



---

## **KARAKTERISTIK TES PENALARAN ILMIAH SISWA SMA MATERI MEKANIKA BERDASARKAN ANALISIS TES TEORI RESPON BUTIR**

**Tasya Fitria Alifa<sup>1</sup>, Taufik Ramlan Ramalis<sup>1</sup>, dan Unang Purwana<sup>1</sup>**

<sup>1</sup> Jurusan Pendidikan Fisika

[tasyafitriaalifa@gmail.com](mailto:tasyafitriaalifa@gmail.com)

**Abstrak:** Hasil studi pendahuluan yang dilakukan berupa wawancara beberapa guru fisika di sekolah menengah atas di kabupaten Pandeglang menunjukkan bahwa meskipun pendidik yang dalam hal ini adalah guru sadar betul akan pentingnya mengetahui karakteristik tes sebelum diberikan kepada siswa, namun hanya beberapa guru yang melakukan analisis tes. Meskipun dilakukan analisis, alat analisis tes yang digunakan adalah analisis klasik. Telah dilakukan penelitian dengan tujuan untuk mengetahui karakteristik tes penalaran ilmiah berdasarkan analisis tes teori respon butir. Tes penalaran ilmiah yang akan dicari karakteristiknya merupakan tes penalaran ilmiah yang dikembangkan oleh peneliti. Metode penelitian yang digunakan dalam penelitian yaitu metode campuran Cresswell. Data kualitatif diperoleh berdasarkan hasil judgement ahli dan data kuantitatif diperoleh dari hasil uji lapangan. Penelitian ini dilakukan dengan subjek penelitian yaitu siswa yang diambil secara acak dari 4 sekolah menengah atas di kabupaten Pandeglang dengan kriteria subjek penelitian yaitu siswa SMA yang sudah mempelajari materi mekanika. Hasil penelitian menunjukkan bahwa karakteristik tes keterampilan penalaran ilmiah yang dikembangkan peneliti memiliki daya pembeda (a) dalam kategori baik dan memiliki tingkat kesukaran tes (b) dalam kategori sedang, serta memiliki faktor tebakan semu (c) yaitu dalam kategori baik. Tes penalaran ilmiah ini memiliki fungsi informasi sebesar 9,16 dengan SEM 0,33. Maka dapat disimpulkan bahwa tes penalaran ilmiah ini cocok diberikan kepada siswa dengan kemampuan tinggi dan tes penalaran ilmiah ini akan reliabel jika diberikan kepada siswa dengan kemampuan sedang sampai sangat tinggi.

**Kata kunci:** Penalaran Ilmiah; Karakteristik Tes; Item Response Theory

## PENDAHULUAN

Fisika merupakan salah satu cabang dari IPA. Dalam proses pembelajaran fisika melibatkan beberapa keterampilan berpikir seperti mengidentifikasi masalah, pembuatan hipotesis, penyelidikan, menentukan variabel, dan analisis data.

Penalaran ilmiah didefinisikan secara luas termasuk keterampilan berpikir yang terlibat dalam penyelidikan, eksperimen, evaluasi bukti, inferensi dan argumentasi yang dilakukan untuk memahami perubahan konseptual atau pemahaman ilmiah. Penalaran ilmiah didefinisikan juga sebagai seperangkat keterampilan penalaran dasar yang umumnya diperlukan bagi siswa untuk berhasil melakukan penyelidikan ilmiah, yang meliputi mengidentifikasi masalah, merumuskan dan menguji hipotesis, memanipulasi dan mengisolasi variabel, dan mengamati serta mengevaluasi konsekuensi. (Bao, 2009). Penelitian yang dilakukan Adey dan Shayer pada tahun 1994 menunjukkan bahwa penalaran ilmiah mempunyai dampak jangka panjang terhadap prestasi akademik siswa (Han, 2013). Lei Bao juga mengungkapkan bahwa penalaran ilmiah berhubungan dengan kemampuan kognitif seperti berpikir kritis dan berkaitan dengan prestasi akademik siswa (Bao, 2013). Oleh karena itu, keterampilan penalaran ilmiah sangat penting untuk dimiliki oleh siswa. Untuk mengukur

keterampilan penalaran ilmiah diperlukan instrumen tes yang baik dan sesuai, untuk itu sangat penting mengetahui karakteristik dari tes penalaran ilmiah yang akan digunakan untuk mengukur keterampilan penalaran ilmiah siswa.

Terdapat enam dimensi keterampilan penalaran ilmiah, meliputi: (1) konservasi massa atau volume, (2) penalaran probabilitas, (3) penalaran proporsional, (4) penalaran korelasional, (5) pengendalian variabel, dan (6) penalaran hipotesis-deduktif.

Dalam penelitian ini, instrumen tes penalaran ilmiah yang digunakan merupakan seperangkat soal pilihan ganda berjumlah 27 butir untuk mengukur ke-enam dimensi penalaran ilmiah yang tersebar disetiap butir soal, yang sebelumnya sudah melalui tahap *judgement* ahli.

Materi mekanika dipilih karena mekanika dianggap sebagai salah satu materi dasar dalam fisika yang harus dipahami terlebih dahulu sebelum memahami materi fisika yang lebih kompleks lainnya. Selain itu, berdasarkan studi literature belum ada yang mengembangkan instrumen tes penalaran ilmiah dengan topic materi mekanika.

Untuk dapat mengukur penalaran ilmiah diperlukan instrumen tes yang baik dan sesuai. Instrumen tes yang baik dan sesuai dapat diketahui melalui karakteristik dari tes itu sendiri dengan cara melakukan analisis baik terhadap butir maupun tes secara keseluruhan. Akan tetapi, permasalahan yang umum terjadi di sekolah adalah bahwa soal-soal yang diberikan kepada siswa tidak melalui tahap analisis tes untuk mengetahui karakteristik dari tes tersebut.

Untuk menganalisis karakteristik tes diperlukan alat

analisis tes. Secara umum terdapat dua alat analisis tes, yaitu analisis tes tes klasik dan teori respons butir. Masing-masing dari alat analisis tes tersebut memiliki keunggulan dan kelemahan. Keunggulan teori tes klasik terletak pada kemudahan pemahaman konsep dan penggunaannya, sehingga sering digunakan. Namun kelemahan teori ini terletak pada hasil estimasi parameter yang tergantung pada karakteristik peserta tes (*group dependent*), dan hasil estimasi kemampuan peserta tergantung pada karakteristik butir (*item dependent*). Berdasarkan uraian diatas, bahwa teori klasik memiliki ciri bahwa karakteristik butir tes tidak dapat dipisahkan dari karakteristik peserta tes.

Di dalam teori respons butir, tidak dikenal kelemahan-kelemahan yang disebutkan di atas. Tujuan utama teori respons butir adalah melepaskan keterpisahan diantara butir uji tes dengan peserta tes. Ciri-ciri teori respons butir adalah 1) karakteristik butir tidak tergantung peserta ujian, 2) skor yang digambarkan peserta ujian tidak tergantung pada tes, 3) merupakan model yang menekankan pada tingkat butir daripada tes, 4) merupakan model yang tidak mensyaratkan secara ketat tes paralel untuk menaksirkan reabilitas, dan 5) merupakan hubungan fungsional antara peserta tes dengan tingkat kemampuan yang dimiliki (Retnawati, 2014).

Terdapat tiga asumsi teori respons butir, yaitu 1) unidimensi, setiap butir tes hanya mengukur satu kemampuan, 2) independensi lokal, tidak ada korelasi antara peserta tes pada butir soal yang berbeda, dan 3) invariansi parameter, karakteristik butir soal tidak bergantung pada distribusi parameter yang menjadi

ciri peserta tes (Hambleton, dkk., 1991).

Dalam analisis tes bentuk dikotomi, terdapat 3 model yang dapat digunakan untuk melakukan analisis dengan menggunakan teori respon butir, yakni model 1 parameter logistikparameter yang digunakan hanya satu, yaitu index  $b_i$  (tingkat kesukaran), 2 parameter logistikdigunakan dua parameter yaitu index  $b_i$  (tingkat kesukaran) dan  $a_i$  (daya beda), dan 3 parameter logistikdisamping index tingkat kesukaran dan daya beda, juga parameter  $c_i$  (*pseudoguessing*) atau tebakan semu (Hambleton, 1985). Pengguna teori ini perlu memilih data yang dianalisis apakah sesuai dengan salah satu dari ketiga model tersebut dengan cara melihat puncak fungsi informasi tertinggi pada kurva fungsi informasi total di setiap model parameter logistik. Dalam penelitian ini model parameter yang sesuai dengan tes adalah model 3 parameter logistik.

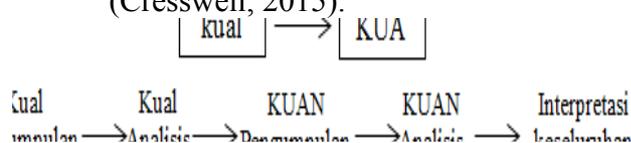
Selain itu, untuk menjelaskan kekuatan suatu butir pada perangkat tes dapat diketahui melalui fungsi informasi total. dengan fungsi informasi juga dapat diketahui tes ini akan reliabel jika di tes-kan pada siswa dengan kemampuan dalam kategori tertentu. Hasil estimasi butir tidak terlepas dengan kesalahan pengukuran. Kesalahan baku (SEM) dalam teori respons butir memiliki hubungan yang erat dengan fungsi informasi. Fungsi informasi dengan SEM mempunyai hubungan dengan berbanding terbalik kuadratik. Semakin besar fungsi informasi

maka SEM semakin kecil dan sebaliknya.

Berdasarkan dari uraian diatas, maka dilakukan penelitian dengan judul “Karakteristik Tes Penalaran Ilmiah Siswa SMA Materi Mekanika Berdasarkan Analisis Tes Teori Respon Butir”.

## METODE PENELITIAN

Metode penelitian yang digunakan yaitu metode penelitian campuran atau *mix methods* (Cresswell, 2015).



**Gambar 1. Desain Penelitian Model Sequential Exploratory Design**

Penelitian karakteristik tes penalaran ilmiah materi mekanika ini merupakan penelitian untuk mengetahui karakteristik tes yang dikembangkan oleh peneliti. Penelitian dilakukan melalui dua tahap, tahap pertama dengan mengumpulkan dan menganalisis data kualitatif sehingga dapat diperoleh rancangan produk atau konstruksi tes keterampilan penalaran ilmiah pada materi mekanika, dan penelitian tahap ke dua dengan mengumpulkan dan menganalisis data kuantitatif (eksperimen) digunakan untuk menguji efektivitas produk.

Data kualitatif yang diperoleh berdasarkan studi pendahuluan dan studi literature, serta berdasarkan catatan hasil *judgement* ahli/pakar. Data tersebut digunakan untuk mengetahui konstruksi tes keterampilan proses sains yang dibuat. Data kuantitatif yang didapat yaitu berdasarkan analisis penilaian pakar/ahli dan analisis hasil uji coba lapangan. Data ini dapat memberikan interpretasi data atau gambaran mengenai konstruksi dan karakteristik

tes keterampilan penalaran ilmiah materi mekanika.

Populasi dalam penelitian ini adalah seluruh siswa SMA yang sudah mempelajari materi mekanika. Sampel dalam penelitian ini adalah respons siswa dari beberapa sekolah menengah atas di kabupaten Pandeglang. Sampel yang digunakan dalam penelitian diambil masing-masing satu kelas dari setiap sekolah yang peneliti jadikan sebagai sampel penelitian, yaitu SMAN 3 Pandeglang berjumlah 33 siswa, SMAN 4 Pandeglang berjumlah 32, SMAN 9 Pandeglang berjumlah 27 siswa, dan SMAN CMBBS berjumlah 21 siswa. Sehingga jumlah total sampel dalam penelitian ini yaitu 113 orang. Analisis tes keterampilan penalaran ilmiah materi mekanika dilakukan berdasarkan model tiga parameter logistik (3-PL) yang dilakukan menggunakan bantuan program eirt versi 2.0.

## HASIL DAN PEMBAHASAN

Hasil pengembangan berupa seperangkat tes keterampilan penalaran ilmiah materi mekanika dengan bentuk tes pilihan ganda dengan 6 dimensi penalaran ilmiah yang tersebar disetiap butir tes. Dimensi penalaran ilmiah yang digunakan, yaitu konservasi massa atau volume, penalaran probabilitas, penalaran proporsional, penalaran korelasional, pengendalian variabel, dan penalaran hipotesis-deduktif.

### Validitas

Validitas angket validasi ahli (isi) dianalisis menggunakan validitas menurut Aiken, untuk mengetahui kesepakatan validasi oleh ahli. Berdasarkan hasil validasi yang dilakukan oleh ahli juga dapat mengetahui apakah diperlukan perombakan total, ada perbaikan atau tes keterampilan proses sains yang

telah disusun dapat digunakan untuk penelitian (tanpa perbaikan).

Pengujian validitas isi dengan menggunakan persamaan yang diusulkan oleh Aiken untuk menghitung indeks validasi berdasarkan hasil penilaian dari para ahli dengan menggunakan persamaan berikut.

$$V = \frac{\sum s}{n(c-1)} \quad (1)$$

dengan,

$$s : r - l_o$$

$l_o$  : skor terendah dalam kategori

(penyekoran)

$c$  : banyaknya kategori yang dapat dipilih rater

$r$  : skor kategori pilihan rater

$n$  : banyaknya rater

Koefisien validitas isi ( $V$ ) memiliki kemungkinan nilai 0 sampai dengan 1, nilai tersebut menunjukkan derajat dari validitas butir. Sebuah butir dapat dianggap valid jika nilai  $V \geq 0,5$  (Suseno, 2014). Interpretasi indeks validasi ahli dapat juga menggunakan kriteria sesuai dengan tabel 1.

**Tabel 2. Interpretasi indeks Aiken'V**

Hasil Validasi	Kriteria
$0,80 \leq V \leq 1,00$	Sangat valid
$0,40 \leq V \leq 0,80$	Validitas sedang
$V \leq 0,40$	Validitas kurang

**Tabel 3. Hasil validasi ahli pada tes penalaran ilmiah berdasarkan indeks Aiken'V**

penilai butir	A	B	C	$\sum S$	V	Interpretasi
1	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
2	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
3	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
4	Skor = 5; s = 4	Skor = 4; s = 3	Skor = 5; s = 4	11	0,9	Sangat Tinggi
5	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
6	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
7	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
8	Skor = 5; s = 4	Skor = 4; s = 3	Skor = 5; s = 4	11	0,9	Sangat Tinggi
9	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
10	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
11	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
12	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
13	Skor = 3; s = 2	Skor = 2; s = 1	Skor = 3; s = 2	5	0,4	Rendah
14	Skor = 3; s = 2	Skor = 2; s = 1	Skor = 3; s = 2	5	0,4	Rendah

penilai butir	A	B	C	$\sum S$	V	Interpretasi
15	= 2 Skor = 5; s = 4	= 1 Skor = 5; s = 4	= 2 Skor = 5; s = 4	12	1	Sangat Tinggi
16	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
17	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
18	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
19	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
20	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
21	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
22	Skor = 4; s = 3	Skor = 5; s = 4	Skor = 5; s = 4	11	0,9	Sangat Tinggi
23	Skor = 4; s = 3	Skor = 4; s = 3	Skor = 5; s = 4	10	0,8	Tinggi
24	Skor = 5; s = 4	Skor = 5; s = 4	Skor = 5; s = 4	12	1	Sangat Tinggi
25	Skor = 4; s = 3	Skor = 4; s = 3	Skor = 4; s = 3	9	1	Sangat Tinggi
26	Skor = 4; s = 3	Skor = 4; s = 3	Skor = 4; s = 3	9	1	Sangat Tinggi
27	Skor = 4; s = 3	Skor = 4; s = 3	Skor = 4; s = 3	9	1	Sangat Tinggi
28	Skor = 2; s = 1	Skor = 1; s = 0	Skor = 3; s = 2	3	0,3	Rendah
29	Skor = 4; s = 3	Skor = 4; s = 3	Skor = 4; s = 3	9	1	Sangat Tinggi
30	Skor = 4; s = 3	Skor = 4; s = 3	Skor = 4; s = 3	9	1	Sangat Tinggi

Berdasarkan hasil *judgement* beberapa ahli dapat diketahui apakah diperlukan perombakan total, ada perbaikan, atau tes keterampilan proses sains yang telah disusun dapat digunakan untuk penelitian (tanpa perbaikan).

Untuk menguji validitas empiris pada tes yang telah dikembangkan adalah dengan analisis modern *item response theory* (IRT). IRT yang digunakan dalam penelitian ini adalah model dikotomi 3PL.

Validasi dari hasil ujicoba digunakan untuk mengetahui estimasi karakteristik tes dan estimasi kemampuan peserta. Hasil estimasi karakteristik tes disajikan sesuai dengan penggunaan model logistik tiga parameter butir yaitu tingkat kesukaran ( $b_i$ ) dan daya beda ( $a_i$ ), dan tebakan semu ( $c_i$ )

Pada kurva karakteristik  $a_i$  proposional terhadap koefisien arah garis singgung (slope) pada titik  $\theta = b_i$ . Butir soal yang memiliki daya beda yang besar mempunyai kurva yang menanjak, sedangkan butir soal yang mempunyai daya pembeda kecil mempunyai kurva yang sangat landai. Secara teoritis, nilai  $a_i$  ini terletak antara  $-\infty$  sampai  $+\infty$ . Pada butir yang baik nilai ini mempunyai hubungan positif dengan performen pada butir dengan kemampuan yang diukur, dan  $a_i$  terletak antara 0 dan 2 (Suseno, 2014).

Nilai kemampuan peserta ( $\theta$ ) terletak di antara  $-4$  dan  $+4$ , sesuai dengan daerah asal distribusi normal. Pernyataan ini merupakan asumsi yang mendasari besar nilai  $b_i$ . Secara teoretis, nilai  $b_i$  terletak

diantara  $-\infty$  dan  $+\infty$ . Suatu butir dikatakan baik jika nilai ini berkisar antara  $-2$  dan  $+2$  (Hambleton, 1985). Jika nilai  $b_i$  mendekati  $-2$ , maka indeks kesukaran butir sangat rendah, sedangkan jika nilai  $b_i$  mendekati  $+2$  maka indeks kesukaran butir sangat tinggi untuk suatu kelompok peserta tes. Dan nilai  $c$  yang masuk dalam kategori baik jika  $c < 1/k$ , dengan  $k$  adalah banyaknya pilihan jawaban.

Secara matematis model logistik tiga parameter (Suwanto, 2011) dapat dituliskan sebagai berikut.

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}} \quad (2)$$

Dengan,

$i$  : 1, 2, 3, ...,  $n$

$P_i(\theta)$  : probabilitas peserta tes yang

memiliki kemampuan  $\theta$  dapat

menjawab butir ke- $i$  dengan benar

$\theta$  : tingkat kemampuan peserta tes

$a_i$  : indeks daya pembeda

$b_i$  : indeks kesukaran butir ke- $i$

$e$  : bilangan natural yang nilainya mendekati 2,718

$n$  : banyaknya butir dalam tes

$c$  : Indeks faktor tebakan semu

Tabel 3. Parameter butir pada model 3PL

Daya beda (a)		
Keterangan	Frekuensi i	Persentase (%)
Baik	23	85,19
Tidak baik	4	14,81
Total	27	100
Tingkat kesukaran (b)		
Keterangan	Frekuensi i	Persentase (%)
Baik	20	74,07
Tidak baik (mudah/suka r)	7	25,93
Total	27	100
Faktor tebakan semu (c)		
Keterangan	Frekuensi i	Persentase (%)
baik	26	96,3
Tidak baik	1	3,7
Total	27	100

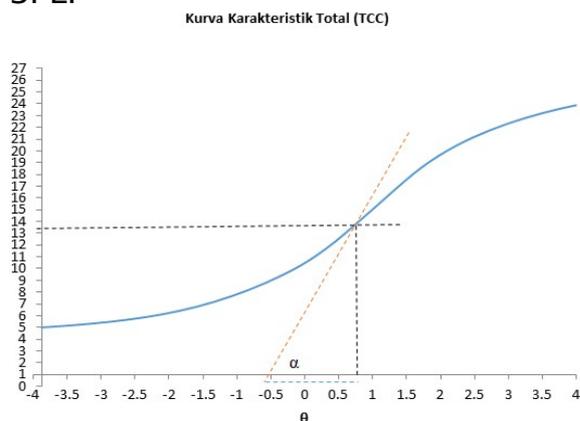
Berdasarkan Tabel 3 diatas, diketahui bahwa untuk daya beda (a) terdapat 24 butir termasuk dalam kategori baik (85,19%) dan 3 butir (14,81%) tidak baik. Tingkat kesukaran (b) terdapat 20 butir termasuk baik (74,07%) dan 7 butir (25,95%) tidak baik (mudah/sukar). Sedangkan untuk faktor tebakan semu (c) terdapat hanya 1 butir soal dari 27 butir yang masuk dalam kategori tidak baik atau sebesar 3,7%.

Tabel 4. Hasil Estimasi Parameter Butir Model 3PL

Butir	Daya pembeda	Kriteria	Tingkat kesukaran n	Kriteria	Tebakan semu n	Kriteria
1	2.280	Tidak Baik	1.609	Baik	0.067	Baik
2	4.585	Tidak Baik	1.393	Baik	0.116	Baik
3	1.850	Baik	0.390	Baik	0.142	Baik
4	0.392	Baik	4.024	Tidak Baik	0.192	Baik
5	2.996	Tidak Baik	0.621	Baik	0.144	Baik
6	3.710	Tidak Baik	0.605	Baik	0.098	Baik
7	1.933	Baik	-0.970	Tidak Baik	0.153	Baik
8	0.454	Baik	1.583	Baik	0.170	Baik
9	0.847	Baik	0.709	Baik	0.137	Baik
10	1.340	Baik	2.144	Tidak Baik	0.126	Baik
11	0.242	Baik	3.991	Tidak Baik	0.200	Baik
12	1.150	Baik	0.651	Baik	0.169	Baik
13	0.449	Baik	1.338	Baik	0.177	Baik

14	1.031	Baik	0.049	Baik	0.148	Baik
15	0.349	Baik	2.748	Tidak Baik	0.188	Baik
16	0.386	Baik	2.971	Baik	0.189	Baik
17	1.787	Baik	1.770	Baik	0.205	Tidak Baik
18	0.840	Baik	0.055	Baik	0.158	Baik
19	0.491	Baik	1.008	Baik	0.186	Baik
20	0.796	Baik	0.447	Baik	0.162	Baik
21	0.941	Baik	1.543	Baik	0.140	Baik
22	0.657	Baik	0.220	Baik	0.157	Baik
23	0.990	Baik	3.265	Tidak Baik	0.153	Baik
24	0.529	Baik	1.240	Baik	0.167	Baik
25	0.759	Baik	0.577	Baik	0.193	Baik
26	1.737	Baik	1.155	Baik	0.144	Baik
27	0.500	Baik	2.023	Tidak Baik	0.174	Baik

Salah satu kelebihan dari teori respon butir yaitu terdapat kurva karakteristik total (TCC). Kurva karakteristik total ini akan menunjukkan estimasi skor yang diperoleh siswa dari hasil mengerjakan 27 butir soal untuk setiap tingkat kemampuan. Berikut kurva karakteristik total hasil analisis teori respon butir model 3PL.



**Gambar 2. Kurva karakteristik total model 3PL**

Partisipan dengan tingkat kemampuan  $-4$  (sangat rendah) akan memperoleh skor 5 dari skor total 27, artinya hanya mampu menjawab benar 5 butir soal dari jumlah total 27 butir soal. Sedangkan partisipan dengan tingkat kemampuan  $4$  (sangat tinggi) akan memperoleh skor 24 dari skor total 27, artinya partisipan dengan kemampuan sangat tinggi akan mampu menjawab benar 24 butir soal dari

jumlah total 27 butir soal. Maka dapat disimpulkan bahwa dalam rentang  $\theta$  sama dengan  $-4$  sampai  $4$ , rentang skor yang diperoleh yaitu 5 sampai 24.

Probabilitas 1 berada pada skor 27, artinya probabilitas 0,5 berada di skor 13,5 pada kurva karakteristik total. Berdasarkan gambar 4.29, dapat diketahui nilai  $b$  dengan cara menarik garis horizontal pada nilai probabilitas 0,5 hingga pada kurva TCC. Kemudian perpotongan antara garis tersebut dan TCC jika ditarik garis vertical merupakan nilai  $b$  dari TCC. Nilai  $b$  atau tingkat kesukaran tes penalaran ilmiah berdasarkan kurva karakteristik total yaitu sebesar 0,698 atau dalam kategori tingkat kesukaran sedang. Nilai  $c$  merupakan *asymptote* dari kurva karakteristik yang merepresentasikan asumsi faktor tebakan semu dari partisipan. Berdasarkan kurva karakteristik total diketahui bahwa nilai  $c$  untuk tes penalaran ilmiah ini berada pada skor 5 artinya pada probabilitas 0,185 artinya berada dalam kategori baik karena  $c$  berada pada nilai kurang dari  $1/k$  yaitu kurang dari 0,2. Sedangkan nilai  $a$  diperoleh dari kemiringan lereng, nilai  $a$  juga bisa diperoleh dari hasil  $\tan \alpha$ . Berdasarkan kurva karakteristik total diperoleh bahwa tes penalaran ilmiah yang

dikembangkan memiliki  $\tan \alpha = 3,52 / 2,04 = 1,725$  artinya daya pembeda (a) tes penalaran ilmiah sebesar 1,725 dalam kategori baik.

### Reabilitas

Reabilitas merupakan derajat keajegan hasil pengukuran pada objek yang sama, jika dilakukan pengukuran mengenai kemampuan seorang siswa, hasil pengukurannya akan sama meskipun penguji dan butir soal berbeda namun memiliki karakteristik yang sama (Widhiarso, 2015).

Keajegan atau kestabilan alat ukur berupa sebuah nilai dapat dilakukan menggunakan perhitungan statistik. Estimasi reabilitas sebuah tes dapat menggunakan analisis model dikotomi dengan diperoleh fungsi informasi dan kesalahan pengukuran atau *standard error measurement* (SEM). Secara matematis, fungsi informasi butir tes dapat dinyatakan dengan menggunakan persamaan berikut.

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad (3)$$

keterangan,

i : 1,2,3,...,n.

$I_i(\theta)$  : fungsi informasi butir ke-i.

$P_i(\theta)$  : peluang peserta dengan kemampuan  $\theta$  menjawab benar butir ke-i.

$P'_i(\theta)$  : turunan fungsi  $P_i(\theta)$  terhadap  $\theta$ .

$Q_i(\theta)$  : peluang peserta dengan kemampuan  $\theta$  menjawab benar butir ke-i.

Nilai fungsi informasi tes akan tinggi jika butir tes memiliki fungsi informasi yang tinggi juga. Secara matematis,  $FI_i(\theta)$  atau fungsi informasi tes dapat dinyatakan dengan persamaan 1.

$$FI(\theta) = \sum_{i=1}^n I_i(\theta) \quad (4)$$

keterangan,

i : 1,2,3,...,n.

$I_i(\theta)$  : fungsi informasi butir ke-i.

$FI(\theta)$  : fungsi informasi tes

Kesalahan penaksiran standar atau kesalahan pengukuran atau *standard error measurement* (SEM) yang memiliki hubungan terbalik dengan fungsi informasi. Semakin besar fungsi informasi maka semakin kecil nilai SEM (Retnawati, 2014). Secara matematis, nilai estimasi SEM dapat dinyatakan dengan persamaa berikut.

$$SEM(\theta) = \frac{1}{\sqrt{FI(\theta)}} \quad (5)$$

keterangan,

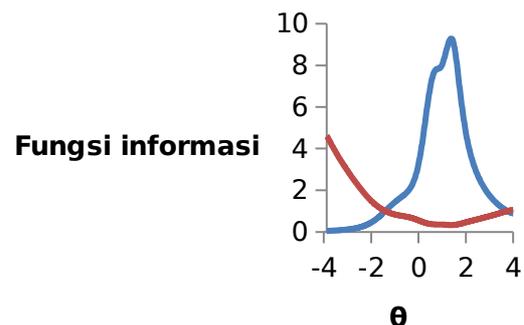
$SEM(\theta)$  : kesalahan baku dalam

pengukuran

$FI(\theta)$  : fungsi informasi tes

$(\theta)$  : kemampuan peserta tes

Hasil dari pengolahan data.hubungan antara fungsi informasi tes dengan SEM pada model 3PL diajikan oleh Gambar 2 berikut.



**Gambar 3. Total fungsi informasi dan SEM model 3PL**

Berdasarkan Gambar 2, Puncak fungsi informasi total model 3PL berada di nilai 9,16 dengan kesalahan penaksiran (SEM) sebesar 0,33 dan teta 1,46, artinya

tes penalaran ilmiah materi mekanika akan reliabel jika digunakan oleh siswa dengan kemampuan sedang. Tes penalaran ilmiah yang dikembangkan yang terdiri dari 27 butir soal dan yang diberikan kepada 113 siswa menunjukkan butir-butir soal akan cocok untuk mengetahui tingkat kemampuan partisipan dengan taraf kemampuan tinggi. Perpotongan antara kedua kurva tersebut juga menunjukkan bahwa butir-butir soal dalam tes penalaran ilmiah dapat untuk mengetahui tingkat abilitas atau tingkat kemampuan dalam rentang -1,33 sampai 3,75 yaitu partisipan dengan kategori kemampuan sedang sampai partisipan berkemampuan sangat tinggi.

## KESIMPULAN

Simpulan yang dapat diambil berdasarkan hasil penelitian yang dilakukan adalah tes keterampilan penalaran ilmiah materi mekanika untuk siswa SMA memiliki karakteristik tes yang baik dan akan reliabel jika diberikan pada siswa dengan tingkat kemampuan sedang.

## DAFTAR PUSTAKA

Lei Bao, \*. T. (2009). Learning and Scientific Reasoning 323(586). *Science AAAS*, 1-9.

- Han, J. (2013). *Scientific Reasoning : Research, Development, and Assessment (Disertasi)*. Ohio State: The Ohio State University.
- Lei Bao, e. (2013). *Reasearch*. Dipetik february 11, 2017, dari iStar Assessment: <http://www.istarassessment.org/>
- Retnawati, H. (2014). *Teori respons butir dan penerapannya*. Yogyakarta: Nuha Medika.
- Hambleton R.K., et. al. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications
- Hambleton, R. S. (1985). *Item response theory*. Boston, MA: Kluwer Inc.
- Creswell, J. W. (2015). *Research Design : Qualitative, Quantitative, and Mix Methods Approach (terjemahan)*. Yogyakarta: Pustaka Pelajar.
- Suseno, M. N. (2014). Pengembangan Pengujian Validitas Isi dan Validitas Konstrak: Interpretasi Hasil Pengujian Validitas. *Seminar Nasional Psikometri* (hal. 70-83). Yogyakarta: Publikasi Ilmiah.
- Suwarto. (2011). Teori Tes Klasik dan Teori Tes Modern 20(1). *Widyatama*, 1-10
- Widhiarso, B. S. (September 2015). *Aplikasi Pemodelan RASCH Pada Assessment Pendidikan*. Cmah: Trisma Komunikasi.

