

BAB III METODE PENELITIAN

3.1 Desain Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini adalah metode *mix methode*. Hasil penelitian berupa pemaparan secara kualitatif dan kuantitatif. Desain penelitian yang digunakan adalah desain penelitian pengembangan (*Research and Development*) oleh Plomp (2013) dan dikombinasikan dengan prosedur pengembangan instrument *three tier diagnostic* oleh Liampa, Malandrakis, Papadopoulou & Pnevmatiko (2019). Pengembangan yang dilakukan berupa pengembangan produk. Menurut Plomp (2013) penelitian pengembangan terdiri dari lima tahap, yaitu 1) investigasi awal (*preliminary investigation*); 2) tahap desain (*the desain phase*); 3) tahap realisasi/konstruksi (*the realization/construction*); 4) tahap evaluasi (*the stage of evaluation*); 5) tahap implementasi (*the stage of implementation*). Pengembangan instrument *three tier diagnostic* terdiri dari tiga fase yaitu 1) mendefinisikan konsep; 2) Mengases miskonsepsi peserta didik; 3) Pengembangan instrument tes.

3.2 Partisipan

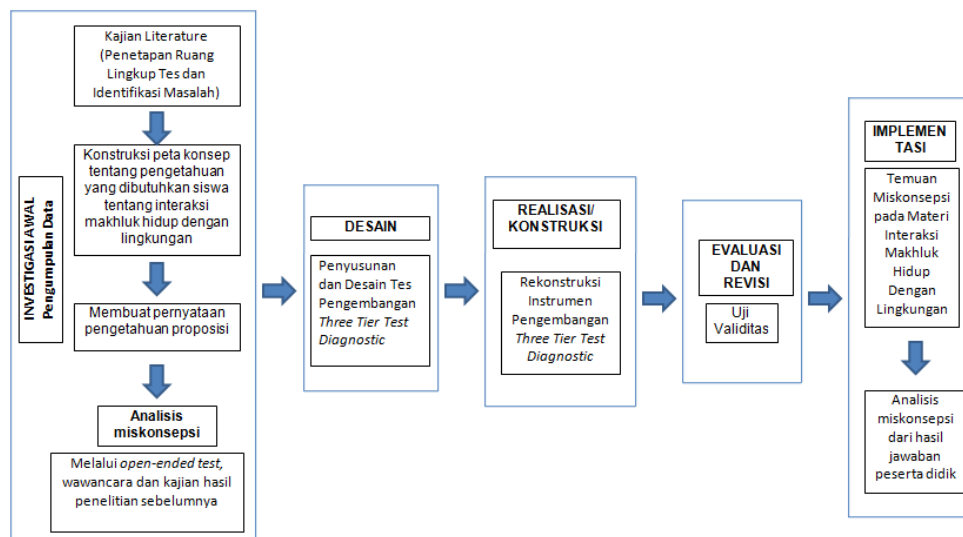
Penelitian ini dilaksanakan di lima SMP dengan kurikulum dan akreditasi yang berbeda. Subjek penelitian terdiri dari 229 peserta didik. Mereka dibagi menjadi tiga kelompok dengan pembagian sebagai berikut :

- a.) Subjek penelitian untuk mengidentifikasi konsep alternatif sebanyak 28 peserta didik melalui *open ended questionnaire* dan 4 peserta didik melalui wawancara di SMP N 29 Bandung
- b.) Subjek penelitian untuk uji coba instrumen *three tier test diagnostic* sebanyak 67 peserta didik di SMPN 1 Guguak, SMPN 1 Suliki dan SMPN 2 Mungka

- c.) Subjek penelitian untuk mengimplementasikan instrumen *three tier test diagnostic* sebanyak 130 peserta didik di SMPN 1 Mungka, SMPN 2 Mungka, SMPN 1 Guguak dan SMPN 1 Suliki

3.3 Prosedur Penelitian

Prosedur penelitian yang akan dilakukan melalui lima tahap, secara ringkas dapat dilihat melalui prosedur berikut :



Gambar 3. 1 Tahapan Pengembangan Instrumen

a.) Investigasi Awal

Pada tahap ini dilakukan pengumpulan data melalui beberapa tahapan, yaitu :

- 1.) Kajian literatur untuk menetapkan ruang lingkup tes

Tahap awal penelitian ini adalah menentukan ruang lingkup dan tujuan tes melalui analisis KI dan KD SMP. Analisis ini bertujuan untuk menentukan indikator yang mewakili konsep pembelajaran yang nantinya akan dikembangkan menjadi butir soal untuk mengidentifikasi miskonsepsi pada materi tersebut.

- 2.) Konstruksi peta konsep tentang pengetahuan yang dibutuhkan oleh siswa tentang interaksi makhluk hidup.
- 3.) Membuat pernyataan proposisi berdasarkan peta konsep yang telah dibuat.

4.) Analisis miskonsepsi

Analisis miskonsepsi bertujuan untuk memperoleh informasi tentang konsep alternatif peserta didik melalui *open ended two tier multiple choise* dan wawancara. Soal-soal pada *open ended two tier multiple choise* merupakan soal pengembangan dari peta konsep dan pernyataan proposisi. Peneliti akan menggunakan jawaban siswa untuk pilihan jawaban pada *three tier test diagnostic* yang akan dikembangkan.

b.) Desain produk (*The design phase*)

Desain produk berupa instrumen *three tier test diagnostic* yang berisi butir soal. Butir soal ini mengacu pada hasil analisis miskonsepsi pada tahapan awal. Hasil dari identifikasi miskonsepsi berupa pola yang terdiri dari konsep benar dan konsep miskonsepsi. Berdasarkan pola tersebut maka peneliti akan mengembangkan *three tier test diagnostic*. Tes ini terdiri dari tiga tingkatan, tingkat pertama terdiri dari empat pilihan jawaban, tingkat kedua terdiri dari alasan pemilihan pada tingkatan pertama dan tingkatan ketiga terdiri dari dua pilihan tingkat keyakinan, yaitu yakin dan tidak yakin.

Pilihan Jawaban Tingkat Pertama

- A. Pilihan jawaban 1
- B. Pilihan jawaban 2
- C. Pilihan jawaban 3
- D. Pilihan jawaban 4

Pilihan Jawaban Tingkat Kedua

- A. Pilihan alasan 1
- B. Pilihan alasan 2
- C. Pilihan alasan 3
- D. Pilihan alasan 4
- E. (Pilihan alasan sendiri)

Pilihan Jawaban Tingkat Ketiga

- A. Yakin
- B. Tidak Yakin

Gambar 3. 2 Desain three tier test diagnostic yang dikembangkan

c.) Realisasi/konstruksi (*The Realization/Construction*)

Selanjutnya peneliti akan mengkonstruksi desain produk sebagai bentuk realisasi instrumen *three tier test diagnostic*. Kontruksi ini dipakai untuk mengidentifikasi miskonsepsi.

d.) Evaluasi dan revisi (*The stage of evaluation and revision*)

Peneliti melakukan uji kualitas *three tier test diagnostic* yang telah dikembangkan. Uji kualitas instrumen dilakukan melalui 4 tahapan yaitu :

1.) Uji validitas

Validitas merupakan acuan yang menunjukkan tingkat kesahihan penelitian untuk mengklaim apa yang diukur sehingga kesimpulan bisa diperoleh dengan akurat (Fraenkel *et all*, 2012; Nkwake, 2015). Uji validitas terdiri dari dua uji yakni uji validitas isi dan uji validitas empiris. Uji validitas isi berupa kesesuaian uji antara konsep dan instrumen yang dikembangkan. Pengujian validitas isi menggunakan metode CVR dan CVI. Uji validitas empiris menggunakan rasch model. 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:

$$\text{CVR} = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

Keterangan

n_e : Jumlah validator yang menyatakan valid

N : Jumlah total validator

Ketentuan :

1. Saat kurang dari $\frac{1}{2}$ dari total responden menyatakan “Ya” maka nilai CVR akan bernilai negatif
2. Saat $\frac{1}{2}$ dari total responden menyatakan “Ya” maka nilai CVR = 0
3. Saat seluruh dari total responden menyatakan “Ya” maka nilai CVR = 1 (hal ini diatur menjadi 0,99 disesuaikan dengan jumlah responden)
4. Saat jumlah responden menyatakan “Ya” lebih dari $\frac{1}{2}$ dari 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. Selain itu dilakukan pula penentuan validitas tes secara keseluruhan dengan menggunakan *Content Validity Index* (CVI). Nilai CVI merupakan rata-rata dari nilai CVR.

Tabel 3. 1 Kategori CVI

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

(Wilson, Pan, & Schumsky, 2012)

Sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Untuk melihat validitas 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 $- 0,2 < ZSTD < + 2,0$ (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 dianalisis dengan cara membandingkan nilai *person reliability* dan *item reliability* yang diperoleh dengan kriteria koefisien reliabilitas. Semakin tinggi nilai item separation yang diperoleh maka semakin baik pengukuran yang dilakukan.

Tabel 3. 2 Kategori Reliabilitas

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

(Sumintono & Widhiarso, 2015)

3.) Daya pembeda

Daya pembeda suatu soal merupakan kemampuan butir soal 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 *nilai separation*.

4.) Indeks kesukaran.

Indeks kesukaran adalah kemampuan tes dalam menjangkau banyaknya subjek tes peserta tes dengan yang mengerjakan 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). Pada pemodelan rasch, Indeks kesukaran dihitung berdasarkan nilai measure logit dan standar deviasi.

Tabel 3. 3 Kriteria *item measure* (kesulitan soal)

Item Measure	Kriteria
measure (nilai logit) > 1,11	Sangat Sulit
$0,00 \leq \text{measure (nilai logit)} \leq 1,11$	Sulit
$-1,11 \leq \text{measure (nilai logit)} \leq 0,00$	Mudah
measure (nilai logit) < -1,11	Sangat mudah

(Sumintono & Widhiarso, 2015)

5.) Implementasi produk (*the stage of implementation*)

Instrumen *three tier test diagnostic* yang telah melalui tahapan evaluasi dan revisi di ujikan kepada peserta didik. *Three tier test diagnostic* akan mengungkapkan miskonsepsi peserta didik pada materi yang diujikan.

3.4 Definisi Operasional Variabel

1) *Three tier test diagnostic* merupakan tes diagnostik berupa pilihan ganda tiga tingkat yang dikembangkan dengan tujuan mengidentifikasi miskonsepsi. Tes ini mempunyai tiga tingkatan, adapun tingkatan pertama berupa konten yang tertuang pada pilihan ganda, tingkatan kedua berupa alasan pemilihan tingkatan pertama dan tingkatan ketiga berupa keyakinan. Tes ini bertujuan untuk mengetahui pemahaman konseptual siswa. Miskonsepsi merupakan ketidaksesuaian pemahaman siswa pada konsep interaksi makhluk hidup dengan lingkungannya. Miskonsepsi ditentukan berdasarkan pola jawaban siswa pada tiap tingkatan *three tier test diagnostic* dengan mengacu kepada

keputusan terhadap pola jawaban peserta didik yang dikembangkan oleh Kaltakci Gurel, Eryilmaz & McDermott (2015).

2) Pengembangan *three tier test* diagnostik interaksi makhluk hidup dengan lingkungannya merupakan instrument tes yang valid, reliabel dan dapat diandalkan untuk memperoleh pemahaman konseptual siswa melalui prosedur pengembangan *research and development* Plom yaitu investigasi awal, desain, realisasi/konstruksi, evaluasi dan implementasi yang dikombinasikan dengan prosedur pengembangan Liampa, Malandrakis, Papadopoulou & Pnevmatiko terdiri dari tiga fase yaitu 1) mendefinisikan konsep; 2) Mengases miskonsepsi peserta didik; 3) Pengembangan instrument tes.

3.5 Instrumen Penelitian

Instrumen pengumpulan data terdiri dari :

1) Lembar validasi ahli dan *reviewer*

Instrumen ini bertujuan untuk mengetahui kevalidan tes diagnostik yang dikembangkan. Ahli dan reviewer akan menilai dan memberi saran agar instrumen layak dan valid. Selanjutnya instrumen yang telah divalidasi akan dianalisis dengan menghitung CVR dan CVI.

2) Lembar identifikasi konsep alternatif peserta didik

Instrumen ini bertujuan untuk mengumpulkan jawaban alternatif peserta didik. Berdasarkan jawaban tersebut bisa diidentifikasi miskonsepsi pada konsep tersebut.

3) Lembar validasi empiris

Instrumen ini berupa *three tier test diagnostic* yang diperoleh dengan memasukkan kosep alternatif peserta didik yang teridentifikasi miskonsepsi dan temuan miskonsepsi hasil kajian literatur pada instrumen *three tier test diagnostic* yang dikembangkan.

4) Lembar identifikasi level konsepsi peserta didik

Lembar ini berisi tentang produk akhir pengembangan instrumen *three tier test diagnostic* yang telah melewati beberapa tahap pengujian. Instrumen ini akan digunakan untuk memetakan kosnsepsi peserta didik.

3.6 Teknik Analisis Data

Data yang dianalisis pada penelitian ini adalah kualitas butir soal dan hasil penggunaan tes. Untuk menganalisis kualitas butir soal menggunakan uji validitas, reliabilitas, daya pembeda dan indeks kesukaran. Sedangkan untuk melihat hasil penggunaan tes dengan menggunakan analisis keputusan terhadap pola jawaban peserta didik.

Tabel 3. 4Keputusan Terhadap Pola Jawaban Peserta Didik

Tingkat Pertama	Tingkat Kedua	Tingkat Ketiga	Keputusan
Benar	Benar	Yakin	<i>Scientific Conceptions</i>
Benar	Benar	Tidak Yakin	<i>Lack of Knowledge</i>
Benar	Salah	Yakin	<i>False Positif</i>
			<i>Rarely Misconceptions</i>
Benar	Salah	Tidak Yakin	<i>Lack of Knowledge</i>
Salah	Benar	Yakin	<i>False Negatif</i>
Salah	Benar	Yakin	<i>False Negatif</i>
		Tidak Yakin	<i>Lack of Knowledge</i>
Salah	Salah	Yakin	<i>Misconceptions</i>
			<i>Rarely Misconceptions</i>
Salah	Salah	Tidak Yakin	<i>Lack of Knowledge</i>

Sumber : (Kaltakci-Gurel, Eryilmaz & McDermott, 2017)

Data dianalisis dengan mengikuti prosedur pengkodean dan penilaian yang diadaptasi dari Arslan, Cigdemoglu & Moseley (2012) dengan ketentuan sebagai berikut :

1. Tingkat Satu (Konten)

Respons yang benar diberi kode 1 dan yang lainnya diberi kode 0.

2. Tingkat Dua (Konten dan Alasan)

Pemberian skor pada tingkat ini dengan mempertimbangkan konten dan alasan, jika konten dan alasan benar, maka item tersebut diberi kode 1, jika salah maka diberi kode 0.

3. Tingkat Ketiga (Konten, Alasan dan Tingkat Kepercayaan)

Jika semua tingkatan konten, alasan dan tingkat kepercayaan benar maka diberikan kode 1, jika tidak maka diberikan kode 0.

4. *Lack of Knowledge*

Jika seorang siswa memilih setidaknya satu jawaban yang salah yang menunjukkan kesalahpahaman untuk konten dan tingkat alasan serta menunjukkan ketidakpastian pada tingkat kepercayaan, maka item itu dikodekan sebagai 1 dan kemungkinan lain dikodekan 0.

Pengkodean '1' menunjukkan bahwa siswa memiliki *Lack of Knowledge* tentang item yang dipertimbangkan.

5. Miskonsepsi

Siswa dinyatakan miskonsepsi ketika item jawaban pada tingkat satu menjawab benar namun alasan salah dengan tingkat keyakinan yakin maka dikodekan sebagai 1, sedangkan jika masuk pada kemungkinan yang lain maka dikodekan dengan 0. Berdasarkan kategori tersebut maka dapat dianalisis pada butir soal yang mana peserta didik mengalami miskonsepsi serta persentasenya. Adapun rumus persamaannya adalah sebagai berikut :

$$P = \frac{S}{N} \times 100 \%$$

Keterangan :

P : Persentase level konsepsi peserta didik

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

N : Jumlah peserta didik yang mengalami miskonsepsi