

## BAB III

### METODE PENELITIAN

#### 3.1. Desain Penelitian

Penelitian ini menggunakan metode penelitian deskriptif dengan menggunakan data kuantitatif. Penelitian deskriptif menggunakan analisis data dengan cara mendeskripsikan atau menggambarkan data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang berlaku untuk umum atau generalisasi (Sugiyono, 2013). Lebih lanjut, Sugiyono (2013) menyatakan bahwa metode ini dapat digunakan jika peneliti hanya ingin mendeskripsikan data sampel, dan tidak ingin membuat kesimpulan yang berlaku untuk populasi di mana sampel diambil.

Penelitian yang dilakukan ini menggunakan data kuantitatif. Data kuantitatif diperoleh saat variabel yang diteliti diukur menggunakan suatu skala yang menunjukkan seberapa besar variabel tersebut muncul dan dilaporkan dalam bentuk angka/skor (Fraenkel, Wallen, & Hyun, 2012). Sugiyono (2013) mengatakan bahwa metode penelitian kuantitatif dapat diartikan sebagai metode penelitian yang berlandaskan pada filsafat positivisme, digunakan untuk meneliti pada populasi atau sampel tertentu, pengumpulan data menggunakan instrumen penelitian, analisis data bersifat kuantitatif/statistik, dengan tujuan untuk menguji hipotesis yang telah ditetapkan.

Penelitian deskriptif menggunakan data kuantitatif ini adalah penelitian non-eksperimen. Artinya, tidak ada pemberian perlakuan atau *treatment* berupa pelaksanaan kegiatan belajar dan hanya menggambarkan kondisi apa adanya. Desain penelitian yang digunakan adalah *one-shot design*. Arikunto (2019) menyatakan bahwa *One-shot design* merupakan desain penelitian yang menggunakan satu kali pengumpulan data. Desain penelitian *one-shot design* digambarkan seperti pada gambar berikut.



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

Keterangan:

X = Pemberian Instrumen Tes Keterampilan Berpikir Tingkat Tinggi

O = Hasil pengolahan data pemberian instrumen

### 3.2. Partisipan

Dalam penelitian ini, pengujian instrumen Keterampilan Berpikir Tingkat Tinggi (HOTS) dilakukan di dua sekolah menengah atas negeri (SMAN) di Kabupaten Bandung Barat. Jumlah partisipan yang menjadi sampel penelitian berjumlah 101 peserta didik. Partisipan yang mengikuti penelitian ini adalah peserta didik SMA kelas X yang telah menerima materi gerak lurus dalam pembelajaran fisika di sekolahnya.

### 3.3. Populasi dan Sampel

Populasi penelitian ini adalah seluruh siswa kelas X di SMA Kabupaten Bandung Barat, dan sampelnya adalah 101 peserta didik kelas X jurusan IPA di SMA Kabupaten Bandung Barat. Pengambilan sampel pada penelitian ini menggunakan teknik *purposive sampling*. Sugiyono (2013) menjelaskan bahwa *purposive sampling* adalah penentuan sampel yang dilakukan dengan pertimbangan tertentu. *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, 2013). Pertimbangan pengambilan sampel dalam penelitian ini adalah siswa SMA kelas X yang sudah mempelajari materi gerak lurus dalam pembelajaran fisika di sekolahnya serta banyaknya kelas yang bisa disediakan pihak sekolah untuk dilakukan penelitian.

### 3.4. Prosedur Penelitian

Prosedur penelitian ini terbagi menjadi beberapa tahapan. Adapun tahapan-tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:

#### 1. Tahap Persiapan

Kegiatan yang dilaksanakan pada tahap persiapan meliputi:

##### a. Studi Pendahuluan

Studi pendahuluan dilakukan untuk memperoleh informasi mengenai masalah yang akan dikaji dalam penelitian.

b. Pembuatan instrumen penelitian

Pembuatan instrumen penelitian yang dilakukan mengacu pada tiap indikator pada masing-masing keterampilan berpikir tingkat tinggi yang diukur pada penelitian sesuai dengan materi gerak lurus di sekolah.

c. Validasi instrumen

Validasi instrumen dilakukan untuk mendapatkan validitas konstruk pada instrumen yang dibuat.

d. Uji coba instrumen

Uji coba instrumen dilakukan untuk mendapatkan validitas isi serta reliabilitas instrumen yang akan digunakan untuk mengambil data penelitian.

e. Analisis uji coba instrumen

Hasil uji coba instrumen dianalisis untuk mengetahui kelayakannya untuk digunakan dalam mengambil data penelitian.

2. Tahap Pelaksanaan

Dalam tahap pelaksanaan, instrumen yang telah diuji coba diberikan ke peserta didik untuk mengambil data penelitian.

3. Tahap Pelaporan

a. Melakukan pengolahan data hasil penelitian

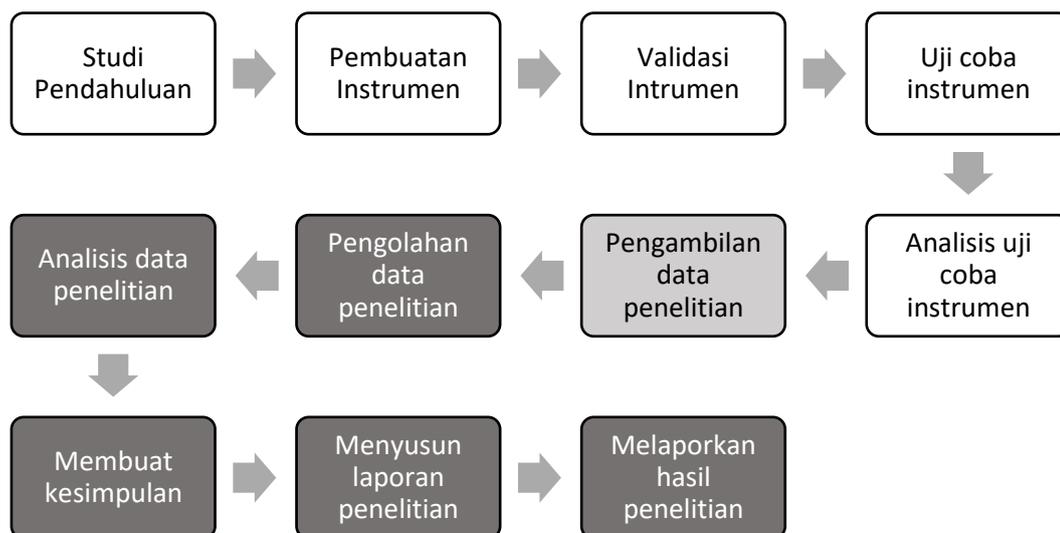
b. Menganalisis data hasil penelitian

c. Membuat kesimpulan berdasarkan hasil analisis data

d. Menyusun laporan penelitian

e. Melaporkan hasil penelitian

Ketiga tahapan tersebut ditampilkan dalam bentuk diagram sebagai berikut:



**Gambar 3. 2.** Diagram Tahapan Penelitian

### 3.5. Instrumen Penelitian

Untuk memperoleh data-data yang diperlukan, maka digunakanlah beberapa instrumen. Instrumen *Higher Order Thinking Skills* (HOTS) digunakan untuk mengukur kemampuan berpikir tingkat tinggi peserta didik. Instrumen ini berupa soal pilihan ganda dengan lima pilihan jawaban (A, B, C, D, dan E) sebanyak 22 soal yang diberikan kepada peserta didik, sebelum digunakan untuk penelitian, instrumen tersebut diuji kelayakannya untuk mengetahui validitas, reliabilitas, daya pembeda dan taraf kesukaran. Sedangkan instrumen yang digunakan dalam analisis IRT adalah *software IRTPro for Student* dan *software ministep*.

#### 3.5.1. Validitas

Validitas menunjukkan kesahihan sebuah instrumen (Arikunto, 2019). Uji validitas yang dilakukan dalam instrumen ini terdiri dari validitas konstruk dan validitas empiris. Penjelasan mengenai masing-masing uji validitas adalah sebagai berikut:

##### a) Validitas Konstruk

Pengujian validitas konstruk adalah pengujian validitas yang dilakukan dengan menggunakan pendapat para ahli yang dalam hal ini disebut *judgment*. Sebuah tes dikatakan memiliki validitas konstruk apabila butir-butir soal tes tersebut mengukur setiap aspek berpikir seperti yang terdapat dalam indikator (Arikunto, 2019). Proses pengujian validitas konstruk dalam penelitian ini melibatkan tiga dosen dari Departemen Pendidikan Fisika FPMIPA UPI

sebagai ahli yang menilai instrumen yang akan digunakan atau yang disebut sebagai validator. Proses validasi ini dilakukan dengan meminta para ahli tersebut menilai validitas instrumen tes yang meliputi empat aspek penilaian, yaitu ketepatan butir soal untuk mengukur indikator, ketepatan butir soal untuk mengukur aspek HOTS (menganalisis, mengevaluasi, dan mencipta), ketepatan butir soal dengan konstruksi soal, dan ketepatan butir soal dalam hal penggunaan bahasa.

Pengolahan data yang didapat dari penilaian ketiga ahli tersebut kemudian diolah menggunakan analisis yang diusulkan oleh Aiken (1985) yang dirumuskan sebagai berikut:

$$V = \frac{\sum(r - l_0)}{n(c - 1)} = \frac{\sum s}{n(c - 1)} \quad (3.1)$$

Keterangan:

- V : Indeks Aiken, indeks kesepakatan ahli mengenai validitas
- r : skor yang diberikan para ahli
- $l_0$  : skor terendah dalam kategori penilaian
- s : skor yang diberikan para ahli dikurangi skor terendah dalam kategori penilaian
- n : banyaknya ahli
- c : banyaknya kategori yang dapat dipilih ahli

Indeks Aiken V memiliki nilai antara 0 sampai 1 yang kemudian diinterpretasikan berdasarkan kriteria pada tabel berikut seperti dalam Retnawati (2016):

**Tabel 3. 1.** Indeks Aiken V

<b>Indeks Aiken V</b>	<b>Kriteria</b>
$V > 0,8$	Validitas Tinggi
$0,4 < V \leq 0,8$	Validitas Sedang
$V \leq 0,4$	Validitas Rendah

Setelah melakukan analisis terhadap validitas isi instrumen tes kemampuan kognitif, maka didapatkan hasil butir soal mana saja yang memiliki validitas rendah, sedang, dan tinggi. Pada butir soal dengan kategori kurang, maka perlu dilakukan perbaikan. Hasil rekapitulasi validitas konstruk instrumen HOTS secara keseluruhan ditampilkan pada tabel berikut:

**Tabel 3. 2.** Nilai V data uji coba

<b>Aspek Penilaian</b>	<b>Rata-rata Nilai V</b>	<b>Interpretasi</b>
<b>Ketepatan butir soal dengan indikator</b>	0,84	Valid (Tinggi)
<b>Ketepatan butir soal untuk mengukur HOTS</b>	0,88	Valid (Tinggi)
<b>Ketepatan konstruksi butir soal</b>	0,91	Valid (Tinggi)
<b>Ketepatan penggunaan bahasa</b>	0,85	Valid (Tinggi)
<b>Rata-rata Nilai V keseluruhan</b>		<b>0,85 (Tinggi)</b>

Hasil validasi tes kemampuan kognitif terhadap keempat aspek menunjukkan bahwa keempat aspek memperoleh rata-rata indeks Aiken V lebih dari 0,8 yang berarti memiliki validitas yang tinggi. Namun, terdapat beberapa butir soal yang perlu direvisi. Berdasarkan uji validitas konstruk ini, terdapat tiga butir soal yang tidak digunakan dalam penelitian ini. Hal tersebut karena para penilai menganggap soal-soal tersebut tidak sesuai jika digunakan untuk mengukur HOTS.

b) Validitas Empiris

Pengujian validitas empiris adalah pengujian validitas yang dilakukan dengan cara menguji coba instrumen yang dibuat kepada peserta didik. Arikunto (2019) mengatakan bahwa instrumen tes akan dikatakan memiliki validitas empiris jika instrumen tersebut sudah diuji dari pengalaman. Proses pengujian dalam penelitian ini dilakukan dengan menguji coba instrumen pada peserta didik kelas X di salah satu SMA di Kabupaten Bandung Barat yang telah mempelajari materi gerak lurus dalam pembelajaran fisika di sekolahnya. Peserta didik yang dilibatkan dalam proses ini adalah sebanyak 64 peserta didik. Data yang diperoleh dari pengujian ini kemudian diolah menggunakan model Rasch seperti dalam pengujian validitas konstruk. Analisis model Rasch ini dibantu dengan menggunakan *software ministep*. Menurut Sumintono dan Widhiarso (2013), butir soal dikatakan valid apabila memenuhi setidaknya dua kriteria berikut:

- Nilai *outfit mean square* (*MnSq*) berada pada rentang  $0,5 < MnSq < 1,5$
- Nilai *outfit Z-Standard* (*ZStd*) berada pada rentang  $-2,0 < ZStd < 2,0$
- Nilai *point measure correlation* (*PtMeaCorr*) berada pada rentang  $0,4 < PtMeaCorr < 0,85$

Jika ketiga kriteria di atas terpenuhi, maka butir soal dinyatakan *outfit* dan dianggap valid. Jika dua dari tiga kriteria terpenuhi, maka butir soal dinyatakan *fit 2* dan juga dianggap valid. Jika hanya satu kriteria yang terpenuhi maka butir soal dinyatakan *fit 1* dan dianggap tidak valid. Jika tidak ada satu pun kriteria yang terpenuhi maka butir soal dinyatakan *misfit* dan dianggap tidak valid.

Berdasarkan hasil uji coba yang telah di analisis menggunakan bantuan *software minitest* diketahui bahwa dari 22 butir soal yang diujikan, tiga butir soal dinyatakan *outfit* dan 19 butir soal sisanya dinyatakan *fit 2*. Sehingga secara keseluruhan 22 butir soal tersebut dinyatakan valid untuk digunakan dalam penelitian. Secara lebih rinci, hasil analisis dari uji instrumen tes untuk setiap butir soal yang ada dalam instrumen yang diujikan ditunjukkan dalam tabel berikut:

**Tabel 3. 3.** Hasil uji validitas data uji coba

Nomor Soal	<i>Outfit MnSq</i>	<i>Outfit ZStd</i>	<i>PtMeaCorr</i>	Kategori	Interpretasi
1	0,89	-0,97	0,41	Outfit	Valid
2	0,79	-0,83	0,42	<i>Outfit</i>	Valid
3	0,95	-0,42	0,29	<i>Fit 2</i>	Valid
4	0,80	-1,00	0,42	<i>Outfit</i>	Valid
5	1,08	0,46	0,12	<i>Fit 2</i>	Valid
6	0,97	-0,11	0,24	<i>Fit 2</i>	Valid
7	1,22	1,94	-0,19	<i>Fit 2</i>	Valid
8	0,90	-0,30	0,20	<i>Fit 2</i>	Valid
9	0,97	-0,46	0,27	<i>Fit 2</i>	Valid
10	1,13	0,99	0,07	<i>Fit 2</i>	Valid
11	0,97	-0,08	0,30	<i>Fit 2</i>	Valid
12	0,93	-0,08	0,19	<i>Fit 2</i>	Valid
13	1,39	1,49	-0,12	<i>Fit 2</i>	Valid
14	0,95	-0,65	0,33	<i>Fit 2</i>	Valid
15	0,73	-0,62	0,35	<i>Fit 2</i>	Valid
16	1,21	1,15	-0,08	<i>Fit 2</i>	Valid
17	1,37	1,61	-0,13	<i>Fit 2</i>	Valid
18	1,23	0,87	-0,01	<i>Fit 2</i>	Valid
19	0,89	-0,77	0,38	<i>Fit 2</i>	Valid
20	0,88	-2,13	0,49	<i>Fit 2</i>	Valid

Nomor Soal	Outfit MnSq	Outfit ZStd	PtMeaCorr	Kategori	Interpretasi
21	1,16	1,15	-0,07	Fit 2	Valid
22	0,84	-0,31	0,31	Fit 2	Valid

Sedangkan validitas tes secara keseluruhan dapat diketahui dari nilai *MnSQ* dan *ZStd* keseluruhan yang didapatkan. Nilai ini menunjukkan kelayakan data instrumen untuk dianalisis menggunakan analisis model *Rasch* yang merupakan salah satu model analisis teori respons butir. Kriteria kelayakan ini seperti yang ada dalam Sumintono dan Widhiarso (2013) adalah 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.

Dari analisis data uji coba yang dilakukan menggunakan *software ministep*, didapatkan nilai *MnSq* dan *ZStd* keseluruhan untuk *person* dan *item* yang ditampilkan dalam tabel sebagai berikut:

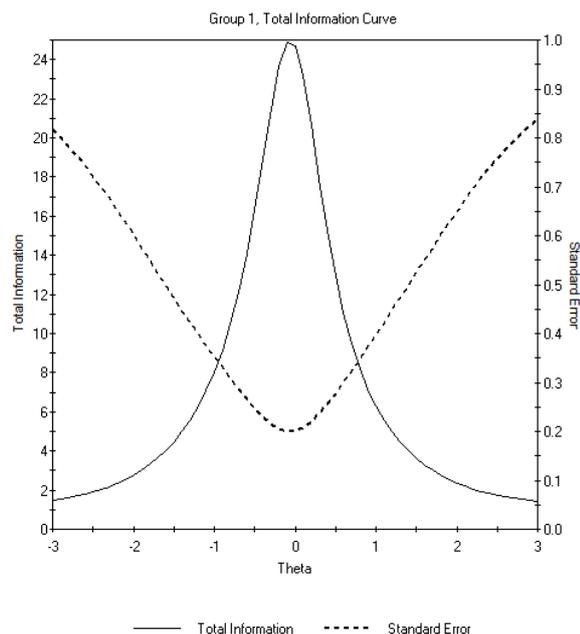
**Tabel 3. 4.** Hasil analisis kelayakan data uji coba

	<i>MnSq</i>		<i>ZStd</i>	
	<i>Infit</i>	<i>Outfit</i>	<i>Infit</i>	<i>Outfit</i>
<b>Person</b>	1,00	1,01	0,0	-0,1
<b>Item</b>	0,99	1,01	-0,1	0,0

Dari nilai-nilai yang didapatkan tersebut, terlihat bahwa baik nilai *MnSq* maupun *ZStd* keseluruhan data uji coba memenuhi kriteria kelayakan data untuk diolah menggunakan teori respons butir (IRT).

### 3.5.2. Reliabilitas

Reliabilitas data uji coba instrumen HOTS gerak lurus dianalisis menggunakan grafik fungsi informasi tes model 2-PL. Analisis ini dilakukan menggunakan bantuan *software IRTPro for Student*. Grafik fungsi informasi yang didapatkan adalah sebagai berikut:



**Gambar 3. 3.** Grafik fungsi informasi model 2-PL data uji coba

Reliabilitas instrumen yang diujikan ditunjukkan dari perpotongan antara kurva informasi total yang digambarkan dengan kurva bersambung dengan kurva *standard error* (SE) yang digambarkan dengan kurva putus-putus. Berdasarkan grafik fungsi informasi model 2-PL di atas, perpotongan antara kurva fungsi informasi dengan kurva kesalahan pengukuran standar (SEM) ada di nilai theta -1 dan 0,8. Nilai theta ini menunjukkan tingkat kemampuan peserta didik yang mengikuti tes. Sehingga, hal ini menunjukkan bahwa berdasarkan data uji coba, instrumen HOTS gerak lurus ini akan reliabel atau bisa diandalkan ketika instrumen tersebut diberikan pada peserta didik dengan kemampuan -1 hingga 0,8. Kategori kemampuan peserta didik (theta) berdasarkan Manfaat & Anasha (2013) ditunjukkan pada tabel berikut:

**Tabel 3. 5.** Interpretasi nilai theta

Theta	Interpretasi kemampuan
2,00 hingga 3,00	Sangat tinggi
1,00 hingga 2,00	Tinggi
-1,00 hingga 1,00	Sedang
-2,00 hingga -1,00	Rendah
-3,00 hingga -2,00	Sangat Rendah

Sehingga berdasarkan data uji coba di atas, instrumen HOTS gerak lurus ini akan lebih efektif ketika digunakan dalam mengukur peserta didik dengan kemampuan sedang.

### 3.5.3. Daya Pembeda

Nilai daya pembeda instrumen HOTS gerak lurus ini dianalisis menggunakan IRT model 2-PL dengan memakai *software IRTPro for Student*. Nilai daya pembeda yang baik berada pada rentang nilai 0 sampai 2. Analisis data uji coba yang dilakukan menunjukkan nilai daya pembeda ( $a$ ) yang ditampilkan dalam tabel sebagai berikut:

**Tabel 3. 6.** Nilai daya pembeda data uji coba

Butir Soal	Nilai daya pembeda ( $a$ )	Interpretasi
1	-0,64	Tidak baik
2	-1,39	Tidak baik
3	1,32	Baik
4	2,30	Tidak baik
5	1,59	Baik
6	2,14	Tidak baik
7	0,95	Baik
8	1,85	Baik
9	-0,36	Tidak baik
10	-1,38	Tidak baik
11	1,71	Baik
12	4,04	Tidak baik
13	2,14	Tidak baik
14	-0,47	Tidak baik
15	3,61	Tidak baik
16	1,44	Baik
17	1,60	Baik
18	1,90	Baik
19	1,18	Baik
20	0,19	Baik
21	1,38	Baik
22	5,01	Tidak baik

Dari data hasil uji coba instrumen ini, 11 butir soal dalam instrumen tersebut dinyatakan memiliki daya pembeda dengan kategori tidak baik. Sedangkan 11 butir soal lainnya memiliki kategori daya pembeda baik sehingga 11 butir soal tersebut

dikatakan dapat membedakan siswa kelompok atas dan kelompok bawah dengan cukup baik.

### 3.5.4. Taraf Kesukaran

Taraf kesukaran butir soal dianalisis menggunakan teori respons butir 2-PL dengan memakai *software IRTPro for Student*. Taraf kesukaran dalam analisis ini ditunjukkan oleh nilai parameter  $b$ . Nilai taraf kesukaran dikategorikan sebagai berikut:

**Tabel 3. 7.** Interpretasi nilai taraf kesukaran

Nilai taraf kesukaran	Interpretasi
$b \leq -2$	Sangat mudah
$-2 < b \leq -1$	Mudah
$-1 < b \leq 1$	Sedang
$1 < b \leq 2$	Sukar
$b > 2$	Sangat sukar

Nilai  $b$  yang diperoleh dari analisis data uji coba 22 soal yang dilakukan menghasilkan data nilai  $b$  yang ditampilkan dengan interpretasinya dalam tabel sebagai berikut:

**Tabel 3. 8.** Nilai taraf kesukaran data uji coba

Butir Soal	Nilai taraf kesukaran ( $b$ )	Interpretasi
1	0,39	Sedang
2	0,18	Sedang
3	-0,19	Sedang
4	-0,11	Sedang
5	-0,16	Sedang
6	-0,12	Sedang
7	-0,26	Sedang
8	-0,14	Sedang
9	0,69	Sedang
10	0,18	Sedang
11	-0,15	Sedang
12	-0,06	Sedang
13	-0,12	Sedang
14	-0,53	Sedang
15	-0,07	Sedang
16	-0,17	Sedang
17	-0,16	Sedang
18	-0,14	Sedang

Butir Soal	Nilai taraf kesukaran (b)	Interpretasi
19	-0,21	Sedang
20	-1,33	Mudah
21	-0,18	Sedang
22	-0,05	Sedang

Dari tabel di atas dapat dilihat bahwa dari analisis data uji coba yang dilakukan, terlihat bahwa hampir semua butir soal memiliki taraf kesukaran sedang. Sedangkan satu butir soal, yaitu soal nomor 20, memiliki taraf kesukaran yang rendah atau mudah dengan nilai  $b = -1,33$ .

### 3.6. Teknik Analisis Data

Teknik analisis data yang akan digunakan dalam penelitian ini perlu ditentukan dengan memilih model analisis mana yang paling sesuai dengan data yang diperoleh. Dalam melakukan analisis menggunakan teori respons butir, perlu dicari tahu terlebih dahulu model logistik yang paling baik berdasarkan data yang diperoleh. Seperti yang telah dijelaskan pada bagian 2.2.2., model logistik yang bisa digunakan berdasarkan Hambleton, Swaminathan, & Rogers, (1991) adalah model logistik satu parameter (1-PL), model logistik dua parameter (2-PL), dan model logistik tiga parameter (3-PL). Dalam menentukan model logistik yang mana yang paling sesuai dengan data yang diperoleh, maka dilihat dari fungsi informasinya. Veerkamp & Berger (1999) mengatakan bahwa fungsi informasi dapat menunjukkan sampai sejauh mana model yang digunakan mampu memberikan informasi. Sehingga semakin tinggi nilai fungsi informasi tes pada sebuah model, semakin tinggi juga informasi yang diberikan model tersebut.

Nilai fungsi informasi butir soal untuk model 1-PL dan 2-PL dapat dihitung menggunakan persamaan yang diberikan oleh Hambleton, Swaminathan, & Rogers, (1991) (persamaan 2.4) dengan nilai fungsi informasi butir soal  $P_i(\theta)$  yang digunakan didapat dari persamaan 2.1 atau 2.2 sesuai dengan model yang digunakan. Sedangkan untuk model 3-PL, persamaan yang digunakan adalah persamaan 2.5. Nilai fungsi informasi tes atau nilai fungsi informasi total dapat dihitung dengan menjumlahkan semua nilai fungsi informasi dari setiap butir soal seperti pada persamaan 3.6.

Analisis yang dilakukan dalam penelitian ini tidak dilakukan secara manual menggunakan persamaan-persamaan tersebut melainkan dilakukan secara otomatis menggunakan bantuan *software IRTPro for Student* dan *ministep*, sehingga nilai fungsi informasi tes yang didapatkan bisa dilihat dari kurva informasi total (*total information curve*), yaitu nilai puncak kurva tersebut. Sehingga dengan membandingkan kurva fungsi informasi total model 1-PL, 2-PL, dan 3-PL, model logistik yang paling baik untuk digunakan dalam menganalisis instrumen Keterampilan Berpikir Tingkat Tinggi pada materi gerak lurus bisa diketahui dari model mana yang memiliki nilai fungsi informasi total paling tinggi.