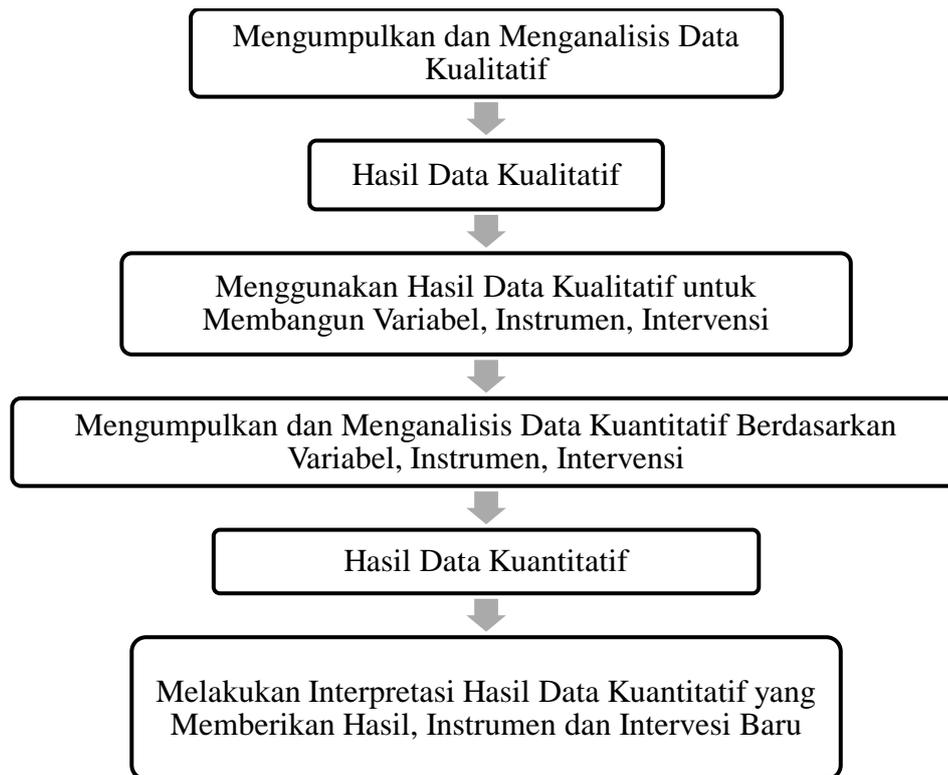


BAB III METODE PENELITIAN

3.1 Metode dan Desain Penelitian

Metode yang digunakan dalam penelitian ini adalah *mix methods*. Metode ini merupakan langkah untuk mengumpulkan, menganalisis, dan menggabungkan dua metode yaitu metode kuantitatif dan kualitatif secara bersamaan untuk menjawab masalah dalam suatu kegiatan penelitian (Cresswell, 2019). Hasil perolehan data dari metode ini akan lebih komprehensif, valid, reliabel dan obyektif (Sugiyono, 2015). Terdapat beberapa desain penelitian dari *mix methods* yaitu *convergent design*, *explanatory sequential design*, dan *exploratory sequential design*. Dari ketiga desain tersebut, desain yang digunakan dalam penelitian ini adalah *exploratory sequential design*. Desain penelitian ini melakukan pengumpulan dan analisis data kualitatif pada fase pertama kemudian pengumpulan dan analisis data secara kuantitatif pada fase kedua sehingga didapatkan hubungan antara data kualitatif dan kuantitatif (Cresswell, 2019). Menurut Tazqia (2022) tujuan dari *exploratory sequential design* adalah mengembangkan suatu produk sehingga desain ini sesuai dengan tujuan penelitian yang akan dilakukan yaitu menghasilkan produk berupa instrumen *four-tier diagnostic test* untuk mengidentifikasi keterampilan berpikir kritis. Alur penelitian dari *exploratory sequential design* ditunjukkan pada Gambar 3



Gambar 3.1 Alur Penelitian *Exploratory Sequential Design*

3.2 Populasi dan Sampel Penelitian

Populasi penelitian yang diambil adalah seluruh siswa kelas XI SMA Negeri dengan sampel sebanyak 268 siswa dari enam sekolah yang berbeda, yaitu dua sekolah yang berlokasi di Kota Bandung dan empat sekolah yang berlokasi di Kabupaten Cianjur. Sampel tersebut ditentukan dengan teknik *purposive sampling* yang dikombinasikan dengan *convenience sampling*. Teknik *purposive sampling* merupakan teknik pengambilan sampel dengan pertimbangan tertentu (Sugiyono, 2015). Sementara teknik *convenience sampling* merupakan teknik pengambilan sampel dimana peneliti memilih responden yang mudah diakses dan tersedia. Dalam penelitian ini, sampel yang ikut dalam penelitian merupakan siswa kelas XI yang telah mempelajari materi suhu dan kalor sedangkan jumlah sampel didasarkan pada ketersediaan dan kemudahan dalam memperoleh sampel melalui bantuan guru mata pelajaran fisika di sekolah untuk merekomendasikan kelas yang bersedia membantu proses penelitian. Pembagian sekolah dan jumlah sampel dapat dilihat pada Tabel 3.1

Tabel 3.1 Pembagian Sekolah dan Jumlah Sampel Penelitian

Nama Sekolah	Jumlah Sampel
Kota Bandung	
SMA A	27
SMA B	33
Kabupaten Cianjur	
SMA C	80
SMA D	61
SMA E	38
SMA F	33

3.3 Instrumen Penelitian

Instrumen yang digunakan dalam penelitian ini berupa lembar wawancara, lembar validasi ahli, instrumen *Two Tier Open Ended Test*, dan instrumen *Four Tier Diagnostic Test*.

3.3.1 Lembar Wawancara

Lembar wawancara merupakan instrumen yang digunakan untuk melakukan studi pendahuluan, khususnya studi lapangan dengan melakukan wawancara bersama guru fisika di SMA Kota Bandung. Dalam lembar wawancara tersebut, terdapat beberapa pertanyaan yang digunakan peneliti untuk mendapatkan informasi mengenai pemahaman siswa terhadap mata pelajaran fisika, kesulitan yang dialami siswa selama proses pembelajaran fisika, pelaksanaan, bentuk, dan media yang digunakan dalam evaluasi pembelajaran, pelaksanaan uji kelayakan terhadap instrumen evaluasi pembelajaran, penggunaan tes diagnostik di sekolah, khususnya mengenai pentingnya keterampilan berpikir kritis bagi siswa, penerapan keterampilan berpikir kritis dalam proses pembelajaran dan bentuk penilaian keterampilan berpikir kritis siswa.

3.3.2 Lembar Validasi Ahli

Lembar validasi ahli merupakan instrumen yang digunakan untuk menilai isi, konstruksi, dan bahasa dari instrumen yang dikembangkan dengan melibatkan validator yang ahli dalam bidang fisika. Lembar validasi terdiri dari 3 aspek penilaian yaitu validitas isi, konstruksi, dan bahasa. Validitas isi dinilai berdasarkan

kesesuaian butir soal terhadap indikator soal, aspek keterampilan berpikir kritis yang ingin dicapai, dan tingkat kognisi siswa. Validitas konstruk dinilai berdasarkan konstruksi dari butir soal yang dikembangkan, baik dari segi kejelasan dalam perumusan butir soal dan pilihan jawaban, maupun gambar/tabel/stimulus yang digunakan. Sedangkan validitas bahasa dinilai berdasarkan penggunaan tata bahasa dan rumusan kalimat pada butir soal yang dapat dipahami. Ketiga validitas tersebut disesuaikan dengan kriteria aspek validitas menurut Puspendik (2011) yang ditunjukkan pada Tabel 3.2

Tabel 3.2 Aspek Penilaian Validitas Ahli

No	Aspek Penilaian
1.	Isi
	a. Butir soal sesuai dengan indikator soal
	b. Butir soal sesuai dengan aspek keterampilan berpikir kritis yang ingin dicapai
	c. Materi sesuai dengan tingkat kognisi siswa
2.	Konstruksi
	a. Pokok soal dirumuskan dengan jelas
	b. Pokok soal tidak mengarah ke jawaban yang benar
	c. Pokok soal tidak mengandung pernyataan negatif-ganda
	d. Butir soal tidak bergantung kepada jawaban soal sebelumnya
	e. Gambar, tabel, atau stimulus dinyatakan dengan jelas dan berfungsi dengan baik
	f. Option homogen dan logis ditinjau dari segi materi
	g. Hanya ada satu jawaban yang benar
h. Panjang rumusan option relatif sama. Jika option tidak sama panjang, telah diurutkan dari yang terpendek ke yang terpanjang atau sebaliknya	
3.	Bahasa
	a. Bahasa yang digunakan sesuai dengan bahasa Indonesia yang baku
	b. Rumusan kalimat soal komunikatif
	c. Ketiadaan makna ganda pada kalimat yang digunakan

Hasil penilaian validasi ahli berupa data kuantitatif dianalisis menggunakan uji *rater (Facet Rasch)* pada *software* MINIFAC dengan memasukkan skor sesuai dengan kriteria yang telah ditentukan. Hasil perhitungan yang didapatkan akan dijadikan bahan pertimbangan dalam menarik kesimpulan terhadap validitas suatu

butir soal. Kriteria dan skor tersebut dapat dilihat pada Tabel 3.3

Tabel 3.3 Kriteria dan Skor Penilaian Validasi Ahli

Kriteria	Keterangan	Skor
V	Valid	3
VR	Valid dengan Revisi	2
TV	Tidak Valid	1

3.3.3 Instrumen *Two Tier Open Ended Test*

Instrumen *Two-Tier Open Ended Test* terdiri dari dua tingkat pertanyaan dimana tingkat pertama merupakan pertanyaan mengenai konten materi yang berbentuk *multiple choice* dengan lima pilihan jawaban dan tingkat kedua merupakan pertanyaan mengenai alasan siswa menjawab pertanyaan pada tingkat pertama yang berbentuk *open ended questions*. Instrumen ini digunakan untuk mengumpulkan jawaban alternatif siswa mengenai konsep suhu dan kalor. Jawaban alternatif tersebut akan digunakan sebagai pilihan alasan ataupun distraktor pada *tier* ketiga. Instrumen *Two Tier Open Ended Test* terdiri dari 20 soal. Format instrumen ini dapat dilihat pada Tabel 3.4

Tabel 3.4 Format Instrumen *Two Tier Open Ended Test*

Tingkat 1	Pertanyaan A. Pilihan jawaban 1 B. Pilihan jawaban 2 C. Pilihan jawaban 3 D. Pilihan jawaban 4 E. Pilihan jawaban 5
Tingkat 2	Alasan siswa memilih jawaban

3.3.4 Instrumen *Four-Tier Diagnostic Test*

Instrumen *Four-Tier Diagnostic Test* terdiri dari empat tingkatan pertanyaan dimana tingkat pertama merupakan pertanyaan mengenai konten materi yang berbentuk *multiple choice* dengan lima pilihan jawaban, tingkat kedua merupakan pertanyaan mengenai tingkat keyakinan siswa dalam memilih jawaban pada pertanyaan tingkat pertama dengan dua tingkat keyakinan (yakin dan tidak yakin),

tingkat ketiga merupakan pertanyaan mengenai alasan siswa memilih jawaban pada pertanyaan tingkat pertama yang berbentuk *multiple choice* dengan lima pilihan jawaban, dan tingkat keempat merupakan pertanyaan mengenai tingkat keyakinan siswa dalam memilih alasan jawaban pada pertanyaan tingkat ketiga dengan dua tingkat keyakinan (yakin dan tidak yakin). Instrumen ini dikembangkan berdasarkan transformasi dari instrumen *Two Tier Open Ended Test* menjadi instrumen *Four Tier Diagnostic Test*. Format instrumen tersebut dapat dilihat pada Tabel 3.5

Tabel 3.5 Format Instrumen *Four Tier Diagnostic Test*

Tingkat 1	Pertanyaan A. Pilihan jawaban 1 B. Pilihan jawaban 2 C. Pilihan jawaban 3 D. Pilihan jawaban 4 E. Pilihan jawaban 5
Tingkat 2	Tingkat keyakinan siswa dalam memilih jawaban A. Yakin B. Tidak Yakin
Tingkat 3	Alasan memilih jawaban A. Pilihan alasan 1 B. Pilihan alasan 2 C. Pilihan alasan 3 D. Pilihan alasan 4 E. Pilihan alasan 5
Tingkat 4	Tingkat keyakinan siswa dalam memilih alasan A. Yakin B. Tidak Yakin

Instrumen ini akan digunakan peneliti dalam dua tahap yaitu uji terbatas dan uji coba. Pada uji terbatas, instrumen ini digunakan untuk melihat estimasi waktu yang dibutuhkan siswa untuk menyelesaikan tes. Sementara pada uji coba, instrumen ini digunakan untuk melihat validitas, reliabilitas, dan parameter pada setiap butir soal dari tes keterampilan berpikir kritis. Contoh soal tersebut dapat dilihat pada Gambar 3.2

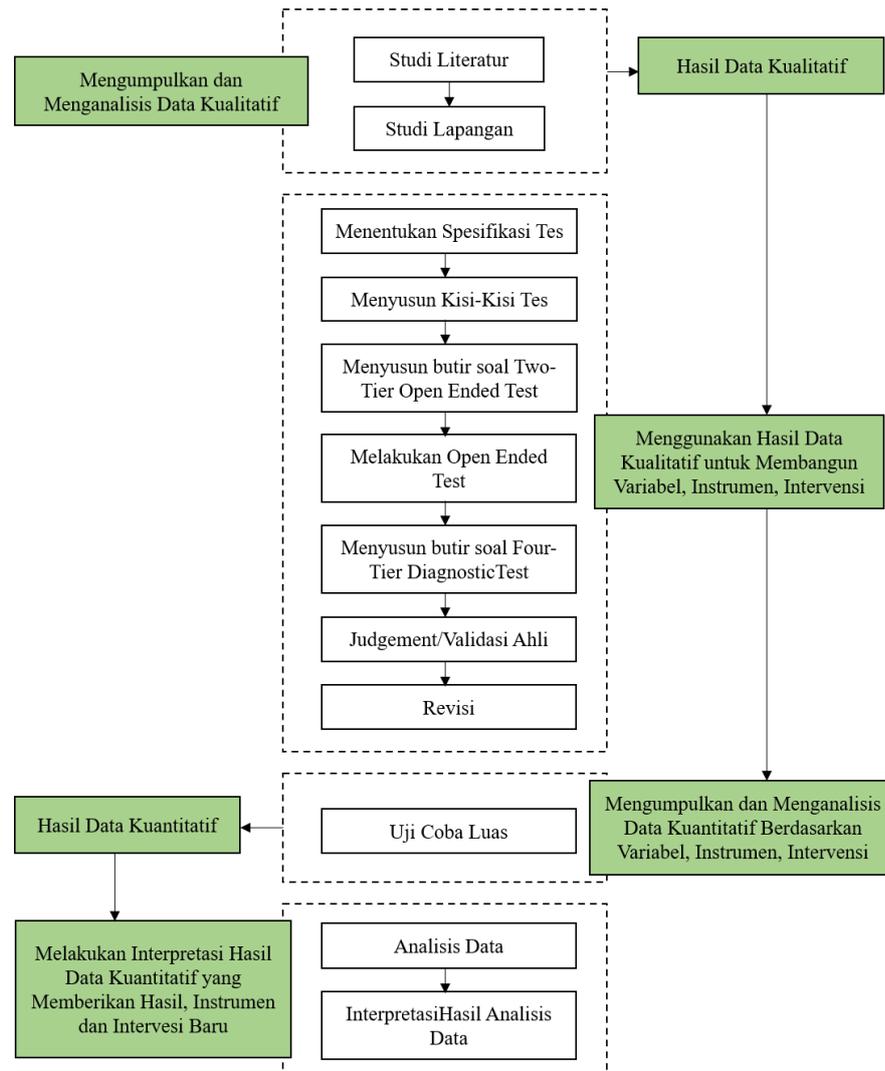
Soal 1

- 1.1 Dea membeli beberapa balon untuk merayakan ulang tahun adiknya. Karena banyak yang harus disiapkan, Dea meniup balon terlebih dahulu sebelum mengerjakan persiapan lainnya. Tempat yang aman untuk menyimpan balon agar tidak mudah meletus adalah...
- Di dalam ruangan
 - Di dalam lemari es
 - Di dalam lemari pakaian yang kosong
 - Di dalam mobil dengan jendela yang tertutup
 - Di ruangan terbuka yang terkena sinar matahari
- 1.2 Apakah anda yakin dengan jawaban tersebut?
- Yakin
 - Tidak Yakin
- 1.3 Alasan Anda memilih jawaban tersebut adalah...
- Ketika suhu lingkungan meningkat, suhu di dalam balon menurun sehingga molekul gas di dalam balon tidak bergerak dan volume balon bertambah
 - Ketika suhu lingkungan meningkat, suhu di dalam balon menurun sehingga molekul gas di dalam balon bergerak lebih cepat dan volume balon bertambah
 - Ketika suhu lingkungan meningkat, suhu di dalam balon menurun sehingga molekul gas di dalam balon bergerak lebih lambat dan volume balon bertambah
 - Ketika suhu lingkungan meningkat, suhu di dalam balon ikut meningkat sehingga molekul gas di dalam balon bergerak lebih lambat dan volume balon bertambah
 - Ketika suhu lingkungan meningkat, suhu di dalam balon ikut meningkat sehingga molekul gas di dalam balon bergerak lebih cepat dan volume balon bertambah
- 1.4 Apakah anda yakin dengan jawaban tersebut?
- Yakin
 - Tidak Yakin

Gambar 3.2 Contoh Instrumen *Four Tier Diagnostic Test*

3.4 Prosedur Penelitian

Prosedur penelitian merupakan suatu rangkaian yang menggambarkan langkah-langkah untuk melakukan kegiatan penelitian. Prosedur penelitian bertujuan mempermudah peneliti dalam menjawab permasalahan penelitian sehingga kegiatan penelitian lebih terarah dan sistematis (Novia, 2018). Prosedur penelitian ini didasarkan pada *exploratory sequential design* yang dipetakan seperti pada Gambar 3.3



Gambar 3.3 Prosedur Penelitian

Tahapan kegiatan pada prosedur penelitian yang didasarkan *exploratory sequential design* dijabarkan sebagai berikut

3.4.1 Mengumpulkan dan Menganalisis Data Kualitatif

Tahap ini dilakukan dengan mengumpulkan data kualitatif melalui studi pendahuluan dengan studi literatur dan studi lapangan. Peneliti melakukan studi literatur dengan mengkaji berbagai literatur yang berkaitan dengan keterampilan berpikir kritis, hal-hal yang harus diperhatikan saat mengembangkan keterampilan berpikir kritis siswa, penyebab rendahnya tingkat keterampilan berpikir kritis siswa, aspek keterampilan berpikir kritis menurut para ahli, instrumen yang digunakan untuk mengidentifikasi keterampilan berpikir kritis siswa, tes diagnostik yang digunakan untuk mengidentifikasi keterampilan berpikir kritis, dan materi fisika

yang akan dijadikan konteks dalam tes. Selain itu, peneliti melakukan studi lapangan melalui wawancara kepada guru mata pelajaran fisika untuk mendapatkan informasi mengenai pemahaman siswa terhadap mata pelajaran fisika, kesulitan yang dialami siswa selama proses pembelajaran fisika, pelaksanaan, bentuk, dan media yang digunakan dalam evaluasi pembelajaran, pelaksanaan uji kelayakan terhadap instrumen evaluasi pembelajaran, penggunaan tes diagnostik di sekolah, khususnya mengenai pentingnya keterampilan berpikir kritis bagi siswa, penerapan keterampilan berpikir kritis dalam proses pembelajaran dan bentuk penilaian keterampilan berpikir kritis siswa. Data tersebut akan digunakan peneliti untuk mengetahui permasalahan penelitian.

3.4.2 Hasil Data Kualitatif

Hasil data kualitatif dari studi pendahuluan diserahkan kepada dosen pembimbing untuk meminta persetujuan, apakah data tersebut dapat digunakan sebagai referensi dalam proses penelitian yang akan dilakukan.

3.4.3 Menggunakan Hasil Data Kualitatif untuk Membangun Variabel, Instrumen, Intervensi

Berdasarkan hasil studi pendahuluan dan persetujuan dosen untuk melanjutkan proses penelitian, dilakukan pembuatan tes keterampilan berpikir kritis. Pada tahap ini, peneliti menentukan beberapa spesifikasi tes dengan menentukan aspek keterampilan berpikir kritis yang hendak diukur. Pada penelitian ini, tes keterampilan berpikir kritis didasarkan pada aspek KBK yang disampaikan oleh Facione (2020). Selain itu, dirumuskan sub-keterampilan dari setiap aspek KBK tersebut. Setelah penentuan spesifikasi tes, peneliti membuat kisi-kisi tes berdasarkan indikator pencapaian kompetensi pada suhu dan kalor yang disesuaikan dengan aspek KBK yang akan diukur. Kisi-kisi tersebut dibuat dalam bentuk tabel yang memuat sejumlah informasi yang akan memudahkan peneliti dalam menyusun butir soal. Kolom pertama merupakan kode indikator pencapaian kompetensi, kolom kedua berisi aspek KBK yang hendak diukur, kolom ketiga berisi sub-keterampilan dari aspek KBK, kolom keempat berisi indikator soal dengan materi fisika yang dipilih sebagai konteks tes, kolom kelima berisi butir soal, dan kolom terakhir berisi kunci jawaban. Kisi-kisi tes yang dibuat oleh peneliti

berpedoman pada Kadarwati (2017) yang mengungkapkan bahwa kisi-kisi tes yang baik harus memenuhi persyaratan berikut: 1) Kisi-kisi harus dapat mewakili isi silabus/kurikulum atau materi yang telah diajarkan secara tepat dan proporsional; 2) Komponen-komponennya diuraikan secara jelas dan mudah dipahami; 3) Materi sebagai konteks dalam tes dapat dibuatkan butir soalnya.

Setelah pembuatan kisi-kisi tes, peneliti menyusun instrumen *two-tier open ended test* dimana *tier* pertama merupakan pertanyaan berupa soal pilihan ganda mengenai konten materi dan *tier* kedua merupakan pertanyaan dalam bentuk *open ended questions* mengenai alasan siswa memilih jawaban pada *tier* pertama. Dengan kata lain, siswa diharuskan untuk memilih satu jawaban yang benar kemudian menjelaskan alasan memilih jawaban tersebut. Instrumen *two-tier open ended test* disebarkan kepada 80 siswa SMA Negeri di kabupaten Cianjur dengan tujuan untuk mengumpulkan jawaban alternatif siswa mengenai konsep suhu dan kalor. Jawaban alternatif tersebut akan digunakan sebagai pilihan alasan ataupun distraktor pada *tier* ketiga.

Tahapan selanjutnya adalah dilakukan pengembangan instrumen dari *two tier open ended test* menjadi *four-tier diagnostic test*. Proses pengembangan dilakukan dengan memasukkan alasan siswa yang didapat dari tahap *open ended tes* pada *tier* ketiga dan menambahkan tingkat keyakinan siswa dalam memilih jawaban yang ditempatkan pada *tier* kedua dan *tier* keempat. Tes keterampilan berpikir kritis yang telah disusun akan dilakukan *judgement* oleh para ahli yang berkompeten dalam bidang fisika dan/atau assesmen. Para ahli mengevaluasi kualitas *item* tes berdasarkan validitas isi, konstruksi, dan bahasa. Hasil yang diperoleh dari tahapan ini adalah data kualitatif dan data kuantitatif. Data kualitatif berupa catatan saran, kritik dan masukan yang akan digunakan sebagai bahan perbaikan pada butir soal tes keterampilan berpikir kritis. Sedangkan data kuantitatif berupa skor yang diberikan para ahli pada setiap aspek untuk setiap butir soal yang dianalisis menggunakan teori respons butir dengan bantuan *software* Minifac (Facet). Bersamaan dengan *judgement* oleh para ahli, instrumen *Four-Tier Diagnostic Test* diuji coba secara terbatas kepada 60 siswa SMA Negeri di kota Bandung. Berdasarkan *judgement* ahli dan uji coba terbatas, peneliti memperbaiki butir soal

yang memiliki kesalahan.

3.4.4 Mengumpulkan dan Menganalisis Data Kuantitatif Berdasarkan Variabel, Instrumen, Intervensi

Setelah dilakukan perbaikan butir soal berdasarkan *judgement* ahli dan uji coba terbatas, instrumen *Four-Tier Diagnostic Test* diuji coba secara luas kepada 132 siswa SMA Negeri di kabupaten Cianjur pada tiga sekolah yang berbeda. Hal ini dilakukan agar data yang diperoleh beragam sehingga hasil yang diperoleh menjadi lebih terpercaya dan dapat menunjukkan bahwa soal yang telah peneliti susun layak digunakan di sekolah manapun. Data yang diperoleh berdasarkan jawaban siswa pada uji coba tersebut kemudian dianalisis menggunakan teori respon butir dengan *Graded Respon Models* berbantuan *software eirt 2.0.3*.

3.4.5 Hasil Data Kuantitatif

Setelah dilakukan analisis hasil uji coba luas dengan *graded response models* (GRM) didapat hasil data kuantitatif

3.4.6 Melakukan Interpretasi Hasil Data Kuantitatif dapat Memberikan Hasil, Instrumen dan Intervensi Baru

Seluruh data yang diperoleh dalam penelitian, baik berupa data kualitatif maupun data kuantitatif yang telah dianalisis kemudian diinterpretasi ke dalam kategori-kategori tertentu menurut teori para ahli. Tahap interpretasi dilakukan dengan tujuan untuk mendapatkan karakteristik tes keterampilan berpikir kritis yaitu validitas, reliabilitas, dan parameter pada setiap butir soal.

3.5 Analisis Data

3.5.1 Data kualitatif

Data kualitatif didapatkan melalui kegiatan wawancara bersama guru fisika pada tahap studi pendahuluan, pelaksanaan *two-tier open ended test* untuk mengumpulkan jawaban alasan yang akan digunakan sebagai pilihan jawaban dan pengecoh pada *option tier* ketiga, catatan perbaikan dari para ahli atau validator ketika melakukan validasi instrumen, dan hasil uji coba terbatas.

Data dari kegiatan wawancara disimpan pada lembar wawancara yang berisi beberapa pertanyaan yang diajukan kepada guru fisika. Lembar pedoman wawancara dapat dilihat pada Lampiran 1. Data tersebut dianalisis secara deskriptif

sehingga diperoleh gambaran umum terkait pemahaman konsep siswa pada mata pelajaran fisika, proses evaluasi pembelajaran, urgensi dan pelaksanaan tes diagnostik, dan penerapan keterampilan berpikir kritis dalam proses pembelajaran.

Data dari *two-tier open ended test* didapatkan dengan membagikan instrumen *two-tier open ended test* dimana tingkat pertama merupakan pertanyaan mengenai konten materi dan tingkat kedua merupakan pertanyaan mengenai alasan siswa menjawab pertanyaan pada tingkat pertama yang berbentuk *open ended questions*. Hasil dari *two-tier open ended test* dianalisis sehingga didapatkan jawaban-jawaban siswa berupa alasan yang digunakan peneliti sebagai pilihan jawaban atau pengecoh pada *tier* ketiga. Data tersebut dapat dilihat pada Lampiran 12.

Data berupa catatan perbaikan dari para ahli atau validator didapatkan dengan memberikan lembar instrumen validasi seperti pada Lampiran 6. Data tersebut dianalisis secara deskriptif sehingga menghasilkan instrumen *four-tier diagnostic test* yang lebih baik berdasarkan hasil saran perbaikan dari para ahli.

Data berupa uji coba terbatas yang diperoleh dengan menyebarkan instrumen *four-tier diagnostic test* kepada siswa. Data tersebut dianalisis secara deskriptif sehingga diperoleh gambaran terkait estimasi waktu yang dibutuhkan siswa untuk menyelesaikan tes.

3.5.2 Data Kuantitatif

3.5.2.1 Analisis Data Hasil Validasi Ahli

Data yang diperoleh dari hasil validasi tiga orang ahli kemudian dianalisis untuk mengetahui kualitas instrumen yang telah dikembangkan. Analisis yang dilakukan adalah uji validitas isi dan konstruksi. Analisis tersebut dapat dijabarkan sebagai berikut.

a. Uji Validitas Isi (*Content Validity*)

Uji validitas isi dilakukan oleh tiga validator yaitu dua dosen ahli dan satu guru fisika terhadap instrumen yang dikembangkan. Aspek yang dinilai dari instrumen *Four-Tier Diagnostic Test* terbagi menjadi tiga kategori yaitu validitas isi, konstruksi, dan bahasa. Setiap ahli menuliskan skala penilaian seperti yang disajikan pada Tabel 3.3 di kolom yang telah disediakan pada lembar validasi. Hasil validasi tersebut dianalisis dengan *uji rater* menggunakan *software* MINIFAC

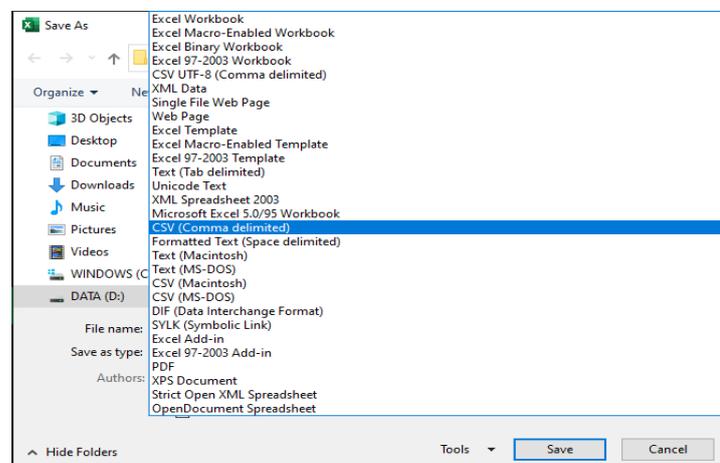
(Facet). Langkah-langkah yang dilakukan untuk mendapatkan nilai *output* dari uji *rater* adalah sebagai berikut

1. Mendownload *software* minifac pada laman web <http://www.winsteps.com/minifac.htm>
2. Memasang *software* minifac pada laptop atau komputer.
3. Memasukkan data hasil validasi ke dalam *Microsoft Excel* dengan format identitas pakar pada kolom pertama (pakar 1,2, dst), nomor item pada kolom kedua, jumlah kriteria pada kolom ketiga, nomor aspek penilaian validitas pada baris pertama dan data hasil validasi ahli pada baris kedua. Contoh format data tersebut dapat dilihat pada Gambar 3.4

No.Pakar	No.Item	Jumlah Kriteria	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14
1	2	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	3	1-14	3	3	3	3	3	3	3	3	2	3	2	3	3	3
1	4	1-14	3	3	3	2	3	3	3	3	2	3	2	3	3	3
1	5	1-14	3	3	2	2	3	3	3	2	3	3	3	3	3	3
1	6	1-14	3	3	3	2	3	3	3	2	3	3	3	3	3	3
1	7	1-14	3	3	3	2	3	3	3	2	3	3	3	3	3	3
1	8	1-14	3	3	3	2	3	3	3	3	2	3	3	3	3	3
1	9	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	10	1-14	3	3	3	2	3	3	3	2	3	2	3	3	3	3
1	11	1-14	3	3	3	2	3	3	3	3	3	3	3	3	3	3
1	12	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	13	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	14	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	15	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	16	1-14	1	3	3	2	3	3	3	2	2	3	3	3	3	3
1	17	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	18	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	19	1-14	2	3	3	2	3	3	3	2	3	3	3	3	3	3
1	20	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	1	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	2	1-14	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	3	1-14	2	3	3	3	3	3	3	3	2	3	3	3	3	3
2	4	1-14	3	3	2	2	3	3	3	3	3	3	3	3	3	3
2	5	1-14	3	3	2	3	3	3	3	3	3	3	3	3	3	3
2	6	1-14	3	3	3	3	3	3	3	3	2	3	3	3	3	3

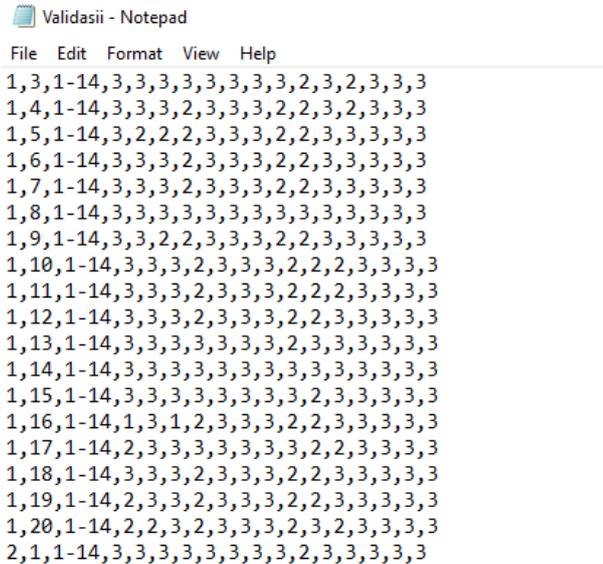
Gambar 3.4 Format Data Hasil Validasi

4. Menyimpan data tersebut dalam bentuk file CSV (*Comma Delimited*) melalui menu *save as* seperti Gambar 3.5



Gambar 3.5 Menyimpan Data dalam Bentuk CSV

5. Mencari file pada layar utama (*desktop*) lalu membuka file tersebut menggunakan Notepad sehingga tampilan data akan terlihat seperti Gambar 3.6



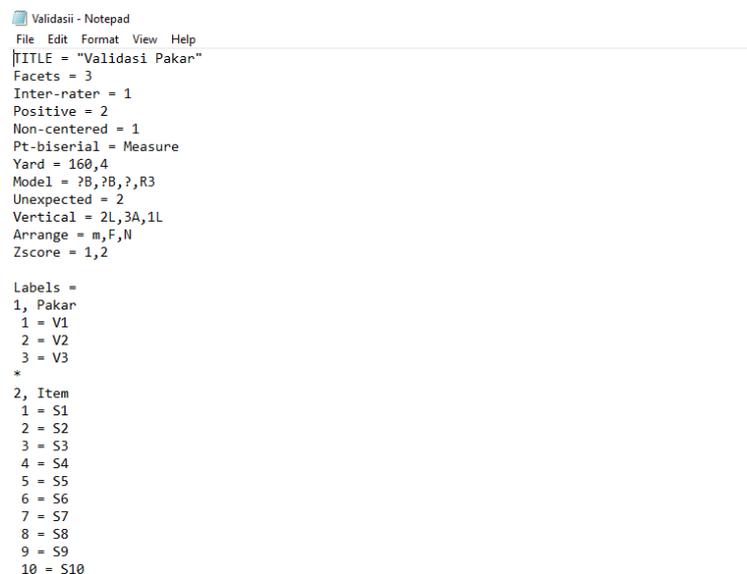
```

Validasi - Notepad
File Edit Format View Help
1,3,1-14,3,3,3,3,3,3,3,2,3,2,3,3,3
1,4,1-14,3,3,3,2,3,3,3,2,2,3,2,3,3,3
1,5,1-14,3,2,2,2,3,3,3,2,2,3,3,3,3,3
1,6,1-14,3,3,3,2,3,3,3,2,2,3,3,3,3,3
1,7,1-14,3,3,3,2,3,3,3,2,2,3,3,3,3,3
1,8,1-14,3,3,3,3,3,3,3,3,3,3,3,3,3,3
1,9,1-14,3,3,2,2,3,3,3,2,2,3,3,3,3,3
1,10,1-14,3,3,3,2,3,3,3,2,2,2,3,3,3,3,3
1,11,1-14,3,3,3,2,3,3,3,2,2,2,3,3,3,3,3
1,12,1-14,3,3,3,2,3,3,3,2,2,3,3,3,3,3,3
1,13,1-14,3,3,3,3,3,3,3,2,3,3,3,3,3,3,3
1,14,1-14,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3
1,15,1-14,3,3,3,3,3,3,3,3,2,3,3,3,3,3,3
1,16,1-14,1,3,1,2,3,3,3,2,2,3,3,3,3,3,3
1,17,1-14,2,3,3,3,3,3,3,3,2,2,3,3,3,3,3
1,18,1-14,3,3,3,2,3,3,3,2,2,3,3,3,3,3,3
1,19,1-14,2,3,3,2,3,3,3,2,2,3,3,3,3,3,3
1,20,1-14,2,2,3,2,3,3,3,2,3,2,3,3,3,3,3
2,1,1-14,3,3,3,3,3,3,3,3,2,3,3,3,3,3,3

```

Gambar 3.6 Tampilan Data pada Notepad

6. Menyalin semua data yang ditampilkan pada Notepad lalu simpan pada kolom data yang terdapat di file khusus berupa baris-baris perintah (*coding*) untuk menganalisis data menggunakan *software* MINIFAC. Format file khusus (*coding*) tersebut dapat dilihat pada Gambar 3.7



```

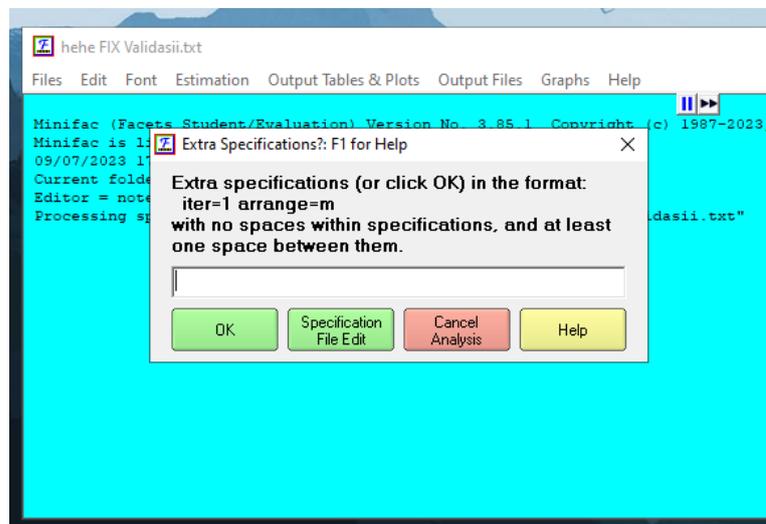
Validasi - Notepad
File Edit Format View Help
|TITLE = "Validasi Pakar"
Facets = 3
Inter-rater = 1
Positive = 2
Non-centered = 1
Pt-biserial = Measure
Yard = 160,4
Model = ?B,?B,?,R3
Unexpected = 2
Vertical = 2L,3A,1L
Arrange = m,F,N
Zscore = 1,2

Labels =
1, Pakar
1 = V1
2 = V2
3 = V3
*
2, Item
1 = S1
2 = S2
3 = S3
4 = S4
5 = S5
6 = S6
7 = S7
8 = S8
9 = S9
10 = S10

```

Gambar 3. 7 Tampilan File *Coding* pada Notepad

7. Menyimpan file *coding* tersebut pada layar utama (*desktop*) lalu menarik file tersebut menuju *software* Minifac sampai muncul tampilan seperti Gambar 3.8



Gambar 3.8 Tampilan Minifac

8. Menekan pilihan OK maka didapatkan hasil analisis yang diinginkan

b. Uji Validitas Konstruk (*Construct Validity*)

Menurut Sumintono & Widhiarso (2014), uji validitas konstruk dikenal dengan unidimensionalitas instrumen (*item unidimensionality*). Unidimensionalitas instrumen merupakan alat ukur yang penting untuk mengevaluasi instrumen yang dikembangkan untuk membuktikan apakah instrumen tersebut mampu mengukur apa yang seharusnya diukur sehingga instrumen dapat dikatakan valid. Nilai unidimensionalitas dapat dilihat melalui hasil pengukuran *variance explained by measure* kemudian diinterpretasikan berdasarkan kriteria *variance explained by measures* yang disampaikan oleh Sumintono & Widhiarso (2015) pada Tabel 3.6

Tabel 3.6 Kriteria Unidimensionalitas

Nilai <i>Variance Explained by Measures (%)</i>	Interpretasi
>20	Terpenuhi
>40	Sesuai
>60	Istimewa

(Sumintono & Widhiarso, 2015)

Selain itu, pengukuran uji validitas instrumen dapat dilihat dari setiap butir

soal melalui nilai pada *Outfit mean square (MNSQ)*, *Outfit Z-standard (ZSTD)*, dan *Point Measure Correlation (PtMeaCorr)*. Butir soal dikatakan fit jika memenuhi kriteria yang ditunjukkan pada Tabel 3.7

Tabel 3.7 Kriteria MNSQ, ZSTD, dan Pt Mea Corr.

Kriteria	Skor	Keterangan
<i>Outfit mean square (MnSQ)</i>	$0,5 < \text{MNSQ} < 1,5$	Diterima
<i>Outfit Z-standard (ZStd)</i>	$-2,0 < \text{ZSTD} < +2,0$	Diterima
<i>PT Measure Correlation (Pt Mean Corr)</i>	$0,4 < \text{PtMea} < 0,8$	Diterima

Hasil dari ketiga nilai tersebut diinterpretasikan berdasarkan kriteria pada Tabel 3.8

Tabel 3.8 Kriteria Nilai *Fit Statistic*

Kriteria Nilai <i>Fit Statistic</i>	Interpretasi
Ketiga kriteria nilai terpenuhi	Sangat Sesuai
Dua dari tiga kriteria nilai terpenuhi	Sesuai
Satu dari tiga kriteria nilai terpenuhi	Kurang Sesuai
Semua kriteria nilai tidak terpenuhi	Tidak Sesuai

3.5.2.2 Analisis Data Hasil Uji Coba Luas

Hasil uji coba luas instrumen keterampilan berpikir kritis dianalisis menggunakan teori respons butir (*Item Response Theory/IRT*) dengan model politomi yaitu *Graded Respon Models*. Tahap pertama dalam menganalisis data hasil uji coba ini adalah menentukan apakah *Graded Respon Models* sesuai untuk mengetahui gambaran karakteristik tes keterampilan berpikir kritis. Pemilihan model tersebut dilakukan dengan melihat nilai *p-value* dan membandingkan *Chi-Kuadrat* hasil perhitungan dengan *Chi-Kuadrat* tabel dengan derajat kebebasan tertentu. Butir dikatakan cocok dengan suatu model ketika nilai *p-value* lebih besar dari 0.05 dan nilai *Chi-Kuadrat* perhitungan lebih kecil dari nilai *Chi-Kuadrat* tabel.

Setelah mengetahui model parameter yang sesuai, langkah selanjutnya adalah mengetahui karakteristik tes secara keseluruhan berupa parameter a yaitu daya

pembeda dan parameter b yaitu tingkat kesukaran melalui kurva karakteristik total atau *total characteristic curve* (TCC). Parameter a (daya pembeda) dapat dilihat melalui kemiringan kurva dimana sudut kemiringan pada kurva didapatkan menggunakan web yang dapat mengukur sudut secara *online*. Semakin besar kemiringannya, maka semakin besar indeks daya pembeda pada tes tersebut. Sementara itu, parameter b (tingkat kesukaran) didapatkan dengan menjumlahkan skor terbesar dan terkecil lalu dibagi dua kemudian menarik garis vertikal pada hasil perhitungan tersebut dan menarik garis yang tegak lurus dari sumbu x pada kurva karakteristik total.

Nilai parameter a (daya pembeda) dikatakan baik atau valid apabila berada pada rentang 0 sampai +2. Jika nilai parameter a berada pada rentang tersebut, maka dapat dikatakan bahwa tes keterampilan berpikir kritis mampu membedakan kemampuan peserta tes dari kemampuan rendah hingga kemampuan tinggi. Sebaliknya, jika nilai parameter a tidak berada pada rentang tersebut, maka dapat dikatakan bahwa tes keterampilan berpikir kritis tidak mampu membedakan kemampuan peserta tes.

Nilai parameter tingkat kesukaran dikatakan baik atau valid apabila berada pada rentang -2 sampai +2. Jika nilai parameter b mendekati -2, maka butir tes tersebut termasuk kategori butir soal yang mudah. Jika nilai parameter b mendekati +2, maka butir tes tersebut termasuk kategori butir soal yang sulit (Hambleton, Swaminathan, & Rogers, 1991). Klasifikasi tingkat kesukaran tersebut dapat dilihat pada Tabel 3.9

Tabel 3.9 Klasifikasi Tingkat Kesukaran (b)

Rentang (b)	Kategori
-2 s/d -1	Mudah
-1 s/d 1	Sedang
1 s/d 2	Sukar

Sementara itu, Harrish (1989) mengatakan bahwa tingkat kesukaran yang baik berada pada rentang -3 sampai +3 dimana jika nilai parameter b mendekati -3, maka butir tes tersebut termasuk kategori butir soal yang mudah dan jika nilai

parameter b mendekati $+3$, maka butir tes tersebut termasuk kategori butir soal yang sulit. Pada penelitian ini, kriteria yang digunakan dalam menentukan tingkat kesukaran tes keterampilan berpikir kritis didasarkan pada klasifikasi tingkat kesukaran yang disampaikan oleh Hambleton (1991).

Selain menentukan parameter a (daya pembeda) dan parameter b (tingkat kesukaran) melalui kurva karakteristik total, karakteristik tes dapat ditunjukkan melalui kurva karakteristik butir. Du Toit (2003) mengatakan bahwa δ_i dapat dilihat dari parameter tahap (*step parameters*), kesulitan tahap (*item step difficulties*), ataupun perpotongan kategori (*category intersections*) dimana δ_i menunjukkan titik pertemuan dua garis probabilitas kategori dalam satu butir. Selain itu, δ_i dinamakan persimpangan kategori (*category response curves intersect*) yang diinterpretasikan sebagai titik pada skala yang bersifat laten dimana dua kategori yang berurutan kurva respons berpotongan. Sehingga dapat disimpulkan bahwa probabilitas individu n dengan skor x pada butir i dapat dilihat dari nilai δ_i yang merupakan titik dimana dua kategori memiliki probabilitas yang sama untuk dipilih oleh level kemampuan tertentu. Interpretasi klasifikasi estimasi kemampuan siswa disajikan pada Tabel 3.10

Tabel 3.10 Klasifikasi Estimasi Kemampuan (θ)

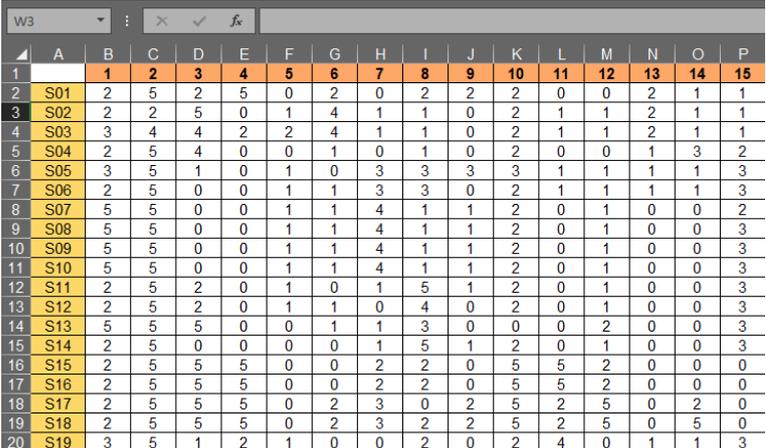
Rentang (θ)	Kategori
-4 s/d -2,5	Sangat rendah
-2,5 s/d -1	Rendah
-1 s/d 1	Sedang
1 s/d 2,5	Tinggi
2,5 s/d 4	Sangat tinggi

Estimasi reliabilitas tes berdasarkan teori respon butir diperoleh melalui fungsi informasi dan *standard error measurement* (SEM). Fungsi informasi memiliki hubungan yang berlawanan dengan SEM. Semakin rendah fungsi informasi sebuah butir, maka semakin besar SEM pada butir tersebut dan begitu juga sebaliknya. Semakin tinggi fungsi informasi sebuah butir, maka semakin kecil SEM pada butir tersebut. Semakin kecil SEM, hasil pengukuran yang didapatkan akan semakin tepat, reliabel, dan dapat dipercaya (Setiawati, F.A. dkk., 2013).

Menurut Hambleton & Swaminathan (dalam Setiawati, F.A. dkk., 2013), penggunaan fungsi informasi lebih akurat jika dibandingkan dengan penggunaan reliabilitas, karena bentuknya bergantung pada butir tes dan mempunyai estimasi kesalahan pengukuran dalam tingkat kemampuan. Kesalahan pengukuran atau *standard error measurement* (SEM) dapat ditentukan melalui perhitungan berdasarkan persamaan $1/\sqrt{I(\theta)}$. Berdasarkan fungsi informasi dan SEM, dapat diketahui apakah tes yang digunakan cocok atau tidak untuk siswa dengan kemampuan rendah, sedang, atau tinggi (Istiyono, 2013).

Hasil data uji coba luas dianalisis dengan bantuan *software eirt 2.0.3* untuk *Graded Respon Models*. Prosedur penggunaan *software eirt* untuk *Graded Respon Models* adalah sebagai berikut.

1. Mengunduh program *set up* “eirt-2.0.0.exe”, pada laman web <https://libirt.psychometricon.net/releases/>
2. Memasang program *eirt* pada laptop atau komputer
3. Memasukkan data hasil uji tes ke dalam *Microsoft Excel* dengan format yaitu kode peserta tes pada kolom pertama, nomor item pada baris pertama dan data hasil tes pada baris kedua dengan tampilan seperti Gambar 3.9



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	S01	2	5	2	5	0	2	0	2	2	2	0	0	2	1	1
3	S02	2	2	5	0	1	4	1	1	0	2	1	1	2	1	1
4	S03	3	4	4	2	2	4	1	1	0	2	1	1	2	1	1
5	S04	2	5	4	0	0	1	0	1	0	2	0	0	1	3	2
6	S05	3	5	1	0	1	0	3	3	3	3	1	1	1	1	3
7	S06	2	5	0	0	1	1	3	3	0	2	1	1	1	1	3
8	S07	5	5	0	0	1	1	4	1	1	2	0	1	0	0	2
9	S08	5	5	0	0	1	1	4	1	1	2	0	1	0	0	3
10	S09	5	5	0	0	1	1	4	1	1	2	0	1	0	0	3
11	S10	5	5	0	0	1	1	4	1	1	2	0	1	0	0	3
12	S11	2	5	2	0	1	0	1	5	1	2	0	1	0	0	3
13	S12	2	5	2	0	1	1	0	4	0	2	0	1	0	0	3
14	S13	5	5	5	0	0	1	1	3	0	0	0	2	0	0	3
15	S14	2	5	0	0	0	0	1	5	1	2	0	1	0	0	3
16	S15	2	5	5	5	0	0	2	2	0	5	5	2	0	0	0
17	S16	2	5	5	5	0	0	2	2	0	5	5	2	0	0	0
18	S17	2	5	5	5	0	2	3	0	2	5	2	5	0	2	0
19	S18	2	5	5	5	0	2	3	2	2	5	2	5	0	5	0
20	S19	3	5	1	2	1	0	0	2	0	2	4	0	1	1	3

Gambar 3.9 Format Data untuk Analisis Eirt

4. Melakukan *block* pada semua data yang akan dianalisis, klik menu eirt pada *toolbar Add- Ins*, kemudian klik *start the assistant* seperti Gambar 3.10

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	S01	2	5	2	5	0	2	0	2	2	2	0	0	2	1	1
3	S02	2	2	5	0	1	4	1	1	0	2	1	1	2	1	1
4	S03	3	4	4	2	2	4	1	1	0	2	1	1	2	1	1
5	S04	2	5	4	0	0	1	0	1	0	2	0	0	1	3	2
6	S05	3	5	1	0	1	0	3	3	3	3	1	1	1	1	3
7	S06	2	5	0	0	1	1	3	3	0	2	1	1	1	1	3
8	S07	5	5	0	0	1	1	4	1	1	2	0	1	0	0	2
9	S08	5	5	0	0	1	1	4	1	1	2	0	1	0	0	3
10	S09	5	5	0	0	1	1	4	1	1	2	0	1	0	0	3
11	S10	5	5	0	0	1	1	4	1	1	2	0	1	0	0	3
12	S11	2	5	2	0	1	0	1	5	1	2	0	1	0	0	3
13	S12	2	5	2	0	1	1	0	4	0	2	0	1	0	0	3
14	S13	5	5	5	0	0	1	1	3	0	0	0	2	0	0	3
15	S14	2	5	0	0	0	0	1	5	1	2	0	1	0	0	3
16	S15	2	5	5	5	0	0	2	2	0	5	5	2	0	0	0
17	S16	2	5	5	5	0	0	2	2	0	5	5	2	0	0	0
18	S17	2	5	5	5	0	2	3	0	2	5	2	5	0	2	0
19	S18	2	5	5	5	0	2	3	2	2	5	2	5	0	5	0
20	S19	3	5	1	2	1	0	0	2	0	2	4	0	1	1	3

Gambar 3.10 Tahap ke-1 Penggunaan Eirt

- Memberi tanda centang pada pilihan *the subject's label are in the first column of the selection* dan *the item's label are in the first line of the selection* kemudian klik *next* seperti Gambar 3.11

Step 1 of 4

Select the range containing the data to analyse.

Selection:

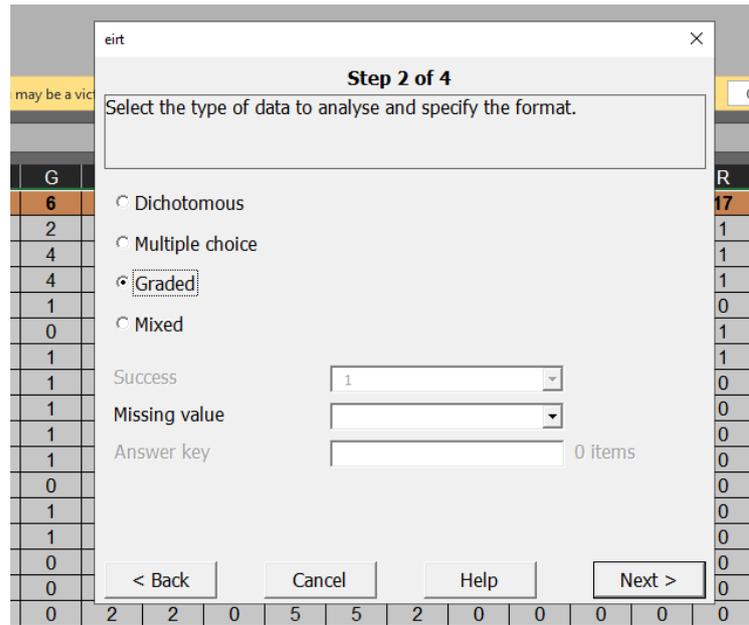
The subjects' label are in the first column of the selection

The items' label are in the first line of the selection

The correction key is in the second line of the selection

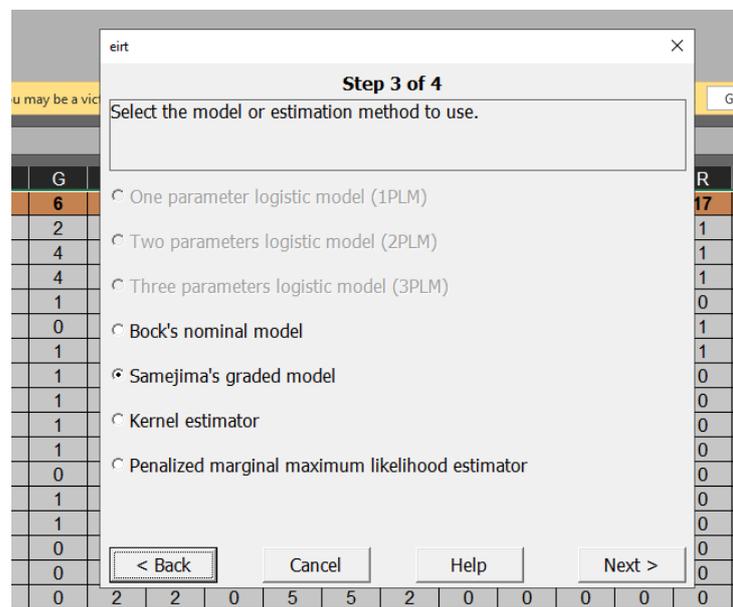
Gambar 3.11 Tahap ke-2 Pemilihan Identitas Data

- Memberi tanda centang pada pilihan *graded* karena bentuk data yang dianalisis merupakan data yang bergradasi kemudian klik *next* seperti Gambar 3.12



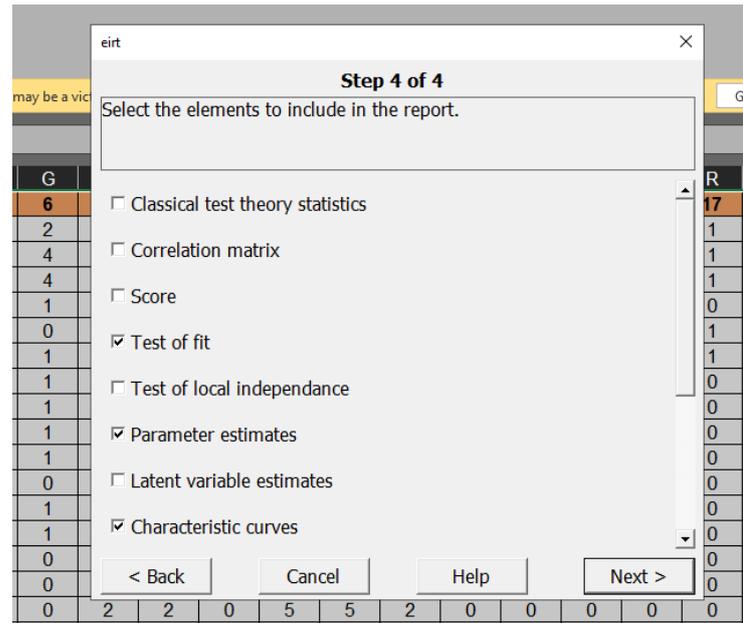
Gambar 3.12 Tahap ke-3 Pemilihan Tipe Data

7. Mencentang metode estimasi *samejima's graded model* kemudian klik *next* seperti Gambar 3.13



Gambar 3.13 Tahap ke-4 Pemilihan Metode Estimasi

8. Memilih elemen-elemen yang ingin diketahui hasil analisisnya kemudian klik *next* seperti Gambar 3.14. Maka hasil analisis akan ditampilkan



Gambar 3.14 Tahap ke-5 Pemilihan Elemen yang akan dianalisis