

Psychometric Properties and Norming of Learning Agility Index: Online Assessment to Measure Fluid Intelligence

Novita Sari

Wahyu Maulana Firdaus

Aswin Januarsjaf

PT Global Talentlytica Indonesia, Jakarta

<https://doi.org/10.24071/suksma.v5i1.7700>

Naskah Masuk 14 Oktober 2023 Naskah Diterima 15 Desember 2023 Naskah Dipublikasikan 31 Januari 2024

Abstract. The Learning Agility Index (LAI) is an online assessment used to assess Fluid Intelligence, the pure intellectual speed and power evaluated by the ability to solve new problems creatively. Using a battery of five sub-tests, LAI can deliver a global trainability quotient (GTQ) indicating how well an individual performs in a training context. This study aims to analyze psychometric properties and norming for LAI. The research data comprised 28,980 subjects from the online test database for employees selection owned by PT Global Talentlytica Indonesia. The research approach used in this study is classical test theory. Psychometric property analysis utilized Cronbach's alpha reliability estimation method and collected validity evidence based on internal structure. The reliability analysis results showed that the five sub-tests had strong reliabilities with Cronbach's alpha values ranging from 0.92 to 0.97. Confirmatory factor analysis confirmed the measurement model of the Learning Agility Index with CFI 0.988, RMSEA 0.022, and SRMR 0.20. The intelligence levels represented by the measurement results were divided into five levels, ranging from "below average" to "far above average." These findings prove that LAI is a reliable and valid online assessment for measuring fluid intelligence in the context of employees selection or evaluation.

Keywords: fluid intelligence; global trainability quotient; psychometric properties; norming; learning agility index

Pendahuluan

Salah satu faktor penting yang mempengaruhi keberhasilan suatu organisasi adalah kinerja karyawannya (Bowen & Ostroff, 2004). Proses rekrutmen karyawan baru menjadi penting untuk dapat

Korespondensi Penulis

Novita Sari, PT Global Talentlytica Indonesia, Jakarta

Email: novitasari.prof@gmail.com

menyeleksi calon karyawan yang memiliki potensi kinerja yang baik di masa depan. Metode yang digunakan dalam seleksi karyawan harus mempertimbangkan kriteria yang dapat memprediksi performa karyawan di masa depan. *General Mental Ability* (GMA) ditemukan konsisten dan valid dalam memprediksi kinerja karyawan (Barros et al., 2014; Salgado & Moscoso, 2019; Tews et al., 2010). Penggunaan GMA dalam proses seleksi dapat dikombinasikan dengan metode lainnya seperti tes integritas dan wawancara terstruktur untuk mendapatkan validitas yang lebih tinggi (Schmidt et al., 2016; Schmidt & Hunter, 1998). Alat ukur yang biasa digunakan dalam proses rekrutmen untuk mengukur GMA adalah tes intelegensi.

Intelegensi merupakan salah satu aspek kunci yang digunakan untuk mengevaluasi potensi individu dalam berbagai konteks, termasuk dalam setting industri. Kecerdasan manusia dibagi menjadi dua yaitu *fluid intelligence* dan *crystallized intelligence*. Tes intelegensi seringkali didasarkan pada dua jenis kecerdasan tersebut. *Crystallized intelligence* merupakan bentuk kecerdasan yang termanifestasi dari sejumlah besar perilaku yang terkait dengan luasnya pengetahuan, pengalaman, kecanggihan, penilaian, keterampilan komunikasi, pemahaman konvensi, dan kapasitas berpikir yang masuk akal (Horn, 1988). Sedangkan *fluid intelligence* merupakan indikasi keterampilan memahami hubungan di antara pola-pola stimulus, menarik kesimpulan dari hubungan, dan memahami implikasinya. Perbedaan antara *crystallized intelligence* dan *fluid intelligence* mungkin dapat menjadi perhatian khusus dalam prosedur seleksi personel dan dalam prediksi kinerja pekerjaan (Beauducel & Kersting, 2002).

Intelligence quotient (IQ) merupakan komponen penting dalam memprediksi prestasi kerja karyawan, tinggi rendahnya IQ berkaitan dengan peningkatan atau penurunan kinerja individu dan organisasi (Murtza et al., 2020). Temuan mengungkapkan bahwa IQ merupakan prediktor kinerja kerja dan kepuasan kerja karyawan. *Crystallized intelligence* terkait erat dengan pengetahuan dan keterampilan yang dipengaruhi Pendidikan dan pengalaman. Sedangkan *fluid intelligence* terkait dengan kemampuan seseorang dalam memahami pola-pola stimulus yang dapat menggambarkan kecepatan seseorang mengumpulkan dan memproses pengetahuan. Menurut Beauducel & Kersting (2002) *crystallized intelligence* dapat menjadi indikator penting mengenai pengetahuan terkait pekerjaan di masa depan. Jika demikian, *fluid intelligence* dapat digunakan untuk mengidentifikasi seberapa cepat seseorang dapat mempelajari dan mempertahankan keterampilan dan prosedur baru (TST Validation Overview, 2006).

PT Global Talentlytica Indonesia telah mengembangkan LAI, sebuah *online assessment* untuk mengukur *fluid intelligence*. PT Global Talentlytica Indonesia merupakan perusahaan penyedia platform online assessment berbasis cloud untuk kebutuhan *Talent Management* terkait *recruitment*, *assessment center*, *online interview* dan konsultasi strategis. Pengembangan LAI merupakan hasil modifikasi dari *Tests for Selection and Training* (TST Validation Overview, 2006). Aspek yang diukur bukanlah akumulasi pengetahuan atau keterampilan, namun menghasilkan *global trainability quotient* (sebagai ukuran kemampuan untuk dilatih) dan laporan tentang area potensi pengembangan individu. LAI terdiri dari lima subtes yang mengukur *speed and accuracy*, *reasoning*, *focus and attention*, *numerical ability* dan *technical problem solving*.

LAI dikembangkan untuk menjawab kebutuhan pemetaan potensi karyawan maupun keperluan analitik sumber daya manusia yang melibatkan aspek potensi individu, khususnya aspek kognitif. PT Global Talentlytica Indonesia menggunakan LAI untuk mengukur *fluid intelligence* dalam proses seleksi karyawan. Hasil uji properti psikometri menggunakan sekitar 26.000 data menunjukkan bahwa LAI merupakan alat ukur yang valid dan reliabel (Talentlytica, 2020). Nilai koefisien Alpha sebesar 0.764 menunjukkan nilai konsistensi internal yang tinggi. Hasil uji CFA menunjukkan bahwa LAI memiliki model fit dengan nilai RMSEA sebesar 0.065 dibawah 0.08, CFI=0.984, TLI=0.950, keduanya memenuhi batas penerimaan akurasi model yaitu 0.95 (Hu & Bentler, 1999).

Penggunaan LAI untuk proses seleksi, sudah lebih dari 100 ribu calon karyawan sejak dikembangkan. Tes yang baik adalah yang dapat mengukur apa yang dirancang untuk diukur seakurat mungkin (Friedenberg, 1995). Properti psikometri tes yang baik adalah reliabel, valid dan memiliki statistik item yang baik. Seiring dengan bertambahnya waktu dan juga subjek yang mengerjakan LAI, karakteristik populasi dapat berubah. Oleh karena itu, uji properti psikometris ulang dan pembuatan norma diperlukan untuk memastikan bahwa tes tetap valid dan relevan untuk populasi saat ini.

Penelitian ini bermaksud untuk melakukan pengujian properti psikometri LAI menggunakan data terbaru yang dimiliki PT. Global Talentlytica Indonesia. Hal ini penting untuk dilakukan untuk mengumpulkan bukti bahwa LAI merupakan alat ukur yang akurat dapat diandalkan dalam mengukur *fluid intelligence* khususnya dalam proses seleksi karyawan. Penggunaan alat ukur yang akurat dan dapat diandalkan dalam proses seleksi karyawan krusial untuk memastikan objektivitas dan efisiensi yang pada akhirnya mendukung pertumbuhan dan pengembangan organisasi. Selain itu penormaan skor LAI juga akan dilakukan untuk melihat level kecerdasan yang dapat diklasifikasikan oleh LAI. Norma skor digunakan untuk membedakan level kecerdasan individu yang dapat dijadikan sebagai salah satu kriteria dalam proses seleksi karyawan. PT Global Talentlytica melakukan pembaharuan norma secara berkala untuk memastikan bahwa norma tersebut tetap relevan untuk populasi saat ini.

Penormaan skor LAI akan membantu dalam mengklasifikasikan level kecerdasan individu. Dengan memiliki kerangka acuan yang jelas, perusahaan dapat memilih calon yang sesuai dengan persyaratan posisi yang ada, memungkinkan efisiensi dalam proses seleksi karyawan. Hasil dari penelitian ini dapat dijadikan acuan dalam meningkatkan kualitas seleksi karyawan, meningkatkan efisiensi proses rekrutmen, dan memberikan manfaat positif bagi perusahaan serta calon karyawan yang berpotensi.

Metode Penelitian

Partisipan penelitian

Penelitian ini melibatkan sampel sebanyak 28,980 data hasil tes LAI yang dipilih dengan cermat dari database PT. Global Talentlytica Indonesia. Pemilihan dilakukan melalui proses *cleaning* pada data

outliers dan data yang memiliki demografi tidak lengkap. Sampel ini mencakup beragam karakteristik yang mencerminkan keragaman populasi tenaga kerja di Indonesia. Rentang usia partisipan dalam penelitian ini berkisar antara 17 hingga 63 tahun, mencakup berbagai tahap karir dan pengalaman kerja. Sampel terdiri dari 16,573 laki-laki (57.1%) dan 12,407 perempuan (42.9%). Data partisipan diperoleh dari seluruh 34 provinsi di Indonesia, menciptakan keterwakilan geografis yang luas. Selain itu, partisipan penelitian memiliki tingkat pendidikan yang beragam, mulai dari pendidikan menengah (SMA) hingga tingkat pendidikan tertinggi (S3). Data pribadi individu dalam penelitian ini dijaga kerahasiaannya sesuai dengan standar etika penelitian dan peraturan privasi yang berlaku.

Instrumen Penelitian

LAI merupakan instrumen penting dalam penelitian ini, dirancang untuk mengukur *fluid intelligence* dalam konteks seleksi karyawan (Talentlytica, 2020). LAI terdiri dari lima subtes yang masing-masing mengukur berbagai aspek kecerdasan yang relevan untuk evaluasi kandidat.

Subtes pertama terdiri dari 60 item (waktu 3 menit), mengukur aspek *speed and accuracy* yaitu seberapa cepat individu dapat bekerja dalam tugas-tugas tertentu serta kemampuan mereka untuk memberikan jawaban yang tepat dalam waktu yang ditentukan. Setiap item terdiri dari empat pasang stimulus, dimana setiap pasang stimulus berisi satu huruf kapital/balok (besar) dan satu huruf kecil. Terdapat lima pilihan jawaban berupa angka 0,1,2,3 dan 4. Subjek diminta untuk menentukan ada berapa pasang stimulus yang memiliki pasangan huruf kapital dan huruf kecil yang sesuai.

Subtes kedua terdiri dari 50 item (waktu 4 menit), mengukur aspek *reasoning* yaitu kemampuan individu dalam menarik kesimpulan dari informasi yang diberikan, menyimpan informasi dalam memori jangka pendek, dan menyelesaikan masalah setelah diberikan instruksi. Hasil yang tinggi pada subtes ini menunjukkan penalaran verbal yang baik. Setiap item terdiri dari *stem* dan opsi jawaban. *Stem* berisikan dua pernyataan perbandingan dan satu pertanyaan. Terdapat tiga pilihan jawaban untuk menjawab pertanyaan tersebut yang disesuaikan dengan pernyataan perbandingan.

Subtes ketiga terdiri dari 72 item (waktu 4 menit), mengukur aspek *focus and attention* yaitu kemampuan individu dalam memusatkan perhatian mereka pada stimulus target dalam jangka waktu tertentu. Setiap item hanya berupa opsi jawaban. Opsi berupa 3 buah huruf yang diletakkan secara berurutan. Jarak antara huruf pertama dan kedua, merentang dari satu sampai empat huruf. Begitu juga jarak antara huruf kedua dan ketiga. Subjek diminta untuk memilih huruf yang memiliki jarak yang lebih banyak.

Subtes keempat terdiri dari 60 item (waktu 4 menit), mengukur aspek *numerical ability* yaitu sejauh mana seorang individu dapat bekerja dengan konsep-konsep kuantitatif. Aspek ini penting untuk pekerjaan yang memerlukan pemahaman dan manipulasi data numerik. Setiap item hanya berupa opsi jawaban berupa 3 buah angka (1-99). Jarak antara angka pertama dan kedua merentang dari 1 sampai 15, begitu juga jarak antara angka kedua dan ketiga. Subjek diminta untuk memilih angka yang memiliki jarak lebih banyak.

Subtes kelima terdiri dari 60 item (waktu 5 menit), mengukur aspek *technical problem solving* yaitu kemampuan individu untuk mempelajari hal baru dengan cepat dan menerapkannya dalam pemecahan masalah teknis yang kompleks. Setiap item terdiri dari *stem* dan opsi. *Stem* berupa tiga pasang bentuk huruf kapital dimana setiap pasang huruf memiliki dua kemungkinan pola, yaitu keduanya sama atau keduanya tidak sama. Terdapat empat pilihan jawaban berupa angka 0,1,2 dan 3. Subjek diminta untuk menentukan ada berapa pasang stimulus yang memiliki pasangan huruf kapital dengan pola yang sama.

Waktu total yang diberikan untuk menyelesaikan tes LAI adalah 20 menit, dengan alokasi waktu yang berbeda untuk masing-masing subtes. Instrumen ini dapat diadministrasikan secara individu atau dalam kelompok, *online* maupun *offline*, sesuai dengan preferensi dan kebutuhan penelitian. Namun saat ini penggunaannya lebih sering secara *online* melalui platform PT Global Talentlytica Indonesia.

Prosedur analisis data

Analisis data dalam penelitian ini akan mencakup serangkaian tahapan untuk mengevaluasi properti psikometri dari LAI dengan pendekatan teori tes klasik. Pertama, akan dilakukan analisis item mencakup penilaian tingkat kesulitan dan daya beda masing-masing item LAI. Tingkat kesulitan merupakan perbandingan dari jumlah jawaban benar pada suatu item dengan jumlah peserta tes (Azis, 2016). Semakin tinggi nilai tingkat kesulitan bermakna semakin banyak peserta yang dapat menjawab item tersebut atau dapat dikatakan item semakin mudah dijawab. Sedangkan daya beda merupakan kemampuan suatu item dalam membedakan kemampuan kelompok peserta dengan skor tinggi dengan kelompok yang mendapat skor rendah.

Selanjutnya, reliabilitas internal instrumen akan diukur menggunakan metode estimasi reliabilitas *alpha cronbach*. Hal ini akan memberikan gambaran tentang konsistensi internal LAI, yaitu sejauh mana item-item dalam alat ukur ini mengukur konsep yang sama. Koefisien *alpha* merupakan batas bawah dari koefisien reliabilitas yang dikenal sebagai koefisien presisi (Crocker & Algina, 2008).

Selain itu, untuk mengevaluasi validitas struktur internal LAI, *Confirmatory Factor Analysis* (CFA) akan digunakan. CFA akan membantu mengidentifikasi apakah struktur faktorial yang diajukan dalam LAI sesuai dengan data yang diamati. Kriteria *goodness of fit* untuk menyimpulkan adanya kecocokan yang relatif baik antara model yang dihipotesiskan dengan data observasi adalah fungsi kecocokan minimum *chi-square* $p > 0,05$; nilai batas mendekati 0,95 untuk TLI, CFI, dan RNI; nilai batasnya mendekati 0,08 untuk SRMR; dan nilai batas mendekati 0,06 untuk RMSEA (Hu & Bentler, 1999).

Proses analisis data ini akan dilakukan menggunakan perangkat lunak JASP versi 0.18.1.0. Hasil dari analisis data ini akan menjadi landasan utama untuk memahami kualitas instrumen LAI dan sejauh mana alat ukur ini dapat diandalkan dalam mengukur *fluid intelligence* calon karyawan. Keseluruhan analisis data ini akan membantu dalam memvalidasi LAI sebagai alat yang akurat dan dapat diandalkan dalam konteks seleksi karyawan.

Penetapan Norma LAI

Proses penetapan norma LAI melibatkan serangkaian langkah yang sistematis untuk memberikan kerangka acuan dalam mengategorikan dan memahami tingkat kecerdasan individu. Berikut adalah langkah-langkah penormaan yang digunakan dalam penelitian ini (Talentlytica, 2020):

- a. Perhitungan Jumlah Benar dan Salah
- b. Kalkulasi Skor dilakukan dengan memperhitungkan jawaban salah. Akan ada pengurangan skor jika jawaban salah.
- c. Perhitungan *z score* setiap subtes: *Z score* dihitung untuk masing-masing subtes sebagai ukuran sejauh mana skor individu berbeda dari rata-rata sampel dalam satuan standar deviasi.
- d. Perhitungan *z-score* GTQ: *Z score* GTQ dihitung dengan mengambil rata-rata *z score* dari semua subtes. Ini menciptakan skor yang mencerminkan tingkat kecerdasan individu secara keseluruhan.
- e. Mean dan Standar Deviasi: Nilai Mean dan Standar Deviasi (SD) untuk perhitungan *z-score* diperoleh dari analisis deskriptif data sampel (Tabel 1).
- f. Perhitungan GTQ per Subtes dan GTQ total: GTQ dihitung dengan rumus (*z score* * 15) + 100. Ini mengubah skor Z menjadi skor GTQ dalam rentang 55 hingga 145.
- g. Pengelompokan Skor GTQ Final: Skor GTQ Final dibulatkan, jika skor GTQ kurang dari 55, maka ditetapkan sebagai 55, dan jika lebih dari 145, maka ditetapkan sebagai 145.
- h. Penentuan Kategori GTQ total (penormaan) dan GTQ per sub tes.

Proses penentuan norma ini memberikan landasan kuat dalam mengategorikan kecerdasan individu dan memahami sejauh mana mereka sesuai dengan kebutuhan posisi yang ada dalam organisasi. Kategori-kategori GTQ ini memberikan panduan yang berharga dalam pengambilan keputusan seleksi karyawan yang lebih akurat dan efektif.

Tabel 1.

Mean dan Standar Deviasi setiap Subtes

Variabel	Mean	SD
Subtes 1	28,24	11.23
Subtes 2	13,14	6,44
Subtes 3	25,16	11,55
Subtes 4	18,35	11.39
Subtes 5	13,01	13,54

Hasil Penelitian dan Pembahasan

Data Demografi

Data demografi responden penelitian di sajikan pada tabel 2. Dari tabel 2 dapat disimpulkan bahwa mayoritas responden adalah laki-laki (57,19%), sementara 42,81% adalah Perempuan. Sebagian besar responden berusia antara 21 hingga 30 tahun (72,86%), mencerminkan kelompok usia yang dominan dalam populasi penelitian. Namun, ada juga sejumlah kecil responden di luar kelompok usia ini, seperti yang berusia di bawah 20 tahun (2,05%) atau di atas 50 tahun (1,47%). Data ini dapat dijadikan salah satu acuan untuk membaca dan menginterpretasikan skor LAI. Dari segi pendidikan, mayoritas responden memiliki tingkat pendidikan Strata 1 (68,87%), dengan sebagian lainnya adalah Diploma (13,66%). Tingkat pendidikan Strata 2 dan Strata 3 memiliki persentase yang lebih kecil, masing-masing sekitar 5,20% dan 0,34%. Hal ini menunjukkan bahwa Sebagian besar calon karyawan yang mengikuti seleksi merupakan lulusan sederajat strata 1. Terakhir, asal provinsi responden juga disajikan, data diperoleh dari 34 provinsi di Indonesia dengan sebagian besar berasal dari DKI Jakarta, Jawa Barat, dan Jawa Timur. Pemahaman karakteristik demografi ini penting dalam penelitian ini, karena perbedaan dalam jenis kelamin, usia, pendidikan, dan asal provinsi dapat memengaruhi hasil penelitian dan generalisabilitas temuan. Dengan demikian, penting untuk mempertimbangkan dampak faktor-faktor demografi ini dalam analisis dan interpretasi hasil penelitian.

Tabel 2.

Karakteristik demografi responden

Variabel	Kategori	Frekuensi	Persentase
Jenis Kelamin	Laki-laki	16573	57,19%
	Perempuan	12407	42,81%
	Total	28980	100,00%
Usia	≤20 tahun	595	2,05%
	21-30	21115	72,86%
	31-40	5170	17,84%
	41-50	1673	5,77%
	≥51	427	1,47%
	Total	28980	100,00%
Pendidikan	SMA/SMK/MA	3458	11,93%
	Diploma	3958	13,66%
	Strata 1	19958	68,87%
	Strata 2	1508	5,20%
	Strata 3	98	0,34%
	Total	28980	100,00%

Provinsi	Aceh	169	0,58%
	Bali	266	0,92%
	Banten	2206	7,61%
	Bengkulu	107	0,37%
	Daerah Istimewa		
	Yogyakarta	663	2,29%
	DKI Jakarta	4685	16,17%
	Gorontalo	34	0,12%
	Jambi	206	0,71%
	Jawa Barat	6901	23,81%
	Jawa Tengah	2123	7,33%
	Jawa Timur	3313	11,43%
	Kalimantan Barat	301	1,04%
	Kalimantan Selatan	314	1,08%
	Kalimantan Tengah	125	0,43%
	Kalimantan Timur	553	1,91%
	Kalimantan Utara	42	0,14%
	Kepulauan Bangka Belitung	56	0,19%
	Kepulauan Riau	213	0,73%
	Lampung	352	1,21%
	Maluku	59	0,20%
	Maluku Utara	22	0,08%
	Nusa Tenggara Barat	136	0,47%
	Nusa Tenggara Timur	174	0,60%
	Papua	113	0,39%
	Papua Barat	30	0,10%
	Riau	1688	5,82%
	Sulawesi Barat	20	0,07%
	Sulawesi Selatan	568	1,96%
	Sulawesi Tengah	100	0,35%
	Sulawesi Tenggara	84	0,29%
	Sulawesi Utara	205	0,71%
	Sumatera Barat	422	1,46%
	Sumatera Selatan	830	2,86%
	Sumatera Utara	1900	6,56%
	Total	28980	100,00%

Analisis item

Tabel 3 menyajikan hasil analisis item-item LAI persubtes. Subtes 1 memiliki nilai tingkat kesulitan berkisar antara 0.64 hingga 0.96 dengan nilai rata-rata sekitar 0.84. Nilai tingkat kesulitan yang relatif tinggi ini mengindikasikan bahwa sebagian besar peserta cenderung berhasil menjawab item-item dalam subtes ini. Sementara itu, daya beda item berkisar antara 0.28 hingga 0.812 dengan nilai rata-rata sekitar 0.50. Daya beda yang moderat ini menggambarkan sebagian besar item dalam subtes ini memiliki kemampuan membedakan antara peserta yang memiliki tingkat kemampuan yang berbeda.

Subtes 2 memiliki nilai tingkat kesulitan berkisar antara 0.50 hingga 0.92 dengan nilai rata-rata sekitar 0.72. Ini menunjukkan bahwa sebagian besar peserta mungkin akan berhasil menjawab item-item dalam subtes ini, meskipun nilai tingkat kesulitan lebih bervariasi dibandingkan dengan Subtes 1. Daya beda item berkisar antara 0.24 hingga 0.58 dengan nilai rata-rata sekitar 0.45. Daya beda ini menunjukkan bahwa beberapa item mungkin memiliki daya beda yang lebih rendah, sehingga perlu diperhatikan untuk pengembangan lebih lanjut.

Subtes 3 memiliki nilai tingkat kesulitan berkisar antara 0.63 hingga 0.95 dengan nilai rata-rata sekitar 0.81. Ini menunjukkan bahwa subtes ini memiliki nilai tingkat kesulitan yang relatif tinggi, dan sebagian besar peserta mungkin akan berhasil menjawab item dalam subtes ini. Namun, daya beda item berkisar antara 0.29 hingga 0.95 dengan nilai rata-rata sekitar 0.41. Ini mengindikasikan bahwa beberapa item mungkin memiliki daya beda yang lebih rendah, sehingga perlu dievaluasi lebih lanjut.

Subtes 4 memiliki nilai tingkat kesulitan berkisar antara 0.18 hingga 0.914 dengan nilai rata-rata sekitar 0.66. Meskipun tingkat kesulitan lebih bervariasi, sebagian besar peserta mungkin akan mampu menjawab item-item dalam subtes ini. Namun, perlu diperhatikan bahwa beberapa item mungkin memiliki nilai tingkat kesulitan yang rendah. Daya beda item berkisar antara (-)0.01 hingga 0.65 dengan nilai rata-rata sekitar 0.40. Ini menunjukkan bahwa beberapa item dalam subtes ini mungkin kurang efektif dalam membedakan antara peserta dengan tingkat kemampuan yang berbeda.

Subtes 5 memiliki nilai tingkat kesulitan berkisar antara 0.28 hingga 0.73 dengan nilai rata-rata sekitar 0.51. Tingkat kesulitan yang moderat ini menunjukkan bahwa sebagian besar peserta dapat berhasil menjawab item-item dalam subtes ini. Daya beda item berkisar antara 0.32 hingga 0.67 dengan nilai rata-rata sekitar 0.51. Daya beda yang relatif tinggi ini mengindikasikan bahwa item-item dalam subtes ini efektif dalam membedakan antara peserta dengan tingkat kemampuan yang berbeda.

Dalam keseluruhan analisis, perlu diperhatikan bahwa beberapa subtes memiliki item-item dengan daya beda yang rendah (Ebel dalam Crocker & Algina, 2008). Hal ini dapat menjadi pertimbangan untuk mengembangkan atau merevisi item-item tersebut agar tes dapat lebih efektif dalam mengukur kemampuan peserta dengan tingkat yang lebih baik. Selain itu, hasil analisis ini juga dapat membantu dalam mengevaluasi kualitas tes dan penggunaannya dalam konteks evaluasi atau pengukuran yang diinginkan.

Tabel 3.*Hasil analisis item LAI*

Variabel	Tingkat Kesulitan	Daya Beda
Subtes 1	0.64-0.96 (Mean=0.84)	0.28-0.812 (Mean=0.50)
Subtes 2	0.50-0.92 (Mean=0.72)	0.24-0.58 (Mean=0.45)
Subtes 3	0.63-0.95 (Mean=0.81)	0.29-0.95 (Mean=0.41)
Subtes 4	0.18-0.914 (Mean=0.66)	(-)0.01-0.65 (Mean=0.40)
Subtes 5	0.28-0.73 (Mean=0.51)	0.32-0.67 (Mean=0.51)

Reliabilitas

Nilai *alpha Cronbach* setiap subtes disajikan pada table 4. Nilai *alpha Cronbach* kelima subtes >0.90. Hasil ini menunjukkan bahwa semua subtes memiliki tingkat reliabilitas sangat baik (Dancey & Reidy, 2017). Ini berarti bahwa item-item dalam semua subtes memiliki konsistensi internal yang kuat dan dapat diandalkan untuk mengukur konstruk yang sama. Tingkat reliabilitas yang tinggi mengindikasikan bahwa perbedaan dalam skor yang ditemukan antara peserta dalam subtes sebagian besar adalah akibat sebenarnya dari perbedaan dalam kemampuan, bukan karena ketidakpastian dalam instrumen pengukuran.

Dalam penelitian atau penggunaan tes, reliabilitas yang tinggi sangat penting, karena memastikan bahwa hasil tes adalah konsisten dan dapat diandalkan. Hasil analisis ini memberikan dukungan kuat untuk keandalan pengukuran. Sebagai hasilnya, hasil dari subtes ini dapat lebih dipercayai dan digunakan secara efektif dalam evaluasi dan pengambilan keputusan.

Tabel 4.*Nilai Alpha Cronbach*

Variabel	Alpha Cronbach
Subtes 1	0.97
Subtes 2	0.94
Subtes 3	0.94
Subtes 4	0.92
Subtes 5	0.96

Validitas

Dalam penelitian ini, peneliti melakukan analisis CFA untuk mengumpulkan bukti validitas berdasarkan struktur internal dari alat ukur yang digunakan, yaitu LAI. Hasil analisis menunjukkan nilai *Chi-Square* sebesar 74 dengan $df=5$ dan nilai signifikan (p) kurang dari 0.001. Meskipun *Chi-Square* adalah ukuran kesesuaian model yang signifikan secara statistik, dalam sampel besar, nilai ini dapat menunjukkan signifikansi statistik bahkan untuk perbedaan kecil antara model yang diusulkan

dan model yang diamati . Jumlah sampel yang besar cenderung mempengaruhi hasil estimasi menjadi signifikan sehingga diartikan sebagai model yang tidak fit (Schumacker & Lomax, 2012). Oleh karena itu, hasil Chi-Square perlu ditafsirkan bersama dengan indikator model fit lainnya.

Nilai *Comparative Fit Index* (CFI) adalah 0.998. CFI mengukur sejauh mana model yang diusulkan sesuai dengan data yang diamati. Nilai CFI mendekati 1 mengindikasikan bahwa model tersebut memiliki tingkat kesesuaian yang sangat baik dengan data. Nilai *Root Mean Square Error of Approximation* (RMSEA) adalah 0.022. RMSEA mengukur sejauh mana model cocok dengan data dengan memperhitungkan kesalahan aproksimasi. Nilai RMSEA yang rendah, seperti yang dihasilkan dalam analisis ini, mengindikasikan bahwa model memiliki tingkat kesalahan yang rendah dalam menggambarkan hubungan antara variabel-variabel dalam konstruk yang diukur (Hu & Bentler, 1999). Nilai *Standardized Root Mean Square Residual* (SRMR) adalah 0.02. SRMR juga mengukur kesesuaian model dengan data, dengan nilai yang lebih rendah menunjukkan kesesuaian yang lebih baik. Hasil SRMR yang rendah ini menunjukkan bahwa model CFA yang diuji memiliki kesesuaian yang sangat baik dengan data yang ada (Hu & Bentler, 1999).

Hasil analisis CFA menunjukkan bahwa model yang digunakan untuk mengukur validitas struktur internal LAI sangat sesuai dengan data yang diamati (Hu & Bentler, 1999). Nilai *Chi-Square*, meskipun signifikan secara statistik, harus ditafsirkan dengan hati-hati dalam sampel besar seperti dalam penelitian ini, dan penekanan harus diberikan pada indikator model fit lainnya. Nilai CFI yang mendekati 1 (0.998), RMSEA yang rendah (0.022), dan SRMR yang rendah (0.02) mengindikasikan bahwa model CFA yang diajukan sangat sesuai dengan data dan mampu dengan baik mengukur konstruk yang diinginkan, yaitu LAI. Hasil ini menunjukkan bahwa LAI adalah alat ukur yang valid dalam mengukur struktur internal yang dimaksud.

Norming

Hasil penormaan skor LAI (GTQ) disajikan dalam Tabel 6 berdasarkan rumus penormaan yang tercantum di Tabel 5, dengan nilai mean (μ) sebesar 100 dan deviasi standar (σ) sebesar 15. Norma skor tersebut memiliki kategori berdasarkan rentang nilai. Penormaan skor LAI (GTQ) ini bertujuan untuk memberikan interpretasi yang lebih jelas terhadap hasil tes. Dengan menggunakan nilai mean (μ) sebesar 100 dan deviasi standar (σ) sebesar 15, penormaan ini memungkinkan untuk mengkategorikan skor peserta ke dalam tingkatan prestasi yang berbeda.

Penting untuk dipahami bahwa penormaan ini mengacu pada distribusi skor dalam sampel referensi yang digunakan, di mana mean dan deviasi standar digunakan sebagai titik referensi. Skor di bawah 100 dianggap sebagai "Dibawah rata-rata," sementara skor di atas 100 dikategorikan sebagai "Diatas rata-rata." Hal ini membantu dalam memberikan gambaran tentang sejauh mana skor peserta berada dalam kaitannya dengan sampel referensi.

Tabel 5.

Rumus penormaan skor LAI (GTQ)

Rentang Nilai	Kategori
$X \leq (\mu - (1.5 \times \sigma))$	Dibawah rata-rata
$(\mu - (1.5 \times \sigma)) < X \leq (\mu - (0.5 \times \sigma))$	Rata-rata bawah
$(\mu - (0.5 \times \sigma)) < X \leq (\mu + (0.5 \times \sigma))$	Rata-rata
$(\mu + (0.5 \times \sigma)) < X \leq (\mu + (1.5 \times \sigma))$	Diatas rata-rata
$(\mu + (1.5 \times \sigma)) < X$	Jauh diatas rata-rata

Hasil penormaan skor seperti ini dapat berguna dalam interpretasi hasil tes. Misalnya, jika seorang peserta memiliki skor 110, itu berarti bahwa skornya di atas rata-rata. Sebaliknya, jika skor peserta adalah 70, maka peserta tersebut dianggap berada di bawah rata-rata dalam konteks sampel referensi yang digunakan. Penting untuk diingat bahwa penormaan skor harus selalu didasarkan pada sampel referensi yang relevan dan sesuai dengan tujuan pengukuran. Norma skor yang digunakan dapat berbeda untuk setiap jenis tes atau alat ukur dan harus dipertimbangkan secara cermat dalam interpretasi hasil.

Tabel 6.

Norma skor LAI (GTQ)

Rentang Nilai	Kategori
55-77	Dibawah rata-rata
78-92	Rata-rata bawah
93-107	Rata-rata
108-122	Diatas rata-rata
122-145	Jauh diatas rata-rata

Kesimpulan dan Saran

Hasil penelitian menunjukkan bahwa LAI merupakan *online assessment* yang memiliki properti psikometri yang baik. Hasil analisis item LAI memberikan wawasan penting terkait dengan konstruksi tes. Subtes-subtes dalam LAI menunjukkan variasi dalam tingkat kesulitan dan daya beda item. Subtes 1 dan Subtes 3, misalnya, memiliki tingkat kesulitan yang relatif tinggi, sementara Subtes 2 dan Subtes 4 memiliki tingkat kesulitan yang lebih bervariasi. Hal ini mencerminkan keragaman dalam item-item yang digunakan dalam tes. Sementara itu, daya beda item-item dalam subtes juga bervariasi, dengan beberapa item yang memiliki daya beda yang lebih rendah. Hasil ini menyoroti kebutuhan untuk mengevaluasi dan mungkin merevisi item-item yang memiliki daya beda rendah untuk meningkatkan

efektivitas tes.

Analisis reliabilitas dengan menggunakan *alpha Cronbach* menghasilkan tingkat reliabilitas yang sangat baik untuk semua subtes LAI, dengan nilai alpha Cronbach di atas 0.90. Hasil ini menunjukkan bahwa item-item dalam semua subtes memiliki konsistensi internal yang kuat dan dapat diandalkan untuk mengukur konstruk yang sama. Ini memberikan keyakinan dalam keandalan pengukuran.

Validitas instrumen ukur sangat penting dalam penelitian. Hasil analisis CFA menunjukkan bahwa model konfirmatori yang digunakan dalam pengukuran struktur internal LAI sangat sesuai dengan data yang diamati. Indikator model fit, seperti CFI, RMSEA, dan SRMR, menunjukkan tingkat validitas yang tinggi. Hasil ini memiliki implikasi penting dalam konteks penelitian. Hasil analisis CFA ini memperkuat dasar empiris penelitian dan memberikan keyakinan dalam validitas pengukuran yang digunakan.

Penormaan skor LAI (GTQ) memberikan kerangka kerja untuk menginterpretasikan skor peserta. Dengan menggunakan nilai mean dan deviasi standar, skor peserta dapat dikategorikan ke dalam level kecerdasan yang berbeda. Hasil penormaan ini membantu dalam memberikan pemahaman yang lebih jelas tentang sejauh mana skor peserta berada dalam kaitannya dengan sampel referensi yang digunakan. Generalisabilitas dan intrepertasi temuan harus mempertimbangkan pada pemahaman karakteristik demografi seperti perbedaan jenis kelamin, usia, pendidikan, dan asal provinsi dapat memengaruhi hasil penelitian.

Dengan demikian, kesimpulan dan analisis yang telah dilakukan dalam penelitian ini memberikan pemahaman yang komprehensif terkait dengan data demografi peserta, karakteristik item tes, reliabilitas, validitas, dan penormaan skor. Semua elemen ini penting dalam menginterpretasikan hasil tes dan memastikan keandalan serta validitas instrumen pengukuran yang digunakan dalam konteks seleksi karyawan atau evaluasi kinerja.

Mengingat variasi dalam tingkat kesulitan dan daya beda item-item dalam subtes, disarankan untuk melakukan evaluasi lebih lanjut terhadap item-item yang memiliki daya beda rendah. Item-item tersebut dapat direvisi atau diganti untuk meningkatkan efektivitas tes. Selanjutnya, penting untuk mempertimbangkan dampak karakteristik demografi responden dalam analisis data. Analisis sub kelompok berdasarkan jenis kelamin, usia, pendidikan, dan asal provinsi dapat memberikan wawasan yang lebih dalam tentang pengaruh faktor-faktor demografi terhadap hasil penelitian.

Analisis validitas dan reliabilitas yang sangat baik memberikan keyakinan dalam pengukuran. Namun, penggunaan tes ini dalam konteks aplikasi praktis juga perlu dievaluasi lebih lanjut untuk memastikan efektivitasnya dalam mengukur kemampuan kandidat dalam situasi sebenarnya. Hasil penormaan skor dapat digunakan dalam interpretasi hasil tes dan pengambilan keputusan. Penting untuk memahami bahwa penormaan ini didasarkan pada sampel referensi tertentu, dan hasil tes harus selalu dianalisis dengan mempertimbangkan konteks penggunaan dan sampel yang relevan.

Ucapan terima kasih: Peneliti mengucapkan terima kasih kepada tim Talent Scientist PT Global Talentlytica Indonesia dan PT Global Talentlytica Indonesia atas kerjasama, kontribusi dan dukungan dalam penelitian ini. Dedikasi dan dukungan Anda telah menjadi faktor kunci kesuksesan penelitian ini. Semoga kerja sama yang baik ini terus berlanjut untuk pencapaian yang lebih besar di masa depan. Terima kasih.

Pendanaan: Penelitian ini didukung secara finansial oleh PT Global Talentlytica Indonesia. Dukungan keuangan yang diberikan oleh PT Global Talentlytica Indonesia telah menjadi fondasi yang memungkinkan penelitian, penulisan, dan publikasi artikel ini. Peneliti menghargai kontribusi yang sangat berharga ini dalam mewujudkan penelitian ini. Semoga kerjasama yang baik ini dapat terus berlanjut untuk penelitian-penelitian selanjutnya.

Konflik kepentingan: Peneliti dengan tulus menyatakan bahwa tidak ada konflik kepentingan yang relevan terkait dengan penelitian ini. Tidak ada konflik kepentingan keuangan, profesional, atau hubungan pribadi yang dapat memengaruhi objektivitas atau integritas penelitian ini. Kami melakukan penelitian ini dengan itikad baik dan transparansi penuh.

Daftar Acuan

- Azis, A. (2016). Analisis tes buatan guru bidang studi matematika kelas v sd 1 katobengke. *Edumatica: Jurnal Pendidikan Matematika*, 6(1), 15–24.
<https://doi.org/10.22437/edumatica.v6i01.2998>
- Barros, E., Kausel, E. E., Cuadra, F., & Díaz, D. A. (2014). Using general mental ability and personality traits to predict job performance in three chilean organizations. *International Journal of Selection and Assessment*, 22(4), 432–438. <https://doi.org/10.1111/ijsa.12089>
- Beauducel, A., & Kersting, M. (2002). Fluid and crystallized intelligence and the berlin model of intelligence structure (BIS). *European Journal of Psychological Assessment*, 18(2), 97–112.
<https://doi.org/10.1027//1015-5759.18.2.97>
- Bowen, D. E., & Ostroff, C. (2004). Understanding hrm-firm performance linkages: the role of the “strength” of the hrm system. *Academy of Management Review*, 29(2), 203–221.
<https://doi.org/10.5465/AMR.2004.12736076>
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Cengage Learning.
- Dancey, C., & Reidy, J. (2017). *Statistics without maths for psychology* (seventh ed). Pearson Education Limited.
- Friedenberg, L. (1995). *Psychological testing: design, analysis, and use*. Allyn and Bacon.
- Horn, J. (1988). *Handbook of multivariate experimental psychology* (Second Ed, pp. 645–682). Springer

- US. https://doi.org/10.1007/978-1-4613-0893-5_2
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Murtza, M. H., Gill, S. A., Aslam, H. D., & Noor, A. (2020). Intelligence quotient, job satisfaction, and job performance: The moderating role of personality type. *Journal of Public Affairs, Special Issue Paper*, 1–12. <https://doi.org/10.1002/pa.2318>
- Salgado, J. F., & Moscoso, S. (2019). Meta-analysis of the validity of general mental ability for five performance criteria: Hunter and hunter (1984) revisited. *Frontiers in Psychology*, 10(OCT), 1–18. <https://doi.org/10.3389/fpsyg.2019.02227>
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274. <https://doi.org/10.1037/0033-2909.124.2.262>
- Schmidt, F. L., Oh, I.-S., & Shaffer, J. A. (2016). The validity and utility of selection methods in personnel psychology: Practical and theoretical Implications of 100 Years. *Fox Business School, Working Pa*(October), 1-. <https://doi.org/10.13140/RG.2.2.18843.26400>
- Schumacker, R., & Lomax, R. (2012). *A beginner's guide to structural equation modeling*. Routledge. <https://doi.org/10.4324/9780203851319>
- Talentlytica. (2020, n.d.). *Laporan Pengujian Validitas dan Reliabilitas TOP*. PT. Global Talentlytica Indonesia. Retrieved from: www.talentlytica.com
- Tews, M. J., Michel, J. W., & Lyons, B. D. (2010). Beyond personality: The impact of GMA on performance for entry/level service employees. *Journal of Service Management*, 21(3), 344–362. <https://doi.org/10.1108/09564231011050797>
- Thomas International. (2006, n.d.). *TST Validation Overview*. (2006). Thomas International Organization. Retrieved from: www.thomasinternational.net.