

BAB III

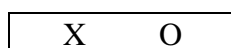
METODE PENELITIAN

3.1 Desain Penelitian

Metode penelitian yang digunakan dalam penelitian ini yaitu metode deskriptif dengan pendekatan kuantitatif. Metode deskriptif merupakan sebuah metode penelitian yang bertujuan untuk membuat deskripsi, gambaran, atau lukisan secara sistematis, faktual dan akurat mengenai fakta-fakta, sifat-sifat serta hubungan antar fenomena yang diselidiki (Nazir, 1988). Sugiyono (2005) menyatakan bahwa metode deskriptif adalah suatu metode yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas.

Pendekatan kuantitatif merupakan pendekatan penelitian yang dilakukan dengan menggunakan angka-angka dari hasil pengamatan fenomena secara objektif yang di analisis dengan menggunakan perhitungan statistik. Hal tersebut sesuai dengan yang dipaparkan oleh Sudjana (2004) bahwa metode penelitian deskriptif dengan pendekatan kuantitatif digunakan apabila bertujuan untuk mendeskripsikan atau menjelaskan peristiwa yang terjadi pada saat ini dalam bentuk angka-angka yang bermakna.

Penelitian deskriptif dengan pendekatan kuantitatif merupakan penelitian non-eksperimental yang artinya tidak ada pemberian *treatment* berupa kegiatan pembelajaran, perubahan-perubahan pada variabel bebas, tetapi menggambarkan suatu kondisi apa adanya (Sukmadinata, 2012). Desain penelitian yang digunakan dalam penelitian ini yaitu *One-Shot Design*. Arikunto (2019) menyatakan bahwa *One-Shot Design* merupakan desain penelitian yang menggunakan satu kali pengumpulan data. Pengumpulan data yang dimaksud ialah pengumpulan data dari hasil tes yang diberikan kepada peserta didik berupa tes pilihan ganda *Higher Order Thinking Skills* (HOTS) pada materi momentum dan impuls. Desain penelitian ini dapat digambarkan seperti pada Gambar 3.1.



Gambar 3. 1 Desain Penelitian *One-Shot Design*

Keterangan:

X : Pemberian tes pilihan ganda

O : Pengolahan data hasil tes dengan menggunakan analisis teori respon butir

3.2 Partisipan

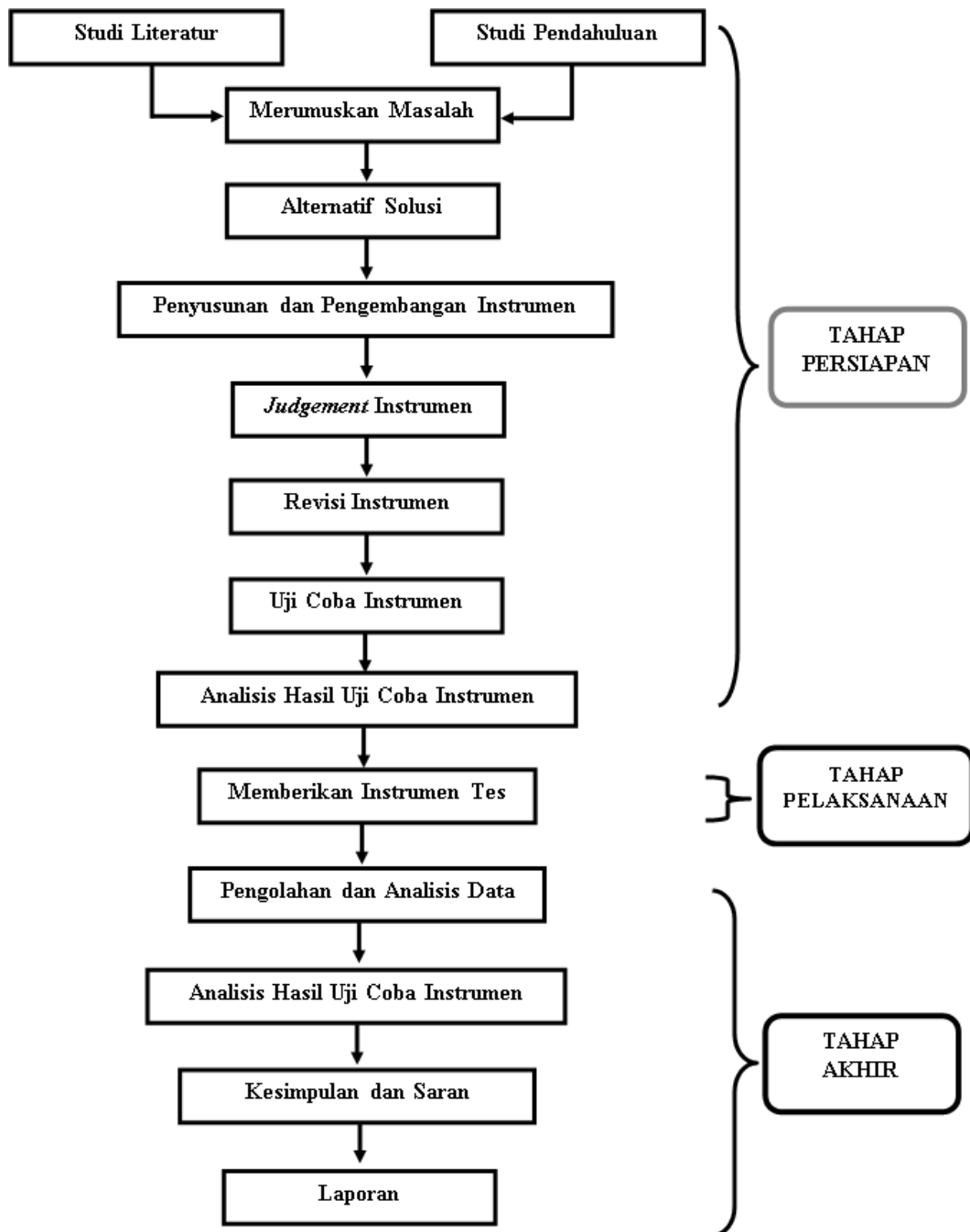
Dalam penelitian ini pengujian instrumen *Higher Order Thinking Skills* (HOTS) dilakukan pada salah satu sekolah menengah atas negeri (SMAN) di Kota Bandung. Adapun jumlah partisipan yang menjadi sampel penelitian berjumlah 122 orang peserta didik yang terdiri dari 53 orang peserta didik laki-laki dan 69 orang peserta didik perempuan. Partisipan yang mengikuti penelitian ini adalah peserta didik SMA kelas XI yang berdasarkan Silabus Fisika SMA Permendikbud No. 37 Tahun 2018 telah menerima materi momentum dan impuls pada semester sebelumnya yaitu semester genap kelas X.

3.3 Populasi dan Sampel

Dalam penelitian ini, yang menjadi populasi penelitian adalah seluruh peserta didik kelas XI pada tahun ajaran 2020/2021 di salah satu SMA Kota Bandung. Sampel dalam penelitian ini berjumlah 122 orang peserta didik kelas XI jurusan IPA. Penarikan sampel pada penelitian ini menggunakan teknik penentuan sampel yang dilakukan dengan pertimbangan tertentu atau dikenal sebagai *purposive sampling*. *Purposive sampling* merupakan salah satu jenis dari teknik *nonprobability sampling*, yaitu teknik pengambilan sampel yang tidak memberi peluang atau kesempatan yang sama bagi setiap anggota populasi untuk dipilih menjadi sampel (Sugiyono, 2017).

3.4 Prosedur Penelitian

Prosedur penelitian yang dilakukan terbagi dalam tiga tahapan, yaitu tahap persiapan, tahap pelaksanaan dan tahap akhir. Prosedur penelitian yang digunakan sebagai acuan dalam pelaksanaan penelitian disajikan dalam Gambar 3.2.



Gambar 3. 2 Prosedur Penelitian

Adapun penjelasan tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:

A. Tahap persiapan

Kegiatan yang dilaksanakan pada tahap persiapan meliputi:

- 1) Melakukan studi literatur dari jurnal-jurnal, buku, skripsi dan tesis,

- 2) Melakukan studi pendahuluan ke sekolah yang akan menjadi tempat penelitian,
- 3) Merumuskan masalah yang akan dikaji dan mencari alternatif solusi dari permasalahan yang telah dirumuskan,
- 4) Mencari alternatif solusi dari permasalahan yang telah dirumuskan,
- 5) Menyusun instrumen penelitian berupa *Higher Order Thinking Skills* (HOTS),
- 6) Melakukan validitas atau *judgement* instrumen,
- 7) Melakukan revisi instrumen berdasarkan hasil *judgement*,
- 8) Melakukan uji coba instrumen penelitian,
- 9) Melakukan analisis hasil uji coba instrumen penelitian.

B. Tahap Pelaksanaan

Kegiatan yang dilaksanakan pada tahap pelaksanaan adalah memberikan instrumen tes kepada peserta didik.

C. Tahap Pelaporan

Kegiatan yang dilaksanakan pada tahap pelaporan meliputi;

- 1) Melakukan pengolahan data hasil penelitian,
- 2) Melakukan analisis data hasil penelitian,
- 3) Menarik kesimpulan penelitian,
- 4) Menyusun dan melaporkan hasil penelitian.

3.5 Instrumen Penelitian

Instrumen digunakan untuk memperoleh data yang diinginkan sesuai dengan tujuan yang dicari dalam sebuah penelitian. Instrumen *Higher Order Thinking Skills* (HOTS) untuk mengukur kemampuan berpikir tingkat tinggi peserta didik. Tes berupa soal pilihan ganda dengan lima pilihan jawaban sebanyak 18 soal yang diberikan kepada peserta didik, sebelum digunakan untuk penelitian, *Higher Order Thinking Skills* (HOTS) terlebih dahulu diuji kelayakannya untuk mengetahui validitas, reliabilitas, daya pembeda dan taraf kesukaran. Adapun penjelasan dari pengujian tersebut adalah sebagai berikut:

a) Validitas

Uji validitas terhadap soal *Higher Order Thinking Skills* (HOTS) terdiri dari validitas konstruk dan validitas empiris.

1) Validitas Konstruk

Sebuah tes dikatakan memiliki validitas konstruk apabila butir-butir soal tes tersebut mengukur setiap aspek berpikir seperti yang terdapat dalam indikator (Arikunto, 2015). Pengujian validitas konstruk dilakukan dengan menggunakan pendapat dari para ahli (*judgement experts*). Dalam penelitian ini, pengujian validitas konstruk melibatkan tiga dosen pendidikan fisika. Proses validasi dilaksanakan dengan meminta ahli menilai validitas instrumen tes yang meliputi empat aspek penilaian, yaitu ketepatan butir soal mengukur indikator, ketepatan butir soal mengukur level kognitif penalaran (menganalisis, mengevaluasi, dan mencipta), ketepatan butir soal dengan konstruksi soal, dan ketepatan butir soal dengan penggunaan bahasa.

Data hasil validasi *Higher order thiking skills* (HOTS) digunakan untuk mencari kesepakatan para ahli untuk menentukan tingkat validitas instrumen yang digunakan. Model logistik satu parameter atau yang terkenal dengan nama model Rasch, merupakan model IRT yang paling sering digunakan. Model ini mengasumsikan bahwa semua aitem mendiskriminasi secara sama serta tidak dapat dijawab dengan benar berdasarkan tebakan (Lord, 1980). Dalam kondisi ukuran sampel yang kecil, estimasi yang dihasilkan model Rasch dimungkinkan lebih akurat jika dibandingkan dengan hasil dari model tiga parameter (Lord, 1980). Model ini juga dinilai memiliki kemudahan dalam pelaksanaannya karena jumlah parameter yang sedikit (Hambleton & Swaminathan, 1985). Sehingga dipilihlah model analisis *rasch* yang merupakan bagian dari teori respon butir dengan menggunakan bantuan *software Minifac (Facet Rasch)*. Menurut Sumintono dan Widhiarso (2013), apabila butir soal memenuhi tiga kriteria yaitu:

- a) Nilai *Outfit mean square* (MnSq) yang diterima adalah: $0,5 < \text{MnSq} < 1,5$.
- b) Nilai *Outfit Z-Standard* (ZStd) yang diterima adalah: $-2,0 < \text{ZStd} < +2,0$.
- c) Nilai *Point Measure Correlation* (PtMeaCorr) yang diterima adalah: $0,4 < \text{PtMeaCorr} < 0,85$.

Berdasarkan ketiga kriteria tersebut, bila semuanya terpenuhi maka butir soal tersebut dapat dikatakan *Outfit* atau valid. Bila hanya dua dari tiga kriteria yang terpenuhi maka butir soal tersebut termasuk kedalam kategori *Fit 2* yang masih dapat dikatakan valid. Bila hanya satu kriteria yang terpenuhi maka butir soal tersebut dikatakan *Fit 1*, dimana butir soal tersebut dinyatakan tidak valid. Begitu juga bila tidak ada kriteria yang terpenuhi, maka butir soal termasuk kedalam kategori *Misfit* atau bisa dinyatakan juga tidak valid. Berdasarkan hasil analisis validitas ahli yang telah dilakukan, dari 30 butir soal yang di validasi oleh para ahli menghasilkan 19 butir soal yang *outfit* atau dinyatakan valid karena memenuhi kriteria yang ada. Sedangkan 11 soal sisanya dinyatakan tidak valid karena masih belum memenuhi kriteria yang ada sehingga 11 butir soal tersebut dibuang. Secara lebih rinci, hasil analisis validasi ahli per butir tes ditunjukkan dalam Tabel 3.1.

Tabel 3. 1 Hasil Pengolahan Validitas Oleh Para Ahli

No. Item	Outfit MnSq	Outfit ZStd	PtMeaCorr	Category	Interpretasi
I2	0.88	-0.2	0.45	Outfit	Valid
I3	0.88	-0.2	0.45	Outfit	Valid
I5	1.28	0.9	0.20	Fit 2	Valid
I1	1.04	0.2	0.43	Outfit	Valid
I7	0.42	-0.6	0.65	Fit 2	Valid
I12	1.89	1.1	-0.09	Fit 1	Tidak Valid
I13	0.42	-0.6	0.65	Fit 2	Valid
I14	0.70	-0.1	0.37	Fit 2	Valid
I16	0.42	-0.6	0.65	Fit 2	Valid
I25	0.70	-0.1	0.37	Fit 2	Valid
I29	0.42	-0.6	0.65	Fit 2	Valid
I4	1.84	0.9	-0.09	Fit 1	Tidak Valid
I6	0.53	0.0	0.34	Fit 2	Valid
I8	0.53	0.0	0.34	Fit 2	Valid

No. Item	Outfit MnSq	Outfit ZStd	PtMeaCorr	Category	Interpretasi
I9	3.71	1.6	-0.32	Fit 1	Tidak Valid
I10	0.31	-0.2	0.54	Fit 2	Valid
I17	0.89	0.3	0.16	Fit 2	Valid
I18	0.89	0.3	0.16	Fit 2	Valid
I19	0.53	0.0	0.34	Fit 2	Valid
I20	0.89	0.3	0.16	Fit 2	Valid
I22	0.89	0.3	0.16	Fit 2	Valid
I24	0.89	0.3	0.16	Fit 2	Valid
I11	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I15	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I21	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I23	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I26	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I27	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I28	Minimum	Minimum	0.00	Miss Fit	Tidak Valid
I30	Minimum	Minimum	0.00	Miss Fit	Tidak Valid

Kelayakan sebuah data untuk diolah menggunakan *rasch* model dapat ditentukan berdasarkan *fit* data keseluruhan tes dengan *rasch* model, kriterianya sebagai berikut:

- a) *Infit-Outfit Meansquare* (MnSq) dalam nilai ideal yaitu 1, atau pada rentang 0.5 sampai 1.5.
- b) *Infit-Outfit Z-Standard* (ZStd) dalam nilai ideal yaitu 0, atau pada rentang -2 sampai 2.

Berdasarkan kriteria tersebut validitas dari sebuah tes secara keseluruhan dapat diketahui apakah sudah memenuhi kriteria kelayakan untuk di analisis menggunakan *rasch* model atau belum. Pada validitas konstruk tes ini didapatkan nilai MnSq dan ZStd secara keseluruhan yang dapat dilihat pada Tabel 3.2.

Tabel 3. 2 Nilai *Meansquare* dan *Z-Standard* Validitas Ahli

	<i>MnSq</i>		<i>ZStd</i>	
	<i>Infit</i>	<i>Outfit</i>	<i>Infit</i>	<i>Outfit</i>
<i>Person</i>	0.99	0.95	0.0	0.00
<i>Item</i>	1.01	0.95	0.2	0.1

Berdasarkan Tabel 3.2 dapat diketahui bahwa nilai MnSq dan ZStd dari keseluruhan data uji coba memenuhi kriteria kelayakan data untuk diolah menggunakan *rasch* model yang merupakan bagian dari teori respon butir.

2) Validitas Empiris (*concurrent validity*)

Instrumen tes akan dikatakan memiliki validitas empiris jika instrumen tersebut sudah diuji dari pengalaman (Arikunto, 2015). Pengujian validitas empiris dilakukan dengan melaksanakan uji coba instrumen tes yang diikuti oleh 62 peserta didik. Uji coba instrumen *higher order thiking skills* (HOTS) diberikan melalui *google form* kepada peserta didik kelas XII di salah satu SMA di Kota Bandung yang sudah mempelajari materi momentum dan impuls, kemudian diolah menggunakan bantuan *software Ministep* yang merupakan media untuk menganalisis butir soal dengan menggunakan *rasch* model yang merupakan bagian dari teori respon butir.

Berasarkan hasil uji coba yang telah di analisis menggunakan bantuan *software ministep* diketahui bahwa dari 19 butir soal yang diujikan, terdapa satu butir soal yang dinyatakan *misfit* karena tidak memenuhi kriteria yang ada sehingga soal tersebut dianggap tidak valid dan tidak dapat digunakan dalam melakukan tes. Secara lebih rinci, hasil analisis dari uji instumen tes per butir soal ditunjukkan dalam Tabel 3.3.

Tabel 3. 3 Hasil Pengolahan Validitas Data Uji Coba Instrumen Tes HOTS

No. Item	Outfit MnSq	Outfit ZStd	PtMeaCorr	Category	Interpretasi
I9	1.52	2.22	0.09	Misfit	Tidak Valid
I14	1.25	1.72	0.18	Fit 2	Valid
I4	1.18	0.90	0.20	Fit 2	Valid
I18	1.12	0.75	0.31	Fit 2	Valid
I6	1.09	0.76	0.29	Fit 2	Valid
I13	1.01	0.13	0.31	Fit 2	Valid
I3	0.99	-0.02	0.36	Fit 2	Valid
I2	1.03	0.25	0.40	Outfit	Valid
I15	1.03	0.24	0.36	Fit 2	Valid
I10	1.02	0.18	0.39	Fit 2	Valid
I7	0.90	-0.44	0.40	Outfit	Valid
I17	0.95	-0.33	0.49	Outfit	Valid

No. Item	Outfit MnSq	Outfit ZStd	PtMeaCorr	Category	Interpretasi
I8	0.79	-0.72	0.42	Outfit	Valid
I1	0.89	-0.64	0.47	Outfit	Valid
I19	0.76	-0.60	0.40	Outfit	Valid
I5	0.89	-0.25	0.41	Outfit	Valid
I11	0.80	-1.09	0.49	Outfit	Valid
I12	0.80	-1.46	0.53	Outfit	Valid
I16	0.83	-1.33	0.54	Outfit	Valid

Validitas data ini akan menjawab argumen tentang kelayakan data untuk diolah menggunakan *rasch* model. Kelayakan ini diukur berdasarkan fit data keseluruhan tes dengan *rasch* model, kriterianya sebagai berikut:

- Infit-Outfit Meansquare* (MnSq) dalam nilai ideal yaitu 1, atau pada rentang 0.5 sampai 1.5.
- Infit-Outfit Z-Standard* (ZStd) dalam nilai ideal yaitu 0, atau pada rentang -2 sampai 2.

Berdasarkan kriteria tersebut validitas dari sebuah tes secara keseluruhan dapat diketahui apakah sudah memenuhi kriteria kelayakan untuk di analisis menggunakan *rasch* model atau belum. Pada validitas empiris tes ini didapatkan nilai MnSq dan ZStd secara keseluruhan yang dapat dilihat pada Tabel 3.4.

Tabel 3. 4 Nilai *Meansquare* dan *Z-Standard* data Uji Coba

	<i>MnSq</i>		<i>ZStd</i>	
	<i>Infit</i>	<i>Outfit</i>	<i>Infit</i>	<i>Outfit</i>
<i>Perso</i>	1.0	0.99	0.0	0.01
<i>Item</i>	1.0	0.99	0.0	0.02

Berdasarkan Tabel 3.2 dapat diketahui bahwa nilai MnSq dan ZStd dari keseluruhan data uji coba memenuhi kriteria kelayakan data untuk diolah menggunakan *rasch* model yang merupakan bagian dari teori respon butir.

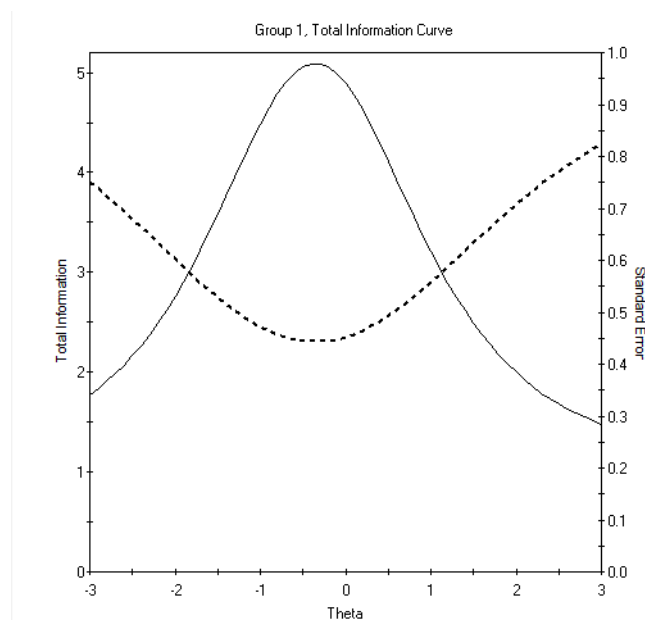
b) Reliabilitas Butir Soal

Aulia Rahman, 2021

ANALISIS KARAKTERISTIK TES HIGHER ORDER THINKING SKILLS (HOTS) PADA MATERI MOMENTUM DAN IMPULS MENGGUNAKAN TEORI RESPON BUTIR

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Hasil analisis karakteristik pada data uji coba menggunakan model 2-PL dengan bantuan *software IRTPro 3 for Student* menghasilkan grafik fungsi informasi tes yang dapat menunjukkan reliabilitas dari tes *higher order thinking skills* (HOTS). Fungsi informasi dapat dilihat pada Gambar 3.3.



Gambar 3.3 Grafik Fungsi Informasi Data Uji Coba

Berdasarkan Gambar 3.3, pada kurva fungsi informasi total model 2-PL, terdapat perpotongan antara kurva fungsi informasi dan kurva kesalahan pengukuran atau SEM yang berada pada rentang skala kemampuan -1.8 hingga 1.1 . Berdasarkan data uji coba tes *higher order thinking skills* (HOTS) tersebut, yang terdiri dari 19 butir soal berbentuk pilihan ganda yang telah dikerjakan oleh 62 peserta didik, akan reliabel atau lebih efektif jika diberikan pada peserta didik dengan rentang kemampuan -1.8 sampai 1.1 . Untuk mendeskripsikan hal tersebut, rentang kemampuan peserta didik (θ) dikategorikan pada Tabel 3.5.

Tabel 3.5 Kategori Kemampuan Penderita Didik

Nilai Kemampuan	Interpretasi Kemampuan
3.00 sampai 2.00	Sangat Tinggi
2.00 sampai 1.00	Tinggi

1.00 sampai - 1.00	Sedang
-1.00 sampai - 2.00	Rendah
-2.00 sampai - 3.00	Sangat Rendah

(Manfaat & Anasha, 2013)

Berdasarkan informasi dari Tabel 3.5, estimasi reliabilitas uji coba tes *higher order thinking skills* (HOTS) lebih efektif dalam mengukur peserta didik dengan kemampuan rendah hingga kemampuan sedang.

c) Taraf Kesukaran

Taraf kesukaran butir soal diperoleh setelah melalui proses pengolahan data menggunakan teori respon butir dua parameter logistik (2-PL) dengan bantuan *software IRTPro for Student*. Taraf kesukaran butir soal dalam *software IRTPro* disebut sebagai parameter *threshold* (b). Dari hasil uji coba instrumen tes sebanyak 19 butir soal yang diberikan, menghasilkan data seperti yang ditunjukkan pada Tabel 3.6.

Tabel 3. 6 Hasil Pengolahan Uji Coba Taraf Kesukaran

Butir Soal	Daya Pembeda	
1	-1.36	Mudah
2	-0.81	Sedang
3	-0.72	Sedang
4	2.93	Sangat Sukar
5	0.72	Sedang
6	-0.76	Sedang
7	15.18	Sangat Sukar
8	1.45	Sukar
9	-0.36	Sedang
10	-0.47	Sedang
11	-1.95	Mudah
12	-0.33	Sedang
13	-0,31	Sedang
14	-0.51	Sedang
15	-1.49	Mudah
16	-0.44	Sedang
17	-0.37	Sedang
18	-0.37	Sedang
19	0.68	Sedang

Berdasarkan hasil pengolahan data, dari 19 butir soal yang diujikan pada 62 peserta didik, diketahui bahwa terdapat tiga butir soal yang

memiliki taraf kesukaran yang dikatakan mudah yaitu butir soal nomor 1 yang memiliki nilai $b = -1.36$, nomor 11 yang memiliki $b = -1.95$, dan nomor 15 yang memiliki $b = -1.49$. Pada butir soal nomor 8 dengan nilai $b = 1.45$, termasuk kedalam kategori sukar. Butir soal yang termasuk kategori sangat sukar adalah butir soal nomor 7 dengan nilai $b = 15.18$, dan nomor 4 dengan $b = 2.93$. Butir soal sisanya yang berjumlah 13 butir soal termasuk kedalam kategori sedang yaitu butir soal nomor 2,3,5,6,9,10,12,13,14,16,17,18 dan 19

d) Daya Pembeda

Setelah instrumen *higher order thinking skills* (HOTS) sebanyak 19 butir soal diolah menggunakan teori respon butir dua parameter logistik (2-PL) dengan bantuan *software IRTPro for student* untuk mengetahui daya pembeda pada masing masing butir soal, didapatkan data seperti yang ditunjukkan dalam Tabel 3.7.

Tabel 3. 7 Hasil Pengolahan Uji Coba Daya Pembeda

Butir Soal	Daya Pembeda	Kategori
1	0.37	Baik
2	0.62	Baik
3	0.70	Baik
4	-0.17	Tidak Baik
5	-0.69	Tidak Baik
6	0.66	Baik
7	-0.03	Tidak Baik
8	-0.35	Tidak Baik
9	1.38	Baik
10	1.07	Baik
11	0.26	Baik
12	1.50	Baik
13	1.59	Baik
14	0.98	Baik
15	0.34	Baik
16	1.15	Baik
17	1.35	Baik
18	1.36	Baik
19	-0.73	Tidak Baik

Berdasarkan Tabel 3.7, daya pembeda dari setiap butir soal memiliki nilai yang bervariasi. Dari 19 butir soal yang diujikan, didapatkan nilai daya pembeda mulai dari -0,73 sampai 1,59 dan tidak semuanya dikatakan

memiliki daya pembeda yang baik. Pada uji coba instrumen tes ini, 14 butir soal memiliki nilai yang termasuk kedalam kategori baik sehingga soal tersebut dapat membedakan peserta didik kelompok atas dan kelompok bawah dengan baik. Sementara sisanya terdapat lima butir soal yang termasuk kedalam kategori tidak baik yaitu nomor 4 dengan nilai daya pembeda sebesar -0.17, butir soal nomor 5 dengan nilai daya pembeda sebesar -0.69, butir soal nomor 7 dengan nilai daya pembeda sebesar -0.03, butir soal nomor 8 dengan nilai daya pembeda sebesar -0.35, dan butir soal nomor 19 dengan nilai daya pembeda sebesar -0.73.

3.6 Teknik Analisis Data

Dalam menentukan model analisis pada teori respon butir yang paling optimal, dilakukan uji kesesuaian model dengan data yang dimiliki untuk mengetahui model analisis mana yang paling optimal dalam menganalisis data yang ada. Pemilihan model analisis yang tidak tepat akan membawa dampak pada timbulnya kesalahan dalam mengestimasi kemampuan individu (Nurchahyo, 2016). Pada dasarnya tidak ada model yang secara sempurna cocok dengan data penelitian (Wiberg, 2004). Dalam hal ini, pada teori respon butir terdapat banyak model yang dapat digunakan. Pada tes *higher order thinking skills* (HOTS) ini yang berupa kumpulan soal berbentuk pilihan ganda dengan penskoran yang dilakukan secara dikotomi yaitu pemberian skor 1 jika peserta didik menjawab benar dan 0 jika peserta didik menjawab salah pada setiap butir soal, maka yang paling tepat digunakan adalah teori respon butir unidimensi. Teori respon butir unidimensi memiliki tiga model analisis, diantaranya yaitu model logistik satu parameter (1-PL) yang sering disebut dengan istilah *rasch*, model logistik dua parameter (2-PL) yang sering disebut dengan istilah *discrimination*, dan model logistik tiga parameter (3-PL) yang sering disebut dengan istilah *guessing*. Adapun persamaan dari ketiga parameter logistik dapat dilihat pada Tabel 2.2.

Menentukan model parameter yang paling optimal dilakukan dengan cara membandingkan nilai fungsi informasi dari setiap model parameter logistik. Kesesuaian model parameter logistik dapat dilihat dari fungsi total informasi tes. Fungsi informasi menunjukkan sejauh mana masing-masing model

mampu memberikan informasi (Veerkamp & Berger, 1999). Semakin tinggi puncak dari fungsi informasi, semakin tinggi pula informasi yang bisa diberikan oleh suatu model. Model yang memiliki puncak fungsi informasi yang paling tinggi adalah model yang sesuai untuk analisis tes (Hakim, 2019). Secara matematis, persamaan fungsi informasi butir didefinisikan sebagai berikut.

$$l_i(\theta) = \frac{[P'_i(\theta)]^2}{P(\theta)Q_i(\theta)}$$

Keterangan :

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

$l_i(\theta)$: Fungsi informasi butir ke-i

$P_i(\theta)$: Peluang peserta dengan kemampuan θ menjawab benar butir i

$P'_i(\theta)$: Turunan fungsi $P_i(\theta)$ terhadap θ

$Q_i(\theta)$: Peluang peserta dengan kemampuan θ menjawab salah butir i

Fungsi informasi butir untuk model logistik tiga parameter logistik (3-PL) dinyatakan dalam persamaan berikut.

$$l_i(\theta) = \frac{2,89a_i^2(1 - c_i)}{[c_i + \exp(Da_i(\theta - b_i))][1 + \exp(-Da_i(\theta - b_i))]}$$

Keterangan :

$l_i(\theta)$: Fungsi informasi butir ke-i

θ : Tingkat kemampuan subjek

a_i : Parameter daya pembeda dari butir ke-i

b_i : Parameter taraf kesukaran butir ke-i

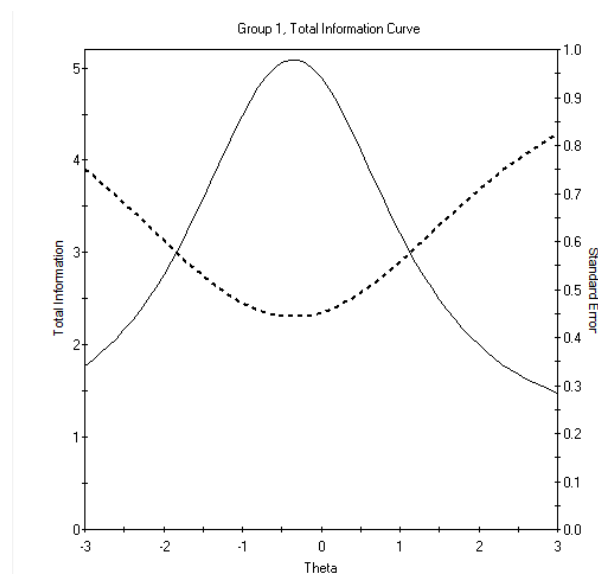
c_i : Indeks tebakan semu butir ke-i

e : Bilangan natural yang nilainya mendekati 2,718

Fungsi informasi tes merupakan penjumlahan dari fungsi informasi butir-butir soal pada tes tersebut (Hambleton & Swaminathan, 1985). Berdasarkan hal tersebut, nilai fungsi informasi tes akan tinggi bila nilai fungsi informasi butir-butir soal pada tes tersebut tinggi pula. Persamaan fungsi informasi tes, secara matematis didefinisikan sebagai berikut:

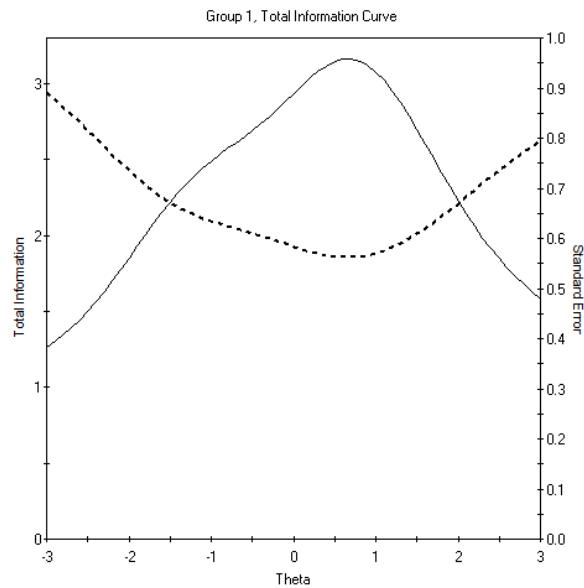
$$I(\theta) = \sum_{i=1}^n l_i(\theta)$$

Data yang telah diperoleh pada penelitian ini dianalisis menggunakan teori respon butir dengan bantuan *software IRTPro 3 for students* dan *software ministep* sebagai alat bantu untuk mengetahui karakteristik instrumen tes. Pada data uji coba terhadap 62 peserta didik, dilakukan analisis terhadap grafik kurva total informasi tes dengan menggunakan model parameter logistik yang berbeda yang kemudian akan dibandingkan nilai total informasinya untuk menentukan model mana yang paling optimal. Model yang digunakan yaitu model 1 parameter logistik, 2 parameter logistik, dan model 3 parameter logistik kemudian didapatkan hasil bahwa pada grafik kurva total informasi tes model 2 parameter logistik yang dapat dilihat pada Gambar 3.4 menunjukkan nilai informasi yang lebih baik jika dibandingkan dengan model 3 parameter logistik yang dapat dilihat pada Gambar 3.5 dan model 1 parameter logistik yang dapat dilihat pada Gambar 3.6.



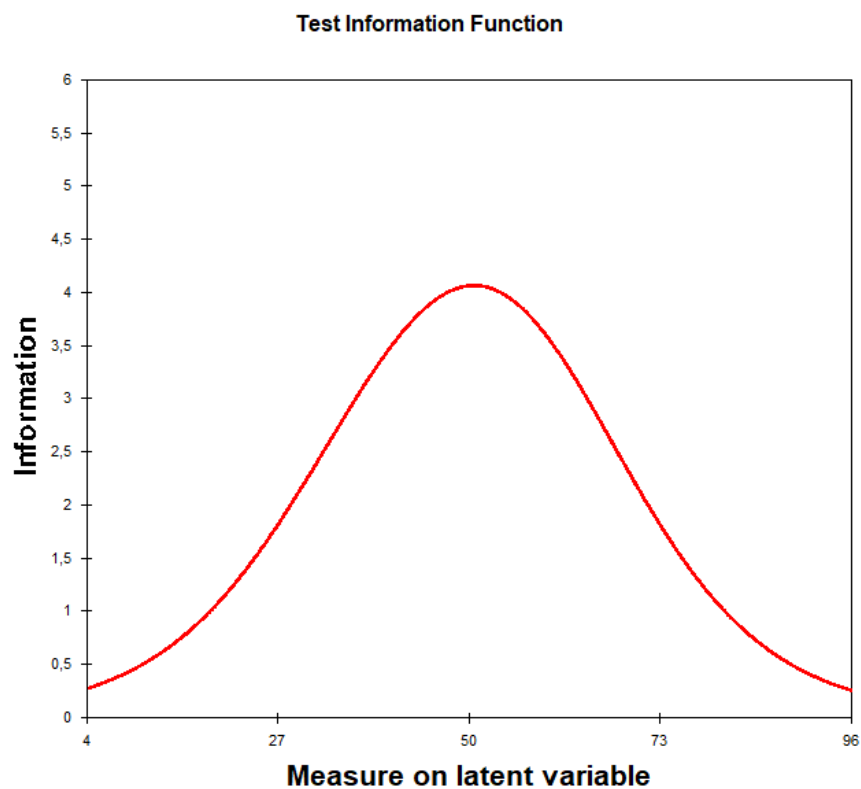
Gambar 3. 4 *Total Information Curve* Data Uji Coba 2 PL

Pada Gambar 3.3 diketahui bahwa secara umum dapat memberikan informasi bahwa tes *higher order thinking skills* (HOTS) dapat mengukur dengan baik khususnya pada kemampuan peserta didik (θ) -0.4 dengan fungsi informasinya 5.09 dan nilai *standard error measurment* 0.4433.



Gambar 3.5 *Total Information Curve* Data Uji Coba 3 PL

Pada Gambar 3.5 yang merupakan *Total Information Curve* Data Uji Coba dengan analisis model (3-PL), diketahui bahwa secara umum dapat memberikan informasi bahwa tes *higher order thinking skills* (HOTS) dapat mengukur kemampuan peserta didik (θ) 0.7 dengan fungsi informasinya 3.16 dan nilai *standard error measurement* 0.5626. Selanjutnya untuk data grafik total informasi data uji coba 1-PL dapat dilihat pada Gambar 3.6.



Gambar 3. 6 *Total Information Curve* Data Uji Coba 3 PL

Pada Gambar 3.6 yang merupakan *Total Information Curve* Data Uji Coba dengan analisis model (1-PL) yang diambil menggunakan *software ministep*,, diketahui bahwa secara umum dapat memberikan informasi dengan fungsi informasinya 4.1. Berdasarkan analisis yang dilakukan, diketahui nilai fungsi informasi untuk model 1-PL yaitu 4.1, untuk model 2-PL yaitu 5.09, dan model 3-PL yaitu 3.16. Sehingga model teori respon butir yang digunakan yaitu model logistik dua parameter (2-PL) yang sering disebut dengan istilah *discrimination* sebagai model analisis yang paling optimal.