keperilakuan dilakukan oleh menunjukkan bahwa semua indikator yang diuji memperoleh nilai Aiken's V antara 0,8 dan 0,9. Hasil ini menunjukkan nilai nilai yang lebih tinggi menunjukkan tingkat kesepakatan yang lebih baik mengenai validitas aitem. Ini menegaskan bahwa indikator-indikator yang digunakan dalam SMAS relevan dan representatif untuk mengukur aspek-aspek metakognisi yang ditargetkan. Sehingga terbentuklah kisi-kisi akhir sebagai berikut:

**Tabel 2. Kisi-Kisi Akhir**

Aspek	Indikator	Aitem (F/UF)	Total
<i>Knowledge of Cognition</i>	Pengetahuan tentang diri sendiri dalam proses pembelajaran	4, 7 (F); 10 (UF)	3
<i>Knowledge of Cognition</i>	Pengetahuan tentang cara menggunakan strategi belajar	2 (F); 13, 15 (UF)	3
<i>Knowledge of Cognition</i>	Pengetahuan tentang kondisi ketika menggunakan strategi	9, 14, 16 (F)	3
<i>Regulation of Cognition</i>	Perencanaan	3, 5, 21 (F)	3
<i>Regulation of Cognition</i>	Strategi manajemen informasi	8, 19, 22 (F)	3
<i>Regulation of Cognition</i>	Pemantauan	1, 11 (F); 23 (UF)	3
<i>Regulation of Cognition</i>	Strategi penelusuran kesalahan	12, 20, 24 (F)	3
<i>Regulation of Cognition</i>	Evaluasi	6, 17 (F); 18 (UF)	3
<b>Total</b>			<b>24</b>

Proses ini memastikan bahwa aitem-aitem yang tersisa memenuhi prinsip penulisan aitem yang baik, yaitu jelas, ringkas, dan mudah dipahami oleh siswa (Clark & Watson, 2019). Aitem-aitem ini juga telah disesuaikan untuk mencerminkan perilaku dan strategi yang terkait dengan dimensi metakognisi yaitu *Knowledge of Cognition* dan *Regulation of Cognition* (Schraw & Dennison, 1994). Dari kisi-kisi akhir, terdapat 24 aitem yang sudah divalidasi aitemnya dengan menggunakan Aiken's V dengan hasil sebagai berikut:

**Tabel 3. Hasil Aiken's V**

Aitem	Aiken's V	Kategori
1	1.00	Valid
2	1.00	Valid
3	1.00	Valid
4	1.00	Valid
5	1.00	Valid
6	1.00	Valid
7	0.92	Valid
8	1.00	Valid
9	1.00	Valid
10	1.00	Valid
11	0.92	Valid
12	1.00	Valid
13	1.00	Valid
14	1.00	Valid
15	0.92	Valid
16	0.92	Valid
17	1.00	Valid
18	1.00	Valid
19	0.92	Valid
20	1.00	Valid
21	1.00	Valid
22	1.00	Valid
23	0.92	Valid
24	1.00	Valid

Dari 52 aitem awal, beberapa aitem dinyatakan gugur karena memiliki nilai  $V$  di bawah ambang batas berdasarkan tabel Aikens  $V$ . Aitem-aitem yang gugur ini perlu direvisi atau dihilangkan dari instrumen, sesuai dengan rekomendasi Penfield dan Giacobbi (2004) yang menyatakan bahwa indeks kuantitatif Aiken's  $V$  memberikan dasar objektif untuk keputusan retensi atau eliminasi aitem Aiken, L. R. (1985). Proses ini memastikan bahwa hanya aitem yang paling relevan dan valid secara isi yang dipertahankan dalam skala akhir. Setelah melakukan validitas isi pada aitem, dilakukanlah uji daya diskriminasi aitem dengan hasil sebagai berikut:

**Tabel 4. Hasil Uji Daya Diskriminasi Aitem**

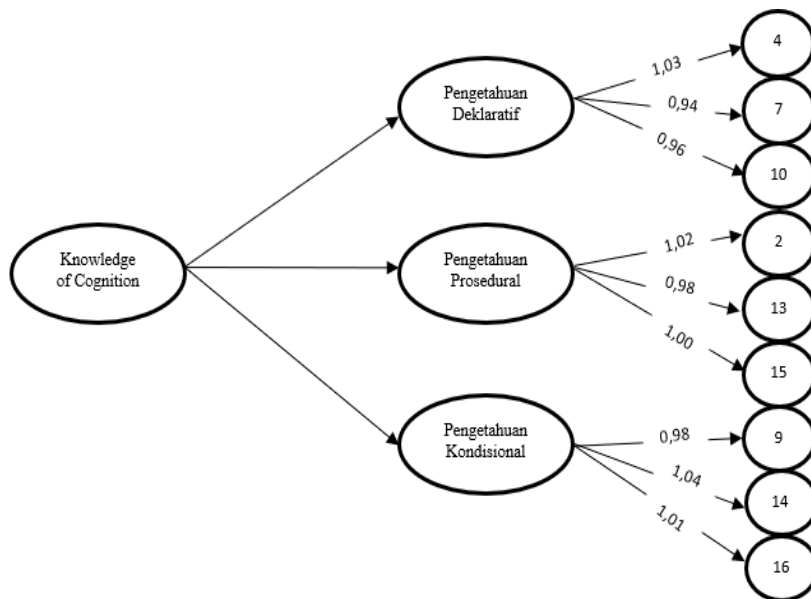
Aitem	<i>Corrected Item-Total r</i>	Kategori
1	0.84	Valid
2	0.83	Valid
3	0.83	Valid
4	0.81	Valid
5	0.84	Valid
6	0.84	Valid
7	0.80	Valid
8	0.82	Valid
9	0.78	Valid
10	0.79	Valid
11	0.84	Valid
12	0.78	Valid
13	0.80	Valid
14	0.83	Valid
15	0.80	Valid
16	0.82	Valid
17	0.86	Valid
18	0.83	Valid
19	0.81	Valid
20	0.83	Valid
21	0.83	Valid
22	0.79	Valid
23	0.76	Valid
24	0.78	Valid

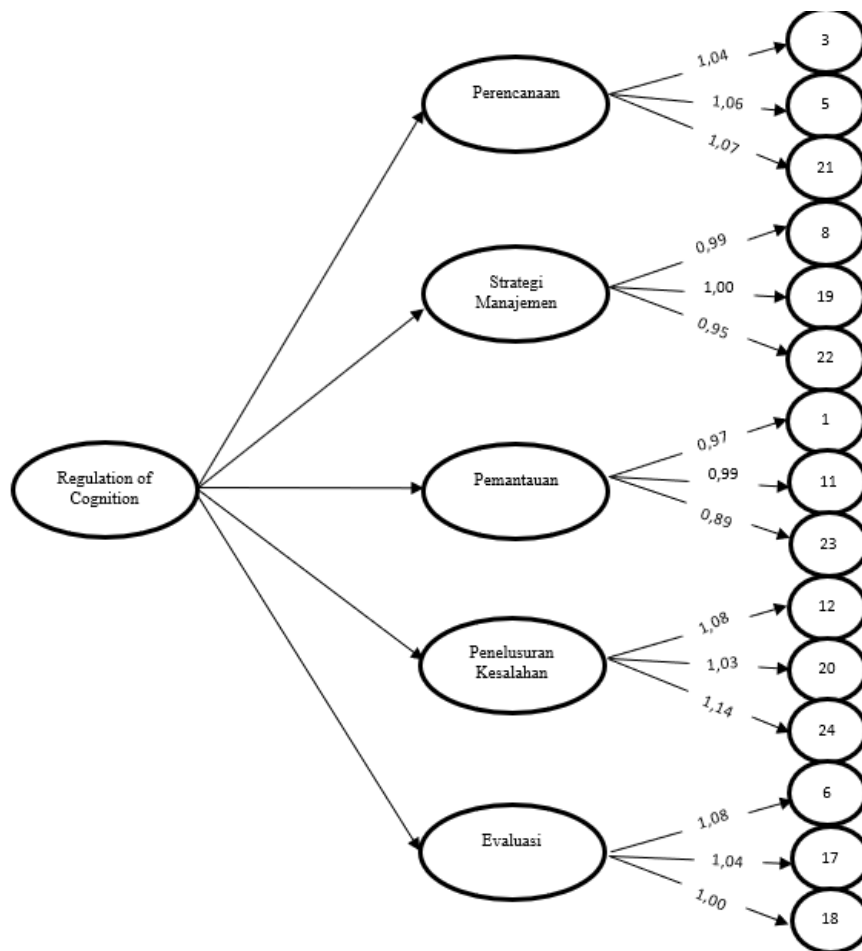
Hasil analisis menunjukkan bahwa semua aitem dalam skala SMAS memiliki nilai korelasi item-total yang dikoreksi di atas 0,3, yang secara umum dianggap sebagai indikasi daya diskriminasi yang baik (Pallant, 2020). Rentang koefisien *corrected item-total correlation* dari 24 item adalah 0,76 hingga 0,86. Ini menunjukkan bahwa semua item memiliki korelasi yang positif dengan total skor, dengan nilai terendah 0,76 dan nilai tertinggi 0,86. Semakin tinggi nilai koefisien, semakin baik kemampuan item dalam membedakan responden berdasarkan karakteristik yang diukur. Dalam rentang ini, sebagian besar item memiliki nilai koefisien yang cukup tinggi yang menunjukkan bahwa aitem-aitem tersebut efektif dalam mengukur konstruk yang sama. Penyetaraan ini penting untuk memastikan bahwa setiap aspek konstruk memiliki representasi yang seimbang dalam skor total, mengurangi potensi bias dalam pengukuran (Clark & Watson, 2019). Aitem yang masih bertahan setelah uji daya diskriminasi aitem, dilakukanlah uji *Explanatory Factor Analysis* (EFA) dengan hasil sebagai berikut:

**Tabel 5. Hasil Uji *Exploratory Factor Analysis* (EFA)**

Aitem	Faktor	Loading
4	1	0.940
7	1	0.697
10	2	0.671
2	1	0.820
13	2	0.905
15	2	0.834
9	1	0.646
14	1	0.775
16	1	0.780
3	1	0.738
5	1	0.847
21	1	0.907
8	1	0.691
19	1	0.748
22	1	0.701
1	1	0.686
11	1	0.557
23	2	0.663
12	1	0.899
20	1	0.775
24	1	1.000
6	1	0.654
17	1	0.791
18	2	0.691

Dalam Tabel 5, hasil *Exploratory Factor Analysis* (EFA) mengidentifikasi dua faktor utama yang konsisten dengan kerangka teoretis metakognisi (Schraw & Dennison, 1994), yaitu *knowledge of cognition* dan *regulation of cognition*. Nilai loading tertinggi pada faktor ini adalah aitem 24 (1.000), diikuti aitem 4 (0.940) dan 21 (0.907), menunjukkan kontribusi kuat terhadap dimensi regulasi kognitif, seperti pemantauan dan evaluasi strategi belajar. Faktor kedua (*knowledge of cognition*) lebih lemah dalam muatan keseluruhan, dengan aitem utama seperti 10 (0.671), 13 (0.905), 15 (0.834), dan 18 (0.691), yang merepresentasikan pengetahuan deklaratif, prosedural, dan kondisional tentang proses belajar. Aitem yang bertahan setelah uji EFA, dilakukanlah uji *Confirmatory Factor Analysis* (CFA) dengan hasil sebagai berikut:





**Gambar 1. Hasil Uji *Confirmatory Factor Analysis* (CFA)**

Gambar diatas secara keseluruhan menunjukkan bahwa semua aitem berkontribusi kuat terhadap konstruk masing-masing faktor dengan faktor *regulation of cognition* menunjukkan muatan lebih dominan. Nilai loading  $\geq 0.50$  memenuhi kriteria konvensional untuk validitas factorial (Hair et al., 2019). Mendukung representasi yang baik dari kesadaran metakognitif pada siswa SMA/SMK Indonesia. Namun, beberapa *estimate* melebihi 1.0 (*Heywood cases*) menandakan potensi ketidakstabilan model. Meskipun demikian, aitem-aitem ini tidak dihapus karena: (1) relevansi teoretis tinggi sebagai indikator kunci regulasi kognitif, (2) inspeksi *modification indices* (MI) dan residual standar menunjukkan nilai rendah tanpa indikasi *misspecification* signifikan, (3) retensi tidak mengganggu fit model keseluruhan atau reliabilitas ( $\Omega=0.980$ ), dan (4) menghapusnya berisiko mengurangi cakupan konstruk, sesuai rekomendasi untuk instrumen pengembangan (Byrne, 2016; Kline, 2016).

**Tabel 6. Indeks Kesesuaian Model**

Indeks	Nilai	Kriteria
CFI	0,943	CFI $\geq 0,90-0,95$
TLI	0,930	TLI $\geq 0,90$
RMSEA	0,079	RMSEA $\leq 0,08$

Nilai CFI sebesar 0,943. Nilai ini menunjukkan *fit* model yang baik, karena mendekati 1 (Byrne, 2016). TLI (*Tucker-Lewis Index*). Nilai TLI sebesar 0,930. Nilai ini juga menunjukkan *fit* model yang baik, mendekati 1 (Marsh et al., 2004). RMSEA (*Root Mean Square Error of Approximation*), Nilai RMSEA sebesar 0,079. Nilai ini menunjukkan *fit* model yang dapat diterima, karena berada di bawah 0,08.

Berdasarkan nilai-nilai *fit* ini, dapat disimpulkan bahwa model teoretis SMAS dengan 24 aitem sesuai dengan data empiris dengan baik. Hasil analisis CFA ini memberikan dukungan kuat bahwa SMAS 24 aitem merupakan alat ukur yang valid dan dapat diandalkan untuk mengukur konstruk metakognitif pada siswa, mencerminkan konsistensi antara model pengukuran teoretis dan data empiris. Validitas konstruk yang kuat ini merupakan

prasyarat untuk alat ukur yang dapat diandalkan dan valid (Marsh et al., 2020). Sehingga reliabilitas pada skala SMAS menunjukkan konsistensi yang sangat tinggi sebagai berikut:

**Tabel 7. Reliabilitas**

Skala	McDonald's Omega ( $\Omega$ )
SMAS (24 aitem)	0,980

Tabel 7 menunjukkan bahwa koefisien McDonald's Omega untuk SMAS dengan 24 aitem adalah 0,980. Nilai ini sangat tinggi, jauh di atas ambang batas yang umumnya diterima (0,70) dan bahkan dianggap sangat baik (0,90 atau lebih tinggi) (Trizano-Hermosilla & Alvarado, 2016). Reliabilitas yang sangat tinggi ini mengindikasikan bahwa SMAS menghasilkan skor yang sangat konsisten dan stabil. Ini berarti bahwa jika skala ini digunakan berulang kali pada kondisi yang relatif sama, hasilnya akan sangat serupa. Konsistensi internal yang luar biasa ini menunjukkan bahwa alat pengukuran tersebut menangkap konstruk yang mendasarinya dengan kesalahan minimal, dan variasi dalam skor yang diperoleh dari SMAS sebagian besar disebabkan oleh perbedaan yang sebenarnya dalam tingkat metakognisi siswa, bukan karena kesalahan pengukuran acak (Mcneish, 2018).

### Diskusi

Temuan penelitian ini menunjukkan bahwa SMAS merepresentasikan struktur dua dimensi metakognisi yaitu *knowledge of cognition* (pengetahuan deklaratif, prosedural, kondisional) dan (perencanaan, manajemen informasi, pemantauan, penelusuran kesalahan, evaluasi). Secara konsisten dan parsimoni, selaras dengan kerangka klasik yang telah divalidasi secara luas pada populasi remaja dan dewasa muda (Schraw & Dennison, 1994; González-Cabañes et al., 2022). Struktur ini tidak hanya menghasilkan model dua faktor yang mapan, tetapi juga menunjukkan bahwa *regulation of cognition* muncul sebagai dimensi yang lebih dominan pada siswa SMA/SMK Indonesia, di mana siswa sering kali memiliki pengetahuan dasar tentang strategi belajar namun kesulitan dalam mengaplikasikannya secara adaptif (Tuononen et al., 2023). Hal ini mencerminkan dinamika perkembangan remaja, di mana kesadaran metakognitif berkembang secara bertahap dari pengetahuan implisit ke regulasi eksplisit, sebagaimana terlihat dalam studi pada siswa sekolah menengah yang menyoroti perbedaan ini berdasarkan prestasi (Nederhand et al., 2020; Al-Balushi et al., 2022).

Lebih lanjut, validitas isi yang tinggi yang diperoleh melalui review validator menegaskan kejelasan indikator SMAS serta relevansi konten terhadap praktik pembelajaran remaja di konteks Indonesia. Aitem disesuaikan dengan kondisi siswa untuk mengurangi bias budaya, menghasilkan daya diskriminasi yang kuat di mana item regulasi membedakan siswa berprestasi tinggi dari rendah lebih baik daripada aitem pengetahuan (Köksal et al., 2021). Temuan ini selaras dengan adaptasi instrumen di negara berkembang, di mana penyesuaian konten lokal meningkatkan sensitivitas terhadap nuansa bahasa dan norma sosial. (Li et al., 2024; Wang et al., 2023).

Kombinasi *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA) dalam penelitian ini menghasilkan bukti yang konvergen mengenai konstruksi laten SMAS, dengan EFA mengidentifikasi beban faktor >0.40 dan CFA mengonfirmasi model tanpa *cross-loading* signifikan. Pendekatan ini konsisten dengan rekomendasi analisis faktor di bidang pendidikan dan psikologi yang menekankan integrasi bifactor model untuk menguji *unidimensionality essential* dan menghindari *overfactoring* (Marsh et al., 2020; Hair et al., 2019; Ondé et al., 2022). SMAS menunjukkan koherensi keseluruhan yang kuat, mirip dengan validasi *Metacognitive Awareness of Reading Strategies Inventory* (MARSII) pada siswa Spanyol di mana struktur serupa mendukung penggunaan skor total untuk asesmen praktis (Ondé et al., 2022).

Dari perspektif model FIT termasuk  $\chi^2$  CFI (>0.95), TLI (>0.95), dan RMSEA (<0.05), Temuan SMAS menunjukkan kecocokan yang sangat baik yang mengindikasikan model yang ringkas. Inspeksi misfit lokal, seperti residual standar dan indeks modifikasi, penting untuk mencegah kesimpulan yang bias oleh satu metrik saja, terutama pada data *self-report* remaja yang rentan terhadap variabilitas respons (Gutiérrez de Blume et al., 2024). Dalam kasus SMAS, residual rendah pada item regulasi menegaskan stabilitas dimensi ini, sementara indeks modifikasi rendah menghindari penyesuaian *post-hoc* yang tidak beralasan, selaras dengan temuan pada validasi *Metacognitive Awareness Scale* (MAS) di Cina (Li et al., 2024).

Kriteria keputusan dan justifikasi teoretis sebelum melakukan penyesuaian model merupakan *best practice* untuk menjaga validitas inferensial dan menghindari *capitalization on chance*, di mana penyesuaian berdasarkan teori lebih disukai daripada empiris semata (Byrne, 2016; Kline, 2016). Pendekatan ini telah terbukti efektif dalam adaptasi instrumen lintas budaya, seperti pada standarisasi MAI internasional di mana invariansi pengukuran dijaga melalui justifikasi serupa, mencegah bias interpretasi antar kelompok (Gutiérrez de Blume et al., 2024; Putnick & Bornstein, 2016). Untuk SMAS, proses ini memastikan bahwa model akhir tidak hanya fit secara statistik, tetapi juga bermakna secara teoretis bagi konteks pendidikan remaja Indonesia.

Implikasi praktis utama dari SMAS adalah integrasinya dalam siklus asesmen intervensi di kelas SMA/SMK, mulai dari pemetaan profil metakognitif kelas untuk mengidentifikasi siswa dengan regulasi rendah seperti yang sering terjadi pada siswa vokasi SMK yang fokus pada keterampilan prosedural hingga penentuan fokus strategi

spesifik (perencanaan tugas, pemantauan pemahaman, evaluasi hasil). Misalnya, guru dapat menggunakan skor SMAS untuk merancang intervensi seperti jurnal reflektif semi-struktural, yang terbukti meningkatkan *regulation of cognition* pada calon guru muda (Alt & Raichel, 2020; Pieschl et al., 2021). Di konteks Indonesia, ini selaras dengan Kurikulum Merdeka yang menekankan pembelajaran berbasis proyek di mana SMAS dapat memandu diferensiasi instruksional untuk mengurangi fluktuasi prestasi (Al-Balushi et al., 2022).

Selain itu, evaluasi perubahan pasca-intervensi melalui SMAS memungkinkan pemantauan longitudinal, seperti pengukuran pra-pasca pelatihan *metacognitive* pada mata pelajaran di mana kesadaran metakognitif berkorelasi positif dengan subketerampilan mendengarkan atau pemecahan masalah (Fu et al., 2023; Markandan et al., 2022). Aplikasi ini tidak hanya meningkatkan efisiensi kelas, tetapi juga memberdayakan konselor sekolah untuk mendeteksi dini siswa berisiko, seperti yang terlihat dalam studi pada siswa Oman di mana pelatihan think-aloud meningkatkan evaluasi diri (Al-Balushi et al., 2022). Dengan demikian, SMAS berpotensi mengurangi ketergantungan pada asesmen formatif tradisional yang kurang sensitif terhadap proses kognitif internal.

Dalam jangka panjang, penggunaan SMAS dapat mendukung kebijakan sekolah terkait literasi belajar dan program pendampingan akademik berbasis data, misalnya melalui integrasi dengan platform digital Kemdikbud untuk tracking nasional *metacognitive awareness*. Ini sejalan dengan rekomendasi global untuk pendidikan berbasis bukti, di mana instrumen seperti SMAS membantu mengatasi kesenjangan akses di daerah pedesaan Indonesia, serupa dengan integrasi computational thinking untuk meningkatkan monitoring dan evaluasi pada siswa (Pallant, 2020; Clark & Watson, 2019; Markandan et al., 2022). Namun, tantangan implementasi seperti pelatihan guru dan akses teknologi perlu diatasi untuk memaksimalkan dampaknya. Secara keseluruhan, diskusi ini menegaskan bahwa SMAS bukan hanya instrumen valid, tetapi juga katalisator untuk transformasi pendidikan remaja di Indonesia, dengan potensi kolaborasi antar-institusi untuk riset lanjut. Dengan mengintegrasikan temuan ini ke praktik, penelitian selanjutnya dapat memperluas dampaknya ke tingkat nasional, memperkaya literatur metakognisi di negara berkembang (Byrne, 2016; Hair et al., 2019).

#### 4. Kesimpulan

Penelitian ini menyajikan skala yang terbangun secara teoretis dan teruji secara empiris untuk menilai kesadaran metakognitif pada siswa SMA/SMK. Instrumen yang dihasilkan menampilkan kejelasan indikator, konsistensi internal, serta kesesuaian struktur yang memadai dengan model metateoretik yang diacu. Skala ini berpotensi digunakan dalam pemetaan kebutuhan belajar dan perancangan intervensi pembelajaran mandiri di sekolah.

Desain sampling *non-probabilitas* dan keterwakilan sekolah yang terbatas membatasi generalisasi temuan. Beberapa indikator menunjukkan kepekaan terhadap konteks bahasa dan budaya sekolah. Analisis lanjutan seperti *uji invariance* antarkelompok (jenis kelamin/jenjang) belum dilakukan. Riset berikutnya disarankan mereplikasi pada sampel yang lebih beragam, melakukan uji stabilitas temporal (*retest*), memeriksa validitas kriteria, dan menguji *invariance*. Pengembangan bentuk singkat (*short-form*) untuk skrining cepat di kelas juga direkomendasikan.

#### 5. Referensi

- Al - Balushi, S. M., Al - Harthy, I. S., Almehrizi, R. S., Ambusaiddi, A. K., Al - Balushi, K. A., Al - Saadi, K. K., Al - Aghbari, M., & Al - Balushi, M. (2022). Metacognitive awareness perceptions of students with high and low scores on TIMSS - like science tests. *International Journal of Cognitive Research in Science, Engineering and Education*, 10(3), 73 – 82.
- Alt, D., & Raichel, N. (2020). Reflective journaling and metacognitive awareness: Insights from a longitudinal study in higher education. *Reflective Practice*, 21(2), 145-158.
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and psychological measurement*, 45(1), 131-142.
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming (3rd ed.)*. Routledge.
- Çini, A., Järvelä, S., Dindar, M., & Malmberg, J. (2023). How multiple levels of metacognitive awareness operate in collaborative problem solving. *Metacognition and Learning*, 18(3), 891–922.
- Clark, L. A., & Watson, D. (2019). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319.
- Curran, P. J. (2016). Detection of careless responding. *Multivariate Behavioral Research*, 51(3-4), 362–364.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911.
- Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, 3(4), 484-501.

- Fu, Y., Wang, M., Min, S., Zhou, S., & Pan, X. (2023). Exploring the relationship between metacognitive awareness and Chinese EFL learners' listening skills. *Frontiers in Psychology, 14*, 1148610.
- González-Cabañes, E., García, T., Álvarez-García, D., Fernández, E., & Rodríguez, C. (2022). Validation of the shortened version of the Metacognitive Awareness Inventory in Spanish university students. *Psicothema, 34*(3), 454–462.
- Gutiérrez de Blume, A. P., Montoya Londoño, D. M., Jiménez Rodríguez, V., Morán Núñez, O., Cuadro, A., Daset, L., Molina Delgado, M., García de la Cadena, C., Beltrán Navarro, M. B., Puente Ferreras, A., Urquijo, S., & Arias, W. L. (2024). Psychometric properties of the Metacognitive Awareness Inventory (MAI): standardization to an international Spanish with 12 countries. *Metacognition and Learning, 19*(3), 793-825.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis (8th ed.)*. Cengage.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling (4th ed.)*. Guilford Press.
- Köksal, Ö., Sodian, B., & Legare, C. H. (2021). Young children's metacognitive awareness of confounded evidence. *Journal of Experimental Child Psychology, 205*, 105080.
- Li, S., Xu, J., Jia, X., Zhao, Y., Liu, X., & Wang, Y. (2024). Translation and psychometric validation of the Chinese version of the metacognitive awareness scale among nursing students. *Frontiers in Psychology, 15*, 1354810.
- Markandan, N., Osman, K., & Halim, L. (2022). Integrating computational thinking and empowering metacognitive awareness in STEM education. *Frontiers in Psychology, 13*, 872593.
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2020). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110.
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological methods, 23*(3), 412.
- Nederhand, M. L., Tabbers, H. K., De Bruin, A. B. H., & Rikers, R. M. J. P. (2020). Metacognitive awareness as measured by second-order judgements among university and secondary school students. *Metacognition and Learning, 15*(3), 351–370.
- Ondé, D., Jiménez, V., Alvarado, J. M., & Gràcia, M. (2022). Analysis of the structural validity of the reduced version of metacognitive awareness of reading strategies inventory. *Frontiers in psychology, 13*, 894327.
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.
- Pieschl, S., Budd, J., Thomm, E., & Archer, J. (2021). Effects of raising student teachers' metacognitive awareness of their educational psychological misconceptions. *Psychology Learning & Teaching, 20*(2), 1–22.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental review, 41*, 71-90.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475.
- Song, J. H., Loyal, S., & Lond, B. (2021). Metacognitive awareness scale, domain specific (MCAS-DS): Assessing metacognitive awareness during Raven's progressive matrices. *Frontiers in Psychology, 11*, 607577.
- Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. *Frontiers in psychology, 7*, 769.
- Tuononen, T., Hyytinen, H., Räisänen, M., Hailikari, T., & Parpala, A. (2023). Metacognitive awareness in relation to university students' learning profiles. *Metacognition and Learning, 18*(1), 37-54.
- Wang, M., Zhang, L. J., & Hamilton, R. (2023). Developing the Metacognitive Awareness of Grit Scale for a better understanding of learners of English as a foreign language. *Frontiers in Psychology, 14*, 1141214.



## CONTEXTUAL VERBAL CREATIVITY TEST (CVCT): A NEW TOOL FOR UNDERSTANDING VERBAL CREATIVITY AMONG STUDENTS OF UNIVERSITY X

Renata V.Gunjaya<sup>1</sup>, Cecilia Ang<sup>2</sup>, Sania Sania<sup>3</sup>, P.Tommy Y. S. Suyasa<sup>4</sup>

<sup>1,2,3,4</sup> Fakultas Psikologi, Universitas Tarumanegara, Jakarta, Indonesia

Email : [tommys@fpsi.untar.ac.id](mailto:tommys@fpsi.untar.ac.id)

p-ISSN: 2337-4845

e-ISSN: 2620-7486



Received	Revised	Accepted	Published
10 March 2025	21 October 2025	27 October 2025	30 October 2025

### Abstract

Creativity is a human intellectual capacity to generate novel and useful ideas in everyday life. Creativity encompasses several domains, including verbal, figural, and kinesthetic creativity. Among these domains, verbal creativity is the type most frequently utilized in daily activities. Although several verbal creativity assessment tools have been developed in Indonesia, an instrument that specifically measures contextual verbal creativity has not yet been available. Therefore, this study aimed to develop the Contextual Verbal Creativity Test (CVCT) and to examine its psychometric properties as a newly designed instrument for assessing verbal creativity with a contextual nature, which can be administered within 10 minutes. During the development process, the researchers constructed a test comprising 20 items, which were administered to 100 undergraduate students from University X. These twenty indicators were evaluated based on four dimensions: fluency, flexibility, originality, and elaboration. Following data collection, reliability and validity analyses were conducted using internal consistency reliability, test-retest reliability, construct validity, and criterion validity. The findings of this research present four psychometric outcomes: (a) the internal consistency reliability; (b) the descriptive results of the measurements; (c) the criterion validity of the CVCT in relation to academic achievement; and (d) the construct validity of the CVCT measurements based on gender.

Keywords: measurement development, test, verbal creativity, contextual, quality education

### Abstrak

Kreativitas merupakan kemampuan intelektual manusia untuk menghasilkan ide baru yang berguna dalam kehidupan sehari-hari. Kreativitas terdiri atas beberapa jenis, yakni kreativitas verbal, figural, dan kinestetik. Diantara beberapa jenis kreativitas tersebut, kreativitas verbal merupakan jenis kreativitas yang banyak digunakan dalam kehidupan sehari-hari. Meskipun beberapa alat ukur kreativitas verbal telah dikembangkan di Indonesia, instrumen yang secara khusus mengukur kreativitas verbal kontekstual belum tersedia. Oleh karena itu, penelitian dilakukan untuk mengembangkan Contextual Verbal Creativity Test (CVCT) dan memperoleh informasi psikometrisnya sebagai alat ukur kreativitas verbal baru yang bersifat kontekstual dan dapat dioperasikan dalam waktu 10 menit. Dalam proses pengembangannya, peneliti menyusun tes yang terdiri dari 20 butir kepada 100 mahasiswa dari Universitas X. Kedua puluh indikator ini diukur berdasarkan empat komponen yang meliputi fluency, flexibility, originality, dan elaboration. Setelah proses pengambilan data, peneliti akan melakukan uji reliabilitas dan validitas menggunakan internal consistency reliability, test-retest reliability, construct validity, dan criterion validity. Temuan dari penelitian meliputi empat informasi psikometris, yaitu: (a) reliabilitas konsistensi internal; (b) gambaran hasil pengukuran; (c) validitas kriteria CVCT terhadap prestasi akademik; dan (d) validitas konstruk pengukuran CVCT berdasarkan jenis kelamin.

Kata Kunci: pengembangan alat ukur, tes, kreativitas verbal, kontekstual, kualitas pendidikan

## 1. Introduction

Creativity is the capacity to generate or cultivate novel products, works, theories, procedures, or thoughts (American Psychology Association, 2023). Creativity refers to the ability to behave or produce a new and adaptive response which relevant and beneficial in everyday life (Peterson & Seligman, 2004). Additionally, a different

interpretation of creativity was formulated from a dynamic standpoint (Walia, 2019). Creativity is conceptualized as an activity responsive to the surrounding environment, instigating the development of productive actions that may deviate from societal conventions and norms, which subsequently generates a novel work. Drawing from these definitions, it can be concluded that the concept of creativity encompasses novelty, usefulness and relevance to everyday life, and reflects a productive activity that often diverges from common or conventional ways of thinking in society.

Building upon this general understanding of creativity, one specific form that plays a significant role in daily life is verbal creativity, as the capacity for flexible and innovative thinking, assessed through the manipulation of words to enhance divergent thinking abilities (Fink et al., 2020). As one of the most common forms of creativity utilized in various aspects of life, verbal creativity holds particular significance since verbal ability is crucial in many sectors, such as art, education, advertising, and business (Efrati et al., 2022; Ehrenreich et al., 2021; Fernández-Monroy et al., 2018; Kartana et al., 2018; Kozbelt, 2020; Lin, 2018; Osowiecka & Kolanczyk, 2018; Suyasa et al., 2025). First, in the arts sector, verbal abilities are utilized in the composition of songs and poems. Second, verbal skills play a crucial role in the teaching and learning process, as they enable teachers to deliver information effectively while helping students articulate their understanding and develop their capacity to produce scholarly works. Third, in the advertising industry, verbal abilities are utilized to produce consumer-attracting content. Fourth, in the business sector, verbal abilities are used to maintain effective collaboration and foster long-term relationships with business partner.

Furthermore, verbal creativity is correlated with academic performance, such as Grade Point Average (GPA) (Scherbakova et al., 2024; Suyasa et al., 2025). Individuals who demonstrate greater fluency, flexibility, elaboration, and originality often perform better in academic tasks. These abilities allow them to demonstrate a high level of work effectivity, adaptability to various viewpoints, a deep level of understanding, and originality in completing academic assignments. Consequently, such qualities contribute to improved academic performance and positively influence students' GPA.

Given the significant role of verbal creativity in practical context, verbal creativity becomes an interesting topic for further exploration. Several instruments have been developed for this purpose, including: (a) the Arabic version of the Torrance Test of Creative Thinking-Verbal (Said-Metwaly et al., 2021); Tes Kreativitas Verbal developed by Munandar (1977); and (c) Alat Ukur Kreativitas Verbal "C" (Kartana et al., 2018). Firstly, the adapted version of the Torrance Test of Creative Thinking (TTCT) (Said-Metwaly et al., 2021) has a duration of 40 minutes and consists of six types of activities: (a) questioning activity involving generating questions about an image; (b) guessing the causes of an event; (c) guessing the effects of an event; (d) creating lists to develop products; (e) seeking new functions for products; (f) specifying possible outcomes for unlikely situations.

Secondly, Tes Kreativitas Verbal constructed by Munandar (1977) consists of six subtests that ask participants to: (a) name words that begin with a few pre-set alphabets (e.g., KA, A,); (b) rearrange letters to create new words; (c) make a sentence consisting of three words from given initials; (d) name an object based on the characteristics given; (e) name the use of an object given; and (f) state the outcome of a certain scenario. Each subtest contains four items and requires a total of 60 minutes to complete. Thirdly, Alat Ukur Kreativitas Verbal "C" constructed by Kartana et al. (2018) consists of 50 words that begin with the letter "C" as the items and requires a total of 10 minutes to complete. The participants are instructed to construct sentences using only words that start with the letter C.

Although the three measurement tools have been developed to assess verbal creativity, these instruments have not yet integrated relevant situational conditions or still lack contextual elements. Albeit participants use as many words as the object requires, their answer will not be contextual. Contextual is defined as the connectedness or connection of a concept in everyday life (Dewey, 1882, as cited in Kirby, 2005). Thus, although the existing measurement tools provide stimuli to prompt responses, these stimuli do not use adjectives as the basis for participants' answers, which made it tend to be less relevant or lack a direct connection to participants' experiences. Contextualized assessment tools, in contrast, yield results that tend to be more accurate in reflecting the experiences and realities encountered by subjects, thereby enhancing the usefulness and applicability of measurement outcomes in daily activities (Ambuehl & Inauen, 2022; Oppong et al., 2023; Schlotzhauer et al., 2025). Ideally, creativity can assist people in producing concepts or responses to problems encountered in everyday life (Cropley, 2020; Harms et al., 2017). It also helps people to view an issue from a different angle and come up with a solution (Weston, 2006). For example, individuals who wish to feel cool despite the extreme heat will use contextual reasoning abilities to operate as a response to existing concerns. Therefore, a contextual verbal creativity instrument was developed in this study to address the limitations of existing measures by incorporating adjective-based stimuli that better reflect participants' real-life experiences.

In contrast to established tools for assessing verbal creativity, the CVCT measurement instrument introduces two innovations: it incorporates a contextual dimension and can be completed within ten minutes. The CVCT consists of twenty items that assess four dimensions of creativity. The elements of creativity are originality, flexibility, fluency, and elaboration (Torrance in Said-Metwaly et al., 2021). The number of ideas that people

possess is related to the first factor, fluency. The second factor, flexibility, is associated with the quantity of a variety of ideas owned by individuals. The third element, originality, refers to a person's capacity to create fresh and creative goods or solutions. The fourth element, elaboration, refers to a person's capacity to add extra details that are pertinent to an idea. Additionally, this study aims to establish comprehensive psychometric information regarding the CVCT by examining: (a) the reliability of internal consistency; (b) the presentation of measurement results (mean and standard deviation); (c) criterion validity; and (d) construct validity, including discriminant evidence.

## 2. Research Method

The Research Method section will describe the number and characteristics of the participants, the design of the Contextual Verbal Creativity Test (CVCT) instrument, the data collection procedure, and the data analysis methods.

### Participants

The research study involved 100 participants with the following characteristics: (a) a mean age of 20.37 years (SD = 1.618); (b) the majority (80%) were female; (c) all were undergraduate students at University X; and (d) they had no prior knowledge of the psychological construct of verbal creativity.

### Instrument and Research Design

**Measurement.** The Contextual Verbal Creativity Test (CVCT) instrument consists of twenty items associated with adjectives. The twenty items are presented in the following Table 1.

**Table 1. Contextual Verbal Creativity Test (CVCT)'s List Items**

No.	Question
1.	<i>Panas</i>
2.	X
3.	X
4.	X
5.	X
6.	X
7.	X
8.	X
9.	X
10.	X
11.	X
12.	X
13.	X
14.	X
15.	X
16.	X
17.	X
18.	X
19.	X
20.	<i>Lembut</i>

These twenty items are presented in the form of adjectives as prompts, accompanied by the following instructions: (1) The tester will read aloud 20 words (adjectives); (2) The testee writes their responses in the designated answer column; (3) In the answer column, the testee is instructed to write words they perceive as related to / associated with / or closely connected to the adjectives read by the tester; (4) The testee is given 30 seconds to respond to each word; (5) The testee may only begin writing when the tester gives the signal "Start" and must stop writing immediately when the tester gives the signal "Stop." These items are assessed based on four aspects: fluency, flexibility, originality, and elaboration.

In the first aspect (Table 2), fluency, the score is measured based on the quantity of responses provided by the testee. A high fluency score is indicated when the testee is able to generate a large number of words related to the test item. Conversely, a low fluency score is indicated when the testee provides only a limited number of related words.

**Table 2. Example of Responses with High and Low Fluency Scores for the Item “Panas”**

Item	Example of a Response with a High Fluency Score	Example of a Response with a Low Fluency Score
<i>Panas</i>	<i>Matahari, Terik, Teh, Kopi, Payung, Laptop, Termos, Api, Keringat, Gosong</i>	<i>Matahari, Terik, Teh</i>
	Total Score : 10	Total Score : 3

In the second aspect (Table 3), flexibility, the score is measured based on the variation in the testee’s responses. A high flexibility score is indicated when the testee is able to express a diverse range of concepts associated with the test item. Conversely, a low flexibility score is indicated when the testee repeatedly uses similar or identical concepts in response to the test item.

**Table 3. Example of Responses with High and Low Flexibility Scores for the Item “Panas”**

Item	Example of a Response with a High Flexibility Score	Example of a Response with a Low Flexibility Score
<i>Panas</i>	<i>Matahari, Teh, Kopi, Laptop, Termos</i>	<i>Matahari, Matahari, Teh, Teh, Laptop, Laptop</i>
	Total Score : 5	Total Score : 3

In the third aspect (Table 4), originality, the score is measured based on the uniqueness of the testee’s responses. A high originality score is indicated when the testee provides words associated with the test item that are unusual or rarely mentioned by other testees. Conversely, a low originality score is given when the testee provides common or frequently mentioned words associated with the test item.

**Table 4. Example of Responses with High and Low Originality Scores for the Item “Panas”**

Item	Example of a Response with a High Originality Score	Example of a Response with a Low Originality Score
<i>Panas</i>	UV Whitening Total Score : 4 (the answer has less than 1% similarity with other testees’ responses)	<i>Matahari</i> Total Score : 1 (the answer has more than 10% similarity with other testees’ responses)

In the fourth aspect (Table 5), elaboration, the score is measured based on the inclusion of additional information relevant to the responses provided by the testee. A high elaboration score is indicated when the testee offers explanations or details related to the concepts they mentioned. Conversely, a low elaboration score is given when the testee provides only brief or single-word responses without further elaboration.

**Table 5. Example of Responses with High and Low Elaboration Scores for the Item “Panas”**

Item	Example of a Response with a High Elaboration Score	Example of a Response with a Low Elaboration Score
<i>Panas</i>	<i>Matahari yang bersinar cerah di pagi hari</i> Total Score : 7	<i>Matahari</i> Total Score : 1

### Data Collection Procedure

Data were collected using a non-probability sampling technique (convenience sampling) after obtaining ethical clearance with approval number 158-TIM/KEPTM/1577/FPsi-UNTAR/V/2023. During the data collection process, the paper-and-pencil method was employed in a classroom setting. Participants used their own writing instruments, while the answer sheets were provided by the researchers. Participants were instructed to follow a structured set of instructions: (i) completed personal information (name, age, and gender); (ii) read the test instructions carefully; (iii) completed a set of practice items; (iv) provided contextually relevant responses to each test item; (v) were given 30 seconds to respond to each of the 20 adjective prompts; and (vi) proceeded to the next item only after the 30-second time limit had elapsed.

### Data Analysis Technique

The analysis was conducted to obtain psychometric information on the Contextual Verbal Creativity Test (CVCT) instrument, including internal consistency reliability, construct validity, and criterion validity. Internal consistency reliability was assessed using Cronbach's Alpha. Criterion validity—examining the relationship between verbal creativity and academic achievement—was analyzed using Spearman's correlation test. Construct validity—assessing differences in verbal creativity based on gender—was tested using the Mann-Whitney U test, as the data were not normally distributed.

### 3. Result and Discussion

Based on the four intended objectives, data processing was conducted to obtain the following information: (a) descriptive statistics (mean and standard deviation); (b) internal consistency reliability; (c) criterion validity with respect to academic achievement; and (d) construct validity (discriminant evidence). The results corresponding to these four aspects are presented in the following paragraphs.

#### Descriptive Analysis

Descriptive analysis was conducted to provide a comprehensive overview of the mean and standard deviation for each of the four aspects of verbal creativity. The descriptive analysis results are presented in Table 6 below.

**Table 6. Descriptive Analysis of the CVCT Measurement Tool**

No.	Variable	Mean	SD	Skewness	Kurtosis	Normality Status
1.	Fluency	7,96	1,865	0,625	-1,535	Normal
2.	Flexibility	6,63	1,470	1,016	-1,058	Normal
3.	Originality	12,45	3,736	2,166	-0,283	Non-normal
4.	Elaboration	8,89	2,593	4,729	6,821	Non-normal

Based on the descriptive analysis, an overview of the four aspects of creativity is presented in Table 6, as follows: (a) Fluency,  $M = 7.96$  words,  $SD = 1.865$  (indicating that individuals were able to generate an average of  $7.96 \pm 1.865$  words within 30 seconds); (b) Flexibility,  $M = 6.63$  words,  $SD = 1.470$  (indicating that individuals produced an average of  $6.63 \pm 1.470$  different categories of words in 30 seconds); (c) Originality,  $M = 12.45$ ,  $SD = 3.736$  (with an average fluency of 7.96 words, the originality score of 12.45 suggests a uniqueness level of approximately 1.56 per word—approaching a score of 2 on a 1–4 scale—which may be considered moderately unique); and (d) Elaboration,  $M = 8.89$ ,  $SD = 2.593$  (with an average fluency of 7.96, the elaboration score of 8.89 indicates that at least one response [ $8.89 - 7.96$ ] was further elaborated).

#### Internal Consistency Reliability

Based on the analysis using Cronbach's Alpha method, the internal consistency reliability of the Contextual Verbal Creativity Test (CVCT) measurement tool was found to be satisfactory ( $\alpha > 0.80$ ) (George & Mallery, 2003). This indicates that the twenty items in the CVCT demonstrate consistent measurement across the four aspects: fluency, flexibility, originality, and elaboration. Detailed coefficients of internal consistency reliability for each aspect are presented in Table 7 below.

**Table 7. Internal Consistency Reliability of the CVCT Measurement Tool**

No.	Aspect	Cronbach's Alpha Standardized
1.	Fluency	0,934
2.	Flexibility	0,898
3.	Originality	0,923
4.	Elaboration	0,951

The twenty items in the CVCT measurement tool consistently yielded high levels of fluency, flexibility, originality, and elaboration. Each aspect demonstrated an internal consistency reliability coefficient greater than 0.80, with elaboration showing the highest reliability, reaching a coefficient of 0.951.

#### Criterion Validity

Through a criterion validity study, the researchers evaluated the validity of the Contextual Verbal Creativity Test (CVCT) by examining its ability to predict a specific criterion. A correlation analysis was conducted between participants' academic achievement—measured by GPA (Grade Point Average)—as the criterion, and the four aspects of verbal creativity (fluency, flexibility, originality, and elaboration). The results of the criterion validity analysis are presented in Table 8 below.

**Table 8. Criterion Validity of the CVCT Measurement Tool**

No.	Variable	M	SD	$r_s$	$p$
1.	Fluency	7,96	1,865	0,26	0,009
2.	Flexibility	6,63	1,471	0,29	0,004
3.	Originality	12,45	3,737	0,27	0,006
4.	Elaboration	8,89	2,593	0,32	0,001

The data analysis using Spearman's correlation test at the 0.05 alpha level revealed significant relationships between CVCT scores and academic achievement ( $M = 3.16$ ,  $SD = 0.036$ ) across all four aspects of verbal creativity: fluency ( $M = 7.96$ ,  $SD = 1.865$ ),  $r_s(98) = 0.26$ ,  $p < 0.01$ ; flexibility ( $M = 6.63$ ,  $SD = 1.471$ ),  $r_s(98) = 0.29$ ,  $p < 0.01$ ; originality ( $M = 12.45$ ,  $SD = 2.737$ ),  $r_s(98) = 0.27$ ,  $p < 0.01$ ; and elaboration ( $M = 8.89$ ,  $SD = 2.593$ ),  $r_s(98) = 0.32$ ,  $p < 0.01$ . These results indicate that the CVCT demonstrates criterion validity in relation to academic performance, meaning it has the capacity to predict participants' academic achievement.

### Construct Validity

A construct validation study was conducted to assess how well the Contextual Verbal Creativity Test (CVCT) measurement tool represents the theoretical framework or construct of verbal creativity. According to de Cassia Nakano, Oliveira, and Zaia (2021), there are no significant differences in creativity levels between males and females. In this study, we employed the Mann-Whitney U test to compare the levels of verbal creativity between male ( $n = 20$ ) and female ( $n = 80$ ) participants. The results of this study align with the findings of de Cassia Nakano et al. (2021), indicating no significant differences in verbal creativity between genders. Therefore, the CVCT can be considered to have construct validity (discriminant evidence) based on gender. The test results are presented in Table 9 below.

**Table 9. Construct Validity of the CVCT Measurement Tool**

No.	Aspect	Male ( $n = 20$ )		Female ( $n = 80$ )		$U$	$p$
		MR	SR	MR	SR		
1.	Fluency	45,55	911	51,74	41,39	701	0,394
2.	Flexibility	44,33	887	52,04	4163	676	0,287
3.	Originality	45,95	919	51,64	4131	709	0,433
4.	Elaboration	44,73	895	51,94	41,56	685	0,320

Notes. MR = Average Rank; SR = Sum of Ranks

At a significance level of 0.05, no differences were found in the levels of verbal creativity between males and females across the four aspects: fluency [ $U = 701$ ,  $p > 0.05$ ], flexibility [ $U = 676$ ,  $p > 0.05$ ], originality [ $U = 709$ ,  $p > 0.05$ ], and elaboration [ $U = 685$ ,  $p > 0.05$ ]. Thus, there are no gender-based differences in any aspect of verbal creativity. These findings indicate that the CVCT measurement tool demonstrates discriminant evidence in construct validity.

## 4. Discussion

The Contextual Verbal Creativity Test (CVCT) is a verbal test consisting of 20 items designed to measure verbal creative ability as a skill that supports daily life. Unlike previous creativity tests, the CVCT is situational and contextual. The term situational refers to something that is about, related to, or appropriate to a situation (Merriam-Webster, n.d.) and to something that is suitable or relevant to a specific situation (Kamus Besar Bahasa Indonesia, n.d.). The inclusion of situational and contextual aspects allows the elaboration component of creativity to be more clearly represented and functional. The situational aspect refers to conditions that occur or are relevant to the present time, while the stimuli used in this test—adjectives—are linked to the current temporal context. For example, in a situational sense, a student feels “hot” when reprimanded by a lecturer, whereas contextually, “Panas” refers to the student's emotional reaction of being upset when corrected.

The Contextual Verbal Creativity Test (CVCT) demonstrated good internal consistency reliability. Based on the reliability analysis using the Cronbach's Alpha method, the obtained coefficient was  $\alpha = 0.86$ . This value indicates that the instrument's internal consistency reliability is considered good, as Pallant (2020) stated that an instrument can be regarded as reliable if it has a Cronbach's Alpha value of at least 0.80. Thus, the CVCT shows an adequate and trustworthy internal consistency in measuring contextual verbal creativity. For future research, to further support internal reliability, it is recommended to test the measurement model using confirmatory factor analysis (CFA) to examine and confirm the proposed factor structure.

For practical purposes, the Contextual Verbal Creativity Test (CVCT) is equipped with percentile norms. However, the development of these norms was limited by the number of participants, which was only 100. This sample size is considered insufficient to serve as a general reference; therefore, future research should expand the study by involving a larger number of participants, ideally exceeding 1,000 (Gregory, 2015). With a larger sample size, the score distribution across the four aspects is expected to approach normality. In the present study, the results of the CVCT measurement showed that two aspects—fluency and flexibility—were normally distributed, whereas the other two aspects were not.

The Contextual Verbal Creativity Test (CVCT) is also equipped with psychometric properties related to criterion validity and construct validity. In terms of criterion validity, this test can predict participants' academic performance. Previous research has shown a strong relationship between verbal creativity and academic achievement, indicating that individuals with higher verbal creativity tend to achieve better academic outcomes. This can be understood as academic success across various subjects—such as Civics, Religion, English, and Family Psychology, etc—often requires verbal processes (Komarudin, 2018). When individuals or students possess fluency, flexibility, elaboration, and originality, they tend to be more productive, adaptive to diverse perspectives, comprehensive, and unique in producing academic work, which ultimately contributes to achieving a higher GPA. To strengthen the criterion validity of the CVCT in future research, two relevant aspects are suggested: concurrent validity and predictive validity. Concurrent validity can be examined through correlations between CVCT scores and other verbal creativity tests with strong psychometric evidence, while predictive validity can be assessed through the ability of CVCT scores to predict individuals' performance in tasks requiring verbal and contextual creative thinking.

In terms of construct validity, this study employed the concept of discriminant evidence. According to de Cassia Nakano et al. (2021), the relationship between gender and creativity is inconsistent, meaning that in some cases males appear more creative, while in others females do, depending on the research methods and participant characteristics. This inconsistency indicates that gender and creativity are distinct or discriminant constructs. Consistent with this notion, the results of the CVCT showed no significant differences in the level of verbal creativity between male and female participants. These findings demonstrate that both genders exhibit comparable levels of creativity, aligning with the theoretical construct suggesting that there is no consistent evidence of gender differences in verbal creativity. Furthermore, from a biological perspective, there is no consistent evidence indicating structural or volumetric differences in the brain's language center (Broca's area) based on sex. The Broca's area, located in the cerebral hemisphere, is known to function as a center for both language production and comprehension (Binkofski & Buccino, 2004).

To further strengthen the construct validity of the CVCT beyond the evidence of discriminant validity discussed above, additional forms of construct validation can be considered. According to Harasty, Double, Halliday, Kril, and McRitchie (1997), proportionally, the Wernicke's area and Broca's area in females are larger than in males, which may explain why females tend to have better language abilities. However, Amunts et al. (1999) found no structural or volumetric differences in area 45 (Broca's area), suggesting that language abilities between males and females are essentially equivalent. Based on these findings, future studies are encouraged to examine other forms of construct validity, such as convergent validity. Convergent evidence refers to the extent to which a measure correlates positively with other instruments that theoretically assess similar constructs. Therefore, the CVCT could be compared or correlated with other established creativity assessments, such as the Torrance Test of Creative Thinking (TTCT), Munandar Verbal Creativity Test, or Verbal Creativity Test "C". Through these comparisons, researchers can obtain convergent evidence that further supports and enhances the construct validity of the CVCT.

## 5. Conclusion

Based on this study, it can be concluded that the CVCT instrument, which consists of 20 adjective items measuring the four aspects of verbal creativity (fluency, flexibility, originality, and elaboration), demonstrates adequate psychometric properties. The psychometric information obtained includes: (a) descriptive statistics (mean and standard deviation) of the CVCT scores from 100 participants; (b) good internal consistency reliability; (d) criterion validity of the CVCT instrument with academic achievement index as the criterion; and (e) construct validity, with gender serving as discriminant evidence.

## 6. References

- Ambuehl, B., & Inauen, J. (2022). Contextualized Measurement Scale Adaptation: A 4-Step Tutorial for Health Psychology Research. *International Journal of Environmental Research and Public Health*, 19(19), 12775. <https://doi.org/10.3390/IJERPH191912775/S1>
- American Psychology Association. (2023). *APA Dictionary of Psychology*. <https://dictionary.apa.org/creativity>
- Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. B., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *The Journal of Comparative Neurology*, 412(2), 319–341. [https://doi.org/DOI:10.1002/\(sici\)1096-9861\(19990920\)412:2<319::aid-cne10>3.0.co;2-7](https://doi.org/DOI:10.1002/(sici)1096-9861(19990920)412:2<319::aid-cne10>3.0.co;2-7)
- Binkofski, F., & Buccino, G. (2004). Motor functions of the Broca's region. *Brain and Language*, 89(2), 362–369. [https://doi.org/DOI: 10.1016/S0093-934X\(03\)00358-4](https://doi.org/DOI: 10.1016/S0093-934X(03)00358-4)
- Cropley, A. (2020). Definitions of creativity. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of Creativity* (3rd ed., Vol. 1, pp. 315–322). Elsevier.
- De Cassia Nakano, T., Oliveira, K. S., & Zaia, P. (2021). Gender differences in creativity: A systematic literature review. *Psicologia: Teoria e Pesquisa*, 37(4). <https://doi.org/DOI: 10.1590/0102.3772e372116>
- Efrati, Y., Gewirtz-Meydan, & Ateret. (2022). Just talk to them! The importance of teacher-pupil communication about child sexual abuse and harassment as a preventive measure: Data from an Israeli sample of pupils and their homeroom teachers. *Sexuality Research & Social Policy: A Journal of the NSRC*, 20(1), 273–286. <https://doi.org/DOI:10.1007/s13178-022-00778-9>
- Ehrenreich, S. E., George, M. J., Burnell, K., & Underwood, M. K. (2021). Importance of digital communication in adolescents' development: Theoretical and empirical advancements in the last decade. *Journal of Research on Adolescence*, 31(4), 928–943. <https://doi.org/https://doi.org/10.1111/jora.12643>
- Fernández-Monroy, M., Martín-Santana, J. D., & Galván-Sánchez, I. (2018). Building successful franchise partnerships: The importance of communication and trust. *Management Decision*, 56(5), 1051–1064. <https://doi.org/https://doi.org/10.1108/md-07-2016-0528>
- Fink, A., Reim, T., Benedek, M., & Grabner, R. H. (2020). The Effects of a Verbal and a Figural Creativity Training on Different Facets of Creative Potential. *The Journal of Creative Behavior*, 54(3), 676–685. <https://doi.org/10.1002/JOCB.402>
- George, D., & Mallery, P. (2003). *SPSS for Windows Step by Step: A Simple Guide and Reference. 11.0 Update* (4th ed.). Allyn & Bacon.
- Gregory, R. J. (2015). *Psychological testing* (7th ed.). Pearson.
- Harasty, J., Double, K. L., Halliday, G. M., Kril, J. J., & McRitchie, D. A. (1997). Language-associated cortical regions are proportionally larger in the female brain. *Archives of Neurology*, 54(2), 171–176. <https://doi.org/DOI: 10.1001/archneur.1997.00550140045011>
- Harms, M., Kennel, V., & Reiter-Palmon, R. (2017). Team Creativity: Cognitive Processes Underlying Problem Solving. *Team Creativity and Innovation*, 61–86. <https://doi.org/10.1093/OSO/9780190222093.003.0004>
- Kamus Besar Bahasa Indonesia. (n.d.). *Situasional*. In Kamus Besar Bahasa Indonesia (KBBI) Daring. Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan Republik Indonesia. <https://kbbi.kemdikbud.go.id/>
- Kartana, D., Setiawan, N. J., & Suyasa, P. T. Y. S. (2018). *Pengembangan alat ukur kreativitas verbal "C" (studi pada mahasiswa jurusan periklanan)*. *Jurnal Muara Ilmu Sosial, Humaniora, dan Seni*, 2(2), 660–670. <https://doi.org/10.24912/jmishumsen.v2i2.2199>
- Kirby, C. C. (2005). *Experience and Inquiry in John Dewey's Contextualism* [Unpublished master dissertation]. University of South Florida.
- Komarudin, D. (2018). Hubungan antara kreativitas dengan prestasi belajar siswa. *Psymphatic: Jurnal Ilmiah Psikologi*, 4(1), 278–288. <https://doi.org/ DOI: 10.15575/psy.v4i1.2196>
- Kozbelt, A. (2020). Music. In M. Runco & S. Pritzker (Eds.), *Encyclopedia of Creativity* (3rd ed., Vol. 1, pp. 196–201). Elsevier.
- Lin, S.-F. (2018). Discovering young children's musical creativity in their everyday life. In Y. Tsubonou, A.-G. Tan, & M. Oie (Eds.), *Creativity in Music Education* (59–72). Springer.
- Merriam-Webster. (2023). *Contextual Definition & Meaning*. <https://www.merriam-webster.com/dictionary/contextual>
- Munandar, S. C. U. (1977). *Creativity and education: A study of the relationships between measures of creative thinking and a number of educational variables in Indonesian primary and junior secondary schools* [Unpublished master thesis]. University of Indonesia
- Oppong, S., Appiah, R., Hapunda, G., & Kheswa, J. G. (2023). Editorial: Contextualizing psychological assessment in Africa: COVID-19 and beyond. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1150387>
- Osowiecka, M., & Kolanczyk, A. (2018). Let's read a poem! What type of poetry boosts creativity? *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01781>
- Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press.
- Said-Metwaly, S., Van den Noortgate, W., & Barbot, B. (2021). Torrance test of creative thinking-verbal, Arabic version: Measurement invariance and latent mean differences across gender, year of study, and academic major. *Thinking Skills and Creativity*, 39. <https://doi.org/10.1016/J.TSC.2020.100768>
- Scherbakova, A., Dumas, D., Acar, S., Berthiaume, K., & Organisciak, P. (2024). Performance and perception of creativity and academic achievement in elementary school students: A normal mixture modeling study. *The Journal of Creative Behavior*, 58(2), 245–261. <https://doi.org/10.1002/JOCB.646>
- Schlotzhauer, A. E., Ng, M. A., & Su, S. (2025). *How to frame the frame of reference: A comparison of contextualization methods*, 40, 385–403. <https://doi.org/10.1007/s10869-024-09953-8>
- Suyasa, P. T. Y. S., Setyawan, I. R., Rahaditya, R., Jessyca, J., Rae, O. B., & Fahditia, A. (2025). Development of verbal creativity test based on creative industry Sectors. *STI Policy and Management Journal*, 10(1), 91–101. <https://doi.org/10.14203/STIPM.2024.413>
- Walia, C. (2019). A Dynamic Definition of Creativity. *Creativity Research Journal*, 31(3), 237–247. <https://doi.org/10.1080/10400419.2019.1641787>
- Weston, A. (2006). *Creativity for critical thinkers*. Creativity for critical thinkers.

# JPPP

Jurnal Penelitian dan Pengukuran Psikologi  
Fakultas Pendidikan Psikologi  
Universitas Negeri Jakarta

