

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pengukuran dalam kegiatan penelitian merupakan rangkaian awal dalam analisis data. Pengukuran menyediakan berbagai kebutuhan yang berkaitan dengan validitas data yang berperan penting dalam kegiatan penelitian khususnya penelitian kuantitatif (Furqon, 2018). Berdasarkan pengukuran tersebut data dalam penelitian dapat dibagi menjadi beberapa kategori berdasarkan sifatnya, yaitu: data nominal yang menunjukkan kategorial, data ordinal yang menunjukkan peringkat, data interval yang menunjukkan kedudukan dan data rasio yang memiliki sifat nol mutlak artinya tidak dapat negatif. Setiap data tersebut dapat diolah sesuai dengan kebutuhan penelitian sebagai bentuk tanggung jawab kegiatan pengukuran dengan melalui tahap validitas.

Validitas data diperlukan dalam suatu tes sebagai proses pemberian informasi berupa ketepatan dan kecermatan pada hasil pengukuran suatu tes. Validitas umumnya digunakan sebagai usaha peneliti dalam menentukan kegunaan dan kelayakan tes tersebut (Hendryadi, 2017). Hal yang paling mendasar dalam pengembangan suatu tes adalah dengan ditemukannya dukungan berupa bukti dan teori bahwa penafsiran tes tersebut telah sesuai dengan tujuan dalam penggunaannya (Mardapi, 2008). Sehingga dengan adanya validitas tersebut dapat menunjukkan apa yang sedang diukur oleh sebuah instrumen atau apa makna yang terkandung dalam sebuah tes. Sebuah tes yang tidak melalui tahap validitas data terlebih dahulu memiliki risiko data dalam tes tersebut tidak valid, sehingga mengakibatkan pada kesalahan dalam pengambilan keputusan dan tindakan yang akhirnya berakibat buruk pada hasil penelitian (Furqon, 2018). Maka, perlu untuk dilakukannya pengukuran pada suatu tes atau instrumen. Pengukuran yang memiliki fokus sebagai bentuk dari tes validitas suatu hasil tes atau instrumen secara umum terdapat 2 yaitu CTT (*Classical Test Theory*) dan IRT (*Item Respons Theory*) (Bassegy, 2016).

CTT atau *Classical Test Theory* (Teori Tes Klasik) dikembangkan dengan tujuan memecahkan masalah terkait kesalahan pengukuran dengan melakukan

pengoreksian ketergantungan yang diamati antar variabel, sehingga dapat meredam akibat dari kesalahan pengukuran tersebut. Model CTT mengarahkan peneliti untuk fokus pada satu sumber kesalahan dengan menggunakan metode komputasi yang berbeda. Selanjutnya, IRT (*Item Response Theory*) dikembangkan dengan tujuan untuk digunakan sebagai salah satu dari metodologi psikometri modern yang diterapkan secara luas sebagai perhitungan dalam penilaian pendidikan dan bidang lain sampai ke penilaian hasil klinis (Huang & Cai 2023). Kedua bentuk pengukuran tersebut merupakan bentuk dari validitas berbagai hasil tes dan pengukuran psikologis yang digunakan sesuai dengan kebutuhan.

Pengukuran psikologis populer digunakan sebagai alat atau tes yang memberikan informasi mengenai berbagai variabel psikologis. Istilah yang populer digunakan untuk pemeriksaan psikologis adalah psikometri. Price (2012) mengemukakan bahwa istilah pengukuran merupakan bentuk dari skor yang dihasilkan oleh individu yang dapat mempresentasikan berbagai karakteristik yang dimiliki oleh dirinya. Definisi pengukuran secara umum dengan pengukuran di dalam psikologi memiliki kesamaan arti, keduanya mengukur sesuatu dengan mengharapkan dapat mendapatkan informasi mengenai sesuatu yang diukur tersebut. Pengukuran psikologis dalam mengukur sesuatu dapat mencakup berbagai macam topik mengenai kemampuan individu setelah melalui tahap validasi dan tes reliabilitas.

Reliabilitas adalah konsep yang menilai sejauh mana hasil pengukuran dari suatu instrumen dapat dipercaya. Hal ini tercermin pada konsistensi hasil pengukuran yang diperoleh dari tes yang diberikan berkali-kali kepada subjek yang sama, dengan syarat bahwa aspek yang diukur tetap konstan. Apabila terdapat perbedaan signifikan dalam hasil pengukuran, maka validitas pengukuran tersebut diragukan dan dapat dianggap tidak reliabel. Reliabilitas dan validitas saling terkait; jika alat ukur tidak mampu mengukur variabel yang dimaksud secara akurat, maka kredibilitas alat ukur tersebut menjadi terbatas.

Salah satu tes psikologi yang telah melalui tahap validitas dan tes reliabilitas yang digunakan untuk melakukan suatu pengukuran tingkat inteligensi individu adalah *Intelligenz Struktur Test* (IST). IST dikembangkan seorang ahli bernama Rudolf Amthauer pada tahun 1953, beliau mengembangkan tes IST sebagai salah satu

bentuk dari alat ukur dalam mengukur konstruk intelegensi seorang manusia. Maka, sebagai alat ukur intelegensi IST memberikan sebuah hasil berupa skor *mean* sebagai bentuk konteks intelegensi seseorang, yang populer dengan sebutan IQ (*Intelligence Quotient*). IST sendiri di dalamnya terdapat sembilan *subtest* yang digunakan dalam mengukur tingkat intelegensi seseorang sebagai bentuk asesmen peserta didik di sekolah.

Asesmen merupakan pengujian yang mengacu pada pengumpulan informasi melalui pengukuran konstruksi psikologis melalui instrumen atau prosedur tertentu, Konstruk yang diukur tersebut mengacu pada fenomena yang ada tetapi tidak dapat diamati secara langsung (Balkin & Junhke, 2018). Fokus dari asesmen tidak terletak pada perolehan skor tes tunggal namun merupakan pengertian lebih luas tentang mengintegrasikan informasi mengenai konseli (Leppma & Jones, 2013). Asesmen menggunakan berbagai metode dan berbagai sumber dalam memperoleh informasi mengenai individu, artinya asesmen tidak terfokus pada bentuk tes saja namun memiliki bentuk non tes dalam prosedurnya. Metode asesmen tersebut mencakup berbagai macam instrumen dan strategi formal dan informal, layaknya tes terstandar maupun tidak terstandar, kuesioner, inventarisasi, daftar periksa, observasi portofolio, penilaian kinerja, skala penilaian, survei, wawancara, dan tindakan lainnya yang dapat memberikan informasi mengenai individu. (Leppma & Jones, 2013), sumber utama informasi mengenai individu tersebut diperoleh dari individu tersebut atau sumber lain yang dapat mencakup anggota keluarga, pasangan, guru, dokter dan profesional lainnya. Asesmen tersebut digunakan tidak hanya secara umum oleh konselor atau profesional lainnya namun diselenggarakan juga di sekolah oleh guru bimbingan dan konseling.

Asesmen di sekolah digunakan sebagai bentuk metode evaluasi yang sebagai bahan informasi konselor sekolah atau guru BK untuk mengetahui dan memahami karakteristik peserta didik (Hays, 2013: 4). Asesmen dalam bimbingan dan konseling digunakan sebagai bentuk dari metode sistematis yang digunakan oleh guru BK dalam memahami karakteristik, lingkungan dan segala macam hubungan yang dimiliki oleh konseli dengan melakukan berbagai teknik dalam memperoleh data baik melalui tes maupun non tes (Kemendikbud, 2021). Asesmen memiliki kebermanfaatan dalam bimbingan dan konseling sebagai penyedia informasi

sehingga guru BK dapat memahami, memberikan tanggapan yang tepat, membuat berbagai perencanaan dalam perkembangan peserta didik dan tepat dalam melakukan evaluasi. Asesmen yang berkualitas baik dapat dijadikan sebagai dasar dalam melakukan berbagai tindakan edukatif yang tepat sehingga program dapat berjalan sesuai dengan ketetapan, sebaliknya tidak akan terlaksananya program bimbingan dan konseling yang komprehensif tanpa adanya asesmen yang baik sebagai dasar pembentukan program (Depdiknas, 2007, 220).

1.2 Identifikasi dan Rumusan Masalah

Penelitian mengenai validitas dan reliabilitas IST sebelumnya telah dilakukan, Penelitian oleh Siregar (2011) meneliti mengenai karakteristik psikometri *subtest* AN pada IST dengan melakukan peninjauan ulang terhadap *subtest* AN tersebut yang digunakan dalam proses seleksi Pusat Penelitian dan Pengabdian pada Masyarakat (P3M) di Fakultas Psikologi Universitas Sumatera Utara. Selanjutnya Sari (2015) menemukan bahwa 45% Item pada *Subtest* RA pada versi IST revisi memiliki indeks kesukaran yang dianggap baik, dan 55% Item sisanya memerlukan peninjauan ulang. Selanjutnya Adinugroho (2016) yaitu pengujian properti psikometrik pada alat tes IST menggunakan teori tes klasik dengan dua parameter pada subtes Form Auswahl (FA). Metode yang digunakan adalah analisis item; validasi *convergent – discriminant* dan uji reliabilitas dengan *split-half*. Hasil penelitiannya menunjukkan bukti empiris untuk mempertahankan seluruh item. Namun, analisis distraktor menunjukkan hasil yang berbeda. Hasil dari penelitian tersebut menghasilkan bahwa setiap pilihan jawaban soal belum dapat mengecoh jawaban peserta tes, maka menuju kesimpulan untuk merevisi pilihan jawaban soal. Secara umum uji validitas dan reliabilitas menunjukkan bahwa FA masih memiliki kemampuan untuk mengukur kemampuan spasial peserta tes khususnya terhadap orang Indonesia. Selanjutnya penelitian oleh Tarigan & Fadillah (2021) mengenai properti psikometri IST pada *subtest* verbal, penelitian tersebut menggunakan metode *Marginal Maksimum Likelihood* dengan pendekatan *Item Response Theory* diketahui bahwa *subtest* SE, WA dan AN masih tergolong baik dengan hasil 71,67 % memiliki kualitas yang cukup baik. Penelitian oleh Agung dan Fitri (2020) mengategorikan setiap subtes ke dalam 3 variasi dari

reliabilitas, yaitu reliabilitas buruk pada subtes SE, WA dan AN, reliabilitas cukup pada subtes GE RA dan FA dan reliabilitas bagus pada subtes ZE WU dan ME.

Pengembangan IST di Indonesia hingga saat ini belum mengalami revisi, sehingga wajar jika item-item dalam soal setiap subtes tidak lagi sejalan dengan perkembangan zaman. Tes inteligensi bertujuan untuk mengukur potensi individu dalam menyelesaikan masalah, beradaptasi terhadap perubahan lingkungan, serta belajar dari pengalaman. Penilaian ini dinyatakan melalui skor yang dibandingkan dengan norma yang ada. Oleh karena itu, norma suatu tes seharusnya diperbarui secara berkala. Hal ini disebabkan oleh perubahan karakteristik atau kemampuan populasi seiring berjalannya waktu. Salah satu perubahan yang diamati adalah peningkatan skor inteligensi selama beberapa dekade (Flynn, 1987). Jika norma dalam sebuah tes tidak pernah diperbaharui, ini dapat menyebabkan hasil skor menjadi lebih tinggi dari yang seharusnya. Ketidakcocokan norma dalam interpretasi hasil tes akan berdampak pada validitas tes itu sendiri. Validitas mengacu pada sejauh mana sebuah tes memiliki ketepatan dan kecermatan dalam mengukur atribut tertentu (Azwar, 2009). Apabila hasil tes tidak mencerminkan kondisi nyata, maka dapat dikatakan tes tersebut tidak valid. Pada akhirnya, jika tes digunakan untuk mengevaluasi calon karyawan, ketidakvalidan tes yang menghasilkan gambaran yang tidak akurat dapat berimplikasi pada keberhasilan dalam proses seleksi, sehingga individu yang diuji mungkin bukanlah yang sesuai dengan kriteria yang diharapkan, akibat adanya kesalahan dalam menafsirkan hasil skor tes.

Pada penelitian ini peneliti memfokuskan pada sub tes SE (*satzergänzung*) yang memiliki arti penyelesaian kalimat atau melengkapi kalimat. Pada subtes ini peserta tes akan melakukan penyelesaian kalimat dengan mengisi kekosongan dan memberikan makna yang logis pada kalimat tersebut. sub tes SE mengungkapkan kemampuan individu dalam kemandirian berpikir, SE juga mengukur keinginan dalam berprestasi, pengambilan keputusan, kemampuan dalam memahami realitas (*common sense*), dan pembentukan penilaian atau pendapat serta kemampuan berpikir konkrit praktis. Manfaat dari kemandirian berpikir pada individu sendiri adalah membentuk ketangguhan mental dan inisiatif serta kreativitas pada individu. Fokus penelitian ditujukan pada sub tes SE untuk melihat apakah hasil pengukuran

dari SE ini masih valid dan reliabel, sehingga apakah sub tes SE sebagai bagian dari IST berfungsi sesuai dengan tujuan awal dari pembentukan IST sebagai tes intelegensi, dengan melihat properti psikometri yang dimiliki oleh sub tes SE. Selain itu, Elvira (2011) melakukan penelitian mengenai karakteristik psikometri IST khususnya pada subtes SE dengan menggunakan teori klasik, ditemukan bahwa secara keseluruhan subtes SE belum memuaskan dan perlu diadakannya perbaikan, karena hanya 9 item soal yang dapat diterima sebagai item yang memiliki kualitas dalam mengukur kemampuan individu. Maka dinilai perlu untuk dilakukannya analisis properti psikometri pada subtes SE untuk melihat apakah subtes SE mampu untuk menyeleksi individu dalam berpikir secara mandiri. Properti psikometri yang dilihat berupa indeks deskriminasi, indeks kesukaran item, indeks reliabilitas, dan indeks validitas. Mengingat bahwa SE digunakan dalam pengukuran kemampuan berpikir mandiri individu, maka penting bahwa tes tersebut valid, reliabel dan norma yang digunakan dalam tes tersebut diperbaharui.

Penelitian ini merupakan penelitian kolaborasi antara dosen dan mahasiswa yang membentuk tim berjumlah 17 orang, yaitu 9 orang mahasiswa dan 8 orang dosen. Penelitian kolaborasi ini menjadi alasan untuk setiap orang memiliki bagiannya tersendiri dalam melakukan identifikasi terkait properti psikometrik masing-masing sub tes dalam IST. Penentuan pembagian tersebut dilakukan dengan cara musyawarah sehingga dalam pembagiannya sehingga dilakukan atas kesepakatan bersama setiap orang yang terlibat dalam penelitian kolaborasi ini. Peneliti memilih sub tes *Satzergänzung* sebagai sub tes pertama dalam *Intelligenz Struktur Test* yang akan di identifikasi terkait properti psikometri sub tes tersebut. IST yang digunakan adalah IST-2005 yang digunakan oleh Laboratorium Bimbingan dan Konseling Universitas Pendidikan Indonesia.

Pokok masalah dari penelitian merupakan : Bagaimanakah hasil properti psikometrik IST subtes SE dalam melakukan asesmen bimbingan dan konseling, maka, untuk memperjelas permasalahan yang akan diteliti masalah pokok penelitian tersebut selanjutnya diturunkan ke dalam beberapa pertanyaan penelitian sebagai berikut :

- 1) Apakah semua soal dalam subtes SE IST-2005 memiliki tingkat kesulitan yang sesuai?

- 2) Apakah semua soal dalam subtes SE IST-2005 memiliki kemampuan membedakan yang baik?
- 3) Apakah semua soal dalam sub tes SE IST-2005 valid untuk dalam melakukan pengukuran?
- 4) Apakah setiap subtes SE IST-2005 memiliki reliabilitas yang cukup untuk mengukur intelegensi?
- 5) Bagaimana norma skor untuk setiap soal di subtes SE IST-2005?

1.3 Tujuan Penelitian

Tujuan umum penelitian adalah menghasilkan properti psikometrik subtes SE IST dengan teori klasik modern untuk asesmen bimbingan dan konseling. Selanjutnya sebagai bentuk khusus dari tujuan umum yang akan menjadi tujuan dari penelitian ini adalah untuk memperoleh hal-hal sebagai berikut :

- 1) Memperoleh bukti-bukti validitas butir soal pada subtes SE IST-2005 berdasarkan teori klasik modern
- 2) Memperoleh hasil daya beda seluruh butir soal pada subtes SE IST-2005 berdasarkan teori klasik modern
- 3) Memperoleh hasil tingkat kesukaran seluruh butir soal pada subtes SE IST-2005 berdasarkan teori klasik modern
- 4) Memperoleh reliabilitas setiap butir soal subtes SE IST-2005 berdasarkan teori klasik modern.
- 5) Tersusunnya norma skor untuk setiap butir soal subtes SE IST-2005.

1.4 Manfaat Penelitian

Penelitian ini dapat memberikan manfaat baik secara teoritis dan praktis. Secara rinci berikut merupakan penjelasan mengenai manfaat dari penelitian ini.

- 1) Manfaat Teoretis

Hasil penelitian dapat memperkaya pengetahuan dan pemahaman mengenai Intelligenz Struktur Test khususnya subtes SE (*Satzergänzung*) setelah dilakukannya uji validitas dan reliabilitas. Hasil penelitian akan menyediakan bukti terkait validitas, reliabilitas dan norma dari subtes SE IST sebagai bukti standarisasi IST dengan menggunakan teori modern klasik.

2) Manfaat Praktis

a) Untuk Bimbingan dan Konseling di Sekolah

Hasil penelitian diharapkan dapat berfungsi sebagai sumber informasi yang berguna dalam melakukan penilaian potensi peserta didik, yang selanjutnya dapat dipertimbangkan dalam melakukan asesmen potensi peserta didik untuk kepentingan proses layanan bimbingan dan konseling.

b) Untuk Penelitian Selanjutnya

Hasil penelitian ini diharapkan dapat menjadi referensi yang bermanfaat bagi penelitian-penelitian mendatang yang memiliki fokus serupa, terutama terkait dengan pengembangan potensi peserta didik di lingkungan sekolah.

c) Untuk Laboratorium Bimbingan dan Konseling

Penelitian ini diharapkan dapat memberikan informasi mengenai validitas dan reliabilitas instrumen yang selama ini digunakan oleh Laboratorium Bimbingan dan Konseling dalam menilai potensi peserta didik, setelah melaksanakan uji berbasis teori klasik dan modern.