

BAB III METODOLOGI PENELITIAN

3.1. Metode dan Desain Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah kuantitatif deskriptif dengan desain penelitian yaitu 3D+1I (*defining, designing, developing, dan implementing*). Desain pada penelitian ini, terdapat empat tahapan, yaitu Tahapan *defining*, mencari literatur penelitian mengenai miskonsepsi pada materi impuls. Tahapan *designing*, merancang instrument tes pada materi impuls dalam bentuk *two-tier open-ended test*. Tahapan *developing*, mengembangkan instrument tes *two-tier open-ended* menjadi instrument tes *four-tier open-ended* melalui proses studi pendahuluan ke siswa dan judgment oleh ahli. Tahap implementasi, menggunakan instrument tes *four-tier open-ended* untuk diberikan kepada siswa.

3.2. Populasi dan Sampel

Populasi penelitian ini adalah siswa kelas X di SMA PGRI 1 Bandung yang diduga memiliki miskonsepsi berdasarkan hasil observasi. Teknik pengambilan sampel yang digunakan adalah *sampling purposive*, yaitu teknik penentuan sampel dengan pertimbangan tertentu. Teknik ini merupakan jenis teknik pengambilan sampel *nonprobability sampling*, yaitu teknik pengambilan sampel yang tidak memberi peluang/kesempatan sama bagi setiap unsur atau anggota populasi untuk dipilih menjadi sampel (Sugiyono, 2012: 84-85). Sampel penelitian ini berjumlah 63 siswa (40 perempuan dan 23 laki-laki) kelas X (MIPA 1 dan MIPA 2).

3.3. Instrumen Penelitian

Instrumen yang digunakan pada penelitian ini adalah berupa soal pilihan ganda bertingkat berbentuk *four-tier-test*. Instrumen tes berbentuk *four-tier test* dapat digunakan untuk mengetahui dan menganalisis miskonsepsi serta benar-benar mendeteksi miskonsepsi tanpa kesalahan dan kurangnya informasi (Gurel, 2015). Berdasarkan literature *four-tier tes* memiliki empat tingkatan, diantaranya 1) Tingkatan pertama berupa pertanyaan soal dan pilihan jawaban, 2) Tingkatan kedua berupa tingkat keyakinan atas jawaban yang dipilih, 3) Tingkatan ketiga berupa

pilihan alasan atas jawaban yang dipilih, 4) Tingkatan keempat berupa tingkat keyakinan atas alasan yang dipilih. Instrumen yang digunakan merujuk pada Caleon. Instrumen *four-tier* yang telah dikembangkan kemudian disebar melalui laman *google form*. *Google form* dipilih sebagai media penyebaran instrument dikarenakan mudah diakses, mudah untuk dimengerti cara penggunaannya, tampilannya mudah dimengerti, waktu menjadi lebih efisien, menghemat penggunaan kertas dan cocok untuk digunakan dimasa pandemik yang menuntun serba Online terutama di bidang Pendidikan (Batubara, 2016). Gambaran instrument *four-tier open-ended* dapat dilihat pada Tabel 3.1.

Tabel 3.1
Gambaran instrument *four-tier open-ended*

<p>7.1. Sebuah bola Softball bermassa 85 gram dilempar mendatar ke kanan dengan kelajuan 20 m/s. Setelah dipukul, bola bergerak ke kiri dengan kelajuan 20 m/s. Berapakah impulsnya yang diberikan oleh pemukul pada bola ...</p> <p>a. - 0,6 N.s b. - 1,7 N.s c. - 2,1 N.s d. - 3,4 N.s e.</p>												
<p>7.2. Tingkat Keyakinan untuk jawaban yang Anda pilih</p> <table style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td><u>Hanya menebak</u></td> <td><u>Sangat Tidak Yakin</u></td> <td><u>Tidak Yakin</u></td> <td><u>Yakin</u></td> <td><u>Sangat Yakin</u></td> <td><u>Benar Yakin</u></td> </tr> </table>	1	2	3	4	5	6	<u>Hanya menebak</u>	<u>Sangat Tidak Yakin</u>	<u>Tidak Yakin</u>	<u>Yakin</u>	<u>Sangat Yakin</u>	<u>Benar Yakin</u>
1	2	3	4	5	6							
<u>Hanya menebak</u>	<u>Sangat Tidak Yakin</u>	<u>Tidak Yakin</u>	<u>Yakin</u>	<u>Sangat Yakin</u>	<u>Benar Yakin</u>							
<p>7.3. Alasan yang tepat atas jawaban anda adalah...</p> <p>a. Impulsnya sangat besar, karena momentum yang diberikan tidak seimbang dengan impuls b. Kecepatannya berbanding terbalik dari awalnya positif menjadi negative c. Bola mendapatkan gaya impulsive dari pemain pada selang waktu yang sangat singkat, benda tersebut akan mengalami perubahan momentum d.</p>												
<p>7.4. Tingkat Keyakinan untuk alasan yang Anda pilih</p> <table style="width: 100%; text-align: center;"> <tr> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> </tr> <tr> <td><u>Hanya menebak</u></td> <td><u>Sangat Tidak Yakin</u></td> <td><u>Tidak Yakin</u></td> <td><u>Yakin</u></td> <td><u>Sangat Yakin</u></td> <td><u>Benar Yakin</u></td> </tr> </table>	1	2	3	4	5	6	<u>Hanya menebak</u>	<u>Sangat Tidak Yakin</u>	<u>Tidak Yakin</u>	<u>Yakin</u>	<u>Sangat Yakin</u>	<u>Benar Yakin</u>
1	2	3	4	5	6							
<u>Hanya menebak</u>	<u>Sangat Tidak Yakin</u>	<u>Tidak Yakin</u>	<u>Yakin</u>	<u>Sangat Yakin</u>	<u>Benar Yakin</u>							

3.4. Prosedur Penelitian

Pada penelitian ini, prosedur penelitian terdiri dari tiga tahap, yaitu tahap perencanaan, tahap pelaksanaan, dan tahap akhir. Ketiga tahap tersebut dapat dijabarkan sebagai berikut

1. Tahap Perencanaan

- a. Melakukan Studi Pendahuluan
 - 1) Studi literasi mengenai penelitian sebelumnya untuk mengetahui hasil miskonsepsi pada materi impuls yang telah ditemukan serta bagian penelitian yang dikembangkan dan observasi ke sekolah.
 - 2) Merumuskan masalah penelitian.
 - 3) Membuat kisi-kisi instrument dan instrument tes diagnostic Impuls dalam format *two-tier test* (Lampiran 1).
 - 4) Studi lapangan yaitu uji instrumen untuk pengembangan dari *two-tier open-ended test* menjadi *four-tier open-ended test* berdasarkan data miskonsepsi siswa menggunakan *google form* (Lampiran 2).
- b. Mengubah bentuk tes diagnostic Impuls dari *two-tier test* menjadi *four-tier test*.
- c. Melakukan *judgement* instrument dalam format *four-tier test* kepada 5 dosen dan 2 guru (Lampiran 4).
- d. Merevisi atau memperbaiki instrument penelitian dan kisi-kisi berdasarkan hasil *judgement* (Lampiran 5).
- e. Menentukan populasi dan sampel penelitian.
- f. Menyiapkan formulir berbasis web menggunakan *google form* (Lampiran 6).

2. Tahap Pelaksanaan

- a. Menyebarkan link *google form* instrument *four-tier test* kepada siswa
- b. Melakukan pengumpulan data dengan memberikan batasan waktu siswa untuk mengerjakan instrument *four-tier test* selama tiga hari
- c. Mengolah data hasil temuan menggunakan aplikasi MINISTEP 4.5.3 untuk kemudian dianalisis kualitas instrumen *four-tier* dan butir soal

- d. Menganalisis kualitas instrumen *four-tier* dan butir soal menggunakan analisis *Rasch Model*

3. Tahap Akhir

- a. Mengolah data hasil temuan menggunakan aplikasi MINISTEP 4.5.3 dengan *Table 1 Variable Map (Wright-Map)*, *Table 13. Item Measure*, *Table 17. Person Measure*
- b. Menganalisis hasil penelitian menggunakan analisis Rasch Model
- c. Memberi kesimpulan hasil penelitian secara kuantitatif dan diperkuat oleh teori yang mendukung
- d. Menyusun laporan hasil penelitian

3.5. Analisis Instrumen Penelitian

3.5.1. Validitas Instrumen *Four-tier Test*

Validitas adalah suatu ukuran yang menunjukkan tingkat-tingkat kevalidan atau keshahihan sesuatu instrument (Arikunto, 2010). Tahap validitas dilakukan dengan dua cara, yaitu validitas konstruk dan validitas isi. Untuk menguji validitas konstruk dapat digunakan pendapat para ahli (*judgment experts*), setelah instrument dikonstruksi tentang aspek-aspek yang hendak diukur berlandaskan teori maka selanjutnya dikonsultasikan dengan para ahli (Arikunto, 2010). Para ahli akan memberi keputusan apakah instrument tersebut dapat digunakan tanpa perbaikan, ada perbaikan, atau dirombak total. Pada penelitian ini menggunakan pendapat dari tujuh judgment ahli/validator, diantaranya ialah 5 dosen ahli dan 2 guru fisika. Proses validasi oleh para ahli dilakukan hingga instrumen *four-tier* dinyatakan layak untuk dijadikan sebagai instrumen penelitian.

Setelah proses revisi selesai langkah selajutnya ialah menghitung validitas isi, untuk mendukung hasil validasi para ahli dilakukan uji validitas yang digunakan ialah validitas isi melalui pendekatan yang diusulkan oleh Lawshe (1975). Validitas isi merupakan modal dasar dalam suatu instrumen penelitian, sebab kesahihah/validitas isi akan menyatakan keterwakilan aspek yang diukur dalam instrumen, dan dipandang dari segi isi instrumen yang diberikan (Yusoff., M, 2013). Pendekatan validitas isi yang dipilih berdasarkan usulan Lawshe (1975) menggunakan rasio validitas isi (CVR) dengan mengukur derajat kesepakatan

para ahli dari satu item dan yang dapat mengekspresikan tingkat validitas isi melalui indikator tunggal yang berkisar dari -1 sampai 1. Hal ini relevan dengan yang dilakukan oleh Setiap penilai / *subject matter experts* (SME) yang terdiri dari panel juri (para ahli) menjawab pertanyaan untuk setiap *item* dengan tiga pilihan jawaban yaitu (1) valid tanpa revisi, (2) valid dengan revisi, (3) tidak valid. Menurut Lawshe, jika lebih dari setengah panelis menunjukkan bahwa item penting/esensial, maka *item* tersebut memiliki setidaknya validitas isi. Formula yang diajukan oleh Lawshe:

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

Keterangan:

n_e = Jumlah panelis yang menjawab valid

N = Jumlah panelis yang memvalidasi

Formula ini menghasilkan nilai-nilai yang berkisar dari +1 sampai -1, nilai positif menunjukkan bahwa setidaknya setengah panelis (SME) menilai item sebagai penting/esensial. Semakin lebih besar CVR dari 0, maka semakin “penting” dan semakin tinggi validitas isinya.

Butir soal yang akan di validasi oleh ahli dinyatakan diterima apabila nilai CVR nya sama atau lebih besar dari nilai kritis. Sedangkan butir item akan ditolak atau diperbaiki apabila nilai CVR item tersebut kurang dari nilai kritis berdasarkan jumlah panelis dan taraf signifikan yang diambil adalah 0,05 seperti Tabel 3.2 berikut!

Tabel 3.2

Nilai Kritis CVR (*one-tailed*)

Jumlah Panelis (N)	Nilai Kritis CVR			
	0,1	0,05	0,025	0,01
5	0,573	0,736	0,877	0,99
6	0,523	0,672	0,800	0,950
7	0,485	0,622	0,741	0,879
8	0,453	0,582	0,693	0,822
9	0,427	0,548	0,653	0,775

10	0,405	0,520	0,620	0,736
----	-------	-------	-------	-------

(Wilson, 2012)

Kemudian setelah seluruh butir soal pada instrument *four-tier* sudah dapat digunakan berdasarkan hasil judgment, maka instrumen diberikan pada siswa. Hasil pengerjaan siswa di uji validitas sebagai bukti bahwa instrument *four-tier* benar-benar mampu mengukur sesuatu yang harus diukur. Uji validitas instrument *four-tier* menggunakan *software* MINISTEP 4.5.3 dengan *output* Table 10 item (*column*): *fit order* yang dianalisis dengan analisis Rasch. Pengukuran uji validitas instrument dilihat berdasarkan nilai *logarithm odd unit (logit)* pada *outfit mean square* (MNSQ), *outfit Z-standard* (ZSTD), dan *point-measure correlation* (PT-MEASURE CORR). Akan tetapi beberapa penelitian mengatakan bahwa *point-measure correlation* hanya digunakan untuk mengetahui daya pembeda dari suatu instrumen (Sabudin, Mansor, Meerah, & Muhammad, 2018; Smiley, 2015). Maka analisis uji validitas pada penelitian ini hanya menggunakan nilai *logarithm odd unit (logit)* pada *outfit mean square* (MNSQ) dan *outfit Z-standard* (ZSTD). Nilai logit merupakan nilai yang dihasilkan melalui perhitungan fungsi logaritma pada *software* MINISTEP dan dengan menggunakan fungsi *logit* ini akan didapatkan mistar pengukuran dengan interval yang sama (Sumintono & Widhiarso, 2015).

Selanjutnya, Nilai *OUTFIT* (MNSQ) dan (ZSTD) dimasukan pada kriteria nilai yang diterima, untuk memeriksa kesesuaian butir soal. Kriteria tersebut dapat dilihat pada tabel 3.3

Tabel 3.3

Kriteria nilai MNSQ dan ZSTD

Outfit	Nilai yang diterima
MNSQ	$0,5 < \text{MNSQ} < 1,5$
ZSTD	$-2,0 < \text{ZSTD} < +0,2$

3.5.2. Realibilitas Butir Soal Instrumen *Four-tier Test*

Suatu tes dikatakan mempunyai tingkat kepercayaan yang tinggi jika tes tersebut menunjukkan hasil yang tetap atau konsisten pada kondisi bagaimanapun (bervariasi). Reliabilitas/tingkat kepercayaan menggunakan Uji validitas instrument *four-tier* menggunakan *software* MINISTEP 4.5.3 dengan *output*

Table 3.1 Summary Statistics dianalisis dengan Rasch Model. Uji reliabilitas pada Rasch memiliki cara penafsiran berbeda dengan reliabilitas dalam teori klasik yang memiliki harga tunggal, sedangkan uji reliabilitas pada Rasch mampu menghasilkan pengukuran berbeda antara individu berkemampuan tinggi dan individu berkemampuan rendah (Sumintono & Widhiarso, 2015). Analisis Rasch dapat menampilkan beberapa nilai realibilitas, diantaranya adalah *person reability*, *item reability*, dan *Cronbach alpha*. *Person reability* menunjukkan konsistensi jawaban peserta didik, *item reability* menunjukkan kualitas item tes, sementara *Cronbach alpha* menunjukkan nilai interaksi Antara person dan item dari instrument *four-tier* secara keseluruhan.

Interpretasi dari nilai *person reability* dan *item reability* dapat dilihat pada Tabel 3.4

Tabel 3.4

Interpretasi nilai person reability dan item reability,

Nilai <i>item reability</i>	Interpretasi
Nilai > 0,94	Istimewa
$0,8 \leq \text{Nilai} \leq 0,90$	Bagus

Nilai <i>person reability</i>	Interpretasi
$0,91 \leq \text{nilai} \leq 0,94$	Bagus
$0,67 \leq \text{nilai} \leq 0,80$	Cukup
Nilai < 0,67	Lemah

(Sumintono & Widhiarso, 2015).

Sedangkan interpretasi nilai *Cronbach alpha* dapat dilihat pada Tabel 3.5

Tabel 3.5

Interpretasi nilai Cronbach alpha

Nilai <i>person reability</i>	Interpretasi
$0,8 > \alpha$	Bagus Sekali
$0,7 \leq \alpha \leq 0,8$	Bagus
$0,6 \leq \alpha \leq 0,7$	Cukup
$0,6 \leq \alpha \leq 0,7$	Jelek

$\alpha < 0,5$	Buruk
----------------	-------

(Sumintono & Widhiarso, 2015).

3.5.3. Tingkat Kesukaran Soal (*Diffulty Index*)

Perhitungan tingkat kesukaran soal digunakan untuk mengukur seberapa besar derajat kesukaran setiap butir soal, jika suatu soal memiliki tingkat kesukaran seimbang, maka dikatakan bahwa soal tersebut baik (Arifin, 2012). Tingkat Kesukaran dapat dianalisis dengan Rasch menggunakan aplikasi MINISTEP 4.5.3 dengan output *Table 1 Variable (Wright maps)* dan output *Table 13 Item Measure* untuk mendapat gambaran mengenai kekuatan butir soal. Sumintono dan Widhiarso (2015) mengungkapkan bahwa penggunaan *Variable (Wright maps)* mampu menghasilkan suatu peta yang menggambarkan sebaran kemampuan siswa/responden dan sebaran tingkat kesulitan soal dengan skala yang sama. Sedangkan *output Table 13 Item Measure* digunakan untuk mengklasifikasikan tingkat kesulitan tiap butir soal berdasarkan nilai logit dan nilai Standar Deviasi yang didapat (Sumintono dan Widhiarso, 2015). Nilai *logit* terbesar menunjukkan tingkat soal yang tinggi, hal ini berkorespondensi dengan kolom *total score*, yaitu menyatakan berapa jumlah jawaban benar. Adapun pada *item measure* terdapat informasi nilai standar deviasi 0,83. Jika nilai ini dikombinasikan dengan rata-rata *logit* maka dapat dikelompok berdasarkan tingkat kesulitannya. $0,0 \text{ logit} + 1\text{SD}$ adalah satu kelompok soal sulit, lebih besar dari $+1\text{SD}$ adalah soal yang sangat sukar; $0,0 \text{ logit} - 1\text{SD}$ adalah satu kelompok soal mudah, lebih besar dari -1SD adalah soal yang sangat mudah (Sumintono & Widhiarso, 2015). Artinya terdapat empat kategori tingkat kesukaran yang dikembangkan oleh Sumintono dan Widhiarso (2015) dan dapat dilihat pada Tabel 3.6.

Tabel 3.6

Interpretasi hasil tingkat kesukaran

Tingkat Kesukaran (TK)	Keterangan
$0,86 < \text{TK}$	Sangat Sukar
$0,00 < \text{TK} \leq 0,86$	Sukar
$-0,86 < \text{TK} \leq 0,00$	Mudah

$$TK \leq -0,86$$

Sangat Mudah

(Sumintono & Widhiarso, 2015).

3.5.4. Daya Pembeda Soal

Perhitungan daya pembeda digunakan untuk mengukur sejauhmana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu (Arifin,Z, 2012). Semakin tinggi indeks diskriminasi daya pembeda maka termasuk soal yang sangat baik, ini berarti bahwa soal tersebut dapat membedakan Antara peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi (Arifin,Z 2012). Untuk menghitung daya pembeda setiap butir soal dapat menggunakan analisis Rasch Model dengan aplikasi WINSTEP 4.5.3 dengan *output Table 10 Item Fit Order* yang dianalisis dengan analisis Rasch. Penggunaan *Item Fit Order* karena didalamnya menampilkan informasi mengenai *Point-measure correlation*. Seperti yang telah dibahas pada bagian validitas bahwa menurut beberapa penelitian, *Point-measure correlation* mampu mengetahui daya pembeda dari suatu instrument. Tabel 3.7 menampilkan interpretasi untuk setiap nilai yang diberikan pada PTMEASURE-AL COOR, tabel ini dikembangkan oleh Smiley (2015).

Tabel 3.7

Interpretasi nilai PTMEASURE-AL COOR

PTMEASURE-AL COOR (ID)	Interpretasi
$0,40 < ID$	Sangat Baik
$0,30 < ID \leq 0,40$	Baik
$0,20 < ID \leq 0,29$	Kurang Baik
$ID < 0,19$	Jelek

(Smiley 2015)

Menurut Smiley (2015) apabila hasilnya bernilai negatif maka itu lebih bermasalah dari nilai nol dan mungkin menunjukkan bahwa item tersebut cacat dalam beberapa hal mendasar dan harus diperiksa apakah dari segi kunci jawaban yang salah, direvisi, atau mungkin dihapus dari tes.

3.6. Analisis Data Temuan Miskonsepsi Siswa

Berdasarkan penelitian Aminudin, dkk (2019) proses analisis data kuantitatif dilakukan dalam dua tahap. Tahap pertama menganalisis data proses siswa menjawab lembar tes *four-tier* melalui kategori konsepsi. Tahap kedua menganalisis data skor konsepsi, skor miskonsepsi, dan skor keyakinan, kemudian membandingkan ketiganya menggunakan *Rasch Analysis*.

3.6.1. Kategori Pemahaman Konsep Impuls Siswa

Gurel (2012) memberikan sebuah cara untuk meidentifikasi kategori pemahaman konsep siswa dan miskonsepsi siswa menggunakan tabel kriteria konsep pada kombinasi jawaban *four-tier test* siswa.

Tabel 3.8
Kategori Konsepsi Versi Gurel, dkk

Conception criteria	Tier I	Tier II	Tier III	Tier IV
Misconception (M)	Wrong	Sure	Wrong	Sure
Understanding (U)	Correct	Sure	Correct	Sure
Partial Understanding (PU)	Correct	Sure	Correct	Not sure
	Correct	Not sure	Correct	Sure
	Correct	Not sure	Correct	Not sure
	Correct	Sure	Wrong	Sur
	Correct	Sure	Wrong	Not sure
	Correct	Not sure	Wrong	Sure
	Correct	Not sure	Wrong	Not sure
	Wrong	Sure	Correct	Sure
	Wrong	Sure	Correct	Not sure
	Wrong	Not sure	Correct	Sure
No Understanding (NU)	Wrong	Sure	Wrong	Not sure
	Wrong	Not sure	Wrong	Sure
	Wrong	Not sure	Wrong	Not sure
Uncodable (UC)	Respondent does not fulfill (response) all or part of tiers in instrument test items.			

(Gurel dkk, 2012)

Namun, Aminudin, dkk (2019) mengembangkan tabel Gurel diatas, dengan membagi level pemahaman siswa yang parsial menjadi *partial positive* dan *partial negative*. Positif parsial (*partial positive*) menunjukkan jawaban benar pada tier 1 dan 3 tapi tidak yakin pada tier 2 atau 4. Dan mengganti pemahaman parsial dengan bagian miskonsepsi spesifik menjadi negatif parsial (*partial negative*) yang menunjukkan jawaban benar hanya pada tier 1 atau 3 dengan variasi level keyakinan. Kategori konsepsi ditunjukkan pada tabel 3.9.

Tabel. 3.9
Pengembangan Kategori Konsepsi Versi Aminudin, dkk

Tier	Kategori																	
	SU		PP				PN						MC	NU				NC
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
1	C	C	C	C	C	C	C	C	C	I	I	I	I	I	I	I	IA	
2	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS		
3	C	C	C	C	I	I	I	I	C	C	C	C	I	I	I	I		
4	S	S	NS	NS	S	S	NS	NS	S	S	NS	NS	S	S	NS	NS		

*SU: Sound Understanding, PP: Partial Positive, PN: Partial Negative, MC: Misconception, NU: No Understanding, NC: No Coding, C: Correct, I: Incorrect, S: Sure, NS: Not Sure, IA: Incomplete Answer

(Aminudin, dkk, 2019)

3.6.2. Analisis Miskonsepsi Siswa Menggunakan Rasch Model

Tahap kedua menganalisis data konsepsi, miskonsepsi, dan tingkat keyakinan, kemudian membandingkan ketiganya menggunakan *Rasch Analysis* (Aminudin, 2019). Berdasarkan penelitian Gurel (2017) menggunakan tabel 3.8 menganalisis miskonsepsi siswa disetiap tier dan disetiap item. Disajikan tabel dibawah ini untuk mengkodekan skor siswa yang menjawab benar dan skor siswa yang miskonsepsi.

Tabel. 3.10
Penilaian Kategori Konsepsi Siswa

Kategori	Coding	Deskripsi
Correct Scores	Hanya Tier 1 (COR1)	Skor "1" jika jawaban siswa benar pada tier pertama dan jika salah mendapatkan skor "0"

	Hanya Tier 3 (COR3)	Skor “1” jika jawaban siswa benar pada tier ketiga dan jika salah mendapatkan skor “0”
	Tier 1 dan Tier 3 (COR1&COR3)	Skor “1” jika jawaban siswa benar pada tier pertama & tier ketiga dan jika salah mendapatkan skor “0”
	Seluruh Tier 1-4 (COR1-4)	Skor “1” jika jawaban siswa benar pada seluruh tier 1-4 dan jika salah mendapatkan skor “0”
Misconception Scores	Hanya Tier 1 (MISC1)	Skor “1” jika jawaban siswa miskonsepsi pada tier pertama dan jika tidak mendapatkan skor “0”
	Hanya Tier 3 (MISC3)	Skor “1” jika jawaban siswa miskonsepsi pada tier ketiga dan jika tidak mendapatkan skor “0”
	Tier 1 dan Tier 3 (MISC1&MISC3)	Skor “1” jika jawaban siswa miskonsepsi pada tier 1 & 3 dan jika tidak mendapatkan skor “0”
	Seluruh Tier 1-4 (MISC1-4)	Skor “1” jika jawaban siswa miskonsepsi pada seluruh tier 1-4 dan jika tidak mendapatkan skor “0”
Confidence Scores	Hanya Tier 1 (CONF1)	Skor “1” jika jawaban siswa pada tier pertama dan jika tidak mendapatkan skor “0”
	Hanya Tier 3 (CONF3)	Skor “1” jika jawaban siswa miskonsepsi pada tier ketiga dan jika tidak mendapatkan skor “0”
	Tier 1 dan Tier 3 (CONF1&CONF3)	Skor “1” jika jawaban siswa miskonsepsi pada tier 1 & 3 dan jika tidak mendapatkan skor “0”

(Gurel dkk, 2017)

Skor jawaban yang benar dihitung berdasarkan kunci jawaban yang telah disediakan. Pada tier pertama, tier ketiga, tier pertama dan ketika diolah dengan menjumlahkan kolom jawaban yang benar dan baris jawaban benar kemudian dibagi oleh jumlah seluruh jumlah respon (Gurel, 2012).

Skor miskonsepsi dihitung berdasarkan item miskonsepsi yang telah disediakan. Pada setiap kategori tier (tier pertama, tier ketiga, tier pertama dan ketiga dan keempat tier) setiap pilihan subjek yang mengindikasikan miskonsepsi dikodekan sebagai 1, dan sebaliknya dikodekan sebagai 0. Kemudian diolah dengan menjumlahkan kolom skor miskonsepsi total masing-

masing subjek berdasarkan hanya pada setiap kategori tier dari tes, sedangkan menjumlahkan baris memberikan jumlah dari tanggapan kesalahpahaman hanya diberikan pada setiap kategori tier oleh semua subjek. Membagi yang terakhir dengan jumlah siswa memberikan proporsi tanggapan kesalahpahaman untuk setiap kategori tier dari setiap item (tingkat kesulitan tingkat pertama) (Gurel, 2012).

Skor Tingkat Keyakinan

Kategori keyakinan (keyakinan Pertama, keyakinan kedua, dan kedua-duanya) skor ini diperoleh dengan menggunakan respons setiap subjek hanya pada tingkat kedua, yang mempertanyakan apakah responden yakin tentang tanggapan mereka terhadap tingkat pertama. Itu diberi kode 1 untuk siswa yang memilih “*Yakin 100%*” atau “*Sangat Yakin*” atau “*Yakin*” sebagai tanggapan, dan 0 untuk siswa yang memilih “*Hanya Menebak*” atau “*Sangat Tidak Yakin*” atau “*Tidak Yakin*” sebagai tanggapan. Menjumlahkan kolom skor kepercayaan diri siswa pada keyakinan pertama, kedua, dan kedua-duanya Untuk menyimpulkannya, ada dua jenis dasar variabel-kuantitatif dan kategorikal. Masing-masing menggunakan jenis analisis dan pengukuran yang berbeda, membutuhkan jenis skala pengukuran yang berbeda: nominal, ordinal, interval, dan rasio (Fraenkel & Wallen, 2000; Gurel dkk, 2012). Dalam hal skala pengukuran, variabel yang dijelaskan di seluruh bagian variabel adalah skala interval dan variabelnya adalah kuantitatif.

Aminudin, dkk (2019) mengembangkan skor yang telah disajikan oleh Gurel, skor ini terdiri dari tiga bagian, yaitu skor untuk pemahaman siswa, skor untuk miskonsepsi, dan skor untuk level keyakinan. Setiap skor dari masing-masing akan didistribusikan pada kategori konsepsi siswa, hasilnya dapat dilihat pada Tabel 3.11

Tabel. 3.11
Skor Kategori Konsepsi Siswa

Kategori	Skor		
	Konsepsi	Miskonsepsi	Keyakinan
Paham Konsep (SU)	4	0	-
Partial Positif (PP)	3	0	-

Partial Negatif (PN)	1	1	-
Tidak Paham Konsep (NU)	0	3	-
Miskonsepsi (MC)	0	4	-
Tidak Terkode (NC)	kosong	kosong	kosong
Yakin (C)	-	-	3
Yakin sebagian (PC)	-	-	1
Tidak Yakin (NC*)	-	-	0

Tahapan selanjutnya, menganalisis data skor konsepsi, skor miskonsepsi, dan skor keyakinan, kemudian membandingkan ketiganya menggunakan *Rasch Analysis*. Perhitungan skor pemahaman, skor miskonsepsi, dan skor keyakinan menggunakan penilaian yang telah dikembangkan oleh Aminudin, dkk. Tier tingkat keyakinan disesuaikan pada empat kategori, *Confidence (C)*, *Partial Confidence (PC)*, dan *No Confidence (NC*)* dan *No Coding (NC)* dan skor kosong untuk kategori *No Coding (NC)*. Hasil penilaian dianalisis dengan Rasch Model menggunakan aplikasi MINISTEP 4.5.3, dengan *Table 13 Item-Measure* dan *Table 17 Person Measure* kemudian dilihat peta sebaran datanya menggunakan *Variable (Wright maps)*.

Tahap terakhir setelah proses membandingkan skor konsepsi, miskonsepsi, dan skor keyakinan selanjutnya menganalisis miskonsepsi siswa setiap konsep yang terdapat dalam butir soal dan disesuaikan dengan hasil penelitian sebelumnya.

