

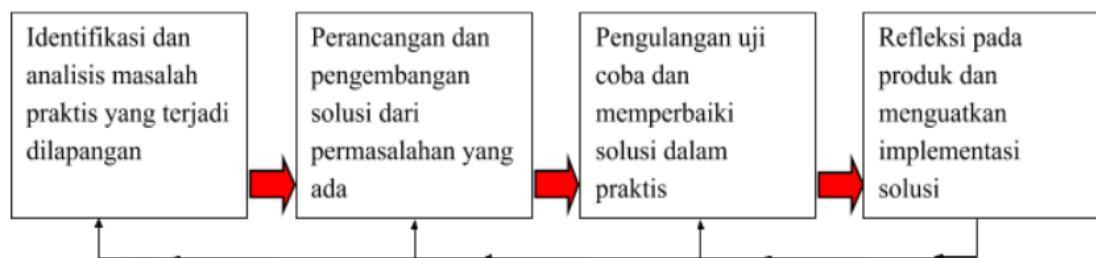
## BAB III

### METODE PENELITIAN

Bagian metode penelitian merupakan pembahasn mengenai metode penelitian yang digunakan dalam penelitian ini yang meliputi: Metode dan desain penelitian, lokasi dan subjek penelitian, prosedur penelitian, jenis data, instrument pengumpulan data, teknik analisis data, analisis data, dan hasil penggunaan tes.

#### 3.1. Metode dan Desain Penelitian

Metode penelitian yang digunakan adalah *Design-Based Research (DBR)*. Metode DBR adalah salah satu metode pengembangan, seperti yang dikemukakan oleh van den Akker (1999) menyatakan bahwa istilah penelitian *design research* dimasukan ke dalam penelitian pengembangan, karena berkaitan dengan pengembangan materi dan bahan pembelajaran. Penelitian berbasis desain biasa digunakan dalam mengembangkan strategi atau model pembelajaran seperti dalam penelitian Brown (1992). Amiel dan Reeves (2008) menjelaskan tahap-tahap pada metode *Design-Based Research*, yaitu sebagai berikut:



Gambar 3. 1. Metode *Design-Based Research*

Terdapat dua model dalam penelitian pengembangan yaitu model produk dan model pengembangan.

##### 3.1.1. Model Produk

Tes Diagnostik *Static Fluid Four-Tier (SFFT)* merupakan produk hasil dari penelitian pengembangan. Tes diagnostik SFFT adalah tes pilihan ganda yang terdiri dari empat tingkatan. Tingkat pertama berupa pilihan ganda tentang konsep ilmiah, tingkat kedua berupa tingkat keyakinan siswa dalam memilih jawaban pada tingkat pertama, tingkat ketiga berupa alasan siswa dalam menjawab pertanyaan

pada tingkat pertama, tingkat keempat berupa tingkat keyakinan siswa dalam memilih jawaban siswa pada tingkat ketiga.

### **3.1.2. Model Pengembangan**

Model yang digunakan adalah model prosedural, yaitu model yang bersifat deskriptif, menggambarkan langkah-langkah secara sistematis untuk menghasilkan produk instrumen yang dapat mengidentifikasi konsepsi alternatif siswa SMA pada materi fluida statis.

### **3.2. Lokasi dan Subjek Penelitian**

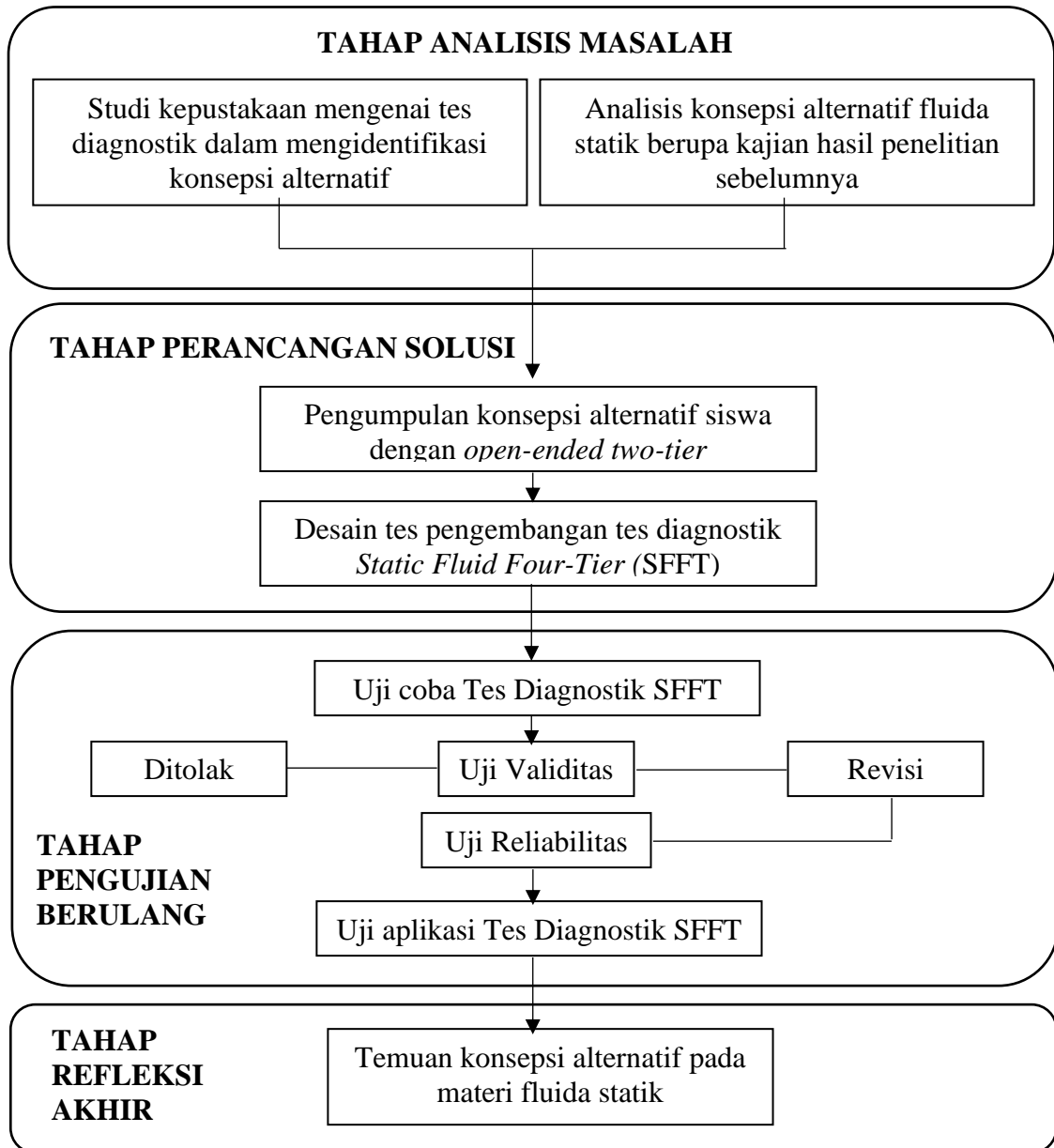
Populasi penelitian dalam penelitian ini adalah siswa SMA kelas XI yang telah mempelajari materi fluida statis. Kelas XI dipilih sebagai populasi karena disesuaikan dengan materi yang akan diteliti. Tes diagnostik SFFT ini mengenai konsep fluida statis yang dipelajari di kelas XI pada semester ganjil, sehingga subjek penelitiannya adalah siswa kelas XI. Populasi terdiri dari SMA yang berlokasi di perbatasan Kabupaten Kuningan dan Kabupaten Brebes.

Subjek penelitian terdiri dari empat SMA yaitu; SMAN 1 Cibingbin, MA Al-Amin, SMAN 1 Banjarharjo, dan SMA Mubarakul Ulum, dibagi menjadi tiga yaitu untuk keperluan identifikasi konsepsi alternatif siswa, subjek uji coba instrumen, dan aplikasi instrumen hasil pengembangan. Subjek penelitian berjumlah 117 siswa dari empat sekolah.

### **3.3. Prosedur Penelitian**

Prosedur penelitian dilakukan dengan mengikuti langkah-langkah pada metode Design-Based Research yaitu tahap analisis masalah, tahap perancangan solusi, tahap pengujian berulang, dan tahap refleksi akhir. Skema penelitian disajikan dalam bentuk bagan pada Gambar 3.2.

Langkah pengembangan tes dapat dijelaskan sebagai berikut:



Gambar 3. 2. Skema Pengembangan Tes Diagnostik SFFT

### 3.4. Jenis Data

Data yang diperoleh, dianalisis, dan digunakan dalam penelitian pengembangan ini terdiri dari 2 yaitu:

#### 3.4.1. Data Kualitatif

Data kualitatif dalam penelitian ini yaitu berupa masukan dari validator mengenai instrumen yang dikembangkan. Masukan dan saran yang diperoleh yaitu berbagai hal terkait dengan kebenaran konsep fluida statis dan konsepsi alternatif yang diidentifikasi, serta ketepatan soal ditinjau dari aspek bahasa, materi, dan konstruksi, dan ketepatan soal dalam mengidentifikasi tingkat konsepsi siswa pada materi fluida statis. Alasan siswa juga akan dianalisis untuk mengetahui alasan mana saja yang teridentifikasi konsepsi alternatif.

### **3.4.2. Data Kuantitatif**

Data kuantitatif dalam penelitian ini adalah data hasil masukan validator, data uji coba instrumen, dan data aplikasi tes pada siswa dengan tujuan mengidentifikasi tingkat konsepsi siswa pada materi fluida statis. Validitas, reliabilitas, tingkat kesukaran, dan daya pembeda didapat dari informasi pada data hasil uji coba instrumen. Sedangkan data dari hasil aplikasi tes digunakan untuk menjelaskan tingkat konsepsi siswa.

## **3.5. Instrumen Penelitian**

Instrumen penelitian merupakan alat yang digunakan untuk mengumpulkan data (Fратиwi, 2017). Dalam menentukan kualitas instrument tes yang dikembangkan, langkah yang harus ditempuh adalah mengestimasi validitas dan reliabilitas tes yang digunakan. Estimasi validitas dan reliabilitas dijelaskan sebagai berikut.

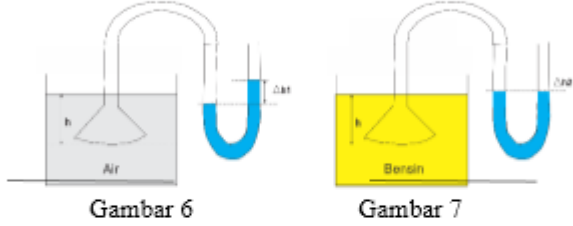
### **3.5.1. Instrumen *Two-tier Open-Ended test***

Pada tahap studi pendahuluan, peneliti menggunakan *open-ended two-tier test* dengan tingkat pertama sebagai pilihan jawaban berupa pilihan ganda dan tingkat kedua merupakan alasan dari jawaban pada tingkat pertama yang berupa *open-ended test*. Instrumen ini digunakan untuk mengumpulkan jawaban alternatif siswa mengenai konsep fluida ststis. Selanjutnya, alasan siswa tersebut dipilih untuk dianalisis alasan mana saja yang teridentifikasi konsepsi alternatif. Jawaban yang teridentifikasi konsepsi alternatif dijadikan sebagai pilihan alasan pada tingkat ketiga pada tes diagnostik *Static Fluid Four-Tier (SFFT)* dan dikombinasikan dengan temuan hasil kajian literature. Instrumen ini terdiri dari 14

soal.. Contoh dari soal *open-ended two-tier test* dapat dilihat pada Gambar 3.3. (Lampiran 1)

**Soal:**

7.1 Dua buah bejana identik diisi air (massa jenis air =  $1000 \text{ kg/m}^3$ ) dan Bensin (massa jenis bensin =  $890 \text{ kg/m}^3$ ) (Giancoli, 2001). Kemudian corong manometer dimasukkan ke dalam bejana berisi air sedalam  $h$  dan terlihat kenaikan zat cair pada pipa setinggi  $\Delta h_1$  seperti pada Gambar 6.



Apabila manometer tersebut dimasukkan kedalam bejana yang berisi minyak sedalam  $h$  juga seperti pada Gambar 7, apa yang akan terjadi dengan  $\Delta h_2$ ?

- $\Delta h_2$  akan lebih tinggi dibandingkan  $\Delta h_1$
- $\Delta h_2$  akan lebih rendah dibandingkan  $\Delta h_1$
- $\Delta h_2$  akan sama tinggi dengan  $\Delta h_1$
- $\Delta h_2$  tidak akan berubah dari keadaan semula
- $\Delta h_2$  tidak dapat ditentukan

7.2 Alasan kamu dalam menjawab pertanyaan 7.1 adalah.....

.....

.....

.....

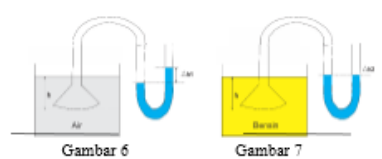
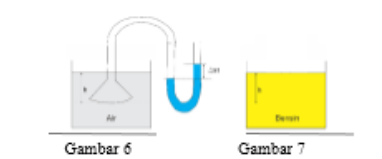
Gambar 3. 3. Instrumen *two-tier open-ended test*

### 3.5.2. Instrumen untuk Validasi Ahli

Pada penelitian ini dilakukan *expert judgment* yang melibatkan para ahli dalam pendidikan fisika dan bidang assessment untuk memvalidasi isi instrumen. *Expert* akan mengisi lembar validasi yang telah disediakan. Lembar validasi ini terdiri dari kolom indikator validasi dan masukan yang diberikan oleh *expert* untuk setiap butir soal. Dari hasil indikator validasi yang telah diisi oleh *expert* kemudian dianalisis secara kuantitatif dengan Rasch menggunakan menu data analisis uji *multi-rater (multi-faceted Rasch measurement)* pada *software* MINIFACT. Hasil dari analisis Rasch akan digunakan untuk menarik kesimpulan pada setiap butir soal apakah valid atau tidak. Instrumen ini disebut sebagai produk 2 yang berjumlah 14 butir soal.

### 3.5.3. Instrumen Validasi Empiris

Instrumen yang digunakan untuk validasi empiris adalah seperangkat tes diagnostik SFFT yang berjumlah 14 butir soal yang merupakan produk 2 dari proses pengembangan. Tes diagnostik ini merupakan hasil revisi dari *expert judgment* dan pengembangan dari *two-tier open-ended* menjadi tes diagnostik SFFT. Hasil perkembangan instrumen dapat dilihat pada Gambar 3.4. (Lampiran 2)

<p><b>Soal:</b> 7.1 Dua buah bejana identik diisi air (massa jenis air = <math>1000 \text{ kg/m}^3</math>) dan Bensin (massa jenis bensin = <math>890 \text{ kg/m}^3</math>) (Giancoli, 2001). Kemudian corong manometer dimasukkan ke dalam bejana berisi air sedalam <math>h</math> dan terlihat kenaikan zat cair pada pipa setinggi <math>\Delta h_1</math> seperti pada Gambar 6.</p>  <p>Apabila manometer tersebut dimasukkan kedalam bejana yang berisi minyak sedalam <math>h</math> juga seperti pada Gambar 7, apa yang akan terjadi dengan <math>\Delta h_2</math>?</p> <ol style="list-style-type: none"> <li><math>\Delta h_2</math> akan lebih tinggi dibandingkan <math>\Delta h_1</math></li> <li><math>\Delta h_2</math> akan lebih rendah dibandingkan <math>\Delta h_1</math></li> <li><math>\Delta h_2</math> akan sama tinggi dengan <math>\Delta h_1</math></li> <li><math>\Delta h_2</math> tidak akan berubah dari keadaan semula</li> <li><math>\Delta h_2</math> tidak dapat ditentukan</li> </ol> <p>7.2 Alasan kamu dalam menjawab pertanyaan 7.1 adalah.....</p> <p><b>Kunci jawaban:</b> 7.1. B</p>	<p><b>Soal:</b> 7.1 Dua buah bejana identik diisi air (massa jenis air = <math>1000 \text{ kg/m}^3</math>) dan Bensin (massa jenis bensin = <math>890 \text{ kg/m}^3</math>) Kemudian corong manometer dimasukkan ke dalam bejana berisi air sedalam <math>h</math> dan terlihat kenaikan zat cair pada pipa setinggi <math>\Delta h_1</math> seperti pada Gambar 6.</p>  <p>Apabila manometer tersebut dimasukkan kedalam bejana yang berisi bensin sedalam <math>h</math> juga seperti pada Gambar 7, apa yang akan terjadi dengan bensin pada manometer (<math>\Delta h_2</math>)...</p> <ol style="list-style-type: none"> <li>tidak dapat ditentukan.</li> <li>akan sama tinggi dengan <math>\Delta h_1</math>.</li> <li>akan lebih tinggi dibandingkan <math>\Delta h_1</math></li> <li>akan lebih rendah dibandingkan <math>\Delta h_1</math>.</li> <li>tidak akan berubah dari keadaan semula.</li> </ol> <p>7.2 Apakah anda yakin dengan jawaban yang anda pilih? a. Yakin b. Tidak yakin</p> <p>7.3 Alasan anda memilih jawaban pada soal 7.1 adalah ...</p> <ol style="list-style-type: none"> <li>tekanan hidrostatis dipengaruhi oleh massa zat cair, semakin besar massa zat cair maka tekanan hidrostatis akan semakin besar.</li> <li>tekanan hidrostatis dipengaruhi oleh massa jenis fluida, semakin besar massa jenis fluida maka tekanan hidrostatis akan semakin kecil.</li> <li>tekanan hidrostatis dipengaruhi oleh massa jenis fluida, semakin besar massa jenis fluida maka tekanan hidrostatis akan semakin besar.</li> <li>tekanan hidrosatis dipengaruhi oleh kedalaman zat cair, semakin dalam zat cair maka tekanan hidrostatis akan semakin besar.</li> <li>lainnya: .....</li> </ol> <p>7.4 Apakah anda yakin dengan jawaban yang anda pilih? a. Yakin b. Tidak yakin</p>
---	---

Gambar 3. 4. Perkembangan *two-tier open-ended* ke *Static Fluid Four-Tier* (SFFT)

### 3.5.4. Instrumen Identifikasi Tingkat Konsepsi

Tes diagnostik SFFT terdiri dari 14 butir soal yang merupakan produk akhir dari proses pengembangan yang dilakukan setelah melewati beberapa tahap pengujian kualitas diantaranya uji validasi, reliabilitas, tingkat kesukaran dan daya beda pembeda. Tes diagnostik SFFT yang telah dikembangkan, kemudian digunakan untuk memetakan tingkat konsepsi siswa.

### 3.6. Teknik Analisis Data

Pengolahan data dilakukan untuk kepentingan penarikan kesimpulan. Pengolahan data meliputi kegiatan penentuan validitas dan reliabilitas dari tes diagnostik SFFT yang dikembangkan. Pengujian validasi, reliabilitas, tingkat kesukaran dan daya pembeda dijelaskan sebagai berikut.

Tanti Septiantini, 2020

**PENGEMBANGAN TES DIAGNOSTIK STATIC FLUID FOUR-TIER (SFFT) UNTUK MENGIDENTIFIKASI KONSEPSI ALTERNATIF SISWA**

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

### 3.6.1. Uji Validasi

Validitas merupakan ukuran yang menunjukkan tingkat ketepatan suatu tes. Tes yang valid merupakan tes yang dapat mengukur apa yang seharusnya diukur. Menurut Satriadi (2019), validasi dilakukan pada dua aspek yaitu *content-releted evidence of validity* dan *criterion-related evidence of validity*. *Content-releted evidence of validity* ini mengacu pada konten dan format instrumen. Isi dan format tes harus sesuai dengan defenisi variable dan sampel subjek yang akan diukur. Aspek lain dari validasi konten adalah format isntrumen, kejelasan pencetakan, kesesuaian bahasa, kejelasan arah, ukuran jenis dan lain sebagainya. *Criterion-related evidence of validity* ini mengacu pada hubungan antar skor yang diperoleh dengan menggunakan instrumen dan skor yang diperoleh dengan menggunakan satu atau lebih instrumen lain. Kekuatan hubungan antara nilai tes dan kinerja kriteria memberikan kriteria terkait bukti validasi.

#### 3.6.1.1. Validasi Isi

Pada tahap validasi isi, pertama-tama item dari tes diagnostik diperiksa dan dinilai oleh *expert* berdasarkan isi dan format di setiap tahap pengembangan tes. Tabel spesifikasi disiapkan untuk konteks dan kasus yang diukur dengan tes dan tabel untuk pilihan kesalahpahaman diperiksa oleh para ahli untuk kecukupan dan kompatibilitasnya dengan item uji. Termasuk didalamnya format, petunjuk arah, angka dan bahasa dinilai oleh para ahli.

Selanjutnya uji validitas dilakukan dengan menggunakan analisis uji *multi-rater (multi-faceted Rasch measurement)* pada *software* MINIFAC untuk dianalisis dengan Rasch. Untuk menghasilkan *output* uji *multi-rater*, langkah yang dilakukan adalah dengan membuat spesifikasi berkas data yang dipunyai dan jenis analisis yang dilakukan. File khusus tersebut tidak lain baris-baris perintah (*coding*) seperti halnya membuat satu program untuk tugas tertentu. *Coding* generik untuk analisis MINIFAC menggunakan program Notepad. (Lampiran 3)

#### 3.6.1.2. Validasi Empiris

Pada tahap ini, soal tes diagnostik SFFT di uji cobakan. Setelah mendapatkan data hasil uji coba tersebut, selanjutnya uji validasi dilakukan dengan menggunakan menu *output 23.Item: dimensionality* pada *software* MINISTEP untuk di analisis dengan Rasch. Validitas juga dikenal dengan istilah *unidimensionalitas* (Sumintono dan Widhiarso, 2014). Uji validitas berdasarkan item dimensionality dengan melihat nilai *raw variance explained by measures*. Persyaratan unidimensionalitas ditunjukkan Tabel 3.1 berikut.

Tabel 3. 1. Interpretasi unidimensionalitas tes

Interpretasi	Skor
Terpenuhi	>20%
Bagus	>40%
Istimewa	>60%

(Sumintono dan Widhiarso, 2014)

### 3.6.2. Uji Reliabilitas

Reliabilitas adalah keajegan atas jawaban yang diberikan pada waktu yang berbeda dan pada waktu yang sama. Suatu tes dikatakan reliabel apabila dapat menghasilkan suatu skor tes yang ajeg dan relatif tidak berubah. Reliabilitas instrumen mempunyai pengertian bahwa suatu instrumen cukup dapat dipercaya atau digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik dan tetap. Sebuah instrumen yang *reliable* berapa kali digunakan dalam penelitian apapun hasilnya akan relatif sama sehingga hasil yang didapatkan dari instrumen tersebut dapat dipercaya.

Uji reliabilitas pada penelitian ini menggunakan *software* MINISTEP 4.4.5. Perangkat lunak MINISTEP ini merupakan sebuah aplikasi pendukung untuk mengetahui kriteria dan kualitas suatu instrumen soal berdasarkan analisis Rasch. Salah satu menu *output* pada *software* ini yaitu *output 3.1 Summary Statistics* yang fungsinya yaitu untuk menampilkan beberapa nilai reliabilitas, diantaranya adalah *person reliability*, *item reliability*, dan *Cronbach alpha*. *Person reliability* menunjukkan konsistensi jawaban siswa, artinya kita dapat mengetahui siswa yang konsisten/ajeg dengan jawaban yang mereka berikan saat menjawab soal tersebut. *Item reliability* menunjukkan kualitas item tes, artinya kita dapat mengetahui kualitas sebuah instrumen yang digunakan dalam penelitian instrumen SFFT. Sementara



*Cronbach alpha* menunjukkan nilai interaksi antara konsistensi jawaban siswa dan item soal dari instrumen *four-tier* secara keseluruhan.

Interpretasi untuk nilai *person reliability* dan *item reliability* dapat dilihat pada Tabel 3.2. Sedangkan interpretasi untuk nilai *Cronbach alpha* dapat dilihat pada Tabel 3.3.

Tabel 3. 2. Interpretasi nilai *person reliability* dan *item reliability*

<b>Nilai person reliability dan item reliability</b>	<b>Interpretasi</b>
$0,94 \leq \text{Nilai}$	Istimewa
$0,91 \leq \text{Nilai} < 0,94$	Bagus Sekali
$0,81 \leq \text{Nilai} < 0,90$	Bagus
$0,67 \leq \text{Nilai} < 0,80$	Cukup
$\text{Nilai} < 0,67$	Lemah

(Sumintono & Widhiarso, 2015)

Tabel 3. 3. Interpretasi nilai *Cronbach alpha*

<b>Nilai Cronbach alpha</b>	<b>Interpretasi</b>
$0,8 \leq \alpha$	Bagus Sekali
$0,7 \leq \alpha < 0,8$	Bagus
$0,6 \leq \alpha < 0,7$	Cukup
$0,5 \leq \alpha < 0,6$	Jelek
$\alpha < 0,5$	Buruk

(Sumintono & Widhiarso, 2015)

Nilai *person reliability* dan *item reliability* yang dipakai adalah *REAL RMSE* karena nilai ini merupakan kondisi terburuk reliabilitas batas bawah berdasarkan instrumen yang dipakai (Sumintono & Widhiarso, 2015).

### 3.6.3. Tingkat Kesukaran

Tingkat kesukaran diperlukan untuk mengetahui distribusi soal dan bertujuan untuk membedakan soal yang mudah, sulit, dan sangat sulit. Tingkat kesukaran tes diagnositik SFFT dapat diketahui dengan menggunakan *software* MINISTEP 4.4.5 pada menu *output Table 1 Variable (Wright) maps* dan *output Table 13 Item Measure* yang dianalisis dengan Rasch. *Output Variable (Wright) maps* pada *software* ini bertujuan untuk mendapat gambaran mengenai kekuatan setiap butir soal (Satriadi, 2019). Sesuai dengan penjelasan Boone & Noltemeyer (2017) bahwa *wright maps* mendukung para peneliti untuk mengukur kekuatan dan kelemahan instrumen, mendokumentasikan penilaian butir tes, membandingkan teori dengan data eksperimen, dan memberikan panduan kepada para peneliti. Sedangkan *Item*

*Measure* digunakan untuk mengklasifikasikan tingkat kesulitan tiap butir soal berdasarkan nilai logit dan nilai Standar Deviasi (SD) yang didapat (Sumintono & Widhiarso, 2015).

Klasifikasi kategori tingkat kesukaran pada instrumen SFFT ini dapat diinterpretasikan seperti pada Tabel 3.4.

Tabel 3. 4. Interpretasi hasil tingkat kesukaran dengan analisis Rasch

<b>Tingkat Kesukaran (TK)</b>	<b>Interpretasi</b>
$0,48 < TK$	Sangat Sukar
$0,00 < TK \leq 0,48$	Sukar
$-0,48 \leq TK \leq 0,00$	Mudah
$TK < -0,48$	Sangat Mudah

(Sumintono & Widhiarso, 2015)

Tingkat kesulitan butir dengan analisis Rasch pada dasarnya sama dengan taraf kesukaran teori tes klasik, yaitu perbandingan antara jumlah jawaban benar dengan jumlah soal yang diujikan. Perbedaannya, pada analisis Rasch nilai peluang diskalakan dengan memasukkan fungsi logaritma, sehingga dapat menghasilkan suatu pengukuran dengan interval yang sama karena terjadi transformasi data *odd ratio* dengan logaritma (Sumintono & Widhiarso, 2015).

#### 3.6.4. Daya Pembeda

Daya pembeda digunakan untuk mengetahui kemampuan soal dalam membedakan siswa dengan kemampuan yang tinggi dan kemampuan yang rendah (Satriadi, 2019). Daya pembeda pada SFFT dapat diketahui menggunakan *software* MINISTEP 4.4.5 dengan menu *output Table 10 Item Fit Order* yang dianalisis dengan Rasch.

Korelasi *point-measure* mengacu pada hubungan antara tingkat kesulitan item tiap individu dan tingkat kesulitan soal secara keseluruhan, dimana nilai satu menunjukkan kemampuan siswa yang tinggi menjawab soal dengan benar dan sebaliknya yang menunjukkan korelasi sempurna menurut Rasch, nilai nol menunjukkan tidak ada hubungan antara respon item dan item secara keseluruhan, sedangkan nilai negatif menunjukkan masalah pada item soal karena sering mendapatkan skor rendah dibanding skor tinggi (Smiley, 2015).

Penggunaan *Point Measure Correlation (Pt Mean Corr)* seperti yang diungkapkan sebelumnya dapat memberikan informasi mengenai daya pembeda dari suatu instrumen soal yang digunakan untuk membedakan kemampuan siswa.

Untuk dapat mengetahui kategori daya pembeda tiap butir soal instrumen tes diagnostic SFFT pada kolom *PT-MEASURE CORR.*, Smiley (2015) memberikan interpretasi untuk setiap nilai yang diberikan seperti yang terdapat pada Tabel 3.5 sebagai berikut.

Tabel 3. 5. Interpretasi nilai *Point Measure Correlation*

<i>Pt Mean Corr</i>	<b>Interpretasi</b>
$0,40 < ID$	Sangat baik
$0,30 \leq ID \leq 0,40$	Baik
$0,20 \leq ID < 0,30$	Kurang baik
$ID < 0,20$	Jelek

Smiley (2015)

### 3.6.5 Deteksi Bias Pengukuran

Menurut Sumintono & Widhiarso (2015) butir maupun instrument pengukuran dapat bersifat bias, yaitu ketika sebuah butir lebih memihak pada salah satu individu dengan karakteristik tertentu dan karakteristik oposisinya justru dirugikan. Contohnya, butir sebuah tes melibatkan gambar berupa salju untuk dikenali kejanggalannya. Bagi anak-anak yang pernah berinteraksi dengan salju soal ini cukup mudah dipahami. Sebaliknya, bagi anak-anak yang tidak berinteraksi dengan salju soal ini sulit dipahami. Butir ini cenderung bias dalam mengukur, yang dalam psikometri dinamakan dengan butir yang terjangkit keberfungsian butir diferensial (*DIF/differential item function*). Adanya butir bias dapat mengganggu validitas skor pengukuran yang didapatkan.

Pemodelan Rasch menyediakan menu untuk memfasilitasi peneliti yang hendak mendeteksi adanya butir-butir yang terjangkit DIF. Pada *software* MINISTEP, informasi mengenai bias butir ini dapat dilihat melalui Tabel 30 Item: DIF, *between/within*. Butir-butir yang memiliki nilai P (PROB.) dibawah 0,05 menunjukkan bahwa butir tersebut terjangkit DIF. Pada table tersebut akan muncul nilai selisih tingkat kesulitan butir ditinjau dari dua sampel yang diuji. Contohnya gender, latar belakang, budaya, atau lokasi demografis.

## 3.7. Analisis Data Hasil Penggunaan Tes

Tes diagnostik SFFT yang dikembangkan akan menghasilkan beberapa pola jawaban siswa yang menunjukkan terjadinya konsepsi alternatif. Setiap jawaban yang diberikan siswa akan ditentukan berdasarkan jawaban yang mereka berikan untuk setiap tingkatan. Tingkatan pertama untuk pilihan jawaban ilmiah, tingkatan kedua untuk tingkat keyakinan siswa dalam menjawab tingkat pertama, tingkatan ketiga untuk alasan memilih jawaban pada tingkatan pertama dan tingkatan keempat untuk tingkatan keyakinan siswa dalam menjawab tingkatan ketiga.

Tingkat konsepsi siswa dikategorikan berdasarkan enam kriteria konsepsi siswa yaitu Paham Konsep (PK), Paham Sebagian (PS), Salah Positif (SP), salah Negatif (SN), Konsepsi Alternatif (KA), dan Non Koding (NK). Kriteria konsepsi ditunjukkan pada Tabel 3.6 berikut:

Tabel 3. 6. Kriteria konsepsi siswa untuk *four-tier*

<b>Kriteria Konsepsi</b>	<b>Tingkat 1</b>	<b>Tingkat 2</b>	<b>Tingkat 3</b>	<b>Tingkat 3</b>
Paham Konsep (PK)	Benar	Yakin	Benar	Yakin
Paham Sebagian (PS)	Benar	Yakin	Benar	Tidak yakin
	Benar	Tidak yakin	Benar	Yakin
	Benar	Tidak yakin	Benar	Tidak yakin
	Benar	Yakin	Salah	Tidak yakin
	Benar	Tidak yakin	Salah	Yakin
	Benar	Tidak yakin	Salah	Tidak yakin
	Salah	Yakin	Benar	Tidak yakin
	Salah	Tidak yakin	Benar	Yakin
	Salah	Tidak yakin	Benar	Tidak yakin
	Salah	Yakin	Salah	Tidak yakin
Salah	Tidak yakin	Salah	Yakin	
Salah Positif (SP)	Benar	Yakin	Salah	Yakin
Salah Negatif (SN)	Salah	Yakin	Benar	Yakin
Konsepsi Alternatif (KA)	Salah	Yakin	Salah	Yakin

<b>Kriteria Konsepsi</b>	<b>Tingkat 1</b>	<b>Tingkat 2</b>	<b>Tingkat 3</b>	<b>Tingkat 3</b>
Non Koding (NK)	Apabila tidak mengisi satu atau lebih item (tingkat)			

(Kaltakci, dkk. 2015)

Untuk menghitung skor masing-masing siswa, maka perlu diberikan skor pada masing-masing kriteria konsepsi siswa. Skor pada setiap konsepsi siswa dapat dilihat pada Tabel 3.7 berikut.

Tabel 3. 7. Skor pada masing-masing kriteria konsepsi

<b>Kriteria Konsepsi</b>	<b>Skor</b>
Paham Konsep (PK)	3
Paham Sebagian (PS)	2
Salah Positif (SP)	1
Salah Negatif (SN)	0
Konsepsi Alternatif (KA)	0
Non Koding (NK)	

(Kurnaz dan Eksi, 2015)

Berdasarkan kategori tersebut, maka data dapat dianalisis untuk menentukan pada butir-butir soal siswa mengalami konsepsi alternatif dan berapa besar persentase siswa yang mengalami konsepsi alternatif.

