**BAB II**

**KAJIAN PUSTAKA**

* 1. **Karakteristik Butir**

Karakteristik butir merupakan parameter butir soal, yang menunjukan kualitas dari butir soal tersebut. Karakteristik butir digunakan untuk menentukan layak atau tidaknya sebuah soal digunkan. Karakteristik butir soal didapatkan dengan melakukan perhitungan dengan menggunakan rumus tertentu terhadap beberapa jawaban/respon dari peserta ujian yang mengerjakan butir soal tersebut. Dalam Teori Responsi Butir karakteristik butir terdiri dari tingkat kesulitan soal (*item difficulty*), daya pembeda soal (*item discrimination*) dan faktor kebetulan menjawab (*pseudo guesing*). Dalam penelitian ini seperti dijelaskan dalam batasan masalah karena yang berhubungan langsung dengan penyajian soal yang *adaptive* hanya 2 parameter logistik saja yaitu tingkat kesulitan soal dan daya pembeda soal. Berikut penjelasan lebih lanjutnya.

* + 1. **Tingkat kesulitan soal / *item difficulty***

Tingkat kesulitan soal atau *item difficulty* adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang bisanya dinyatakan dalam bentuk indeks. Range indeks tersebut biasanya berkisar antara 0,00 – 1,00. Soal dengan indeks kesukaran 0,00 berarti tidak ada peserta tes yang berhasil menjawab dengan benar, sebalik nya soal dengan indeks kesukaran 1,00 berarti semua peserta tes menjawab benar butir soal tersebut. Jika di tuliskan dalam rumus :

……. (2.1)

|  |
| --- |
| b = tingkat kesulitan soal |
| Np = jumlah siswa yang menjawab benar |
| N = jumlah jawaban siswa |

Klasifikasi tingkat kesulitasn soal menurut Nitko(1996)[4] seperti yang di tulis dalam bukunya dibagi menjadi 3, bisa di lihat pada table 2.1.

**Tabel 2.1 klasifikasi tingkat kesulitan soal**

|  |  |
| --- | --- |
| **Range Kesukaran / *b*** | **Klasifikasi soal** |
| 0,00 – 0,30 | Sukar |
| 0,31 – 0,70 | Sedang |
| 0,71 – 1,00 | Mudah |

* + 1. **Daya Beda Butir / *item discrimination***

Daya beda butir soal adalah nilai indeks yang menunjukan tingkat kemampuan butir soal untuk membedakan kelompok peserta ujian yang berkemampuan tinggi dengan dan berkemampuan rendah. Sebuah soal yang baik harus bisa membedakan mana siswa yang bisa dan mana siswa yang tidak bisa. Jadi butir soal yang memiliki daya beda yang baik akan terjawab benar oleh siswa yang memiliki kemampuan tinggi dan sebaliknya akan terjawab salah oleh siswa yang memiliki tingkat kemampuan rendah. Seperti tingkat kesulitan soal, daya beda juga di cari dengan menggunakan perhitungan dari analisa respon dari peserta ujian sebelumnya. Daya beda dapat di hitung dengan rumus sebagai berikut :

……….(2.2)

|  |
| --- |
| a = daya beda soal |
| BA= jumlah jawaban benar pada kelompok atas |
| BB = jumlah jawaban benar pada kelompok bawah |
| N = jumlah jawaban |

Untuk klasifikasi daya beda, pada penelitian ini penulis mengacu pada penelitian Crocker & Algina (1986)[5] yang dipaparkan pada tabel 2.2

**Tabel 2.2 klasifikasi daya beda soal**

|  |  |
| --- | --- |
| **Range Daya beda / *b*** | **konklusi** |
| b ≥ 0,4 | baik |
| 0,3 ≤ b < 0,4 | cukup baik |
| 0,2 ≤ b < 0,3 | cukup |
| b < 0,2 | buruk |

Pembagian kelompok baik atas maupun bawah diambil dari 25% - 33%[6] dari jumlah peserta yang mengerjakan butir soal tersebut yang memiliki nilai atau skor tertinggi untuk kelompok atas dan nilai atau skor terendah untuk kelompok bawah. Pada hakikatnya range daya beda adalah tidak negatif hal yaitu butir soal tersebut tidak dijawab benar justru lebih banyak oleh kelompok bawah. Hal ini berarti butir soal yang memiliki daya beda negatif justru tidak mencerminkan tingkat kemampuan siswa. Berbeda dengan tingkat kesulitan soal yang berbanding terbalik dengan tingkat kemampuan siswa, daya beda soal sebanding dengan tingkat kemampuan siswa, semakin besar daya beda soal maka semakin layak butir soal tersebut untuk disajikan.

* 1. **Teori Responsi Butir**

Pada pengujian teori klasik seperti *paper base test* dengan model urutan soal yang tetap akan menurunkan validitas dari ujian, misalnya kemungkinan *cheating* dari peserta tes. Selain itu karena keajegan tersebut maka nilai dari ujian tidak bisa menggambarkan kemampuan maksimal dari peserta ujian. Dalam ujian teori klasik / *classical test theory* dengan kondisi keajegan seperti diatas taraf sukar butir bergantung / *dependent* pada kemampuan responden. Pada peserta ujian yang memiliki tingkat kemampuan tinggi butir soal yang sama akan lebih mudah dibanding peserta dengan kemampuan yang lebih rendah. Jika untuk tujuan tes konvensional tentu tidak masalah, hanya saja dengan perkembangan teori dan ilmu dalam evaluasi pendidikan, sebuah ujian juga harus bisa mencerminkan kemampuan dari siswa dan bisa menggambarkan perkembangan kemampuan peserta tes dari waktu ke waktu.

Dengan perkembangnya asusmsi diatas dalam dunia evaluasi pendidikan, diawali tahun 1916 oleh Binet dan Smon. Kemudian pada tahun 1936, penelitian Ricahard menemukan keterkaitan antara karakteristik butir dengan perhitungan estimasi kemampuan peserta ujian. Baru kemudian pada 1952, ditemukan rumus model perhitungan peluang terjawab benar suatu butir soal berdasarkan analisis karakteristik butir soal yang kemudian disebut dengan “Parameter Logistik” yang terdiri dari tingkat kesulitan soal (*item difficulty*), daya beda (*item discrimination*) dan tebakan semu (*pseudo guessing*). Kemudian teori analisis tersebut yang kemudian dikenal dengan nama Teori