

Nomor: 143/BK/07/S/2024

**EVALUASI PROPERTI PSIKOMETRI *INTELLIGENZ-STRUKTUR-TEST* VERSI
2005 SUBTES *MERKAUFGABEN* (ME) MENGGUNAKAN *CLASSICAL TEST
THEORY***

SKRIPSI

diajukan untuk memenuhi sebagian syarat memperoleh gelar Sarjana Pendidikan
Bidang Ilmu Bimbingan dan Konseling



oleh:

Nabila Az-zahra

2000180

PROGRAM STUDI SARJANA BIMBINGAN DAN KONSELING

FAKULTAS ILMU PENDIDIKAN

UNIVERSITAS PENDIDIKAN INDONESIA

2024

LEMBAR HAK CIPTA

**EVALUASI PROPERTI PSIKOMETRI *INTELLIGENZ-STRUKTUR-TEST*
VERSI 2005 SUBTES *MERKAUFGABEN (ME)* MENGGUNAKAN
*CLASSICAL TEST THEORY***

oleh

Nabila Az-zahra

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Pendidikan pada Bidang Keilmuan Bimbingan dan Konseling

© Nabila Az-zahra

Universitas Pendidikan Indonesia

Juli 2024

Hak Cipta dilindungi Undang-Undang.

Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian,
dengan dicetak ulang, menyalin, atau cara lainnya tanpa izin penulis.

Nabila Az-zahra, 2024

***EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES
MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY***

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

LEMBAR PENGESAHAN

NABILA AZ-ZAHRA
NIM 2000180

EVALUASI PROPERTI PSIKOMETRI *INTELLIGENZ-STRUKTUR-TEST* VERSI 2005 SUBTES *MERKAUFGABEN (ME)* MENGGUNAKAN *CLASSICAL TEST* *THEORY*

disetujui dan disahkan oleh

Pembimbing I



Dr. Eka Sakti Yudha, M.Pd.
NIP. 198308292010121004

Pembimbing II



Dr. Sudaryat Nurdin Akhmad, M.Pd.
NIP. 196306301995121001

Mengetahui

Ketua Program Studi Bimbingan dan Konseling
Fakultas Ilmu Pendidikan
Universitas Pendidikan Indonesia



Dr. Ipah Saripah, M.Pd.
NIP. 197710142001122001

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi dengan judul "**Evaluasi Properti Psikometri *Intelligenz-Struktur-Test* versi 2005 subtes *Merkaufgaben* menggunakan Teori Klasik**" beserta seluruh isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi apabila di kemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Bandung, 23 Juli 2024



Nabila Az-zahra
NIM 2000180

KATA PENGANTAR

Upaya untuk memahami kecerdasan manusia telah lama menjadi subjek yang menarik bagi para akademisi. *Intelligenz-Struktur-Test* versi 2005 dalam domain ini muncul sebagai instrumen yang signifikan, khususnya subtes *merkaufgaben* (ME), yang dirancang dengan cermat untuk menyelidiki konstruksi memori yang tidak terpisahkan dari proses kognitif. Kebenaran dari kesimpulan yang diambil dari subtes ini didasarkan pada ketangguhan psikometriknya sebuah atribut yang diperhatikan oleh *classical test theory* (CTT). CTT, dengan konsepsi dikotomisnya tentang skor yang diamati-memisahkan skor yang benar dari skor kesalahan-memberikan latar belakang analitis yang ketat untuk evaluasi ini. CTT menyatakan bahwa skor yang diamati (X) adalah jumlah dari skor sebenarnya (T) dan skor kesalahan (E), yang secara matematis direpresentasikan sebagai $X=T+E$. Kerangka teori vital dalam mengkalibrasi tes untuk mencapai reliabilitas dan validitas, memastikan bahwa tes tersebut tidak hanya mencerminkan konstruk yang diminati, tetapi juga tahan terhadap perubahan-perubahan kesalahan pengukuran.

Peneliti berusaha untuk melakukan evaluasi sistematis terhadap properti psikometri IST-2005 subtes ME melalui lensa CTT. Penyelidikan ilmiah ini dimotivasi oleh keharusan untuk memperkuat reliabilitas subtes dan memvalidasi konstruksinya secara ketat, dengan demikian menegaskan kapasitasnya untuk secara akurat mengukur apa yang diukur subtes. Dorongan untuk penelitian ini berakar pada komitmen untuk memajukan ilmu psikometri, memperjuangkan pengembangan alat asesmen yang kuat secara empiris dan sempurna secara metodologis.

Selain itu, skripsi ini tidak hanya mencerminkan ketertarikan penulis pada bidang psikometri dan pengukuran psikologis tetapi juga komitmen penulis sebagai calon sarjana program studi Bimbingan dan Konseling untuk mengaplikasikan kemampuan dalam asesmen psikologis akan prinsip-prinsip pengukuran dalam praktik bimbingan dan konseling. Penulis menyadari bahwa seorang guru BK memegang peranan penting dalam menginterpretasikan hasil tes psikologis untuk mendukung pengembangan dan optimalisasi potensi siswa.

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Penyusunan skripsi ini merupakan langkah awal dalam perjalanan akademis penulis di Universitas Pendidikan Indonesia, dan penulis berharap bahwa hasil karya ini dapat memperkaya khasanah ilmu Bimbingan dan Konseling serta membawa manfaat yang luas bagi Pendidikan di Indonesia.

UCAPAN TERIMA KASIH

Segegap syukur peneliti haturkan kepada Allah SWT, Tuhan semesta alam. Berkat anugerah dan kasih sayang-Nya, peneliti berhasil menuntaskan skripsi yang berjudul “**Evaluasi Properti Psikometri *Intelligenz-Struktur-Test* versi 2005 subtes *Merkaufgaben* menggunakan Teori Klasik**”. Selama proses penyelesaian, peneliti dihadapkan pada situasi yang membutuhkan lebih dari sekadar keahlian dan pengetahuan. Dalam momen-momen kritis tersebut, dukungan tiada henti peneliti terima dari berbagai pihak, yang tidak hanya memberikan semangat tetapi juga sumber daya yang substansial, telah menjadi fondasi yang memungkinkan peneliti untuk melangkah lebih jauh dan mengatasi rintangan dengan keteguhan hati. Oleh karena itu, dengan rasa hormat dan rasa syukur yang mendalam, peneliti ingin menyampaikan rasa terima kasih yang setinggi-tingginya kepada:

1. Para Dosen Pembimbing, yaitu Dr. Eka Sakti Yudha, M. Pd., dan Drs. Sudaryat Nurdin Akhmad, M.Pd., yang dengan penuh dedikasi dan kesabaran, telah memberikan waktu, arahan, dukungan, dan motivasi selama proses penyusunan skripsi berlangsung.
2. Dr. Ipah Saripah, M.Pd selaku Ketua Program Studi Bimbingan dan Konseling yang telah memberikan bantuan dan kemudahan berupa memberikan izin kepada peneliti untuk melaksanakan penelitian.
3. Dr. Setiawati, M.Pd., selaku Kepala Laboratorium Bimbingan dan Konseling dan para stafnya, yang dengan murah hati telah menyediakan dukungan finansial dan data hasil skor testi yang pernah mengikuti IST rentang 2018-2024, memungkinkan penelitian ini dilaksanakan dengan standar yang tinggi.
4. Bapak dan Ibu Dosen Program Studi Bimbingan dan Konseling FIP UPI, yang telah membimbing dan menyebarkan ilmu yang dimiliki dari awal perkuliahan, hingga selesainya penulisan skripsi.
5. Staf Tata Usaha Program Studi Bimbingan dan Konseling FIP UPI yang telah membantu peneliti dalam administrasi perkuliahan khususnya pada proses persiapan sidang skripsi berlangsung.
6. Kedua orang tua tercinta serta kakak dan kedua adik yang terkadang menimbulkan kegemasan tersendiri namun memiliki hati yang penuh perhatian,

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

ucapan terima kasih yang tulus atas kesabaran dalam menghadapi kekacauan pikiran peneliti. Kasih sayang, doa, dukungan materi, kekuatan, dan keyakinan yang diberikan telah menjadi sumber inspirasi dan motivasi untuk menyelesaikan penulisan skripsi.

7. Tim Evaluasi IST-2005 atas solidaritas dalam berbagi pengetahuan, kesedihan, pengalaman, memberikan motivasi dan inspirasi yang memicu ide baru dalam penulisan skripsi.
8. Rekan seperjuangan BK-2020, yang dengan setia dan tanpa pamrih, telah memberikan dukungan emosional yang tak tergantikan, menjadi sumber kekuatan dan kebahagiaan yang konstan dalam perjalanan penyelesaian skripsi.
9. Warga kosan bu Ayu yang tanpa lelah mendukung

Peneliti berharap bahwa kontribusi yang berharga dari semua pihak akan tercatat dalam tulisan ilmiah ini, tidak hanya sebagai bagian dari keberhasilan tersusunnya skripsi, tetapi juga sebagai bukti nyata dari kolaborasi dan komunitas ilmiah yang dinamis dan saling mendukung. Semoga Allah SWT senantiasa membalas dengan kebaikan atas semua yang telah diberikan kepada peneliti.

Bandung, 23 Juli 2024



Nabila Az-zahra
NIM 2000180

ABSTRAK

Nabila Az-zahra. 2000180. Evaluasi Properti Psikometri *Intelligenz-Struktur - Test* versi 2005 subtes *Merkaufgaben* menggunakan Teori Klasik. Bimbingan dan Konseling. Fakultas Ilmu Pendidikan. Universitas Pendidikan Indonesia

Kualitas properti psikometri yang baik pada tes psikologi vital dalam mengungkap atribut yang diukur dengan akurat. Penelitian ini menyelidiki properti psikometri dan menetapkan norma baru untuk *merkaufgaben* (ME), yaitu subtes tugas memori dalam *intelligenz-struktur-test* (IST). Metode *ex-post facto* dengan pengambilan sampel kuota, digunakan untuk menganalisis data dari 22.654 testi yang berasal dari berbagai strata Pendidikan di Jawa Barat. Evaluasi tes klasik subtes ME digunakan untuk menguji analisis item, validitas, reliabilitas, dan menghitung acuan norma terkini. Hasil analisis item menunjukkan bahwa tidak ada item yang perlu dieliminasi atau direvisi, dengan pola kesukaran, daya pembeda dan fungsi distraktor yang cukup efektif. Hanya saja perlu sedikit perbaikan karena hanya 22 yang berfungsi efektif dan jawaban kosong melebihi ambang batas 10%. Subtes ME juga terindikasi memiliki validitas dan reliabilitas yang baik, dengan korelasi moderat dengan subtes lain dan kesalahan pengukuran yang relatif kecil (16.2%). Temuan ini mendukung penetapan norma baru yang didasarkan pada pengambilan sampel yang representatif.

Kata Kunci: *merkaufgaben*, *intelligenz-struktur-test*, properti psikometri, inteligensi, teori klasik

ABSTRACT

Nabila Az-zahra. 2000180. *Evaluation Psychometric Properties of the Intelligenz-Structure-Test version 2005 subtest Merkaufgaben using Classical Test-Theory. Guidance and Counseling. Faculty of Education. Indonesian University of Education*

Psychometric properties of psychological tests must be of good quality, in order to accurately reveal the attributes they are intended to measure. This study investigated the psychometric properties and established new norms for merkaufgaben (ME), the memory task subtests of the intelligenz-struktur-test (IST). An ex-post facto method with quota sampling was used to analyze data from 22,654 testees from various educational strata in West Java. Classical test evaluation of ME subtests was used to test item analysis, validity, reliability, and calculate the current norm reference. The results of item analysis showed that there were no items that needed to be eliminated or revised, with moderately effective patterns of difficulty, differentiating power and distractor function. However, there was a need for slight improvement as only 22 were functioning effectively and blank answers exceeded the 10% threshold. The ME subtest also indicated good validity and reliability, with moderate correlations with other subtests and a relatively small measurement error (16.2%). These findings support the establishment of new norms based on representative sampling.

Keywords: *merkaufgaben, intelligenz-struktur-test, psychometric properties, intelligence, classical test theory*

DAFTAR ISI

LEMBAR HAK CIPTA	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH.....	vi
ABSTRAK	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi dan Rumusan Masalah	7
1.3 Tujuan Penelitian	8
1.4 Manfaat Penelitian	8
1.4.1 Manfaat Teoretis	8
1.4.2 Manfaat Praktis	9
1.5 Struktur Organisasi Skripsi	9
BAB II KAJIAN TEORI.....	10
2.1 Paradigma dalam Evaluasi Tes	10
2.2 Konsep Teori Klasik	11
2.2.1 Definisi dan Asumsi Teori Klasik.....	11
2.2.2 Parameter Teori Klasik	13
2.3 Evaluasi Properti Psikometri.....	16
2.3.1 Konsep Validitas	16
2.3.2 Konsep Reliabilitas	20
2.3.3 Standardisasi Norma	26
2.4 <i>Intelligenz-Struktur-Test</i>	33
2.4.1 Konsep <i>Intelligenz-Struktur-Test</i>	33

Nabila Az-zahra, 2024

**EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES
MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

2.4.2 Landasan Teori IST: <i>Primary Mental Abilities</i>	37
2.4.3 Konsep Subtes <i>Merkaufgaben</i>	39
2.4.4 Bentuk Norma ME IST-2005.....	44
2.5 Penelitian Terdahulu	45
2.6 Posisi Evaluasi Properti Psikometri Tes untuk Asesmen di Sekolah.....	48
BAB III METODOLOGI PENELITIAN.....	50
3.1 Paradigma dan Pendekatan Penelitian	50
3.2 Desain Penelitian.....	51
3.3 Populasi dan Partisipan	51
3.4 Variabel Penelitian dan Definisi Operasional.....	53
3.5 Instrumen Penelitian	53
3.6 Prosedur Penelitian	54
3.7 Teknik Analisis Data.....	55
3.8 Isu Etik dalam Standardisasi subtes ME IST-2005.....	59
BAB IV HASIL DAN PEMBAHASAN	60
4.1 Analisis Item subtes ME	60
4.1.1 Analisis Taraf Kesukaran Item subtes ME.....	60
4.1.2 Analisis Daya Pembeda Item subtes ME	62
4.1.3 Distraktor subtes ME	64
4.2 Validitas subtes ME IST-2005	66
4.2.1 Validitas Diskriminan	66
4.2.2 Validitas Konstruk	67
4.3 Reliabilitas subtes ME IST-2005	70
4.4 Perbandingan Norma subtes ME IST-2005 dengan IST-70	71
4.5 Keterbatasan Penelitian.....	75
BAB V KESIMPULAN DAN REKOMENDASI.....	77
5.1 Kesimpulan	77
5.2 Rekomendasi.....	78
DAFTAR PUSTAKA	81

DAFTAR TABEL

Tabel 2.1 Tiga Prosedur Validitas dalam Teori Klasik	17
Tabel 2.2 Sumber Bukti Validitas	18
Tabel 2.3 Formula Estimasi Konsistensi Internal	23
Tabel 2.4 Macam-macam Pedoman Ukuran Sampel	27
Tabel 2.5 Kemampuan Khusus dalam IST	35
Tabel 2.6 Tujuh Struktur Kemampuan Khusus PMA	37
Tabel 2.7 Tujuh Faktor Utama Thurstone yang Selaras dengan subtes IST-70... 38	
Tabel 2.8 Norma IST-2005 ME berdasarkan Usia Setara	44
Tabel 2.9 Formula Transformasi Skor	45
Tabel 2.10 Penelitian Terdahulu IST	45
Tabel 3.1 Distribusi berdasarkan Wilayah dan Jenjang Pendidikan	52
Tabel 3.2 Definisi Operasional Properti Psikometri Variabel Penelitian.....	53
Tabel 3.3 Contoh Item pada Subtes ME.....	54
Tabel 3.4 Analisis subtes ME IST-2005 melalui Teori Klasik	57
Tabel 3.5 Kategorisasi Norma IST subtes ME.....	59
Tabel 4.1 Taraf Kesukaran Item dan Daya Pembeda Subtes ME	60
Tabel 4.2 Perbandingan Hasil Koefisien Korelasi Butir-butir Item subtes ME ...	63
Tabel 4.3 Analisis Distraktor terhadap Pilihan Jawaban subtes ME.....	64
Tabel 4.4 Hasil Pengujian Validitas Diskriminan terhadap Subtes ME.....	66
Tabel 4.5 Perbandingan Uji-t kelompok Sekolah Dasar dan Menengah.....	67
Tabel 4.6 Perbandingan Norma Usia Setara subtes ME IST-2005 dan IST-70 ...	72
Tabel 4.7 Norma Usia Setara Subtes ME IST-2005	74

DAFTAR GAMBAR

Gambar 2.1 Dua Model Analisis Instrumen.....	10
Gambar 2.2 Faktor M teori PMA Thurstone	43
Gambar 3.1 Prosedur Evaluasi Tes dalam Teori Klasik.....	55

DAFTAR LAMPIRAN

Lampiran 1. Administrasi Penelitian.....	91
Lampiran 1.1 Surat Keputusan (SK) Pengangkatan Dosen Pembimbing.....	91
Lampiran 1.2 Surat Izin Pengambilan Data Hasil IST subtes Merkaufgaben....	92
Lampiran 1.3 Lembar Bimbingan Skripsi.....	94
Lampiran 1.3.1 Pembimbing I: Dr. Eka Sakti Yudha, M.Pd	94
Lampiran 1.3.2 Pembimbing II: Drs. Sudaryat Nurdin Akhmad, M.Pd	95
Lampiran 1.4 Surat Rekomendasi Uji Plagiarisme	96
Lampiran 1.5 Surat Rekomendasi Ujian Sidang	97
Lampiran 2. Hasil Pengolahan Data menggunakan Teori Klasik.....	98
Lampiran 2.1 Sebaran Data Partisipan berdasarkan Demografis dan Jenjang Pendidikan.....	100
Lampiran 2.2. Pengolahan Analisis Item.....	102
Lampiran 2.2.1. Taraf Kesukaran dan Daya Beda Butir subtes ME	102
Lampiran 2.2.2. Distraktor	102
Lampiran 2.3 Pengolahan Uji Validitas.....	103
Lampiran 2.4. Pengolahan Uji Reliabilitas menggunakan Koefisien KR ₂₀	103
Lampiran 2.5 Pengolahan Pembuatan Norma Persentil IST-2005	104
Lampiran 2.5.1 Transformasi Skor Mentah subtes ME menjadi Skor Z.....	104
Lampiran 2.5.2 Transformasi Skor Z menjadi Skala Standar Norma IST-70 ..	104
Lampiran 2.5.3 Transformasi Skor Z menjadi Skor Persentil	105
Lampiran 3. Dokumentasi Penelitian.....	109
Lampiran 4. Anggota Tim Evaluasi IST-2005.....	112
Lampiran 5. Riwayat Hidup.....	113

DAFTAR PUSTAKA

- ABKIN. (2005). *Kode Etik Bimbingan dan Konseling*. Bandung: Pengurus Besar
- Adams, D., Chuah, K. M., Sumintono, B., & Mohamed, A. (2022). Students' readiness for e-learning during the COVID-19 pandemic in a South-East Asian university: a Rasch analysis. *Asian Education and Development Studies*, 11(2), 324-339. doi: <https://doi.org/10.1108/aed-s-05-2020-0100>.
- Adams, D., Tan, M. H. J., & Sumintono, B. (2021). Students' readiness for blended learning in a leading Malaysian private higher education institution. *Interactive Technology and Smart Education*, 18(4), 515-534. doi: <https://doi.org/10.1108/itse-03-2020-0032>.
- Adinugroho, I. (2016). Pengujian properti psikometrik intelligenz struktur test subtes kemampuan spasial dua dimensi (form auswahl): Studi pada dua SMA swasta di Jakarta. *Jurnal Ilmiah Psikologi MANASA*, 5(2), 165-180.
- Adiwimarta, S. S., Darmojuwono, S., & Hastrick, E.S. (2011). *Kamus Universal Langenscheidt* (Jerman). Jakarta: Katalis.
- AERA, APA, & NCME. (1985). *Standards for educational and psychological testing*. Washington: American Psychological Association.
- AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. New York: American Psychological Association.
- Agatha, M., Budiman, N., & Nadhirah, N. A. (2023). Etika Penggunaan Tes Psikologi Dalam Bimbingan Dan Konseling. *Science and Education Journal (SICEDU)*, 2(2), 273-285.
- Agung, I. M., & Fitri, A. R. (2016). *Analisis Psikometri dan Standardisasi Norma pada Tes Intelligensi Struktur Test (IST) pada Mahasiswa UIN Sultan Syarif Kasim Riau (Psychometric analysis and norm standardization on Intelligent Structure Test (IST) UIN Sultan Syarif Kasim undergraduate students)*. Research report. Pekanbaru: LP2M UIN Suska Riau.
- Agung, I. M., & Fitri, A. R. (2020). Analisis Psikometri Intelligenz Struktur Test (IST) pada Mahasiswa. *Psikobuletin: Buletin Ilmiah Psikologi*, 1(1), 1-10. doi: <http://dx.doi.org/10.24014/pib.v1i1.8174>.
- Akmal, N., Widyastuti., & Nur, H., (2021). Psychometric Properties Analysis of IST (Intelligenz Struktur Test). *Indonesian Journal of Educational Studies*, 24 (2), 126-142.
- Alghamdi, N. H., Killian, M. L., Aitha, B., Pohlig, R. T., & Silbernagel, K. G. (2019). Quantifying the dimensions of achilles tendon insertional area using ultrasound imaging a validity and reliability study. *Muscle Ligaments and Tendons Journal*, 09(04), 544. doi: <https://doi.org/10.32098/mltj.04.2019.09>.
- Aliyu, A. A., Bello, M. U., Kasim, R., & Martin, D. (2014). Positivist and Non-Positivist Paradigm in Social Science Research: Conflicting Paradigms or

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

- Perfect Partners? *Journal of Management and Sustainability*, 4(3), 79-95. doi: <https://doi.org/10.5539/jms.v4n3p79>.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks.
- Almat, N. S., Aliya, M. S., Zhanna, U. T., & Gulmira, D. S. (2023). The Relationship Between Social Intelligence And IQ: A Psychometric Analysis. *The Open Psychology Journal*, 16(1), 1-6. doi: <https://doi.org/10.2174/18743501-v16-e230120-2022-78>.
- Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention*. Heidelberg: Springer.
- Amthauer, R. (1971). *Intelligenz-Struktur-Test 70 (I-S-T 70)*. Göttingen: Hogrefe
- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2005). Test strukturnya inteligence IST 2000 R. Praha: Testcentrum, 78(27), 1-19.
- Anastasi, A. (1968). *Psychological testing* (Edisi Ketiga). Toronto: Macmillan.
- Anastasi, A., Urbina, S. (1997). *Psychological testing* (Edisi Ketujuh). Upper Saddle River: Prentice Hall.
- Antony, M. M., & Barlow, D. H. (2020). *Handbook of assessment and treatment planning for psychological disorders* (Edisi Kedua). New York: Guilford.
- Aryadoust, V., Tan, H. A. H., & Ng, L. Y. (2019). A Scientometric review of Rasch measurement: The rise and progress of a specialty. *Frontiers in psychology*, 10(1), 1-16. doi: 10.3389/fpsyg.2019.02197.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, 2(1), 89-195.
- Ayanwale, M.A., Chere-Masopha, J., Morena, M.C., (2022). The Classical Test or Item Response Measurement Theory: The Status of the Framework at the Examination Council of Lesotho 21. *IJLTER*, 21(8), 384-406. doi:10.26803/ijlter.21.8.22.
- Azizah, A. B., Damayanti, D., & Agustin, R. A. (2020). Pengaruh Inteligensi terhadap Keberhasilan Peserta Didik. *PANDAWA*, 2(1), 70-76.
- Azwar, S. (2015). *Dasar-dasar Psikometri*. Yogyakarta: Pustaka Pelajar
- Barclay, D., Thompson, R., & Higgins, C. (1995). The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Use as an Illustration. *Technology Studies*, 2(2), 285-309.
- Baron, J. (2018). Individual mental abilities vs. The world's problems. *Journal of Intelligence*, 6(2), 23. doi: 10.3390/jintelligence6020023.
- Baumgartner, T. A. (2009). Tutorial: Calculating percentile rank and percentile norms using SPSS. *Measurement in Physical Education and Exercise Science*, 13(4), 227-233. doi: <https://doi.org/10.1080/10913670903262769>.

- Bawono, B. A. (2008). *Uji Aspek-Aspek Psikometri Subtes Merkaufgaben dari Baterai Intelligenz Struktur Test* [Skripsi]. Jakarta: Universitas Katolik Atma Jaya.
- Beauducel, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences*, 30(6), 977-994. doi: [https://doi.org/10.1016/s0191-8869\(00\)00087-8](https://doi.org/10.1016/s0191-8869(00)00087-8).
- Beauducel, A., Liepmann, D., Horn, S., & Brocke, B. (2010). *Intelligence structure test*. Göttingen: Hogrefe.
- Bichi, A. A. (2016). Classical Test Theory: an introduction to linear modeling approach to test and item analysis. *International Journal for Social Studies*, 2(9), 27-33. doi: 10.26643/ijss.v2i9.6690.
- Blose, N. P., & Hadebe-Ndlovu, B. N. (2020). Grade Ten Teachers' Understanding of Multiple Intelligences in Teaching Physical Science in the Rural Context. *Alternation*, 27(2), 189-209. doi:10.29086/25195476/2020/v 27n2a 11.
- Bond, T. G., and Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (Edisi Ketiga). New York: Routledge.
- Bond, T. G., Yan, Z., & Heene, M. (2020). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (Edisi Keempat). New York: Routledge.
- Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences*, 115(26), 6674-6678. doi:10.1073/pnas.1718793115.
- Bryan, V. M., & Mayer, J. D. (2020). A meta-analysis of the correlations among broad intelligences: Understanding their relations. *Intelligence*, 81, 101469. doi: <https://doi.org/10.1016/j.intell.2020.101469>
- Bühner, M., Ziegler, M., Krumm, S., & Schmidt-Atzert, L. (2006). Ist der IST 2000 R Rasch-skalierbar?. *Diagnostica*, 52(3), 119-130.
- Buonaccorsi, J. P. (2010). *Measurement error: models, methods, and applications*. Boca Raton: Chapman & Hall.
- Cai, L., Choi, K., Hansen, M., & Harrell, L. (2016). Item response theory. *Annual Review of Statistics and Its Application*, 3, 297-321.
- Calculator.net. (2015). Sample Size Calculator [online]. Diakses dari <https://www.calculator.net/sample-size-calculator.html>.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Carter, P. (2005). *The Complete Book of Intelligence Test: 500 Exercises to Improve, Upgrade and Enhance Your Mind Strength*. Chichester: John Wiley & Sons.
- Chadha, N. K. (2009). *Applied psychometry*. Singapore: Sage Publication.

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

- Choa, E., & Chun, S. (2018). Originators of reliability coefficients: A historical review of the originators of reliability coefficients including Cronbach's Alpha. *조사연구*, 19(2), 73-104.
- Coaley, K. (2014). *An introduction to psychological assessment and psychometrics* (Edisi Kedua). London: Sage Publication.
- Cochran, W. G. (1977). *Sampling techniques* (Edisi Ketiga). New York: John Wiley & Sons.
- Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York: Academic Press.
- Cohen, R. J., & Swerdlik, M. E. (1996). *Psychological testing and assessment: An introduction to tests and measurement* (Edisi Ketujuh). McGraw Hill.
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(7), 1-9.
- Cresswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (Edisi Keempat). Lincoln: Pearson.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches*. (Edisi Kelima). New York: SAGE Publications.
- Crocker. L. & Algina. J. (2008). *Introduction to classical and modern test theory* (Edisi Kedua). New York: Holt, Rinehart and Winston.
- Culligan, B. (2011). Item response theory, reliability and standard error. Aoyama Gakuin Women's Junior College.
- Dhakar, A., & Bobrin, B. D. (2023). *Cognitive deficits*. StatPearls Publishing.
- DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, 2(2), 4.
- Dickerson, B. C., Sperling, R. A., Hyman, B. T., Albert, M. S., & Blacker, D. (2007). Clinical prediction of Alzheimer disease dementia across the spectrum of mild cognitive impairment. *Archives of general psychiatry*, 64(12), 1443-1450.
- Drummond, R. J., Sheperis, C. J., & Jones, K. D. (2015). *Assessment procedures for counselors and helping professionals* (Edisi Kedelapan). Pearson
- Dunn, G. (1989). *Design and analysis of reliability studies: the statistical evaluation of measurement errors*. London: Edward Arnold.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Associates.
- Ermis-Demirtas, H. Establishing Content-Related Validity Evidence for Assessments in Counseling: Application of a Sequential Mixed-Method

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

- Approach. *Int J Adv Counselling*, 40(1), 387–397 (2018). doi: <https://doi.org/10.1007/s10447-018-9332-4>.
- Eysenck., H., J. (1996). *Intelligenz-Test*. Hamburg: Rowohlt
- Fatkhudin, A., Surarso, B., & Subagio, A. (2016). Item Response Theory Model Empat Parameter Logistik Pada Computerized Adaptive Test. *Jurnal Sistem Informasi Bisnis*, 4(2), 121-129. doi: <https://doi.org/10.21456/vol4iss2pp121-129>.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2), 171-191.
- Freeman, F. N. (1925). What Is Intelligence?. *The School Review*, 33(4), 253–263. doi: <https://doi.org/10.1086/438138>.
- Freese, H., Krüger, H., & Wolters, B. (1997). Langenscheidts Universal Wörterbuch Englisch. German.
- Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: An introduction* (Edisi Kedua). Thousand Oaks, CA: Sage Publications.
- Gajjar, S., Sharma, R., Kumar, P., & Rana, M. (2014). Item and test analysis to identify quality multiple-choice questions (MCQs) from an assessment of medical students of Ahmedabad, Gujarat. *Indian Journal of Community Medicine: Official Publication of Indian Association of Preventive & Social Medicine*, 39(1), 17-20.
- Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data?. *MIS quarterly*, 36(3), 981-1001.
- Goodhue, D., Lewis, W., & Thompson, R. (2006, January). PLS, small sample size, and statistical power in MIS research. Dalam *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, 8(1), 202b-202b.
- Gorsuch, R. L. (1983). *Factor Analysis* (Edisi Kedua). Hillsdale: Lawrence Erlbaum.
- Gottfredson, L. S. (2004). Schools and the g factor. *The Wilson Quarterly*, 28(3), 35-45.
- Grace, Y. Y., Delaigle, A., & Gustafson, P. (Penyunting). (2021). *Handbook of measurement error models*. Boca Raton: CRC Press.
- Grady, C. L., & Craik, F. I. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology*, 10(2), 224–231. doi: [https://doi.org/10.1016/s0959-4388\(00\)00073-8](https://doi.org/10.1016/s0959-4388(00)00073-8).
- Green, S. B. (1991). How Many Subjects does it take to do a Regression Analysis. *Multivariate Behavioral Research*, 26(3), 499–510. https://doi.org/10.1207/s15327906mbr2603_7.
- Gregory, R. J. (2011). *Tes psikologi: Sejarah, prinsip, dan aplikasi* (Edisi Keenam). Terjemahan oleh Amitya Kumara & Mikael Seno. Jakarta: Erlangga.

- Gregory, R. J. (2011). *Tes psikologi: Sejarah, prinsip, dan aplikasi* (Edisi Keenam). Terjemahan oleh Amitya Kumara & Mikael Seno. Jakarta: Erlangga.
- Guilford, J. P dan Benjamin Frucher. (1978). *Fundamental Statistics in Psychology and Education* (Edisi Kedelapan). Tokyo: Mc Graw-Hill
- Guttman, L. (1945). *A basis for analyzing test-retest reliability. Psychometrika*, 10(4), 255–282. doi:10.1007/bf02288892.
- Haberlandt, K. (1999). *Human Memory: exploration and application..* New York: Allyn and Bacon.
- Habsy, B. A., Mujiono, A. P., Halmahera, A. D. S., Rohmawati, L. I. S., Nikmah, L. A., Hilman, L., & Mariatiningsih, M. (2023). Menelaah profil pelajar pancasila dan perwujudannya dalam Pendidikan yang berpihak pada peserta didik. *Jurnal Pendidikan Tambusai*, 7(3), 32079-32093.
- Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate Data Analysis* (Edisi Kedelapan). United Kingdom: Cengage Learning.
- Hakiki, A. W., Fitri, A. R., & Agung, I. M. (2018). Analisis Properti Psikometri Subtes Merkaufgaben (ME) dengan Rasch Model. *Jurnal Psikologi*, 14(1), 40. <https://doi.org/10.24014/jp.v14i1.4900>.
- Haladyna, T. M., & Downing, S. M. (1988). Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education*, 1(1), 51-78.
- Halvorsen, M. B., Kildahl, A. N., Kaiser, S., Axelsdottir, B., Aman, M. G., & Helverschou, S. B. (2024). Applicability and Psychometric Properties of General Mental Health Assessment Tools in Autistic People: A Systematic Review. *Journal of Autism and Developmental Disorders*, 1-14. doi: <https://doi.org/10.1007/s10803-024-06324-3>.
- Hambleton. R. K. & Swaminathan. H. (1985). *Item response theory*. Boston: Kluwer Nijhoff Publisher.
- Hambleton. R. K., Swaminathan. H, & Rogers. H. J. (1991). *Fundamentals of item response theory*. California: Sage Publications.
- Harris, R. J. (1975). *A primer of multivariate statistics*. New York: Academic Press.
- Harry A. King (1978) A Nomogram to Assist in Planning Surveys of Small (N < 2,000) Populations, *Research Quarterly. American Alliance for Health, Physical Education and Recreation*, 49(4), 552-557. doi: <http://dx.doi.org/10.1080/10671315.1978.10615570>.
- Hartono, W., Hadi, S., Rosnawati, R., & Retnawati, H. (2023). Exploration of Diagnostic Testing Instruments: Validity, Reliability, and Item Characteristics. *Pegem Journal of Education and Instruction*, 13(3), 386-394. doi: <https://doi.org/10.47750/pegegog.13.03.39>.

- Harvill, L. M. (1991). Standard error of measurement: an NCME instructional module on. *Educational Measurement: issues and practice*, 10(2), 33-41. doi: <https://doi.org/10.1111/j.1745-3992.1991.tb00195.x>.
- Hayes, A. M., Lauer, L. T., Kao, A. E., Sun, S., Klug, M. E., Tsan, L., ... & Kanoski, S. E. (2024). Western diet consumption impairs memory function via dysregulated hippocampus acetylcholine signaling. *Brain, behavior, and immunity*, 118, 408-422.
- Hendryadi, H. (2021). Editorial Note: Uji Validitas Dengan Korelasi Item-Total?. *Jurnal Manajemen Strategi dan Aplikasi Bisnis*, 4(2), 315-320. doi: <https://doi.org/10.36407/jmsab.v4i2.404>.
- Heppner, P. P., Kivlighan, D. M., & Wampold, B. E. (2008). *Research design in counseling* (Edisi Ketiga). Pacific Grove: Brooks/Cole.
- HIMPSI. (2010) *Kode Etik Psikologi Indonesia*. Jakarta: Pengurus Pusat Himpunan Psikologi Indonesia.
- Hingorjo, M. R., & Jaleel, F. (2012). Analysis of one-best MCQs: The difficulty index, discrimination index, and distractor efficiency. *JPMMA. The Journal of the Pakistan Medical Association*, 62(2), 142-147.
- Houser, R. A. (2019). *Counseling and educational research: Evaluation and application* (Edisi Keempat). London: Sage.
- Hox, J. (2010). *Multilevel analysis*. New York: Routledge.
- Hox, J., & McNeish, D. (2020). Small samples in multilevel modeling. Dalam Van de Schoot, R., & Miocević, M. (Penyunting). *Small sample size solutions: A guide for applied researchers and practitioners*. New York: Taylor & Francis.
- Hubley, A. M., & Zumbo, B. D. (2011). Validity and the Consequences of Test Interpretation and Use. *Social Indicators Research*, 103(2), 219–230. doi: <https://doi.org/10.1007/s11205-011-9843-4>.
- Hughes, D. J. (2018). Psychometric validity: Establishing the accuracy and appropriateness of psychometric measures. Dalam *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development*, (p. 751-779). Hoboken: John Wiley & Sons
- Ilievová, L., Demuthova, S., Vavrova, M., & Poliaková, N. (2021). Analysis of the Intelligence Structure in Nursing Students. *Revista Romaneasca pentru Educatie Multidimensionala*, 13(2), 206-222.
- Isaac, S., & Michael, W. B. (1981). *Recommended sample sizes for two different precision levels. Handbook in Research and Evaluation* (Edisi Kedua). San Diego: EdITS Publishers.
- Jabrayilov, R., Emons, W. H., & Sijtsma, K. (2016). Comparison of classical test theory and item response theory in individual change assessment. *Applied psychological measurement*, 40(8), 559-572. <https://doi.org/10.1177/014662161666640>.

- Jabrayilov, R., Emons, W. H., & Sijtsma, K. (2016). Comparison of classical test theory and item response theory in individual change assessment. *Applied psychological measurement, 40*(8), 559-572. <https://doi.org/10.1177/014662161666640>.
- Jack R. Frankel dan Norman E. Wallen. (2011). *How to Design and Evaluate Research in Education* (Edisi Kedelapan). New Work: McGraw Hill.
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling, 10*(1), 128-141.
- Janti, S. (2014). Analisis Validitas dan Reliabilitas dengan Skala Likert terhadap Pengembangan IS/IT dalam Penentuan Pengambilan Keputusan Penerapan Strategic Planning pada Industri Garmen. Dalam *prosiding SNAST*, (p. 155-160). Yogyakarta: Universitas AKPRIND Indonesia.
- Javed, S., & Khan, F. N. (2010). Multiple Intelligences in Communication Classroom. *JISR management and social sciences & economics, 8*(1), 189-202. doi: <https://doi.org/10.31384/jirmsse/2010.08.1.14>.
- Jørgensen, M. B., Konge, L., & Subhi, Y. (2018). Contrasting groups' standard setting for consequences analysis in validity studies: reporting considerations. *Advances in Simulation, 3*(1). doi: <https://doi.org/10.1186/s41077-018-0064-7>.
- Joyce, B. R., Weil, M., & Calhoun, E. (2003). *Models of teaching* (Edisi Kelima). Boston: Allyn and Bacon.
- Kerlinger, F.N. & Lee, H.B. (2000). *Foundations of Behavioral Research* (Edisi Keempat). Harcourt College Pub.
- Khoiri., (2021) Buku Statistika Konseptual dan Aplikatif Perspektif Manajemen. Semarang: SeAP.
- Kline, R. B. (2016). *Perinciples and Practice of Structural Equation Modeling*. New York: The Guilford Press.
- Kline, T. J. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks: Sage.
- Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal, 28*(1), 227-261.
- Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal, 28*(1), 227-261.
- Kramer, H, K., (2009). *Kamus Bahasa Indonesia Indonesisch-Deutsches Wörterverzeichnis*. Heidelberg: Nepal Research
- Kreft, I. G., & De Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage.

- Krejcie, R. V., & Morgan, D. W. (1970), Determining sample size for research activities, *Educational and psychological measurement*, 30(1), 607–610. doi: <https://doi.org/10.1177/001316447003000308>.
- Kuder, G. F., & Richardson, M. W. (1937). *The theory of the estimation of test reliability*. *Psychometrika*, 2(3), 151–160. doi:10.1007/bf02288391
- Kudiya, K., Sumintono, B., Sabana, S., Sachari, A. (2018). Batik artisans' judgment of batik wax quality and its criteria: An application of the Many-Facets Rasch model. Dalam *Pacific Rim Objective Measurement Symposium (PROMS) 2016 Conference Proceedings: Rasch and the Future*, 27-37. Springer Singapore.
- Kurniawan, R., & Pradipta, A. W. (2023). Taraf kesukaran dan Daya Beda Butir Soal Ujian Akhir Semester Matakuliah Penelitian Pendidikan. *Jurnal Pendidikan*, 11(2), 234–341. doi: <https://doi.org/10.36232/Pendidikan.v11i2.3999>.
- Kuvaas, B., Buch, R., & Dysvik, A. (2020). Individual variable pay for performance, controlling effects, and intrinsic motivation. *Motivation and Emotion*, 44(2020), 525-533. doi: <https://doi.org/10.1007/s11031-020-09828-4>.
- Kyriazos, T. A. (2018). Applied Psychometrics Sample Size and Sample Power Considerations in Factor Analysis (EFA, CFA) and SEM in General. *Psychology*, 9(1), 2207-2230. doi:10.4236psych.2018.98126.
- Lachin, J. M. (2004). The role of measurement reliability in clinical trials. *Clinical Trials*, 1(6), 553-566. doi: <https://doi.org/10.1191/1740774504cn0570a>.
- Laliyo, L. A. R., Tangio, J. S., Sumintono, B., Jahja, M., & Panigoro, C. (2020). Analytic Approach of Response Pattern of Diagnostic Test Items in Evaluating Students' Conceptual Understanding of Characteristics of Particle of Matter. *Journal of Baltic Science Education*, 19(5), 824-841. doi: <https://doi.org/10.33225/jbse/20.19.824>.
- Lamers, S. M., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. (2011). Evaluating the psychometric properties of the Mental Health Continuum-Short Form (MHC-SF). *Journal of clinical psychology*, 67(1), 99–110. doi: 10.1002/jclp.20741.
- Leedy, P. D., & Ormrod, J. E. (2005), *Practical research: Planning and design* (Edisi Kedelapan). Upper Saddle River: Prentice Hall.
- Lenhard, A., Lenhard, W. & Gary, S. (2019). Continuous norming of psychometric tests: A simulation study of parametric and semi-parametric approaches. *PLoS ONE*, 14(9), e0222279. doi: <https://doi.org/10.1371/journal.pone.0222279>.
- Lenhard, W., & Lenhard, A. (2020). Improvement of Norm Score Quality via Regression-Based Continuous Norming. *Educational and Psychological Measurement*, Online First, 1-33. doi: <https://doi.org/10.1177/0013164420928457>.

- Linacre, J. M. (1994). *Many Facet Rasch Model* (Edisi Kedua). Chicago: Mesa Press.
- Ling Lee, W., Chinna, K., & Sumintono, B. (2021). Psychometrics assessment of HeartQoL questionnaire: A Rasch analysis. *European Journal of Preventive Cardiology*, 28(12), e1-e5. doi:10.1177/2047487320902322.
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis Validitas, Reliabilitas, Taraf kesukaran dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan.
- Magno, C. (2009). Demonstrating the difference between classical test theory and item response theory using derived test data. *The International Journal of Educational and Psychological Assessment*, 1(1), 1–11.
- Mahmud, J. (2017). Item response theory: a basic concept. *Educational Research and Reviews*, 12(5), 258-266. doi: <https://doi.org/10.5897/err2017.3147>.
- Maryati, M., Prasetyo, Z. K., Wilujeng, I., & Sumintono, B. (2019). Measuring Teachers' pedagogical Content Knowledge Using Many-Facet Rasch Model. *Jurnal Cakrawala Pendidikan*, 38(3), 452-464. doi: <https://doi.org/10.21831/cp.v38i3.26598>.
- Matthews, N., Mattingley, J. B., & Dux, P. E. (2022). Media-multitasking and cognitive control across the lifespan. *Scientific Reports*, 12(1). doi: <https://doi.org/10.1038/s41598-022-07777-1>.
- Maxwell, S. E. (2000). Sample size and multiple regression. *Psychological Methods*, 5(1), 434-458.
- Meeker, W. Q., Escobar, L. A., & Pascual, F. G. (2022). *Statistical methods for reliability data*. John Wiley & Sons.
- Messick, S. (1989). *Educational measurement* (Edisi Ketiga). New York: Macmillan.
- Millsap, R. E. (1998). Group differences in regression intercepts: Implications for factorial invariance. *Multivariate Behavioral Research*, 33(3), 403-424.
- Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2017). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica*, 30(1). doi: <https://doi.org/10.1186/s41155-016-0057-1>.
- Murphy, K. R. & Davidshofer, C. O. (2005). *Psychological Testing: principals and applications* (Edisi Keenam). New Jersey: Prentice-Hall
- Musa, A., Shaheen, S., & Ahmed, A. (2018a). Distractor analysis of multiple-choice questions: A descriptive study of physiology examinations at the Faculty of Medicine, University of Khartoum. *Khartoum Medical Journal*, 11(1), 1444-14453. doi: <https://doi.org/10.53332/kmj.v11i1.619>.
- Musa, A., Shaheen, S., Elmardi, A., & Ahmed, A. (2018b). Item difficulty & item discrimination as quality indicators of physiology MCQ examinations at the

- Faculty of Medicine Khartoum University. *Khartoum Medical Journal*, 11(2), 1477-1486. doi: <https://doi.org/10.53332/kmj.v11i2.610>.
- Nakku, V. B., Agbola, F. W., Miles, M. P., & Mahmood, A. (2020). The interrelationship between SME government support programs, entrepreneurial orientation, and performance: A developing economy perspective. *Journal of Small Business Management*, 58(1), 2-31. doi: <https://doi.org/10.1080/00472778.2019.1659671>.
- Neuman, W. L. (2013). *Social research methods: Qualitative and quantitative approaches* (Edisi Ketujuh). Boston: Pearson.
- Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P. (2022). HUBUNGAN ANTARA VALIDITAS ITEM DENGAN DAYA PEMBEDA DAN TINGKAT KESUKARAN SOAL PILIHAN GANDA PAS. *Natural Science Education Research*, 4(3), 249–257. doi:10.21107/nser.v4i3.8682.
- Oosterhuis, H. E., van der Ark, L. A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, 23(2), 191-202. doi: <https://doi.org/10.1177/1073191115580638>.
- Park, M., & Liu, X. (2021). An investigation of item difficulties in energy aspects across biology, chemistry, environmental science, and physics. *Research in Science Education*, 51, 43-60. doi: <https://doi.org/10.1007/s11165-019-9819-y>.
- Permana, M. Z. (2017). *Panduan Praktis Personality Assessment*. Jakarta: Raih Asa Sukses.
- Pfeifer, R., & Scheier, C. (2001). *Understanding Intelligence*. First MIT Press.
- Prah, J. K., Amoah, S., Ocansey, D. W., Arthur, R., Walker, E., & Obiri-Yeboah, D. (2019). Evaluation of urinalysis parameters and antimicrobial susceptibility of uropathogens among out-patients at University of Cape Coast Hospital. *Ghana medical journal*, 53(1), 44-51.
- Rahi, S. (2017). Research design and methods: A systematic review of research paradigms, sampling issues and instruments development. *International Journal of Economics & Management Sciences*, 6(2), 1-5.
- Rahman, A. F., Widyastuti., & Ridfah. A., (2020). *Penyusunan Norma Intelligenz Struktur Test (Ist-70) berdasarkan Kelompok Jenjang Pendidikan SMP di Wilayah Makassar* [Skripsi]. Makassar: Universitas Negeri Makassar.
- Rahmawati, E. (2014). Evaluasi karakteristik psikometri intelligenz struktur test (IST). *Proceeding: Seminar Nasional Psikometri, Universitas Negeri Sumatera*, 270-282.
- Raja, V. L. (2020). Test Item Analysis of Reading Comprehension Examination Faculty of Teachers and Training Education. *KAIROS*, 4(1), 52-65.
- Raosoft, I. (2010). Raosoft Sample Size Calculator [online]. Diakses dari <http://www.raosoft.com/samplesize.html>.

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

- Reise, S. P., & Revicki, D. A. (Eds). (2014). *Handbook of item response theory modeling: Applications to typical performance assessment*. Routledge.
- Ridho, A. (2007). Karakteristik Psikometrik Tes berdasarkan Pendekatan Teori Tes Klasik dan Teori Respon Item. *INSAN Media Psikologi*, 9(2), 1-20.
- Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological science*, 29(8), 1358-1369. doi: 10.31234/osf.io/kymhp.
- Robinson, S. E., Roth, S. L., Keim, J., Levenson, M., Flentje, J. R., & Bashor, K. (1991). Nurse burnout: Work related and demographic factors as culprits. *Research in nursing & health*, 14(3), 223-228.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3-13.
- Roscoe, J. T. (1975). *Fundamental research statistics for the behavioral sciences* (Edisi Kedua). New York: Holt Rinehart and Winston.
- Rost, D.H. (2013). *Handbuch Intelligenz*. Weinheim: Beltz.
- Roth-Walter, F., Berni Canani, R., O'Mahony, L., Peroni, D., Sokolowska, M., Vassilopoulou, E., & Venter, C. (2024). Nutrition in chronic inflammatory conditions: bypassing the mucosal block for micronutrients. *Allergy*, 79(2), 353-383.
- Rusland, S. L., Jaafar, N. I., Sumintono, B., & Foroudi, P. (2020). Evaluating knowledge creation processes in the Royal Malaysian Navy (RMN) fleet Personnel conceptualization, participation and differences. *Cogent Business & Management*, 7(1), 1-25. doi:10.108023311975.2020.1785106.
- Sabatini, S., Cosentino, S., Chapman, S., Ballard, C., Brooker, H., Corbett, A., & Stephan, B. C. M. (2023). Cognitive trajectories: exploring the predictive role of subjective cognitive decline and awareness of age-related changes for cognitive functioning. *Frontiers in Psychiatry*, 14(1). doi: <https://doi.org/10.3389/fpsy.2023.1270798>
- Santrock, J, W. (2005). *A tropical approach to lifespan Development* (Edisi Kedua). New York: McGraw Hill.
- Saputra, G. W., Rivai, M. A., Su'udah, M., Wulandari, S. L. G., Dewi, T. R., Fitroh. (2017). Pengaruh teknologi informasi terhadap kecerdasan (intelektual, spiritual, emosional dan sosial) studi kasus: Anak-anak. *Jurnal Sistem Informasi*, 10(2), 77-88.
- Saputra, M. F. A., Siregar, S. A., & Izdihar, Z. N. (2017). Social effects of digital pornography. *Bulletin of Social Informatics Theory and Application*, 1(2), 79-85.
- Sari, P. N. (2010). Hubungan status gizi dengan tingkat kecerdasan intelektual (*intelligence quotient* – iq) pada anak usia sekolah dasar ditinjau dari status sosial-ekonomi orang tua dan tingkat Pendidikan ibu. (*Skripsi diterbitkan*). Surakarta: Fakultas Kedokteran Universitas Sebelas Maret.

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

- Schmidt-Atzert, L., Amelang, M., Fydrich, T., & Moosbrugger, H. (2012). *Psychologische Diagnostik* (Edisi Keenam). Berlin: Springer.
- Schneider, F., Horowitz, A., Lesch, K. P., & Dandekar, T. (2020). Delaying memory decline: different options and emerging solutions. *Translational psychiatry*, *10*(1), 13.
- Schubert, A. L., Löffler, C., Wiebel, C., Kaulhausen, F., & Baudson, T. G. (2024). Don't waste your time measuring intelligence: Further evidence for the validity of a three-minute speeded reasoning test. *Intelligence*, *102*, 101804.
- Schweizer, K. (Ed.). (2006). *Leistung und Leistungsdiagnostik*. Berlin: Springer
- Sekaran, U., & Bougie, R. (2016). *Research Methods For Business: A Skill Building Approach* (Edisi Ketujuh). United Kingdom: John Wiley & Sons Ltd.
- Setiawan, B., Panduwangi, M., & Sumintono, B. (2018). A Rasch analysis of the community's preference for different attributes of Islamic banks in Indonesia. *International Journal of Social Economics*, *45*(12), 1647-1662.
- Shafiee, E., Farzad, M., MacDermid, J. C., Beygi, A. S., Vafaei, A., & Farhoud, A. (2020). Cross-cultural adaptation and measurement properties of the patient-rated tennis elbow evaluation for the persian language. *Hand Therapy*, *25*(2), 56-62. doi: <https://doi.org/10.1177/1758998320910177>.
- Shearer, C. B. (2020). Multiple intelligences in gifted and talented education: lessons learned from neuroscience after 35 years. *Roepfer Rev.* *42*(1), 49–63. doi: 10.1080/02783193.2019.1690079.
- Sijtsma, K. (2011). Introduction to the measurement of psychological attributes. *Measurement*, *44*(7), 1209-1219. doi: 10.1016/j.measurement.2011.03.019.
- Sirodj, D. A. N. (2018). Analisis Kualitas Item Intelligence Structure Test (IST) Melalui Metode Item Response Theory (IRT). *Schema: Journal of Psychological Research*, *4*(2), 98-107.
- Solichin, M. (2017). Analisis daya beda soal, taraf kesukaran, validitas butir tes, interpretasi hasil tes dan validitas ramalan dalam evaluasi Pendidikan. *Dirasat: Jurnal Manajemen dan Pendidikan Islam*, *2*(2), 192-213.
- Son, A. L. (2019). Instrumentasi Kemampuan Pemecahan Masalah Matematis: Analisis Reliabilitas, Validitas, Taraf kesukaran Dan Daya Beda Butir Soal. *Gema Wiralodra*, *10*(1), 41–52.
- Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201–292. doi:10.2307/1412107.
- Spearman, C. (1927). *The Abilities of Man: Their Nature and Measurement*. Macmillan and Co.
- Spoors, J. M. (2023). *Fundamentals of Psychological Assessment and Testing*. Routledge.
- Springer, D. W., Abell, N., & Hudson, W. W. (2002). Creating and validating rapid assessment instruments for practice and research: part 1. *Research on Social*

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Work Practice, 12(3), 408-439. doi:
<https://doi.org/10.1177/1049731502012003005>.

- Stănescu, M., & Tomescu, G. (2020). The Relationship between Dance and Multiple Intelligences of Institutionalised Children: A Theoretical Framework for Applied Research. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 11(4Sup1), 167-184. doi:10.18662/brain/11.4Sup1/163.
- Sternberg, R. J. (1997). The concept of intelligence and its role in lifelong learning and success. *American psychologist*, 52(10), 1030.
- Suhr, D. D. (2006). *Exploratory or Confirmatory Factor Analysis*. Cary: SAS Institute Inc.
- Sulistiana, D. (2014). *Adaptasi dan Standarisasi Multidimensional Aptitude Battery-II sebagai Tes Intelegensi bagi Siswa SMA [Skripsi]*. Bandung: Universitas Pendidikan Indonesia
- Sumintardja, E. & Rismiyati EK. (1997). *Konsep Dasar serta Strategi Pemahaman Psikodiagnostik*. Jakarta: PT. Charoen Pokphand Indonesia.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trim komunikata.
- Supratiknya, A. (2014). *Pengukuran psikologis*. Yogyakarta: Universitas Sanata Dharma.
- Suryani, Y. M. (2018). Aplikasi Rasch Model dalam Mengevaluasi Intelligenz Structure Test (IST). *Psikohumaniora: Jurnal Penelitian Psikologi*, 3(1), 73-100. doi: <https://doi.org/10.21580/pjpp.v3i1.2052>.
- Swan, J., Nivel, E., Kant, N., Hedges, J., Atkinson, T., & Steunebrink, B. (2022). *The Road to General Intelligence*. Springer.
- Taber, K. S. (2017). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. doi: <https://doi.org/10.1007/s11165-016-9602-2>.
- Tarigan, M., & Fadillah, F. (2021a). Analisis Item Response Theory Raven's Coloured Progressive Matrices pada Sampel Anak Usia Dini. *PSIKODIMENSIA*, 20(2), 158-169. doi: <https://doi.org/10.24167/psidim.v20i2.3101>.
- Tarigan, M., & Fadillah, F. (2021b). Properti psikometri struktur inteligensi IST subtes verbal (Satzergaenzung, Wortauswahl, dan Analogien) berbahasa Indonesia. *Jurnal Muara Ilmu Sosial, Humaniora, dan Seni*, 5(1), 63-72. doi: <https://doi.org/10.24912/jmishumsen.v5i1.9623.2021>.
- Tarigan, M., & Fadillah, F. (2022). Properti Psikometrik Intelligenz Struktur Test Subtes Kemampuan Numerik (Rechenaufgaben dan Zahlen Reihen). *Intuisi: Jurnal Psikologi Ilmiah*, 13(2), 155-170. doi: <https://doi.org/10.15294/intuisi.v13i2.31839>.

Nabila Az-zahra, 2024

EVALUASI PROPERTI PSIKOMETRI INTELLIGENZ-STRUKTUR-TEST VERSI 2005 SUBTES MERKAUFGABEN (ME) MENGGUNAKAN CLASSICAL TEST THEORY

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

- Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions a descriptive analysis. *BMC medical education*, 9(1), 1-8. doi: <https://doi.org/10.1186/1472-6920-9-40>.
- Terman, L. M., & Whitmire, E. D. (1921). Age and grade norms for the National intelligence tests, Scales A and B. *The Journal of Educational Research*, 3(2), 124-132.
- Thompson, C. B., & Panacek, E. A. (2007). Research study designs: Non-experimental. *Air medical journal*, 26(1), 18-22. doi: <https://doi.org/10.1016/j.amj.2006.10.003>.
- Thorndike, L. L. (1984). *Introduction to educational psychology*. St Paul: West Publishing Company.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Tighe, J., McManus, I., Dewhurst, N.G. et al. (2010) The standard error of measurement is a more appropriate measure of quality for postgraduate medical assessments than is reliability: an analysis of MRCP(UK) examinations. *BMC Med Educ* 10, 40. doi:10.1186/1472-6920-10-40.
- Troche, S. J., Makowski, L. M., Pahud, O., & Rammsayer, T. H. (2024). Working memory updating as a mediator of the relation between temporal resolution power and psychometric intelligence. *Personality and Individual Differences*, 220, 112479. doi: 10.1016/j.paid.2023.112479.
- Tuckman, H. P. (1975). Teacher effectiveness and student performance. *The Journal of Economic Education*, 7(1), 34-39.
- Van Zile-Tamsen, C. (2017). Using Rasch analysis to inform rating scale development. *Research in Higher Education*, 58(8), 922-933.
- Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond g: Putting multiple intelligences theory to the test. *Intelligence*, 34(5), 487-502.
- Waterhouse, L. (2006). Inadequate evidence for multiple intelligences, Mozart effect, and emotional intelligence theories. *Educational psychologist*, 41(4), 247-255. doi: https://doi.org/10.1207/s15326985ep4104_5.
- Wedman, J., & Lyrén, P. E. (2019). Methods for examining the psychometric quality of subscores: A review and application. *Practical Assessment, Research, and Evaluation*, 20(1), 21.
- Widhiarso, W., & Sumintono, B. (2016). Examining response aberrance as a cause of outliers in statistical analysis. *Personality and Individual Differences*, 98(1), 11-15. doi: <https://doi.org/10.1016/j.paid.2016.03.099>.
- Wood, A. H., Chappell, H. F., & Zulyniak, M. A. (2022). Dietary and supplemental long-chain omega-3 fatty acids as moderators of cognitive impairment and Alzheimer's disease. *European journal of nutrition*, 1-16.

- Woolfolk, Anita E dan Nicolich, Lorraine McCune (1984). *Educational psychology for teachers*. New Jersey: Prentice Hall.
- Yadav, V., Jain, R., Mittal, M. L., Panwar, A., & Lyons, A. (2019). The impact of lean practices on the operational performance of SMEs in India. *Industrial Management & Data Systems*, 119(2) 317-330. doi: <https://doi.org/10.1108/imds-02-2018-0088>.
- Yang, K., & Banamah, A. (2014). Quota Sampling as an Alternative to Probability Sampling? An Experimental Study. *Sociological Research Online*, 19(1), 1–11. doi:10.5153/sro.3199.
- Yin, X., (2023). Art Education in Colleges and Universities Based on the Theory of Multiple Intelligence. *ITSPOA*, 6(8), 142-148. doi: <https://doi.org/10.31058/j.edu.2023.68018>.
- Yu, C. H. (2005). *Test-Retest Reliability: Encyclopedia of Social Measurement*. Amsterdam: Elsevier.
- Zanon, C., Hutz, C. S., Yoo, H., & Hambleton, R. K. (2016). *An application of item response theory to psychological test development*. *Psicologia: Reflexão e Crítica*, 29(1). doi: <https://doi.org/10.1186/s41155-016-0040-x>.