



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

TINJAUAN PUSTAKA

A. Karakteristik Psikometri

Karakteristik pengukuran psikometri adalah perbandingan antara atribut yang diukur dengan alat ukurnya, hasilnya dinyatakan secara kuantitatif, dan hasilnya bersifat deskriptif (Azwar, 2009). Karakteristik pengukuran yang pertama ialah perbandingan atribut dengan alat ukur, artinya benda atau dimensi yang diukur merupakan subjek pengukuran dan objek pengukuran adalah dimensi yang diukur, alat ukur hanya diketahui apabila atribut yang hendak diukur telah diketahui lebih dahulu. Sifat karakteristik psikometri dinyatakan dalam bentuk angka dan hanya memberikan sebatas angka tanpa interpretasi lebih jauh.

Pendekatan yang digunakan pada penelitian ini menggunakan pendekatan modern atau *item response theory*. Keunggulan pendekatan IRT berkaitan dengan *probabilitas* peserta untuk menjawab benar suatu *item* bergantung pada kemampuan peserta tes, skor memiliki *standard error of measurement* yang berbeda, *item* yang sedikit berkualitas baik dapat lebih *reliable, unbiased estimates* dari karakteristik *item* dapat diperoleh dari sampel yang tidak representatif, skor tes diinterpretasi berdasarkan isi *item*, skala interval, jenis *item* campuran dapat mengarahkan skor tes pada model pengukuran yang tepat (Embretson & Reise, 2000).

Pada pendekatan *item response theory* atau teori respon butir terdapat beberapa model pengukuran. Model paling sederhana dalam teori respon butir adalah



model logistik 1 parameter yang juga dikenal sebagai model Rasch (Embretson, 2000). Model Rasch menggunakan parameter b atau kesulitan aitem untuk membedakan antar aitem. Pemodelan Rasch mengakomodasi pendekatan probabilitas dalam memandang atribut sebuah objek ukur yang bersifat deterministik sehingga mampu mengidentifikasi objek ukur secara lebih cermat, skor murni yang bebas dari *error* pengukuran, data interval, tahan terhadap data hilang, dan pengukuran secara objektif (Sumintono & Widhiarso, 2015). Analisis yang digunakan dalam pemodelan Rasch menggunakan program *winstep*.

Karakteristik psikometri menggunakan pendekatan *item response teory* terdiri dari beberapa hal. Dalam hal ini dapat dilihat melalui:

1. Reliabilitas

Reliabilitas adalah sifat konsisten dalam pengukuran (Gregory, 2010). Anastasi dan Urbina (1997) menyatakan bahwa suatu tes dapat dikatakan reliabel apabila tes mampu memberikan hasil yang konsisten meskipun tes diberikan nilai dan skor atau diberikan pada waktu yang berbeda. Reliabilitas menjelaskan seberapa jauh pengukuran yang dilakukan berkali-kali akan menghasilkan informasi yang sama (Sumintono & Widhiarso, 2013).

Reliabilitas dalam pemodelan Rasch didasarkan pada konsistensi internal yang menekankan pada konsistensi pada butir-butir yang terdapat dalam instrumen. Reliabilitas dapat dilihat dari hasil *summary statistic* dengan melihat koefisien *cronbach alpha KR-20* (mengukur reliabilitas, yaitu interaksi antara *person* dan aitem secara keseluruhan). Kriteria nilai yang digunakan ialah < 0.5 (buruk); $0.5 - 0.6$



(jelek); 0.6 – 0.7 (cukup); 0.7 – 0.8 (bagus); dan > 0.8 (bagus sekali) (Sumintono & Widhiarso, 2015).

2. Unidimensionalitas

Unidimensionalitas adalah ukuran yang penting untuk mengevaluasi apakah instrumen yang dikembangkan mampu mengukur apa yang seharusnya diukur (Sumintono & Widhiarso, 2013). Independensi lokal juga berhubungan dengan jumlah variabel trait yang mendasari aitem. Independensi lokal akan menunjukkan unidimensionalitas ketika model teori respon butir terdiri dari parameter subjek hanya pada satu dimensi. Namun, independensi lokal juga bisa dicapai dengan data multidimensi jika masing-masing dimensi memiliki parameter subjek ataupun dengan data yang mana setiap aitem saling beketergantungan.

Pada pemodelan Rasch, analisis unidimensionalitas menggunakan analisis komponen utama (*Principal Component Analysis*) dari residual, yaitu mengukur sejauh mana keragaman dari instrumen mengukur apa yang seharusnya diukur (Sumintono & Widhiarso, 2015). Dalam analisis faktor, analisis ini untuk melihat validitas konstruk. Hasil analisis memiliki persyaratan terhadap unidimensionalitas minimal 20% dapat terpenuhi; apabila lebih 40% artinya lebih bagus, apalagi bila lebih 60% artinya istimewa.

3. Indeks Kesukaran Aitem (b)

Parameter kesukaran aitem merupakan parameter lokasi yang menunjukkan posisi *item characteristic curve* (ICC) dalam skala kemampuan. Bila nilai b lebih besar, lebih besar kemampuan yang diperlukan peserta untuk dapat menjawab aitem



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dengan benar, yang berarti aitem lebih sulit. *Item characteristic curve* (ICC) pada aitem yang lebih sulit berada disebelah kanan atau semakin tinggi kemampuan yang dibutuhkan untuk menjawab aitem dengan benar, letak KKA (Karakteristik Kurva Aitem) semakin ke kanan. Semakin mudah aitem atau semakin rendah kemampuan letak KKA semakin ke kiri (Hambleton dkk, 1991). Menurut Hambleton dan Swaminathan (1985) nilai b bergerak dari -2 logit sampai dengan $+2$ logit. Apabila mendekati -2 logit semakin mudah dan mendekati $+2$ logit akan semakin sulit.

Pengelompokan tingkat kesukaran butir aitem dapat dilakukan dengan menggunakan nilai Standar deviasi (Sumintono & Widhiarso, 2015). Menurut Sumintono dan Widhiarso (2015) lebih besar dari $+1SD$ adalah kelompok soal yang tergolong sangat sukar; 0.0 logit $+1SD$ adalah kelompok soal yang tergolong soal sukar; 0.0 logit $-1SD$ adalah kelompok soal yang tergolong soal mudah; dan lebih kecil dari $+1SD$ adalah kelompok soal yang tergolong sangat mudah.

4. *Differential Item Functioning* (DIF)

Analisis aitem dalam IRT dapat dilakukan menggunakan DIF untuk mendeteksi adanya bias respon yang disebabkan perbedaan karakteristik antara satu kelompok dengan kelompok yang lain. Aitem dikatakan terdeteksi DIF apabila peserta tes dengan karakteristik yang sama, namun dari kelompok berbeda, memiliki peluang yang berbeda dalam menjawab atau merespon aitem tersebut. Hal ini diperlukan untuk mengetahui apakah aitem-aitem yang diberikan mempunyai bias dalam kategori responden tertentu atau tidak. Bias dalam aitem dapat diketahui



berdasarkan nilai signifikansi aitem yang berada dibawah 5% (0,05) pada tabel PROB. pada program *winstep* (Sumintono & Widhiarso, 2015).

5. Aitem Fit Order

Hal ini dilakukan untuk melihat dan memberikan informasi mengenai rasio penjawab aitem dengan benar dan banyaknya penjawab aitem. Dalam pemodelan Rasch dengan menggunakan program *winstep* dapat dilihat dari tabel *item fit order*. Untuk memeriksa aitem mana yang *fit* dan *misfit* dapat dilihat melalui nilai INFIT dan MNSQ dengan kriteria: nilai MNSQ (*Outfit Mean Square*) yang diterima ($0.5 < \text{MNSQ} < 1,5$); nilai ZSTD (*Outfit Z-Standard*) yang diterima ($-2.0 < \text{ZSTD} < +2.0$) dan nilai Pt Mean Corr (*Point Measure Correlation*) dengan nilai $0.4 < \text{Pt Measurre Corr} < 0.85$ (Sumintono & Widiarso, 2013).

Dari hal-hal tersebut, maka dapat disimpulkan IRT (*Item Response Theory*) merupakan suatu pendekatan yang lebih baru dan sistematis dibandingkan dengan pendekatan teori klasik. Analisa yang dilakukan menggunakan respon jawaban subjek. Hal ini tentunya dapat mempermudah dan dapat melihat suatu *item* baik atau tidak. Pendekatan IRT juga mampu menyeleksi *item*. Hal yang paling *simple* melalui pemograman ialah menggunakan pemodelan *rasch* untuk melakukan analisa terhadap respon subjek.

B. Subtes *Merkaufgaben* (ME) pada IST

1. Pengertian Subtes ME

Merkaufgaben berasal dari bahasa Jerman yang terdiri atas dua kata yaitu, *merk* ialah ingat dan *aufgaben* yang berarti tugas (Adiwimarta, Darmojuwono, &



Hastrick, 2011) dan pada IST disebut subtes ME. Berdasarkan asal kata, *Merkaufgaben* pada IST dimaksudkan sebagai subtes yang meminta subjek untuk mengingat. Subtes ME (*Merkaufgaben*) merupakan subtes kelima dari tes IST yang berarti kemampuan memperhatikan atau mencamkan, kemampuan menyimpan atau mengingat kata-kata yang telah dipelajari, menyimpan lama dalam ingatan dan daya ingat atau ingatan. Pada subtes ME, aspek yang diukur adalah apakah ingatan subyek dapat dipercaya atau tidak, selain itu subtes ini berkaitan dengan perhatian, konsentrasi, dan ingatan seseorang tajam atau tumpul (Sumintardja & Rismiyati, 1997).

Melalui tes IST, maka dapat dilihat kemampuan mengingat seseorang dalam bentuk pemberian hafalan kata-kata. Kata-kata tersebut diberikan waktu dan juga tergantung pada kesiapan testi dalam melakukan tes. Sistem ingatan terbagi atas tiga, yaitu: (1). ingatan sensori, yakni tempat sementara penyimpanan informasi, daya ingatannya hanya sekitar satu detik; (2). Ingatan jangka pendek, yakni ingatan yang dapat bertahan selama 15-25 detik. Dan ingatan jangka panjang, yakni penyimpanan informasi yang relatif permanen serta kapasitasnya juga terbatas (Sarwono, 2012).

2. Karakteristik Subtes ME

Subtes ME ini terdiri dari 20 soal yang memiliki batasan waktu dalam pengerjannya. Hierarki struktur sistem kognitif terdapat tiga domain, yaitu: *perceiving*, kemampuan kognitif dalam memanipulasikan objek secara langsung; *conceptualizing*, kemampuan analisa yang didukung oleh kemampuan verbal dan kemampuan *reasoning*; dan *symbolizing*, representasi dari sejumlah objek atau



gagasan yang dihasilkan dalam pemikiran seseorang. Pada subtes ME, tergolong pada sistem kognitif *perceiving*, yakni terkait pada hasil tes yang menggali terkait memori atau daya ingat. Memori tersebut terbagi atas *memory for design*, *associative memory*, dan *memory span*. Pada subtes ME tergolong pada *associative memory* (Sumintardja & Rismiyati, 1997)

C. Kerangka Berpikir

Beberapa penelitian mengemukakan bahwa subtes *Merkaufgaben* (ME) memiliki karakteristik yang kurang baik sehingga perlu direvisi. Sebagaimana penelitian yang telah dilakukan oleh Rahmawati (2014) yang menemukan bahwa hanya 6 dari 20 aitem subtes ME yang termasuk kategori baik dan perlu dilakukan revisi dilihat dari indeks diskriminasi aitem dengan rentang nilai $0 < a > 2$. Aitem-aitem tersebut ialah aitem nomor 157, 158, 159, 161, 166, dan 172. Sehingga perlu dilakukan revisi terhadap aitem yang termasuk kategori tidak baik.

Selanjutnya, penelitian yang dilakukan oleh Agung dan Fitri (2016) menunjukkan bahwa subjek lebih cenderung menjawab aitem-aitem yang termasuk kategori tidak baik, seperti pada nomor 160, 162, 163, dan 164. Skor subtes ME yang diperoleh subjek termasuk kategori rendah hanya 6,38%, skor kategori sedang 91,72%, dan skor kategori tinggi sebesar 1,90%. Penelitian yang dilakukan Bawono (2008) menunjukkan hasil yang berbeda. Pengujian yang dilakukan terhadap subtes ME menggunakan pendekatan klasik yang menyatakan bahwa ME masih reliabel



dalam mengukur kemampuan mengingat seseorang. Subtes ME yang memiliki daya diskriminasi yang tergolong baik dan reliabilitas dengan rentang nilai 0.71 – 0.95.

Penemuan-penemuan di atas membuat peneliti tertarik dalam melakukan revisi subtes ME karena telah memiliki data yang lebih lengkap dan pengujian psikometri yang dilakukan belum secara keseluruhan. Peneliti berasumsi meskipun aitem subtes ME termasuk kategori baik dari segi tingkat kesukaran, mungkin saja permasalahan pada aitem subtes ME terdapat pada karakteristik psikometri lainnya. Penelitian ini terbagi atas dua tahapan. Tahap pertama, peneliti akan melakukan seleksi terhadap aitem subtes ME dengan menganalisis menggunakan model Rasch. Analisis yang dilakukan menggunakan program *Winstep* dengan melihat aitem fit dan misfit pada *aitem fit order*, bias terhadap jenis kelamin, dan indeks kesukaran. Pada tahap selanjutnya, peneliti akan melakukan revisi dan uji coba terhadap aitem subtes ME versi revisi kemudian dilakukan analisis ulang, sehingga menghasilkan aitem dengan karakteristik psikometri yang baik dan berjalan sesuai fungsinya, yaitu mengukur daya ingat seseorang.

D. Hipotesis

Hipotesis dalam penelitian ini ialah terdapat perubahan nilai koefisien dari karakteristik psikometri soal subtes ME (*Merkaufgaben*) pada IST sebelum dan sesudah revisi.