

Tes Kepribadian *Big Five* dengan *Two-Alternative Forced Choice* pada Karyawan BUMN: Pembuktian Validitas Berdasarkan Struktur Internal

Atsal Nurprayoga¹, Moh. Abdul Hakim²

^{1,2}Fakultas Psikologi, Program Studi Psikologi, Universitas Sebelas Maret
e-mail: [*2m.a.hakim@staff.uns.ac.id](mailto:m.a.hakim@staff.uns.ac.id)

Received: 17th May 2025/ Revised: 19th June 2025/ Accepted: 29th December 2025

Abstract. *Personality assessment plays an important role in maintaining the quality of human resources (HR) within companies. In Indonesia, its implementation still faces challenges, including the limited availability of tools aligned with current developments and the existence of test trading practices. To address these issues, state-owned enterprise (BUMN) PT X developed a Big Five personality test using the two-alternative forced choice (2AFC) method. This study aimed to examine validity evidence based on the test's internal structure using confirmatory factor analysis (CFA). The participants consisted of 2211 BUMN employees at the Officer level, following a response check procedure. Results indicated that the test demonstrated good internal structural validity, shown by acceptable model fit across all dimensions based on Chi-square, SRMR, RMSEA, CFI, and TLI indices. However, residual covariance was identified in four of the five dimensions: openness to experience, conscientiousness, agreeableness, and neuroticism. These findings provide a foundation for the company to validate the use of the test, though regular evaluation is recommended to maintain the test's relevance and accuracy in workplace applications. Further research should review the facets (indicators) and construct structure to enhance the measurement model and address residual covariance.*

Keywords: *confirmatory factor analysis; big five personality; internal structure validity*

Abstrak. Asesmen kepribadian berperan penting dalam menjaga kualitas sumber daya manusia (SDM) di lingkungan perusahaan. Di Indonesia, praktik asesmen kepribadian masih menghadapi sejumlah tantangan, seperti keterbatasan alat ukur yang relevan dengan perkembangan zaman serta adanya praktik jual-beli alat tes. Guna menghadapi tantangan tersebut, perusahaan BUMN PT X mengembangkan alat tes kepribadian *Big Five* dengan metode *two-alternative forced choice* (2AFC). Penelitian ini bertujuan untuk menguji bukti validitas berdasarkan struktur internal alat tes tersebut. Teknik analisis data menggunakan analisis faktor konfirmatori (CFA). Subjek penelitian terdiri dari 2211 karyawan BUMN pada level jabatan Officer setelah melalui prosedur pemeriksaan respons (*response check*). Hasil analisis menunjukkan bahwa alat tes memiliki bukti validitas struktur internal yang baik, ditunjukkan oleh kecocokan model (*model fit*) pada seluruh dimensi berdasarkan indeks Chi-square, SRMR, RMSEA, CFI, dan TLI. Namun, ditemukan *residual covariance* pada empat dari lima dimensi, yaitu *openness to experience*, *conscientiousness*, *agreeableness*, dan *neuroticism*. Temuan ini dapat digunakan perusahaan sebagai dasar validasi penggunaan alat tes, namun tetap perlu dilakukan evaluasi berkala agar hasil tes relevan dan akurat dalam konteks

pekerjaan di lapangan. Penelitian lanjutan perlu meninjau kembali faset (indikator) dan struktur konstruk guna memperbaiki model serta mengatasi *residual covariance*.

Kata kunci: *analisis faktor konfirmatori; big five personality; validitas struktur internal*

Setiap perusahaan termasuk Badan Usaha Milik Negara (BUMN) memerlukan asesmen kepribadian untuk memprediksi perilaku karyawan, meningkatkan kinerja, serta merancang strategi penempatan yang sesuai untuk kesejahteraan kerja dan kepuasan karir (Rothman & Coetzer, 2003; Lounsbury et al., 2004; Ones et al., 2005). Karakteristik kepribadian dapat mempengaruhi keterlibatan kerja, hubungan antar karyawan, dan kemampuan mengelola perusahaan. Penelitian Marlina (2017) menunjukkan bahwa kepribadian dan hubungan antar karyawan berpengaruh signifikan terhadap keterlibatan kerja di BUMN. Selain itu, Riyadiningsih & Sundari (2014) menemukan bahwa tipe kepribadian sangat terkait dengan produktivitas karyawan dan kinerja organisasi. Maka dari itu, asesmen kepribadian menjadi salah satu penilaian yang penting dalam dunia kerja.

Dalam lingkup BUMN, karakteristik kepribadian menjadi salah satu penilaian yang digunakan, misalnya dalam proses manajemen talenta untuk posisi Direksi BUMN. Berdasarkan Peraturan Menteri Badan Usaha Milik Negara Republik Indonesia Nomor PER-3-/MBU/03/2023 tentang “Organ dan Sumber Daya Manusia Badan Usaha Milik Negara”, Pasal 23 mengatur bahwa talenta yang dipilih untuk posisi Direksi harus melalui asesmen yang mencakup penilaian standar kompetensi, kualifikasi profesional (*professional qualification*), dan karakter (*traits*). Asesmen karakter kepribadian meliputi keterbukaan (*Openness*), kesadaran (*Conscientiousness*), ekstrasversi (*Extraversion*), mudah akur atau mudah bersepakat (*Agreeableness*), dan toleransi pada tekanan (*Stress Tolerance*). Karakteristik kepribadian yang digunakan didasarkan pada teori lima faktor kepribadian (*Big Five Personality Traits*) oleh Goldberg (1990). Proses asesmen termasuk metodologi dan hasil, dilaksanakan oleh Lembaga Profesional yang memenuhi syarat administrasi dan syarat teknis yang telah ditetapkan oleh Menteri. Penggunaan penilaian karakteristik kepribadian di BUMN tidak hanya dalam manajemen talenta Direksi, juga bisa pada proses rekrutmen, promosi, dan pengembangan sumber daya manusia, sesuai kebijakan dan kebutuhan masing-masing perusahaan BUMN. Penyesuaian standar penilaian kepribadian dapat disesuaikan dengan kebijakan pada setiap perusahaan BUMN, namun tetap mengacu pada ketentuan yang diatur dalam Peraturan Menteri BUMN.

Di sisi lain, pelaksanaan asesmen kepribadian di Indonesia memiliki tantangan dan hambatan terutama pada alat tes psikologi. Suwartono (2016) menyatakan bahwa berbagai alat tes psikologi yang tersedia tidak sesuai dengan perkembangan zaman, baik dari segi tes maupun standarisasi dari alat tes tersebut. Tantangan lainnya adalah adanya jual-beli alat tes psikologi dan soal psikotes yang tersebar di masyarakat awam (Rahmatullah, 2023). Adanya masalah ini menyebabkan hasil tes yang diberikan tidak mengukur atau memberikan hasil yang sesuai dari yang semestinya. Salah satu upaya untuk mengatasi hambatan tersebut adalah dengan membuat alat tes. Pembuatan alat tes membutuhkan kerja sama dari banyak ahli bidang seperti psikolog, peneliti, *psychometrician*, dan praktisi lapangan (Suwartono, 2020). Keunggulan dari pembuatan alat tes adalah dapat menyesuaikannya sesuai dengan kebutuhan dari apa yang hendak diukur.

Salah satu perusahaan BUMN yaitu PT X mengembangkan alat tes kepribadian yang digunakan untuk rekrutmen, pemetaan, dan promosi jabatan menggunakan teori kepribadian *Big Five*. Teori kepribadian ini menjadi dasar yang banyak digunakan dalam penelitian dan penggunaan asesmen kepribadian karena dapat digunakan dalam konteks lintas budaya. Terdapat penelitian oleh Rollock & Lui (2016) menguji keberlakuan lintas budaya dari model kepribadian *Big Five* dengan menggunakan NEO Five-Factor Inventory (NEO-FFI) pada dua kelompok mahasiswa, yaitu Euro-Amerika dan Internasional Asia. Mengadopsi pendekatan serupa seperti yang dilakukan oleh McCrae & Costa (1997), penelitian ini bertujuan untuk mengetahui apakah struktur lima faktor kepribadian dapat terbentuk secara konsisten pada dua latar budaya yang berbeda. Melalui analisis *Exploratory Structural Equation Modeling* (ESEM), hasilnya menunjukkan bahwa struktur lima faktor terkonfirmasi secara konsisten pada kedua kelompok, yang mengindikasikan bahwa dimensi dasar dalam model *Big Five* dapat dikenali dan dibentuk serupa lintas budaya. Selain itu, studi meta analisis oleh Barrick & Mount (1991) tentang *Big Five* dan kinerja kerja, menunjukkan bahwa *Big Five* berkaitan signifikan dengan berbagai kinerja di berbagai jenis pekerjaan. Sehingga, penggunaan *Big Five* masih digunakan karena mampu memprediksi berbagai hasil penting kehidupan sehari-hari terutama di dunia kerja seperti performa kerja, kepuasan hidup, dan hubungan interpersonal.

Alat tes kepribadian ini unik karena menggunakan metode *two-alternative forced choice* (2AFC) dibandingkan *rating scale* salah satunya adalah skala Likert yang umum digunakan

pada alat tes kepribadian. Akan tetapi terdapat kekurangan dari penggunaan skala Likert adalah rentan terjadi berbagai jenis bias salah satunya adalah *response bias* (bias respons) (Moors et al., 2014). Bias respons adalah jenis kesalahan survei yang terjadi ketika responden secara sengaja atau tidak sengaja mengubah, menyensor, atau memihak baik opini, keyakinan, dan pikiran yang sebenarnya saat menjawab pertanyaan (Graeff, 2005). Akibatnya jawaban yang diberikan responden tidak akurat dalam menggambarkan kondisi responden yang sebenarnya. Sedangkan metode 2AFC merupakan metode dengan format *forced choice* dimana responden dihadapkan pada dua pernyataan dan diminta untuk memilih salah satu yang paling menggambarkan diri responden. Selain itu, 2AFC memiliki kelebihan yaitu tidak rentan terhadap bias respons seperti *acquiescence response style*, *central tendency*, dan *extreme response style* (Brown, 2015; Supratiknya, 2014). Oleh karena itu, metode 2AFC dipilih dalam penggunaan alat tes kepribadian oleh BUMN PT X terutama untuk mengatasi bias respons.

Alat tes kepribadian *Big Five* yang digunakan oleh BUMN PT X, dikembangkan oleh Hakim et al. (2024). Alat tes kepribadian ini telah melalui pembuktian validitas berdasarkan isi oleh pakar (*expert judgment*), uji coba alat tes, dan pembuktian validitas berdasarkan struktur internal. Pembuktian validitas berdasarkan isi dilakukan oleh dua *Subject Matter Expert* (SME) dengan menggunakan rumus Aiken's V menghasilkan skor V 1 setelah revisi pada putaran kedua. Selanjutnya, uji coba alat tes melibatkan 20 mitra peneliti untuk mengumpulkan responden yang tersebar di 33 provinsi Indonesia. Kriteria responden yang ditetapkan adalah Warga Negara Indonesia (WNI) yang berusia minimal 25 tahun dan didapatkan total responden sejumlah 1071 responden. Pembuktian validitas berdasarkan struktur internal dilakukan dengan penguguran butir tes dan uji model menunjukkan kesesuaian (model fit indices) yang memuaskan. Dengan demikian, alat tes kepribadian yang dikembangkan dinyatakan layak dan valid untuk digunakan oleh BUMN PT X. Meskipun alat tes kepribadian yang dikembangkan oleh (Hakim et al., 2024) telah dikembangkan dan digunakan, pembuktian validitas perlu dilakukan secara berkala.

Uji validitas memiliki tujuan untuk melihat seberapa tepat alat pengukuran dengan apa yang hendak diukur. Terdapat penelitian terdahulu yang melakukan pembuktian validitas berdasarkan struktur internal (uji validitas konstruk) pada alat tes kepribadian *Big Five*, diantaranya dilakukan oleh Ahya & Siaputra (2021) melakukan validasi BFI-2 versi Bahasa Indonesia. BFI-1 (*Big Five Inventory*) merupakan alat ukur kepribadian yang berdasar

dari teori kepribadian *Big Five* yang dikembangkan melalui pendekatan leksikal dan dirancang untuk mengukur lima dimensi utama kepribadian. Sedangkan BFI-2 merupakan versi terbaru BFI-1 yang dikembangkan, dimana tidak hanya mengukur pada level dimensi melainkan hingga level faset. Hasil penelitian menunjukkan bahwa analisis CFA model BFI-2 memiliki kesesuaian model yang lebih baik dibandingkan BFI-1. Terdapat juga penelitian oleh Akhtar (2018) bertujuan mengadaptasi skala TIPI (*Ten-Item Personality Inventory*) ke dalam versi Indonesia dan menguji validitas serta reliabilitasnya. TIPI adalah instrumen singkat menggunakan dasar teori *Big Five* dan terdiri dari sepuluh item, masing-masing merupakan pasangan kata sifat (Gosling et al., 2003). Hasil penelitian menunjukkan bahwa TIPI versi Indonesia memiliki karakteristik psikometrik yang dapat diterima, dengan reliabilitas test-retest yang memadai serta validitas konstruk konvergen dan diskriminan yang baik terhadap *Big Five Inventory* (BFI). Alat tes kepribadian seperti BFI-1, BFI-2, TIPI, dan alat tes *Big Five* lainnya umumnya menggunakan skala Likert atau *rating scale* sebagai metode respons. Hingga saat ini, belum ditemukan penelitian yang membuktikan validitas berdasarkan struktur internal terhadap alat tes kepribadian *Big Five* dengan metode *two-alternative forced choice* (2AFC), khususnya pada karyawan BUMN. Oleh karena itu, penelitian ini bertujuan untuk melakukan pembuktian validitas berdasarkan struktur internal terhadap alat tes kepribadian *Big Five* dengan metode 2AFC yang dikembangkan oleh BUMN PT X.

Metode

Partisipan Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari hasil administrasi BUMN PT X. Subjek yang digunakan dalam penelitian ini adalah karyawan BUMN PT X yang bergerak di bidang transportasi pada level jabatan Officer. Teknik pengambilan sampel yang digunakan adalah *non-probability* dengan pendekatan *purposive sampling*. *Purposive sampling* merupakan pendekatan dengan cara menentukan sampel menggunakan pertimbangan tertentu (Amin et al., 2023). Sampel yang digunakan pada penelitian ini merupakan karyawan BUMN PT X sebanyak 3989 karyawan yang berasal dari kantor pusat dan 20 kantor cabang. Setelah dilakukan pemeriksaan respons (*response check*), jumlah data yang digunakan dalam analisis berjumlah 2211 karyawan.

Instrumen Penelitian

Penelitian menggunakan tes kepribadian *Big Five* yang dikembangkan oleh Hakim et al. (2024) dengan metode *two-alternative forced choice* (2AFC). Format 2AFC merupakan format tes yang mengharuskan responden untuk memilih dua pernyataan yang menggambarkan sifat yang bertolak belakang (2 arah). Pernyataan disusun dalam bentuk situasi atau kondisi yang merepresentasikan setiap dimensi dan faset kepribadian yang diukur. Terdapat 5 dimensi atau konstruk yang diukur (*Openness to experience*, *Conscientiousness*, *Extraversion*, *Agreeableness*, dan *Neuroticism*). Salah satu contoh 2 pasang pernyataan adalah pada dimensi *Openness to Experience* yang merepresentasikan keterbukaan individu terhadap hal-hal baru disusun dalam format *two-alternative forced choice* (2AFC). Item tersebut terdiri atas dua pernyataan yang saling berlawanan, yaitu: “Saya senang mencoba hal baru” dan “Saya lebih nyaman dengan hal yang sudah saya kenal.” Responden diminta untuk memilih salah satu pernyataan yang paling sesuai dengan dirinya. Masing-masing dimensi memiliki 6 faset dan masing-masing faset terdiri dari 4 pasang pernyataan. Dengan demikian total secara keseluruhan alat tes kepribadian memiliki 120 pasang pernyataan.

Tabel 1

Daftar Kode Dimensi dan Faset Tes Kepribadian Big Five

Dimensi	Definisi	Faset	Kode Item
<i>Openness to Experience</i>	Dimensi yang mengukur seberapa jauh individu dalam menggali hal baru.	<i>Fantasy</i>	O1
		<i>Intellect</i>	O2
		<i>Aesthetics</i>	O3
		<i>Feelings</i>	O4
		<i>Actions</i>	O5
		<i>Values</i>	O6
<i>Conscientiousness</i>	Dimensi yang mengukur kedisiplinan / keteraturan diri individu.	<i>Competence</i>	C1
		<i>Orderliness</i>	C2
		<i>Dutifulness</i>	C3
		<i>Achievement-Striving</i>	C4
		<i>Self-discipline</i>	C5
		<i>Cautiousness</i>	C6
<i>Extraversion</i>	Dimensi yang mengukur ketertarikan individu	<i>Friendliness</i>	E1
		<i>Gregariousness</i>	E2
		<i>Assertiveness</i>	E3

Dimensi	Definisi	Faset	Kode Item
	berinteraksi dengan individu lain (hubungan interpersonal).	<i>Activity level</i>	E4
		<i>Excitement-seeking</i>	E5
		<i>Cheerfulness</i>	E6
<i>Agreeableness</i>	Dimensi yang mengukur kepercayaan individu terhadap orang lain.	<i>Trust</i>	A1
		<i>Morality</i>	A2
		<i>Cooperation</i>	A3
		<i>Altruism</i>	A4
		<i>Modesty</i>	A5
		<i>Tender-mindedness</i>	A6
<i>Neuroticism</i>	Dimensi yang mengukur stabilitas emosi individu.	<i>Anxiety</i>	N1
		<i>Anger/hostility</i>	N2
		<i>Depression</i>	N3
		<i>Self-Consciousness</i>	N4
		<i>Impulsiveness</i>	N5
		<i>Recklessness</i>	N6

Prosedur Penelitian

Penentuan sampel dalam penelitian ini menggunakan teknik total population sampling, dengan jumlah data awal sebanyak 3989 karyawan BUMN PT X yang mengikuti asesmen kepribadian. Data diambil melalui *computer-based test* (CBT), yang dilaksanakan pada periode 26 Februari hingga 4 Juni 2024 dan melibatkan karyawan dari kantor pusat serta 20 kantor cabang. Teknik pengambilan sampel selanjutnya disesuaikan secara purposif, yaitu hanya menyertakan karyawan yang mengikuti asesmen kepribadian. Proses skoring dilakukan dengan sistem dua pilihan (*two-alternative forced choice*) yang menggunakan kategori jawaban 0 dan 1, di mana peserta memilih salah satu dari dua pernyataan yang paling sesuai dengan dirinya; skor 1 diberikan jika pernyataan sesuai dengan faset yang diukur, dan skor 0 jika sebaliknya. Sebelum analisis data dilakukan, respons peserta diperiksa melalui *response check* untuk menilai konsistensi jawaban dengan menambahkan 10 pasang pernyataan yang tersebar dalam tiap dimensi kepribadian. Hanya data dengan tingkat konsistensi minimal 90% yang digunakan dalam analisis, sedangkan sisanya dieliminasi untuk menjaga validitas data. Selanjutnya, uji validitas berdasarkan struktur internal dilakukan menggunakan teknik *Confirmatory Factor Analysis* (CFA) dengan bantuan perangkat lunak Jamovi versi 2.3.16.

Analisis Data

Penelitian ini menggunakan metode *Confirmatory Factor Analysis* (CFA) untuk menguji validitas berdasarkan struktur internal pada alat tes kepribadian *Big Five* yang dikembangkan oleh Hakim et al. (2024) dan digunakan oleh BUMN PT X. Analisis dilakukan dengan bantuan perangkat lunak Jamovi versi 2.6.13 yang menggunakan teknik estimasi Maximum Likelihood (ML), merupakan metode estimasi umum dalam *Structural Equation Modeling* (SEM) (Hair et al., 2019). Setiap dimensi *Big Five* dianalisis secara terpisah karena secara teoritis bersifat orthogonal (independen) (Saucier, 2002). Penilaian model fit dilakukan melalui beberapa indeks seperti χ^2 (*Chi-Square*), CFI (*Comparative Fit Index*), TLI (*Tucker-Lewis Index*), SRMR (*Standardized Root Mean Square Residual*), dan RMSEA (*Root Mean Square Error of Approximation*), adanya beberapa indeks ini karena χ^2 sensitif terhadap ukuran sampel. Nilai-nilai yang digunakan sebagai acuan kelayakan model adalah sebagai berikut: *factor loading* ≥ 0.30 ; $\chi^2 \geq 0.05$; CFI dan TLI ≥ 0.90 ; SRMR ≤ 0.08 ; serta RMSEA ≤ 0.06 (Hair et al., 2019; Wang & Wang, 2020). Meskipun data yang dihasilkan metode *two-alternative forced choice* (2AFC) bersifat dikotomis (0 dan 1), penjumlahan skor pada tiap faset atau dimensi menghasilkan skor total yang bersifat kuantitatif dan diperlakukan sebagai data interval, karena setiap kenaikan skor mencerminkan peningkatan yang sebanding dalam tingkat karakteristik psikologis, sehingga memungkinkan penggunaan analisis statistik parametrik seperti CFA.

Hasil

Uji *Confirmatory Factor Analysis*

Pembuktian validitas berdasarkan struktur internal dilakukan menggunakan *confirmatory factor analysis* (CFA) yang dianalisis secara terpisah untuk masing-masing dimensi, mengingat model *Big Five* secara teoritis bersifat orthogonal atau independen (Saucier, 2002), sehingga analisis gabungan berisiko menghasilkan korelasi antarkonstruksi yang tidak diharapkan. Hubungan antara dimensi dan faset dalam masing-masing dimensi dianalisis melalui nilai *factor loading*, yang digunakan sebagai salah satu dasar dalam pembuktian validitas struktur internal (Hair et al., 2019; Wang & Wang, 2020), dengan batas nilai *factor loading* ≥ 0.30 dianggap memadai. Penilaian model juga mencakup *overall model fit* dengan sejumlah indeks: χ^2 ($p \geq 0.05$), CFI (≥ 0.90), TLI (≥ 0.90), SRMR (≤ 0.08), dan RMSEA ($\leq$

0.06). Apabila seluruh nilai indeks tersebut memenuhi batas yang ditetapkan, maka model dapat dikatakan memiliki tingkat kecocokan (model fit) yang baik terhadap data.

Namun, hasil CFA awal menunjukkan bahwa empat dari lima dimensi *Big Five* (*openness to experience, conscientiousness, agreeableness, dan neuroticism*) memiliki model yang tidak fit (model misfit). Menurut Brown (2015), model misfit dapat disebabkan oleh jumlah faktor, pemilihan indikator, atau kesalahan pengukuran. Oleh karena itu, dilakukan modifikasi model (*respecification*) untuk meningkatkan model fit, antara lain melalui peninjauan ulang indikator dengan cara penambahan atau penghapusan, serta penambahan korelasi residual (*residual covariances*) berdasarkan *modification indices* (MI). Modifikasi ini harus memenuhi sejumlah ketentuan, seperti setiap dimensi memiliki minimal dua hingga tiga faset, terdapat indikator yang tidak saling berbagi korelasi error, dan setidaknya satu indikator bebas dari korelasi error dengan indikator lain, serta didukung oleh dasar teoritis dan empiris (Kline, 2016).

1. Dimensi *Openness To Experience*

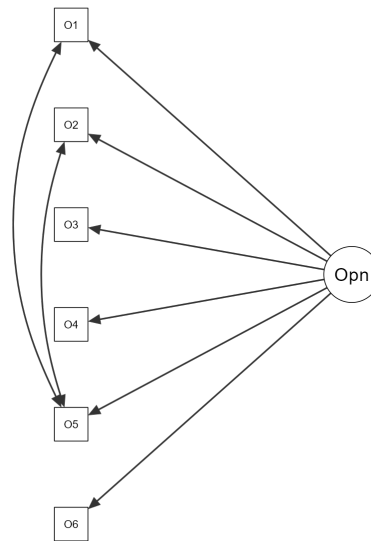
Tabel 2

Uji Confirmatory Factor Analysis Awal Dimensi Openness to Experience

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	p
0.785	0.641	0.044	0.083	143	9	< 0.001

Hasil analisis *confirmatory factor analysis* (CFA) pada dimensi *openness to experience* (lihat Tabel 2) menunjukkan bahwa model tidak fit, dengan nilai $\chi^2 = 143$, $df = 9$, $p < 0.0001$; CFI = 0.785, TLI = 0.641, SRMR = 0.0444, dan RMSEA = 0.0820. Meskipun seluruh faset signifikan ($p < 0.001$), hanya tiga faset yang memenuhi nilai *factor loading* ≥ 0.30 , yakni *fantasy* (O1), *intellect* (O2), dan *actions* (O5), dengan *factor loading* tertinggi pada *actions* (O5) 0.604 dan terendah pada *feelings* (O4) sebesar 0.164. Untuk meningkatkan model fit, dilakukan modifikasi dengan menambahkan *residual covariance* antar pasangan faset O1 dan O5 (MI = 91.34) serta O2 dan O5 (MI = 34.28), tanpa menghapus faset demi menjaga keselarasan konstruk (lihat Gambar 1).

Gambar 1. *Path Diagram Dimensi Openness to Experience*



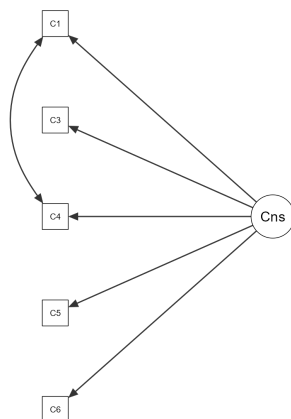
2. Dimensi *Conscientiousness*

Tabel 3

Uji Confirmatory Factor Analysis Awal Conscientiousness

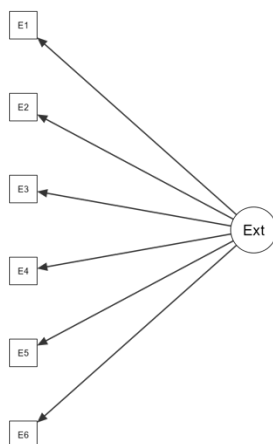
Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	<i>p</i>
0.324	-0.126	0.121	0.216	934	9	< 0.001

Hasil analisis *confirmatory factor analysis* (CFA) pada dimensi *conscientiousness* (lihat Tabel 3) menunjukkan model tidak fit, dengan $\chi^2 = 934$, $df = 9$, $p < 0.001$, CFI = 0.324, TLI = -0.126, SRMR = 0.121, dan RMSEA = 0.216. Dari enam faset, lima signifikan ($p < 0.001$) dan satu tidak signifikan ($p = 0.312$), serta hanya dua faset yang memenuhi batas *factor loading* ≥ 0.30 , yaitu *competence* (C1) dan *achievement-striving* (C4). *Factor loading* tertinggi adalah C4 sebesar 0.600, sedangkan terendah adalah *orderliness* (C2) sebesar 0.066. Modifikasi model dilakukan dengan menghapus faset C2 karena memiliki nilai *factor loading* yang sangat rendah, selain itu juga menambahkan *residual covariance* antara pasangan faset C1 dan C4 (MI = 156.86) sebagaimana ditunjukkan dalam path diagram (lihat Gambar 2).

Gambar 2. *Path Diagram Dimensi Conscientiousness*3. Dimensi *Extraversion***Tabel 4***Uji Confirmatory Factor Analysis Awal Extraversion*

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	<i>p</i>
0.976	0.959	0.032	0.069	104	9	< 0.001

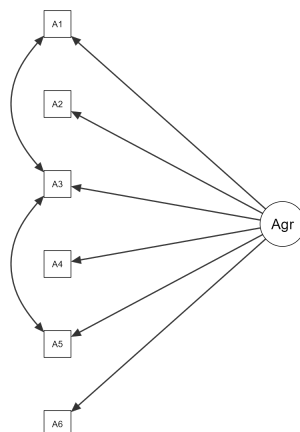
Hasil *confirmatory factor analysis* (CFA) pada dimensi *extraversion* (lihat Tabel 4) menunjukkan model fit, dengan $\chi^2 = 104$, $df = 9$, $p < 0.001$, CFI = 0.976, TLI = 0.959, SRMR = 0.0329, dan RMSEA = 0.0691. Seluruh faset signifikan ($p < 0.001$), dan lima faset memenuhi batas *factor loading* ≥ 0.30 , yaitu E1 hingga E5. *Factor loading* tertinggi terdapat pada faset *gregariousness* (E2) sebesar 0.906 dan terendah pada faset *cheerfulness* (E6) sebesar 0.195. Karena model telah fit, tidak dilakukan modifikasi meskipun satu faset berada di bawah nilai ambang batas.

Gambar 3. *Path Diagram Dimensi Extraversion*

4. Dimensi *Agreeableness***Tabel 5***Uji Confirmatory Factor Analysis Awal Agreeableness*

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	<i>p</i>
0.889	0.814	0.040	0.076	125	9	< 0.001

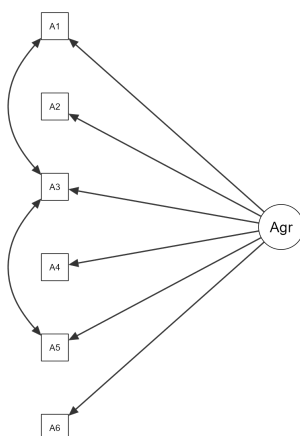
Hasil *confirmatory factor analysis* (CFA) pada dimensi *agreeableness* (lihat Tabel 5) menunjukkan model tidak fit, dengan $\chi^2 = 125$, $df = 9$, $p < 0.001$; CFI = 0.889, TLI = 0.814, SRMR = 0.0407, dan RMSEA = 0.0764. Sebagian besar faset signifikan ($p < 0.001$), kecuali *modesty* (A5) $p = 0.006$, dan empat dari enam faset memenuhi batas *factor loading* ≥ 0.30 , yaitu A1, A2, A4, dan A6. Nilai tertinggi pada *altruism* (A4) sebesar 0.689 dan terendah pada *modesty* (A5) sebesar 0.073. Modifikasi dilakukan tanpa menghapus faset dan dengan menambahkan *residual covariance* pada pasangan faset *cooperation* (A3) dengan A5 (MI = 58.90) dan *trust* (A1) dengan A3 (MI = 46.01) sebagaimana ditunjukkan dalam path diagram (lihat Gambar 4).

Gambar 4. Path Diagram Dimensi *Agreeableness*5. Dimensi *Neuroticism***Tabel 6***Uji Confirmatory Factor Analysis Awal Neuroticism*

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	<i>p</i>
0.909	0.849	0.046	0.095	192	9	< 0.001

Hasil *confirmatory factor analysis* (CFA) pada dimensi *neuroticism* (lihat Tabel 6) menunjukkan model belum fit, dengan $\chi^2 = 192$, $df = 9$, $p < 0.001$, CFI = 0.909, TLI = 0.849, SRMR = 0.0467, dan RMSEA = 0.0958. Seluruh faset signifikan ($p < 0.001$) dan memenuhi batas *factor loading* ≥ 0.30 , dengan nilai tertinggi pada *self-consciousness* (N4) sebesar 0.666 dan terendah pada *recklessness* (N6) sebesar 0.310. Modifikasi dilakukan dengan tidak dilakukan penghapusan faset, namun dilakukan dengan menambahkan *residual covariance* antara faset N5 dengan N6 (MI = 123.68), sebagaimana ditunjukkan dalam path diagram (lihat Gambar 5).

Gambar 5. Path Diagram Dimensi Neuroticism



Uji *Confirmatory Factor Analysis* Pascamodifikasi

Setelah dilakukan modifikasi model melalui penghapusan atau penambahan indikator serta penambahan *residual covariances*, analisis CFA diulang untuk menilai peningkatan indeks kecocokan model (fit indices) dan memastikan representasi konstruk yang lebih sesuai.

1. Dimensi *Openness To Experience*

Tabel 7

Uji CFA Pascamodifikasi Dimensi Openness to Experience

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	p
0.975	0.946	0.017	0.031	22.6	7	0.002

Hasil analisis CFA pascamodifikasi dimensi *openness to experience* (lihat Tabel 7) menunjukkan nilai $\chi^2 = 22.6$, $df = 7$, $p = 0.002$, dengan indeks kecocokan model yang memenuhi kriteria CFI = 0.975, TLI = 0.946, SRMR = 0.017, dan RMSEA = 0.031, sehingga dimensi ini dapat dikatakan memiliki model yang fit. Analisis *factor loading* menunjukkan seluruh faset signifikan ($p < 0.001$), dengan empat faset memenuhi batas minimum (≥ 0.30), yaitu *intellect* (O2), *aesthetics* (O3), *feelings* (O4), dan *values* (O6). *Factor loading* tertinggi terdapat pada faset O3 sebesar 0.501, sedangkan terendah pada faset *actions* (O5) sebesar 0.224.

2. Dimensi *Conscientiousness*

Tabel 8

Uji CFA Pascamodifikasi Dimensi Conscientiousness

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	p
0.982	0.954	0.016	0.035	15.2	4	0.004

Hasil analisis CFA pascamodifikasi dimensi *conscientiousness* (lihat Tabel 8) menunjukkan nilai $\chi^2 = 15.2$, $df = 4$, $p = 0.004$, dengan indeks kecocokan model yang memenuhi kriteria CFI = 0.982, TLI = 0.954, SRMR = 0.016, dan RMSEA = 0.036, sehingga dimensi ini dapat dikatakan memiliki model yang fit. Analisis *factor loading* menunjukkan seluruh faset signifikan ($p < 0.001$), kecuali satu faset yang dieliminasi karena $p = 0.312$. Tiga faset memenuhi batas minimum *factor loading* (≥ 0.30), yaitu *dutifulness* (C3), *self-discipline* (C5), dan *cautiousness* (C6). *Factor loading* tertinggi terdapat pada faset C3 sebesar 0.579, sedangkan terendah pada *competence* (C1) sebesar 0.175.

3. Dimensi *Agreeableness*

Tabel 9

Uji CFA Pascamodifikasi Dimensi Agreeableness

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	p
0.988	0.974	0.014	0.028	19.9	7	0.006

Hasil analisis CFA pascamodifikasi dimensi *agreeableness* (lihat Tabel 9) menunjukkan nilai $\chi^2 = 19.9$, $df = 7$, $p = 0.006$, dengan indeks kecocokan model yang memenuhi kriteria

CFI = 0.988, TLI = 0.974, SRMR = 0.014, dan RMSEA = 0.028, sehingga dimensi ini dapat dikatakan memiliki model yang fit. Analisis *factor loading* menunjukkan lima faset signifikan ($p < 0.001$), kecuali *modesty* (A5) dengan $p = 0.081$. Empat faset memenuhi batas *factor loading* (≥ 0.30), yaitu *trust* (A1), *morality* (A2), *cooperation* (A3), *altruism* (A4), dan *tender-mindedness* (A6). *Factor loading* tertinggi terdapat pada faset A4 sebesar 0.707, sedangkan terendah pada A5 sebesar 0.045.

4. Dimensi *Neuroticism*

Tabel 10

Uji CFA Pascamodifikasi Dimensi Neuroticism

Fit Measure					Exact Fit	
CFI	TLI	SRMR	RMSEA	χ^2	df	p
0.970	0.944	0.027	0.058	68.2	8	< 0.001

Hasil analisis CFA pascamodifikasi dimensi *neuroticism* (lihat Tabel 10) menunjukkan nilai $\chi^2 = 68.2$, $df = 8$, $p = 0.006$, dengan indeks kecocokan model yang memenuhi kriteria CFI = 0.970, TLI = 0.944, SRMR = 0.027, dan RMSEA = 0.058, sehingga model dapat dikatakan fit. Seluruh faset signifikan ($p < 0.001$) dan memenuhi batas *factor loading* (≥ 0.30), dengan nilai tertinggi pada *self-consciousness* (N4) sebesar 0.666 dan terendah pada *recklessness* (N6) sebesar 0.310.

Diskusi

Penelitian ini bertujuan untuk melakukan pembuktian validitas berdasarkan struktur internal alat tes kepribadian *Big Five* dengan *two-alternative forced choice* pada sampel karyawan BUMN, menggunakan analisis *Confirmatory Factor Analysis* (CFA). Hasil penelitian menunjukkan bahwa seluruh dimensi pada alat tes kepribadian *Big Five* memenuhi pembuktian validitas berdasarkan struktur internal dengan analisis CFA setelah dilakukan modifikasi model. Hal ini selaras dengan temuan Hilliard et al. (2022), yang mengembangkan alat ukur berbasis gambar dan membuktikan keberadaan lima dimensi kepribadian melalui validitas konvergen terhadap IPIP-NEO-120. Meskipun menggunakan pendekatan berbeda, kedua penelitian menunjukkan konsistensi struktur lima faktor, memperkuat validitas konstruk teori *Big Five*. Temuan ini juga didukung oleh penelitian Akhtar & Azwar (2019) yang mengadaptasi IPIP-BFM-50 versi Indonesia dan melaporkan bahwa lima dimensi kepribadian memiliki struktur faktor yang sesuai melalui analisis eksploratori. Demikian

pula, Ahya & Siaputra (2021) dalam validasi BFI-2 pada mahasiswa Indonesia menunjukkan bahwa masing-masing dimensi memiliki reliabilitas dan validitas konstruk yang baik. Secara keseluruhan, kesamaan temuan ini memperkuat bukti bahwa struktur lima dimensi *Big Five* tetap stabil termasuk dalam konteks pengukuran kepribadian pada karyawan BUMN menggunakan format *two-alternative forced choice*. Penelitian juga menunjukkan analisis *factor loading* pascamodifikasi menunjukkan bahwa sebagian besar faset signifikan ($p < 0.001$) dan memenuhi batas ≥ 0.30 (Wang & Wang, 2020), dengan beberapa faset menunjukkan kontribusi kuat (≥ 0.50) dalam merepresentasikan dimensi yang diukur. Seperti faset *aesthetics* (O3) pada dimensi *openness to experience* (0.501), *dutifulness* (C3) pada *conscientiousness* (0.579), *gregariousness* (E2) pada *extraversion* (0.906), *altruism* (A4) pada *agreeableness* (0.707), dan *self-consciousness* (N4) pada *neuroticism* (0.666).

Ditemukan *residual covariances* pada keempat dimensi pada analisis awal yang menyebabkan model tidak fit. *Residual covariance* mengacu pada residual dua variabel yang tidak dapat dijelaskan sepenuhnya dan menunjukkan adanya hubungan tambahan antara dua faset atau lebih dalam satu dimensi. Menurut Cole et al. (2007), mengabaikan *residual covariance* akibat desain dapat menyebabkan kesalahan spesifikasi model dan menghasilkan variabel laten yang tidak akurat dalam merepresentasikan konstruk yang diukur. Analisis *residual covariance* menunjukkan adanya hubungan tambahan antar faset dalam beberapa dimensi. Pada dimensi *openness to experience*, ditemukan hubungan antara faset *fantasy* (O1) dan *actions* (O5), serta antara faset *intellect* (O2) dan *actions* (O5), yang mencerminkan bahwa pemikiran baru yang muncul dari imajinasi (*fantasy*) dan kecerdasan dapat memengaruhi tindakan individu. Pada dimensi *conscientiousness*, terdapat antara faset *competence* (C1) dan orientasi pada pencapaian (*achievement-striving*, C4), menunjukkan bahwa individu yang kompeten cenderung lebih mampu mencapai tujuan yang ditetapkan. Pada dimensi *agreeableness*, hubungan ditemukan antara faset *cooperation* (A3) dengan *trust* (A1) serta dengan *modesty* (A5), mengindikasikan bahwa kerja sama yang efektif memerlukan rasa saling percaya dan sikap rendah hati antar individu. Sementara itu, pada dimensi *neuroticism*, terdapat hubungan antara faset *impulsiveness* (N5) dan *recklessness* (N6), yang menunjukkan bahwa individu yang impulsif lebih rentan bertindak ceroboh.

Penggunaan metode *two-alternative forced choice* (2AFC) pada alat tes kepribadian *Big Five* bertujuan mengurangi bias respons yang umum terjadi pada skala Likert. Namun

demikian, metode ini memiliki beberapa keterbatasan. Menurut Brown (2015), format *forced choice* menyulitkan interpretasi skor absolut karena hanya merepresentasikan preferensi relatif antar pernyataan, bukan tingkat aktual dari konstruk yang diukur. Meskipun pendekatan penskoran modern seperti Thurstonian IRT dapat digunakan untuk memetakan preferensi tersebut ke dalam skala laten, kualitas pengukuran sangat dipengaruhi oleh desain butir. Jika pasangan pernyataan tidak cukup kontras atau tidak seimbang, estimasi menjadi tidak akurat. Selain itu, karakteristik struktural *forced choice* seperti *circular dependence* (ketergantungan antar butir) dan *trade-off respons* (pertukaran preferensi) juga dapat menurunkan konsistensi jawaban. Penggunaan metode 2AFC menuntut perancangan item dan analisis yang cermat agar menghasilkan pengukuran yang valid. Oleh karena itu, dalam penyusunan pasangan pernyataan, item-item dirancang agar saling bertolak belakang namun tetap merepresentasikan karakteristik masing-masing faset, guna meminimalkan *circular dependence* dan *trade-off respons* yang melekat pada format 2AFC.

Temuan penelitian ini menunjukkan bahwa tiap dimensi pada alat tes kepribadian *Big Five* dengan metode *two-alternative forced choice* terbukti mengukur lima dimensi yang hendak diukur. Tiap dimensi memiliki model fit yang ditunjukkan dari indeks ketepatan model *overall model fit* (Chi-square, SRMR, RMSEA, CFI, dan TLI). Analisis *factor loading* untuk melihat kontribusi faset yang dominan juga dilakukan. Kemudian, ditemukan *residual covariance* yang menyebabkan adanya misfit. Penelitian ini memiliki keterbatasan belum mengeksplorasi lebih lanjut penyebab timbulnya residual covariances. Sehingga penelitian selanjutnya disarankan untuk meninjau ulang validitas indikator (faset) dan struktur konstruk pada penelitian selanjutnya dan melakukan uji validitas konstruk kembali.

Kesimpulan

Hasil penelitian menunjukkan bahwa alat tes kepribadian *Big Five* yang digunakan oleh BUMN PT X memiliki validitas berdasarkan struktur internal yang baik setelah dilakukan modifikasi model (*model respecification*). Seluruh dimensi menunjukkan model fit yang memadai berdasarkan indikator. Dengan demikian, format *two-alternative forced choice* yang dikembangkan dapat diandalkan dalam konteks asesmen kepribadian karyawan di lingkungan BUMN. Analisis juga mengungkapkan bahwa setiap dimensi memiliki satu faset dominan dengan kontribusi tertinggi terhadap variabel laten, antara lain: *aesthetics* (O3) pada

dimensi *openness to experience* (0.501), *dutifulness* (C3) pada *conscientiousness* (0.579), *gregariousness* (E2) pada *extraversion* (0.906), *altruism* (A4) pada *agreeableness* (0.707), dan *self-consciousness* (N4) pada *neuroticism* (0.666). Meski demikian, terdapat beberapa faset dengan *factor loading* di bawah ideal dan ditemukannya *residual covariances* pada empat dari lima dimensi, yang menunjukkan adanya korelasi antar indikator yang belum sepenuhnya dijelaskan oleh konstruk.

Saran

Berdasarkan temuan tersebut, disarankan agar penelitian lanjutan meninjau kembali validitas indikator dan struktur konstruk untuk meningkatkan presisi pemodelan dan mengidentifikasi potensi korelasi residual yang masih tersembunyi. Bagi BUMN PT X, hasil ini dapat dijadikan dasar bahwa alat tes yang digunakan telah terbukti valid dalam mengukur lima dimensi kepribadian. Namun, perusahaan tetap disarankan melakukan evaluasi berkala terhadap hasil tes serta menyesuaikan interpretasinya dengan kebutuhan kerja di lapangan, mengingat keberadaan *residual covariance* dapat memengaruhi ketepatan interpretasi hasil asesmen.

Referensi

- Ahya, A., & Siaputra, I. B. (2021). Validasi big five inventory-2 (BFI-2) untuk indonesia: Belum sempurna tetapi valid dan reliabel mengukur kepribadian. *Jurnal Psikologi Ulayat*. <https://doi.org/10.24854/jpu458>
- Akhtar, H. (2018). Translation and validation of the ten-item personality inventory (TIPI) into bahasa indonesia. *International Journal of Research Studies in Psychology*, 7(2). <https://doi.org/10.5861/ijrsp.2018.3009>
- Akhtar, H., & Azwar, S. (2019). Indonesian adaptation and psychometric properties evaluation of the big five personality inventory: IPIP-BFM-50. *Jurnal Psikologi*, 46(1), 32. <https://doi.org/10.22146/jpsi.33571>
- Amin, N. F., Garancang, S., & Abunawas, K. (2023). *Konsep umum populasi dan sampel dalam penelitian*. 14, 15–31. <https://journal.unismuh.ac.id/index.php/pilar/article/view/10624>
- Rahmatullah, B. (2023). Analisis pelanggaran kode etik psikologi (Studi Kasus Putusan Nomor 463/Pdt.G/2013/PN.JKT.SEL). *JUSTICES: Journal of Law*, 2(2), 72–83. <https://doi.org/10.58355/justices.v2i2.72>
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44 (1), 1–26. <https://doi.org/https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>

- Brown, A. (2015). Personality assessment, forced-choice. In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (pp. 840–848). Elsevier Inc. <https://doi.org/10.1016/B978-0-08-097086-8.25084-8>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, 12(4), 381–398. <https://doi.org/10.1037/1082-989X.12.4.381>
- Goldberg, L. R. (1990). An alternative “Description of personality”: The big-five factor structure. *Journal of Personality and Social Psychology*, 59 (6), 1216–1219. <https://doi.org/https://doi.org/10.1037/0022-3514.59.6.1216>
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)
- Graeff, T. R. (2005). Response bias. *Encyclopedia of Social Measurement*, 3(Dalam K.Kempf-Leonard (Ed.)), 411–418. <https://doi.org/https://doi.org/10.1016/B0-12-369398-5/00037-2>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning. www.cengage.com/highered
- Hakim, M. A., Kusuma, N. K. L., Prasetyani, R. N. N., Pradana, R. G., Anggrani, F. K., Saniatuzzulda, R., & et al. (2024). *Laporan final pekerjaan: Pengembangan alat ukur dan sistem pendukung asesmen kompetensi PT X*.
- Hilliard, A., Kazim, E., Bitsakis, T., & Leutner, F. (2022). Measuring Personality through Images: Validating a Forced-Choice Image-Based Assessment of the Big Five Personality Traits. *Journal of Intelligence*, 10(1). <https://doi.org/10.3390/jintelligence10010012>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.
- Lounsbury, J. W., Park, S. H., Sundstrom, E., Williamson, J. M., & Pemberton, A. E. (2004). Personality, career satisfaction, and life satisfaction: Test of a directional model. *Journal of Career Assessment*, 12(4), 395–406. <https://doi.org/10.1177/1069072704266658>
- Marlina, N. (2017). Pengaruh kepribadian dan hubungan antar karyawan terhadap keterlibatan kerja karyawan pada bumh pt taspen (persero) cabang pontianak dan pt jasa raharja (persero) cabang pontianak. *Jurnal Manajemen Update*. <https://jurnal.untan.ac.id/index.php/ejmfe/article/view/20662>
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509–516. <https://doi.org/10.1037/0003-066X.52.5.509>
- Moors, G., Kieruj, N. D., & Vermunt, J. K. (2014). The effect of labeling and numbering of response scales on the likelihood of response bias. *Sociological Methodology*, 44(1), 369–399. <https://doi.org/10.1177/0081175013516114>
- Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Personality at work: Raising awareness and correcting misconceptions. In *Human Performance* (Vol. 18, Issue 4, pp. 389–404). https://doi.org/10.1207/s15327043hup1804_5

- Riyadiningsih, H., & Sundari S. (2014). Tipe kepribadian karyawan dalam meningkatkan kinerja organisasi. *Dalam Proceedings of Seminar Nasional Dan Call for Paper (Sancal 2014) : Research Methods and Organizational Studies*, 388–396.
- Rollock, D., & Lui, P. P. (2016). Measurement invariance and the five-factor model of personality: Asian international and euro american cultural groups. *Assessment*, 23(5), 571–587. <https://doi.org/10.1177/1073191115590854>
- Rothman, S., & Coetzer, E. P. (2003). The Big Five personality dimensions and job performance. *Journal of Industrial Psychology*, 29 (1), 68–74. <https://doi.org/https://doi.org/10.4102/sajip.v29i1.88>
- Saucier, G. (2002). Orthogonal markers for orthogonal factors: The case of the Big Five. *Journal of Research in Personality*, 36(1), 1–31. <https://doi.org/10.1006/jrpe.2001.2335>
- Supratiknya, A. (2014). *Pengukuran psikologis*. Universitas Sanata Dharma.
- Suwartono, C. (2020). Alat tes psikologi konteks Indonesia: Tantangan psikologi di era MEA. *Jurnal Psikologi Ulayat*, 3(1), 1–6. <https://doi.org/10.24854/jpu37>
- Wang, J., & Wang, X. (2020). *Structural equation modelin: Applications using Mplus* (2nd ed.). Wiley.