

# BAB I

## PENDAHULUAN

Bab satu membahas latar belakang penelitian dengan mengungkap urgensi evaluasi properti psikometri *intelligenz-struktur-test* versi 2005 subtes *merkaufgaben* (ME), mengidentifikasi isu masalah penelitian sebagai dasar alasan penelitian dilakukan, tujuan penelitian dituangkan dengan menargetkan hasil yang signifikan dan aplikatif, serta manfaat atau signifikansi penelitian bagi pihak terkait. Bab ini juga memaparkan struktur organisasi skripsi, yang mencakup pembahasan secara sistematis dari awal hingga akhir penelitian.

### 1.1 Latar Belakang

Inteligensi, atau kecerdasan, adalah fondasi yang menentukan keberhasilan Pendidikan dan pengembangan potensi individu. Kemampuan ini memengaruhi proses belajar dan adaptasi, seperti yang diungkapkan oleh Spearman (1904; 1927) dengan teori faktor tunggal yang menyatakan inteligensi sebagai kemampuan umum yang memengaruhi kinerja mental pada berbagai jenis tugas. Tidak sejalan, Thurstone (1938) menekankan bahwa inteligensi terdiri dari beberapa kemampuan kognitif yang terpisah. Lebih lanjut, Guilford (1972) menekankan pada proses berpikir konvergen dan divergen sebagai inti dari inteligensi, sementara Sternberg (1997) memandang inteligensi sebagai kemampuan mental yang diperlukan untuk adaptasi dengan membentuk dan memilih setiap konteks lingkungan. Keempat definisi tersebut memandang inteligensi sebagai perpaduan antara kemampuan beradaptasi, proses kognitif dan kesadaran diri yang memungkinkan individu untuk menavigasi dan berinteraksi secara efektif dengan lingkungannya.

Inteligensi dalam Pendidikan memainkan peran kunci, terutama untuk menunjang proses pembelajaran. Studi menemukan bahwa durasi Pendidikan memiliki efek yang konsisten dalam meningkatkan inteligensi di berbagai kelompok usia dan kategori kognitif, menyoroti pentingnya Pendidikan sebagai metode untuk meningkatkan kemampuan intelektual (Ritchie & Tucker-Drob, 2018). Inteligensi berpengaruh dalam keberhasilan akademis dan adaptasi sosial siswa (Habsy, et al., 2023). Keberhasilan itu didukung dengan dimilikinya kemampuan berpikir kritis, pemecahan masalah dan tugas-tugas memori yang

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

merupakan komponen esensial dalam Pendidikan abad ke-21 (Baron, 2018; Azizah, et al., 2020). Selain itu, inteligensi berkorelasi dengan kemampuan untuk memanfaatkan peluang belajar dan menghadapi tantangan akademis (Magdalena, 2021). Dengan demikian, pengukuran tes inteligensi menjadi vital untuk mengidentifikasi kebutuhan Pendidikan individu dan merancang intervensi yang tepat untuk mendukung perkembangan intelektual siswa.

Salah satu alat ukur konstruk inteligensi yang telah diadaptasi dalam konteks Indonesia adalah *intelligenz-struktur-test* (IST). IST adalah tes inteligensi yang dikembangkan oleh Rudolf Amathuer di Frankfurt, Jerman pada tahun 1953 yang terdiri dari sembilan subtes yang masing-masing memiliki karakteristik item yang unik (Rahmawati, 2014). IST termasuk dalam kategori *multiple aptitude batteries test*, yaitu tes yang terdiri dari serangkaian subtes, di mana setiap subtes dibuat untuk mengukur aspek kemampuan khusus pada manusia atau yang disebut dengan *aptitude* (Anastasi & Urbina, 1997; Coaley, 2014, p. 39). IST juga merupakan alat ukur jenis *power test* yang sering digunakan baik di lingkungan Pendidikan maupun pekerjaan meskipun usianya sudah lebih dari 40 tahun dan tes inteligensi yang paling direkomendasikan oleh biro penyelenggara tes (Bawono, 2008; Rahmawati, 2014).

Pandangan Amthauer mengenai inteligensi banyak didasari oleh teori faktor, baik itu teori bifaktor, teori multifaktor, model struktur inteligensi Guilford, maupun teori hierarki faktor (Permana, 2017, p. 39). Beberapa dasar teori tersebut yang dijadikan landasan teori IST adalah teori multifaktor *Primary Mental Abilities* (PMA). PMA adalah teori inteligensi yang memaparkan bahwa manusia memiliki tujuh kemampuan dasar yang relatif independen satu sama lain, yaitu *verbal comprehension* (V); *word fluency* (W); *number* (N); *space* (S); *associative memory* (M); *perceptual speed* (P); *general reasoning* (I) (Thurstone, 1938, p. 100-102). Berbagai kemampuan ini adalah elemen-elemen intelektual yang unik dan tidak hanya bergantung pada satu faktor umum inteligensi (Guilford, 1972, p. 129).

Salah satu faktor abilitas dasar primer adalah faktor memori (M) (Thurstone, 1938). Faktor memori mengevaluasi kecepatan dan ketepatan mengingat informasi. Faktor memori dapat dinilai melalui IST subtes *merkaufgaben* (ME), yang

mengukur kemampuan mengingat *mnemonik* dan mereproduksi informasi memori jangka pendek (Amthauer, 1971). Pentingnya faktor memori dalam evaluasi kognitif individu terlihat dalam pengaruhnya terhadap proses berpikir dan pemecahan masalah (Buhner et al., 2006). Dalam konteks yang lebih luas, memori berperan dalam pengambilan keputusan sehari-hari, pemrosesan informasi, dan penilaian relevansi pengalaman masa lalu, yang semuanya memiliki dampak signifikan dalam lingkungan Pendidikan, pekerjaan, dan masyarakat (Bawono, 2008). Oleh karena itu, penelitian ini mengambil subtes *merkaufgaben* (ME) sebagai fokus utama untuk di evaluasi properti psikometrinya.

Untuk mengevaluasi properti psikometri subtes ME, maka guru BK perlu terampil mengadministrasikan, menskor, menafsirkan dan melaporkan hasil pengumpulan data untuk keperluan layanan bimbingan dan konseling (ABKIN, 2005, p. 17). Ini tercantum dalam Standar Kompetensi Konselor Indonesia (SKKI) kompetensi ke-empat, yaitu menguasai konsep dan praksis asesmen (ABKIN, 2005, p. 13). Guru BK dituntut melakukan pengukuran terhadap atribut psikologis siswa untuk menginformasikan proses pengambilan keputusan (HIMAPS, 2010). Hal ini menjadi dasar untuk menyusun program bimbingan dan konseling yang bertujuan untuk optimalisasi potensi siswa, karena perolehan data yang akurat dan dapat diandalkan merupakan prasyarat penting dalam membentuk program bimbingan. Mengingat hal tersebut, ketersediaan alat asesmen yang terstandarisasi dalam bidang bimbingan dan konseling menjadi signifikan (Sulistiana, 2014, p. 5). Implementasi instrumentasi dalam bimbingan dan konseling mengemban fungsi pemahaman, esensial untuk analisis level awal inteligensi siswa, mendukung stakeholder dalam pengembangan intervensi Pendidikan yang tepat (Agatha, et al., 2023).

Peran guru BK menjadi vital dalam konteks ini, mengingat mereka adalah penghubung antara hasil evaluasi subtes ME dan aplikasinya dalam Pendidikan. Guru BK, dengan keahlian dalam asesmen dan intervensi Pendidikan, dapat memanfaatkan informasi dari subtes ME untuk mengidentifikasi siswa yang mungkin memerlukan dukungan tambahan dalam mata pelajaran yang menuntut kemampuan mengingat. Kolaborasi erat antara guru BK dan guru mata pelajaran

memungkinkan integrasi strategi pembelajaran yang memanfaatkan kekuatan memori siswa dan pengembangan metode yang inovatif untuk mengatasi kesulitan belajar yang teridentifikasi.

Interpretasi hasil subtes ME oleh guru BK memiliki implikasi yang luas, memberikan wawasan mendalam tentang potensi kesulitan belajar, kemampuan retensi dan pengambilan informasi, serta strategi pembelajaran yang digunakan oleh siswa. Hasil asesmen ini juga dapat menunjukkan adanya masalah emosional yang mempengaruhi kemampuan memori siswa, yang memerlukan intervensi dan dukungan yang tepat. Informasi dari hasil asesmen memori ini penting tidak hanya untuk konteks akademik, tetapi juga untuk optimalisasi perkembangan pribadi siswa secara keseluruhan. Dengan demikian, informasi ini menjadi fondasi yang kuat dalam proses perancangan program dan penanganan siswa selama kegiatan konseling, memungkinkan guru BK untuk menentukan pendekatan yang paling efektif untuk diaplikasikan dalam mendukung siswa.

Kontinuitas antara peran pemahaman guru BK atas subtes memori sebagai salah satu subtes dalam alat ukur menjadi jelas ketika mempertimbangkan sejarah adaptasi dan revisi *Intelligenz-Struktur-Test (IST)*. Sejak diperkenalkannya IST-70 pada tahun 1973, alat ini telah mengalami serangkaian adaptasi dan revisi. Revisi awal terjadi pada tahun 1999, yang kemudian berkembang menjadi IST 2000, dan terakhir adalah IST 2000R pada tahun 2007 (Amthaeur et al., 2005, p. 3-5; Amelang & Schmidt-Atzert, 2006, p 55; Sirodj, 2018). Di Indonesia, adaptasi pertama dilakukan oleh Fakultas Psikologi Universitas Padjadjaran Bandung pada tahun 1973 (Rahmawati, 2014; Agung & Fitri, 2020). Hingga akhirnya direvisi Laboratorium Bimbingan dan Konseling, Fakultas Ilmu Pendidikan Universitas Pendidikan Indonesia (Lab BK FIP UPI), yang diberi nama IST-2005 (Sudaryat, 2005). Kendati revisi yang terjadi, IST-2005 belum mengalami revisi terhadap item maupun norma selama hampir dua dekade. Padahal versi ini masih digunakan Lab BK FIP UPI untuk mengukur minat dan bakat peserta tes (testi) di kawasan Indonesia Barat, khususnya di provinsi Jawa Barat.

Interval waktu yang panjang sejak pembaruan terakhir norma IST-2005 telah menyebabkan sebagian item-item yang diukur diketahui secara luas oleh publik,

dan beberapa di antaranya kini kurang relevan dengan konteks zaman ini, yang menyoroti perlunya evaluasi alat ukur (Akmal, et al., 2021; Tarigan & Fadillah, 2021, 1a; 1b). Literatur menunjukkan bahwa norma tersebut mungkin telah kehilangan keberlakuannya dalam jangka waktu lima hingga sepuluh tahun, yang mendukung ide bahwa proses penormaan ulang perlu menjadi prosedur yang rutin, bukan sekadar pengecualian (Anastasi & Urbina, 1997, p. 201; Gregory, 2011, p. 45). Meski telah muncul versi lokal dari norma IQ, ketiadaan referensi norma standar untuk skor inteligensi IST menimbulkan tantangan substansial, yang berpotensi menyebabkan salah tafsir terhadap hasil yang diperoleh (Agung & Fitri, 2016). Karena norma yang selama ini digunakan di Indonesia tidak tepat untuk menggambarkan persentase IQ karena adanya perbedaan karakteristik kemampuan individu Indonesia maupun Jerman, sehingga perlu adanya penyusunan kembali norma IST-2005 berdasarkan populasi di Indonesia maupun populasi di suatu provinsi.

Guna memperbaharui norma maka perlu melihat semua butir item subtes ME berkualitas. Hal ini dapat dicapai dengan melakukan analisis butir item yang komprehensif, sehingga memungkinkan identifikasi item-item yang valid dan reliabel (Anastasi, 1968). Terdapat dua metodologi utama untuk melakukan analisis butir item, yaitu teori klasik atau *classical test theory* (CCT) dan teori modern atau *item response theory* (IRT) (Anastasi, 1968; Nunnally, 1978). CCT adalah metode tradisional yang menangani satu sumber kesalahan pada satu waktu dengan menerapkan formula yang spesifik (Chang Ho, 2005). CCT dikenal karena keunggulan dalam kesederhanaan dan efisiensinya (Crocker & Algina, 2008). Berakar pada matematika dasar seperti rata-rata dan korelasi, CCT memfasilitasi pemahaman konseptual bagi para ahli dan peneliti (Magno, 2009). Selain itu, CCT telah menjadi fondasi dalam psikometri selama lebih dari satu abad, menunjukkan keandalannya dalam berbagai situasi pengujian (Ayanwale, et al., 2022). Meskipun IRT menawarkan pendekatan yang lebih kompleks dan mungkin lebih tepat dalam beberapa situasi, CCT tetap relevan, terutama ketika sumber daya terbatas atau ketika kesederhanaan dan kejelasan menjadi prioritas utama dalam penelitian psikometri (Zanon, et al., 2016).

Berbagai studi telah dilakukan terkait psikometri terhadap subtes ME, baik CTT maupun IRT. Penelitian Bawono (2008) mengkonfirmasi validitas dan reliabilitas subtes ME dalam mengukur konstruksi memori, khususnya memori jangka panjang, dengan menerapkan CTT. Rahmawati (2014) menunjukkan bahwa subtes ME memiliki daya pembeda yang tinggi ( $>2.00$ ) dan peluang penebak yang rendah ( $<0.35$ ), menandakan kualitas psikometri yang kuat. Namun, melalui metode yang sama, ditemukan item nomor 157 tidak konsisten dengan item lain dalam mengukur daya ingat, yang menunjukkan kebutuhan untuk revisi dan item tersebut tidak mampu mengukur daya ingat mahasiswa (Sirodj, 2018). Agung dan Fitri (2016) menemukan bahwa meskipun skor subtes ME umumnya berada dalam kategori sedang, beberapa item memerlukan perbaikan. Hakiki, et al., (2018) melalui IRT-1PL ( $N_1=293$  dan  $N_2=124$ ) berhasil mengatasi masalah misfit dan bias pada subtes ME yang direvisi, dengan koefisien reliabilitas yang baik. Suryani (2018) melalui IRT-1PL, menemukan analisis tingkat kesesuaian butir item ME tergolong baik, namun ditemukan bias pada 9 item (157, 160, 161, 162, 163, 165, 168, 169, dan 172). Rahman, et al., (2019) berhasil menghasilkan norma baru menggunakan CTT dan metode *trimmed mean* untuk jenjang Pendidikan SMP yang disesuaikan dengan kelompok usia 14-16 tahun di wilayah Makassar. Terakhir, Akmal, et al., (2021) melalui CTT ( $N=1112$ ), menemukan semua butir item pada subtes ME memenuhi syarat untuk diterima.

Studi terdahulu mengungkap bahwa pola riset properti psikometri subtes ME di Indonesia selama ini mengandalkan satu pendekatan saja dan kebanyakan tidak menghitung properti psikometri standardisasi norma. Kendati, ada penelitian yang membuat norma dengan kategori kelompok jenjang Pendidikan dan usia. Norma yang digunakan hanya sebatas pada jenjang Pendidikan SMP, variasi umurnya terbatas, tidak ada data perhitungan jelas terkait validitas dan reliabilitas, serta normanya hanya terkhusus wilayah Makassar. Selain itu, studi pendahuluan mengungkap bahwa evaluasi properti psikometri IST-2005, yang digunakan oleh Lab BK FIP UPI selama dua dekade, belum pernah dianalisis ulang. Berdasarkan studi tersebut, peneliti tertarik untuk mengevaluasi properti psikometri IST-2005

subtes *merkaufgaben* (ME) menggunakan teori klasik, untuk melihat apakah butir-butir item masih valid, reliabel dan memiliki standardisasi norma.

## 1.2 Identifikasi dan Rumusan Masalah

Penelitian dalam bidang psikometri di Indonesia, khususnya yang berkaitan dengan subtes ME, mengalami stagnasi yang cukup signifikan dalam evaluasi psikometri. Penelitian terdahulu lebih banyak mengandalkan analisis butir item tanpa melakukan penormalan secara komprehensif (Buwono, 2008; Rahmawati, 2011; Agung & Fitri, 2016). Hal ini menimbulkan kekhawatiran mengenai validitas dan reliabilitas dari ukuran yang dihasilkan, terutama ketika digunakan untuk keputusan Pendidikan yang penting (Wedman & Lyrén, 2019, p. 2), terutama dalam asesmen inteligensi di sekolah. Selain itu, norma ME yang dikembangkan oleh Rahman dkk. (2019) terbatas pada sampel siswa sekolah menengah dari satu kota dan tidak memiliki hasil validitas dan reliabilitas yang baik, sehingga mempertanyakan keterwakilan dan kemampuan generalisasinya.

Keterbatasan dalam pengembangan norma ini tidak hanya berdampak pada validitas pengukuran, tetapi juga pada keadilan dan kesetaraan penilaian. Norma yang tidak mencerminkan keragaman demografi dan budaya Indonesia dapat menimbulkan bias dalam penilaian, yang secara tidak adil memengaruhi hasil tes (Morgado et al., 2017, p. 13). Oleh karena itu, diperlukan pendekatan yang lebih holistik dan inklusif dalam pengembangan norma, yang memperhitungkan perbedaan tingkat Pendidikan dan usia peserta tes (Hughes, 2018). Untuk menjawab permasalahan tersebut, penelitian ini bertujuan untuk melakukan evaluasi secara komprehensif terhadap properti psikometri dari subtes *merkaufgaben* (ME) IST-2005 yang digunakan oleh Lab BK FIP UPI, menggunakan teori klasik atau CTT. Untuk memperjelas masalah penelitian, maka dirumuskan empat pertanyaan penelitian sebagai berikut:

- 1) Bagaimanakah kualitas item IST-2005 subtes *merkaufgaben* (ME) berdasarkan hasil analisis diskriminasi item, mencakup taraf kesukaran, daya pembeda, dan efektivitas distraktornya?

- 2) Apakah butir-butir item IST-2005 subtes *merkaufgaben* (ME) masih valid untuk mengukur memori testi?
- 3) Apakah butir-butir item IST-2005 subtes *merkaufgaben* (ME) memiliki reliabilitas yang memadai untuk mengukur memori testi?
- 4) Bagaimana bentuk norma baru berdasarkan variasi usia setara IST-2005 subtes *merkaufgaben* (ME)?

### 1.3 Tujuan Penelitian

Penelitian ini secara umum bertujuan untuk mengevaluasi properti psikometri IST-2005 subtes *merkaufgaben* (ME) menggunakan teori klasik. Secara lebih khusus, penelitian ini bertujuan untuk mencapai tujuan-tujuan berikut:

- 1) Mengkaji kualitas psikometri IST-2005 subtes *merkaufgaben* (ME), melalui analisis diskriminasi item untuk menentukan taraf kesukaran, daya pembeda dan efektivitas distraktor butir-butir item.
- 2) Memvalidasi keabsahan butir-butir item IST-2005 subtes *merkaufgaben* (ME), untuk memastikan setiap item masih memiliki kapasitas yang valid
- 3) Menilai reliabilitas interitem IST-2005 subtes *merkaufgaben* (ME), untuk menjamin bahwa alat ukur konsisten dan dapat diandalkan.
- 4) Mengembangkan norma persentil baru IST-2005 subtes *merkaufgaben* (ME) yang disesuaikan berdasarkan usia pertahun.

### 1.4 Manfaat Penelitian

Eksplorasi IST-2005 dalam konteks Pendidikan terus berkembang di Jawa Barat menyediakan berbagai manfaat yang luas. Manfaat-manfaat ini mencakup aspek teoretis hingga penerapan praktis yang konkret.

#### 1.4.1 Manfaat Teoretis

Penelitian ini memberikan sumbangan pemikiran dalam kajian psikometri IST-2005. Sebagai rujukan baru, penelitian ini mengundang para akademisi untuk mengeksplorasi dan memvalidasi penggunaan IST-2005 subtes *merkaufgaben* (ME) dalam konteks Jawa Barat, memperluas pemahaman tentang pengukuran bakat dan peminatan. Khususnya, bagi dunia bimbingan dan konseling, penelitian

ini menawarkan perspektif segar mengenai kesesuaian IST-2005 subtes *merkaufgaben* (ME) dengan kebutuhan dan karakteristik peserta tes (testi) saat ini.

#### **1.4.2 Manfaat Praktis**

##### **1) Bagi Laboratorium Bimbingan dan Konseling**

Temuan diharapkan dapat menjadi panduan praktis dalam menilai kelayakan butir-butir item IST-2005 subtes *merkaufgaben* (ME). Dengan analisis yang mendalam, Lab BK FIP UPI dapat menentukan item mana yang perlu direvisi atau diganti, serta mengembangkan norma baru yang lebih relevan.

##### **2) Bagi Peneliti Selanjutnya**

Temuan diharapkan berpotensi menjadi landasan bagi penelitian masa depan yang bertujuan untuk mengembangkan bakat testi. Dengan referensi ini, peneliti lain dapat merancang studi yang lebih terfokus untuk mengoptimalkan proses seleksi dan penempatan testi sehingga memaksimalkan potensi mereka.

#### **1.5 Struktur Organisasi Skripsi**

Struktur organisasi skripsi dalam penelitian ini terdiri atas lima bab dengan beberapa subbab di dalamnya. Bab I berisikan uraian pendahuluan, kemudian dilanjutkan dengan subbab rumusan permasalahan, tujuan penelitian, manfaat penelitian, dan struktur organisasi skripsi. Kajian teoretis yang berisi paradigma teori klasik, konsep dasar tentang IST dan subtes ME, serta temuan dari penelitian terdahulu yang relevan dengan topik subtes ME tertuang di Bab II. Bab ini juga membahas evaluasi properti psikometri, termasuk validitas, reliabilitas, standardisasi norma. Metode penelitian tertuang secara terperinci di Bab III, yang terdiri dari desain penelitian, populasi dan partisipan, instrumen penelitian, prosedur penelitian, teknik analisis data, dan isu etik dalam standardisasi tes. Adapun Bab IV berisikan diskusi bahasan dari hasil temuan selama penelitian. Hasil temuan tersebut diolah, dianalisis dan hasil analisis dijelaskan di Bab ini. Bab V atau penutup berisikan kesimpulan, menjawab rumusan masalah penelitian serta mengajukan rekomendasi yang bermanfaat bagi peneliti selanjutnya dan Lab BK FIP UPI.