

BAB III METODE PENELITIAN

3.1 Desain Penelitian

Desain penelitian yang digunakan adalah desain penelitian pengembangan (*Research and Development*) Plomp (2013) yang dikombinasikan dengan prosedur pengembangan tes diagnostik *four-tier* yang dikembangkan oleh Pesman & Eryilmaz (2010). Pengembangan yang dilakukan berorientasi pada pengembangan produk.

Secara garis besar, prosedur pengembangannya terdiri dari lima tahap yaitu 1) investigasi awal (*preliminary investigation*); 2) tahap desain (*the design phase*); 3) tahap realisasi/ konstruksi (*the realization/construction*); 4) tahap evaluasi dan revisi (*the stage of evaluation and revision*); serta 5) tahap implementasi (*the stage of implementation*). Kemudian dikombinasikan dengan prosedur pengembangan tes diagnostik *four-tier* yang terdiri dari tiga tahap yaitu 1) Melakukan wawancara dan kajian literatur mengenai miskonsepsi pada materi gelombang dan optik; 2) pengumpulan informasi mengenai konsep alternatif peserta didik; serta 3) mengembangkan dan mengelolah instrumen tes diagnostik *four-tier*. Prosedur penelitian secara lengkap dapat dilihat pada gambar 3.1.

3.2 Lokasi dan Subjek Penelitian

Penelitian dilaksanakan di tiga SMP yang berlokasi di Kecamatan Lembang, Kabupaten Bandung Barat, Provinsi Jawa Barat. Alasan pemilihan tiga sekolah karena ketiga sekolah tersebut homogen baik dari segi kurikulum maupun akreditasi. Subjek dalam penelitian ini adalah peserta didik SMP kelas VIII semester genap tahun ajaran 2018/2019 yang telah mempelajari materi gelombang dan optik.

Subjek penelitian dibagi menjadi tiga yaitu untuk keperluan identifikasi konsep alternatif peserta didik, subjek uji coba instrumen, dan untuk implementasi instrumen tes diagnostik *four-tier* hasil pengembangan. Subjek penelitian untuk keperluan identifikasi konsep alternatif terdiri dari 45 peserta didik dengan rincian untuk setiap sekolah berjumlah 15 orang (kriterianya yaitu masing-masing 5 peserta didik dengan nilai ulangan harian tinggi, sedang, dan rendah). Informasi nilai

ulangan harian diperoleh dari hasil studi dokumentasi nilai ulangan harian peserta didik dan wawancara dengan pendidik IPA yang bersangkutan. Subjek penelitian untuk keperluan uji coba instrumen terdiri dari 247 peserta didik yang tersebar di SMPN 3 dan SMPN 4. Sedangkan subjek penelitian untuk keperluan implementasi instrumen tes diagnostik *four-tier* hasil pengembangan terdiri dari 51 peserta didik di SMPN 2.

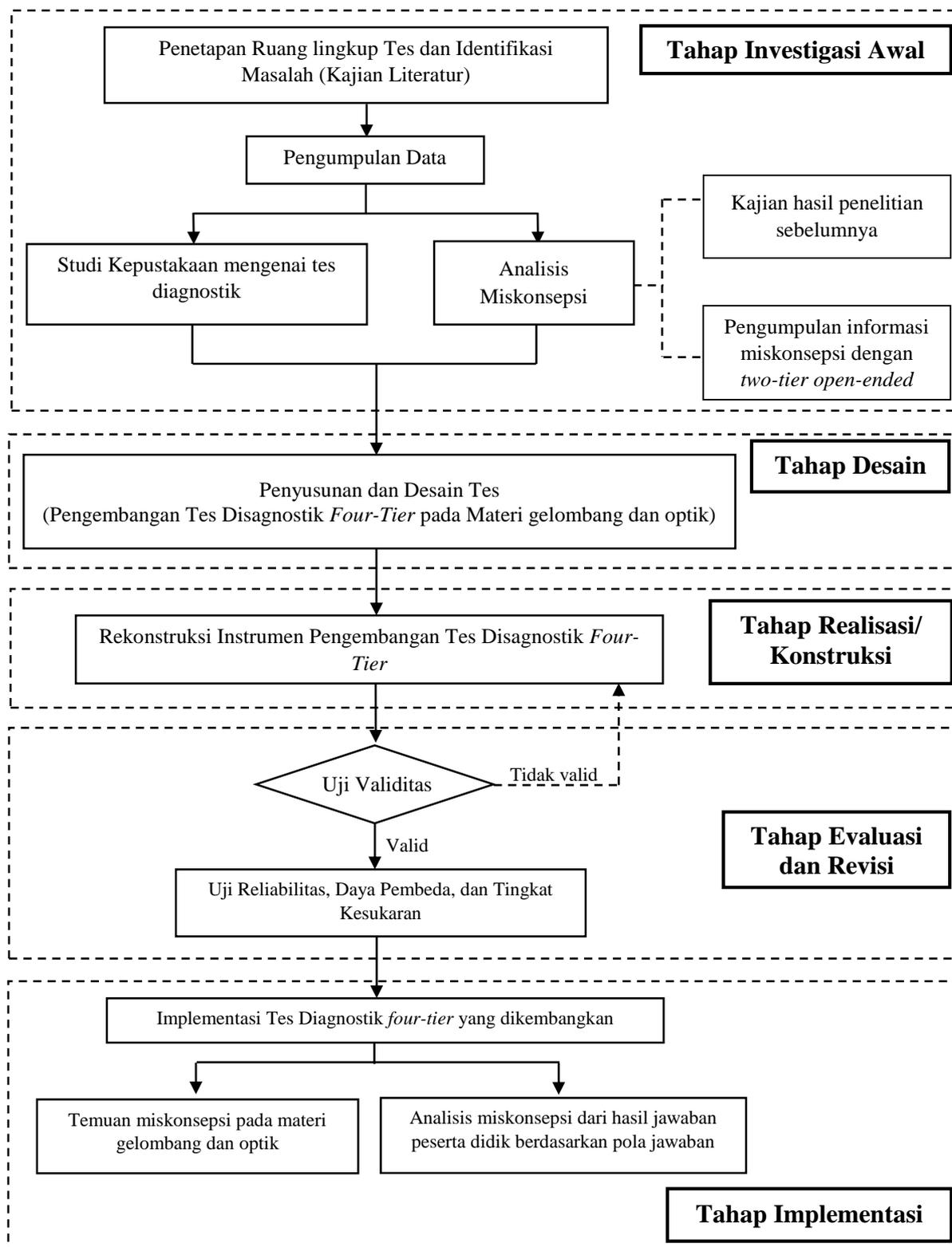
3.3 Definisi Operasional

Adapun definisi operasional dalam penelitian ini yaitu:

- 1) Pengembangan tes diagnostik adalah suatu proses yang dilakukan untuk memperoleh tes diagnostik yang dapat dinyatakan valid dan layak melalui prosedur penelitian pengembangan Plomp yang dikombinasikan dengan prosedur pengembangan tes diagnostik oleh Pesman & Eryilmaz.
- 2) Tes diagnostik *four-tier* adalah tes untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik. Tes ini berupa pilihan ganda empat tingkat yang terdiri dari pilihan jawaban pada tingkat pertama, tingkat keyakinan terhadap pilihan jawaban pada tingkat kedua, pilihan alasan pada tingkat ketiga, dan tingkat keyakinan terhadap pilihan alasan pada tingkat keempat.
- 3) Miskonsepsi adalah pemahaman konsep peserta didik terkait materi gelombang dan optik yang tidak sesuai dengan pemahaman konsep yang benar menurut para ahli.
- 4) Materi gelombang dan optik merupakan salah satu materi dalam pembelajaran IPA yang membahas terkait getaran, gelombang, bunyi, sifat-sifat cahaya, pembentukan bayangan, dan alat optik.

3.4 Prosedur Penelitian

Prosedur pelaksanaan penelitian dilakukan melalui lima tahapan yaitu 1) investigasi awal (*preliminary investigation*); 2) tahap desain (*the design phase*); 3) tahap realisasi/ konstruksi (*the realization/construction*); 4) tahap evaluasi dan revisi (*the stage of evaluation and revision*); serta 5) tahap implementasi (*the stage of implementation*). Skema penelitian disajikan dalam bentuk bagan di bawah ini:



Gambar 3.1 Prosedur Penelitian Pengembangan Plomp yang Dikombinasikan dengan Prosedur Pengembangan Tes Diagnostik *Four-Tier*

1) Tahap investigasi awal (*preliminary investigation*)

Pengembangan tes diagnostik *four-tier* dimulai dengan tahap investigasi awal (*preliminary investigation*). Tahap ini terdiri dari 2 langkah, sebagai berikut:

a) Penentuan ruang lingkup dan tujuan tes

Penetapan ruang lingkup tes dan tujuan tes pada penelitian ini dilakukan melalui analisis KI dan KD SMP yang terdapat pada kurikulum 2013. Berdasarkan hasil analisis diperoleh beberapa label konsep diantaranya getaran, gelombang, bunyi, sifat-sifat cahaya, pembentukan bayangan, dan alat optik. Analisis terhadap kurikulum 2013 bertujuan untuk menentukan indikator yang mewakili setiap label konsep yang akan dikembangkan menjadi butir soal untuk mengidentifikasi miskonsepsi pada materi tersebut.

b) Pengumpulan data

Proses pengumpulan data pada penelitian ini dilakukan melalui studi kepustakaan dan analisis miskonsepsi. Hasil dari studi kepustakaan atau kajian literatur mengenai tes diagnostik *four-tier*, serta miskonsepsi pada materi gelombang dan optik telah dijelaskan sebelumnya pada kajian teori. Analisis miskonsepsi dilakukan untuk memperoleh informasi mengenai konsep alternatif peserta didik. Pada proses tersebut peserta didik diberikan soal dalam bentuk *open ended two-tier multiple choice*. Berdasarkan soal tersebut peserta didik diminta memilih jawaban yang benar dan menuliskan alasan memilih jawaban tersebut. Alasan yang diberikan peserta didik akan ditabulasi untuk selanjutnya digunakan untuk pilihan alasan pada tes diagnostik *four-tier* yang dikembangkan.

2) Tahap desain (*the design phase*)

Hasil dari identifikasi miskonsepsi pada tahap investigasi awal (*preliminary investigation*) akan dianalisis untuk selanjutnya dikembangkan menjadi butir soal. Identifikasi miskonsepsi akan melahirkan sebuah pola yang terdiri dari konsep yang benar dan konsep yang terdapat miskonsepsi di dalamnya. Berdasarkan pada pola tersebut akan dikembangkan menjadi sebuah tes diagnostik berbentuk pilihan ganda tipe *four-tier*. Pada tingkat pertama terdiri empat pilihan jawaban, tingkat ketiga terdiri dari empat pilihan alasan, tingkat kedua dan keempat merupakan tingkat

keyakinan terhadap jawaban yang diberikan pada tingkat pertama dan ketiga. Tingkat keyakinan ini terdiri dari dua pilihan yaitu yakin dan tidak yakin.

<p>Soal ke-<i>n</i></p> <p>Butir Soal Pilihan Ganda</p> <p><i>n.1</i> Pilihan Jawaban</p> <p>A. (pilihan jawaban 1)</p> <p>B. (pilihan jawaban 2)</p> <p>C. (pilihan jawaban 3)</p> <p>D. (pilihan jawaban 4)</p> <p><i>n.2</i> Tingkat keyakinan terkait pilihan jawaban</p> <p>A. Yakin B. Tidak Yakin</p> <p><i>n.3</i> Pilihan alasan</p> <p>A. (pilihan alasan 1)</p> <p>B. (pilihan alasan 2)</p> <p>C. (pilihan alasan 3)</p> <p>D. (pilihan alasan 4)</p> <p><i>n.4</i> Tingkat keyakinan terkait pilihan alasan</p> <p>A. Yakin B. Tidak Yakin</p>
--

Gambar 3.2 Desain Tes Diagnostik *Four-Tier* yang Dikembangkan

3) Tahap realisasi/ konstruksi (*the realization/construction*)

Setelah membuat desain instrumen tes diagnostik *four-tier*, peneliti mengkonstruksi desain instrumen tes diagnostik tersebut. Konstruksi instrumen tes diagnostik *four-tier* yang dikembangkan disesuaikan dengan desain yang telah dijelaskan pada tahap sebelumnya. Konstruksi tersebut merupakan bentuk realisasi instrumen tes diagnostik *four-tier* hendak digunakan oleh peneliti untuk mengidentifikasi miskonsepsi peserta didik pada materi gelombang dan optik.

4) Tahap evaluasi dan revisi (*the stage of evaluation and revision*)

Tahap ini dilakukan dengan melakukan uji kualitas tes diagnostik *four-tier* yang dikembangkan. Tahap evaluasi dan revisi dilakukan melalui 4 tahapan yaitu uji validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Uji validitas yang dilakukan adalah uji validitas isi (*content validity*) dan uji validitas konstruk

(*construct validity*). Uji validitas isi (*content validity*) berupa uji kesesuaian antara konsep dan instrumen yang dikembangkan. Metode pengujian validitas yang digunakan adalah metode CVR dan CVI yang merupakan sebuah pendekatan validitas isi untuk mengetahui kesesuaian item dengan domain yang diukur berdasarkan *judgement* para ahli. Para ahli yang terlibat dalam validasi ini terdiri dari tiga orang. Setiap ahli memiliki keahlian masing-masing dalam bidang pendidikan fisika, *assessment* dan evaluasi. Validasi ini juga melibatkan tiga *reviewer* yaitu pendidik IPA di Sekolah Menengah Pertama. Pelibatan pendidik sebagai *reviewer* dimaksudkan agar diperoleh masukan yang berkaitan dengan penerapan produk ini di lapangan dengan pertimbangan bahwa mereka lebih mengetahui kondisi dan kemampuan peserta didik. Untuk uji validitas konstruk, uji reliabilitas, daya pembeda, dan tingkat kesukaran digunakan analisis dengan menggunakan rasch model.

Keunggulan dari Rasch model dibanding metode lainnya yaitu kemampuan melakukan prediksi terhadap data yang hilang (*missing data*), yang didasarkan pada pola respon yang sistematis. Metode lain biasanya mengestimasi data yang hilang dengan nilai nol (0), sedangkan Rasch model akan menghasilkan prediksi kemungkinan nilai terbaik dari data yang hilang tersebut. Dengan demikian data yang diperoleh seolah-olah sebagai data yang lengkap dan lebih akurat dalam analisis statistik selanjutnya (Sumintono & Widhiarso, 2015).

5) Tahap implementasi (*the stage of implementation*)

Setelah melalui tahap evaluasi dan revisi, instrumen tes diagnostik *four-tier* yang dikembangkan kemudian di implementasikan kepada peserta didik untuk mengungkap miskonsepsi yang dialami peserta didik pada materi gelombang dan optik. Jawaban yang diberikan peserta didik akan dianalisis untuk mengetahui miskonsepsi yang dialami. Hasil implementasi juga akan memberikan informasi terkait analisis kelebihan dan kekurangan tes diagnostik *four-tier* yang dikembangkan.

3.5 Instrumen Penelitian

Instrumen penelitian adalah alat yang digunakan untuk mengumpulkan data yang sedang diteliti. Dalam penelitian ini digunakan beberapa instrumen sebagai berikut:

1) Lembar validasi ahli dan *reviewer*

Lembar validasi ahli dan *reviewer* ini berfungsi sebagai instrumen penelitian yang bertujuan mengetahui kriteria kevalidan tes diagnostik yang dikembangkan. Lembar validasi ini akan diberikan kepada ahli dan *reviewer* yang berkompeten dalam menilai dan memberi saran. Ahli ini memiliki keahlian dibidang asesmen dan pendidikan fisika. *Reviewer* merupakan pendidik IPA di tiga SMP. Dari hasil masukan dan saran yang diberikan oleh ahli dan *reviewer* akan dianalisis secara kuantitatif untuk menghitung *Content Validity Ratio* (CVR) dan *Content Validity Index* (CVI). Hasil perhitungan akan menjadi pertimbangan untuk menarik sebuah kesimpulan terhadap tes diagnostik yang dikembangkan apakah dinyatakan valid atau tidak.

2) Lembar indentifikasi konsep alternatif peserta didik (*two-tier open ended*)

Instrumen ini digunakan untuk mengumpulkan jawaban alternatif peserta didik mengenai materi gelombang dan optik. Jawaban tersebut akan ditabulasi kemudian di analisis alasan-alasan mana yang teridentifikasi miskonsepsi. Jawaban peserta didik yang teridentifikasi miskonsepsi akan dijadikan sebagai pilihan alasan pada *tier* ketiga dalam perangkat tes diagnostik *four-tier*.

3) Lembar validasi empiris

Instrumen validasi empiris merupakan seperangkat tes diagnostik *four-tier* yang diperoleh dengan memasukkan konsep alternatif peserta didik yang teridentifikasi miskonsepsi dan temuan miskonsepsi hasil kajian literatur pada *tier* ketiga tes diagnostik yang dikembangkan. Tes diagnostik ini merupakan transformasi dari instrumen *two-tier open ended* menjadi tes diagnostik *four-tier* dan hasil revisi dari ahli.

4) Lembar identifikasi level konsepsi peserta didik

Instrumen ini merupakan produk akhir dari proses pengembangan yang dilakukan setelah melewati beberapa tahap pengujian yaitu uji validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Instrumen tes diagnostik *four-tier* yang telah dikembangkan akan digunakan untuk memetakan level konsepsi peserta didik.

3.6 Teknik Analisis Data

Pada penelitian ini, teknik analisis data yang digunakan terbagi menjadi dua bagian yaitu kualitas butir soal dan hasil penggunaan tes. Kualitas butir soal

meliputi uji validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Sedangkan pengukuran hasil penggunaan tes menggunakan analisis keputusan terhadap pola jawaban peserta didik.

1) Uji validitas

Validitas merupakan hal yang paling penting untuk dipertimbangkan ketika mempersiapkan atau memilih instrumen untuk digunakan. Validitas telah didefinisikan sebagai hal yang mengacu pada kesesuaian, kebenaran, keberagaman, dan kegunaan dari inferensi spesifik yang peneliti buat berdasarkan pada data yang mereka kumpulkan. Validasi adalah proses pengumpulan dan analisis bukti untuk mendukung kesimpulan tersebut (Fraenkel, dkk. 2011).

Validitas tes diagnostik *four-tier* ditentukan dengan beberapa teknik secara kuantitatif dan kualitatif. Pertama-tama, item dari tes diagnostik ini diperiksa dan dinilai oleh para ahli dan *reviewer* berdasarkan isi dan format. Uji validitas dilakukan dengan menggunakan metode *Content Validity Ratio (CVR)*. Perhitungan nilai CVR dilakukan berdasarkan persamaan Lawshe (1975) sebagai berikut :

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \quad (3.1)$$

Keterangan :

n_e : Jumlah validator yang menyatakan valid

N : Jumlah total validator

Ketentuan:

- Saat kurang dari $\frac{1}{2}$ dari total responden menyatakan “Ya” maka nilai CVR akan bernilai negatif.
- Saat $\frac{1}{2}$ dari total responden menyatakan “Ya” maka nilai CVR = 0.
- Saat seluruh responden menyatakan “Ya” maka nilai CVR = 1 (hal ini diatur menjadi 0,99 disesuaikan dengan jumlah responden)
- Saat jumlah responden yang menyatakan “Ya” lebih dari $\frac{1}{2}$ jumlah total responden maka nilai CVR = 0 – 0,99.

Hasil perhitungan CVR setiap butir soal kemudian dibandingkan dengan nilai CVR minimum untuk menentukan valid atau tidaknya soal tersebut. Kriteria nilai CVR minimum untuk jumlah validator enam orang adalah lebih besar atau sama

dengan 0,672 ($CVR \geq 0,672$). Selain itu, dilakukan pula penentuan validitas tes secara keseluruhan dengan menggunakan metode *Content Validity Index* (CVI). Nilai CVI merupakan rata-rata dari nilai CVR. Kriteria nilai CVI sebagai berikut:

Tabel 3.1
Kriteria Hasil Perhitungan CVI

Rentang	Kategori
$CVI \geq 0,68$	Sangat Sesuai
$0,34 \leq CVI \leq 0,67$	Sesuai
$CVI < 0,34$	Tidak Sesuai

(Wilson, dkk. 2012)

Sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Untuk melihat validitas konstruk terkait kualitas butir soal digunakan rasch model. Uji validitas ini dilakukan dengan bantuan program *winsteps*. Hal yang dilihat adalah berdasarkan kriteria nilai *Outfit Mean Square* (*MNSQ*) yaitu $0,5 < MNSQ < 1,5$ (semakin mendekati 1,00 maka kualitas butir soal semakin baik) dan nilai *Outfit Z-Standard* (*ZSTD*) yaitu $-2,0 < ZSTD < +2,0$ (semakin mendekati nilai 0,0 maka kualitas butir soal semakin baik) (Sumintono & Widhiarso, 2015).

2) Reliabilitas

Reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen yang berkenaan dengan pertanyaan, apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Uji reliabilitas butir soal dilakukan dengan menggunakan Rasch model. Pada rasch model, reliabilitas digambarkan dengan nilai *person reliability*, *item reliability*, dan *item separation*. Biasanya reliabilitas yang rendah dikarenakan sampel terlalu sedikit sehingga variasi hirarki pada penggaris logit hanya sedikit (Sumintono & Widhiarso, 2015). Reliabilitas dianalisis dengan cara membandingkan nilai *person reliability*, dan *item reliability* yang diperoleh dengan kriteria koefisien reliabilitas pada Tabel 3.2. Sedangkan semakin tinggi nilai *item separation* yang diperoleh maka semakin baik pengukuran yang dilakukan (Sumintono & Widhiarso, 2015).

Tabel 3.2
Kriteria Koefisien Reliabilitas

Rentang	Kategori
$0,94 \leq r < 1,00$	Reliabilitas istimewa
$0,91 \leq r < 0,94$	Reliabilitas bagus sekali
$0,80 \leq r < 0,91$	Reliabilitas bagus
$0,67 \leq r < 0,80$	Reliabilitas cukup
$0,00 \leq r < 0,67$	Reliabilitas lemah

(Fisher, 2007)

3) Daya pembeda

Daya pembeda suatu soal merupakan kemampuan butir soal itu untuk membedakan peserta didik dengan kemampuan tinggi dan peserta didik dengan kemampuan rendah. Informasi mengenai daya pembeda sangat diperlukan untuk memperbaiki kualitas soal. Semakin tinggi daya pembeda soal berarti semakin mampu soal tersebut membedakan kemampuan peserta didik dalam memahami materi. Daya pembeda soal dilihat berdasarkan kriteria nilai *Point Measure Correlation (Pt Mean Corr)* (Alagumalai, dkk. 2005).

Tabel 3.3
Kriteria Daya Pembeda

Nilai <i>Point Measure Correlation</i>	Kriteria
$Pt\ Mean\ Corr \geq 0,4$	Sangat bagus
$0,3 \leq Pt\ Mean\ Corr < 0,4$	Bagus
$0,2 \leq Pt\ Mean\ Corr < 0,3$	Cukup
$Pt\ Mean\ Corr < 0,2$	Tidak mampu mendeskriminasi

4) Tingkat kesukaran

Tingkat kesukaran adalah kemampuan tes dalam menjangkau banyaknya subjek peserta tes yang dapat mengerjakan dengan benar. Jika banyak subjek peserta tes yang tidak dapat menjawab dengan benar, maka tingkat kesukaran tes tersebut tinggi, sebaliknya jika hanya sedikit dari subjek yang tidak dapat menjawab dengan benar maka tingkat kesukarannya rendah (Sumintono & Widhiarso, 2015). Analisis tingkat kesukaran soal menggunakan Rasch model berdasarkan nilai *measure*. Tabel 3.4 menunjukkan kriteria tingkat kesukaran.

Tabel 3.4
Kriteria Tingkat Kesukaran

Nilai <i>measure</i> (P)	Kriteria
$P > 1,0$	Sangat sulit
$0,0 < P \leq 1,0$	Sulit
$-1,0 \leq P \leq 0,0$	Mudah
$P < -1,0$	Sangat mudah

(Sumintono & Widhiarso, 2015)

5) Analisis data hasil implementasi tes diagnostik *four-tier*

Tes diagnostik *four-tier* yang dikembangkan akan menghasilkan beberapa pola jawaban peserta didik yang menunjukkan terjadinya miskonsepsi, sebagaimana ditunjukkan pada Tabel 3.5. Setiap jawaban yang diberikan peserta didik akan ditentukan berdasarkan pola jawaban yang mereka berikan untuk setiap tingkat.

Tabel 3.5
Keputusan Terhadap Pola Jawaban Peserta Didik

Tingkat Pertama	Tingkat Kedua	Tingkat Ketiga	Tingkat Keempat	Keputusan
Benar	Yakin	Benar	Yakin	SC
Benar	Yakin	Benar	Tidak yakin	LK
Benar	Tidak yakin	Benar	Yakin	LK
Benar	Tidak yakin	Benar	Tidak yakin	LK
Benar	Yakin	Salah	Yakin	FP
Benar	Yakin	Salah	Tidak yakin	LK
Benar	Tidak yakin	Salah	Yakin	LK
Benar	Tidak yakin	Salah	Tidak yakin	LK
Salah	Yakin	Benar	Yakin	FN
Salah	Yakin	Benar	Tidak yakin	LK
Salah	Tidak yakin	Benar	Yakin	LK
Salah	Tidak yakin	Benar	Tidak yakin	LK
Salah	Yakin	Salah	Yakin	MSC
Salah	Yakin	Salah	Tidak yakin	LK
Salah	Tidak yakin	Salah	Yakin	LK
Salah	Tidak yakin	Salah	Tidak yakin	LK

(Gurel, dkk. 2017)

Keterangan:

SC: Scientific Conception; *LK*: Lack of Knowledge; *FN*: False Negatif; *FP*: False Positif; *MSC*: Misconception

Ketentuan:

- Jawaban benar pada tingkat pertama dan tingkat ketiga diberi nilai 1 sedangkan jawaban salah diberi nilai 0.
- Jawaban yakin pada tingkat kedua dan tingkat keempat diberi nilai 1 sedangkan jawaban tidak yakin diberi nilai 0.
- Pengkategorian kombinasi jawaban yang diberikan peserta didik ke dalam level-level konsepsi, mengharuskan keempat *tier* pada tes diagnostik *four-tier* direspons atau diisi oleh peserta didik. Salah satu diantara *tier* yang tidak direspons menyebabkan ketidakakuratan dalam menempatkan peserta didik pada level konsepsi yang ada. Sehingga untuk menghindari yang demikian, setiap butir soal layak dianalisis jika keempat *tier* diisi. Jika terdapat peserta didik yang tidak menjawab keseluruhan soal maka pada analisis jawaban peserta didik dikode dengan istilah *error* dan tidak diikuti dalam perhitungan untuk menentukan jumlah dan persentase pada masing-masing level konsepsi.

Berdasarkan kategori tersebut, maka dapat dianalisis pada butir soal yang mana peserta didik mengalami miskonsepsi dan seberapa besar persentasenya, dengan menggunakan persamaan sebagai berikut:

$$P = \frac{S}{N} \times 100 \% \quad (3.2)$$

Keterangan :

P : Persentase level konsepsi peserta didik

S : Jumlah peserta didik yang teridentifikasi MSC/SC/LK/FP/FN

N : Jumlah peserta didik yang mengikuti tes

(Prodjosantoso, dkk. 2018)

Sebuah miskonsepsi dianggap signifikan apabila miskonsepsi yang dialami peserta didik sebesar 10% atau lebih dari jumlah keseluruhan peserta didik yang diteliti (Caleon & Subramaniam, 2010b).