

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Penelitian ini merupakan penelitian pengembangan dengan model pengembangan yang digunakan dalam penelitian ini menggunakan pendekatan *Research and Development (R&D)*. R&D merupakan metode penelitian yang digunakan untuk mengembangkan atau memvalidasi produk-produk yang digunakan dalam pendidikan dan pembelajaran (Sugiyono, 2020). Data akan dikumpulkan melalui Google Forms kepada peserta didik. Data yang diperoleh akan dianalisis menggunakan model Rasch, yang memberikan gambaran tentang tingkat pemahaman peserta didik serta mengidentifikasi perbedaan antara peserta didik laki-laki dan perempuan. Pada penelitian ini tidak diberikan intervensi kepada peserta didik, sehingga pengambilan data pada peserta didik menjadi efektif.

Penelitian ini termasuk ke dalam kategori penelitian pengembangan karena penelitian ini menghasilkan sebuah produk instrumen tes. Model pengembangan merupakan dasar untuk mengembangkan produk yang akan dihasilkan. Model pengembangan yang digunakan dalam penelitian ini adalah bentuk 4D yaitu *Define, Design, Development, dan Disseminate*. Menurut Endang M. dalam (Rokhimawan, 2016), tahapan dalam prosedur pengembangan penelitian adalah :

1. *Define* (pendefinisian) pada tahap ini dilakukan untuk menganalisis kurikulum, menganalisis karakteristik pada peserta didik, menganalisis materi, serta merumuskan tujuan.
2. *Design* (perancangan) pada tahap ini dilakukan perancangan yang meliputi ; pengumpulan referensi materi, pembuatan kisi-kisi materi serta pembuatan butir soal dengan bentuk *five-tier-test*.
3. *Development* (pengembangan) pada tahap ini instrumen *five tier test open ended* dikembangkan menjadi *five tier test close ended* berdasarkan hasil

studi pendahuluan alasan siswa pada *tier* ketiga. Selanjutnya instrumen dilakukan tahap uji validasi pada guru, dosen, dan peserta didik. Hasil uji validasi kemudian digunakan sebagai revisi sehingga butir soal yang dikembangkan benar-benar layak untuk digunakan.

4. *Disseminate* (penyebarluasan) pada tahap ini instrumen yang dinyatakan layak akan disebar dan diujicobakan kepada peserta didik dalam jumlah besar untuk memperoleh respon.

3.2 Partisipan Penelitian

Partisipan yang digunakan dalam penelitian ini yaitu peserta didik kelas XI peminatan MIPA yang telah mempelajari materi kalor dan termodinamika di SMA negeri di Kabupaten Karawang. SMA Negeri yang digunakan pada saat pengambilan data studi pendahuluan berbeda dengan sekolah yang digunakan saat uji coba instrumen skala besar. Partisipan saat pengambilan data studi pendahuluan terdiri dari 10 siswa laki-laki dan 15 siswa perempuan dengan rentang usia 15 – 17 tahun. Partisipan diambil dari salah satu kelas yang sudah mendapatkan pembelajaran mengenai topik kalor dan termodinamika.

3.3 Populasi dan Sampel

1. Populasi

Populasi merupakan sekumpulan obyek yang akan diteliti. Populasi merupakan sebuah obyek/subyek yang memiliki kualitas dan karakteristik tertentu yang ditetapkan oleh seorang peneliti untuk dapat dipelajari dan ditarik kesimpulannya. Populasi bukan sekedar jumlah yang ada pada obyek/subyek yang dipelajari, tetapi meliputi seluruh karakteristik/sifat yang dimiliki oleh subyek atau obyek yang diteliti itu (Sugiyono, Statistika Untuk Penelitian, 2019). Pada penelitian ini, populasi yang digunakan yaitu siswa kelas XI MIPA SMA Negeri dari tiga SMA Negeri yang berada di Kabupaten Karawang.

2. Sampel

Sampel adalah sebagian atau wakil populasi yang diteliti (Arikunto dalam (Hendri, 2015)). Menurut Roscoe dalam Sugiyono (2019) menyatakan bahwa ukuran sampel yang layak dalam penelitian adalah

antara 30 sampai dengan 500, selain itu bila sampel dibagi dalam kategori maka jumlah sampel setiap kategori minimal 30, bila dalam penelitian akan melakukan analisis dengan *multivariate* maka jumlah anggota sampel minimal 10 kali dari jumlah variabel yang diteliti.

Untuk menentukan sampel dalam penelitian, terdapat berbagai teknik sampling yang digunakan. Penentuan sampel yang digunakan dalam penelitian ini dilakukan dengan teknik *purposive sampling*. Menurut Sugiyono (2019), *purposive sampling* merupakan teknik penentuan sampel dengan pertimbangan tertentu. Dalam penelitian ini, pertimbangan dalam pengambilan sampel yaitu siswa yang sudah mempelajari materi kalor dan termodinamika dari ketiga SMA yang berbeda. Penentuan jumlah sampel digunakan tabel penentuan jumlah sampel dengan interval kepercayaan (*confidence interval*) 90% atau taraf kesalahan 10% yang dikembangkan oleh *Isaac* dan *Michael* (dalam Sugiyono, 2019).

Populasi siswa kelas XI SMA Negeri dari tiga SMA Negeri yang ada di Kabupaten Karawang berjumlah 700 siswa, maka jumlah sampel minimal berdasarkan tabel penentuan jumlah sampel yang dikembangkan oleh *Isaac* dan *Michael* dengan taraf kesalahan 10% adalah $N_T = 195$ siswa. Untuk menentukan jumlah sampel minimal yang dibutuhkan dari masing-masing sekolah sebagai berikut :

$$N_s = \frac{N}{\sum N} N_T$$

N_s = Jumlah sampel siswa yang dibutuhkan

N = Jumlah siswa pada tiap jenis sekolah

$\sum N$ = Jumlah seluruh siswa

N_T = Jumlah sampel berdasarkan tabel

Jumlah sampel siswa yang diambil dari masing-masing SMA Negeri ditampilkan pada Tabel 3.1 berikut.

Tabel 3.1 Jumlah Sampel pada Masing-masing SMA

Jenis Sekolah	Jumlah Siswa IPA Kelas XI (N)	Jumlah Sampel (Ns)
---------------	-------------------------------	--------------------

SMAN A	280	78
SMAN B	245	68
SMAN C	175	49
Total	700	195

Pada penelitian ini, terdapat dua proses pengambilan data yang terdiri dari studi pendahuluan dan pengambilan data utama untuk penelitian. Pada pelaksanaan studi pendahuluan, peneliti melibatkan sebanyak 25 siswa SMA kelas XI (15 siswa perempuan dan 10 siswa laki-laki) yang berasal dari salah satu SMA Negeri di Kabupaten Karawang. Kemudian pada proses pengambilan data utama untuk mengidentifikasi miskonsepsi dan penyebab miskonsepsi, peneliti melibatkan sebanyak 195 siswa SMA kelas XI (105 siswa perempuan dan 90 siswa laki-laki) yang berasal dari ketiga SMA Negeri di Kabupaten Karawang.

3.4 Instrumen Penelitian

Instrumen yang digunakan dalam penelitian ini adalah instrumen tes diagnostik miskonsepsi berupa soal lima tingkat (*five-tier test*) serta lembar validasi instrumen. Berikut penjelasan dari instrumen yang akan digunakan.

3.4.1 Instrumen Tes Diagnostik Miskonsepsi

Tes yang digunakan dalam penelitian ini merupakan tes diagnostik tingkat lima atau *five-tier-test* pada materi kalor dan termodinamika sebanyak 15 soal. Dalam *Five-Tier-Test*, pada *tier* pertama berupa pertanyaan pilihan ganda dengan terdapat beberapa pilihan jawaban, *tier* kedua berupa tingkat keyakinan peserta didik dalam memilih jawaban pada *tier* pertama, *tier* ketiga berupa alasan terkait jawaban pada *tier* pertama, *tier* keempat berupa tingkat keyakinan peserta didik dalam memilih jawaban pada *tier* ketiga, serta *tier* kelima berupa sumber darimana siswa mendapatkan jawaban pada *tier* dan *tier* ketiga. Instrumen *five-tier test* yang dibuat sudah di uji coba terbatas kepada siswa kelas XI MIPA yang sudah mempelajari materi kalor dan termodinamika. Instrumen ini diberikan kepada siswa secara online melalui *Google Formulir* yang berisi tentang produk akhir pengembangan instrumen

five-tier-test yang telah melewati beberapa tahap validasi, dengan batas waktu yang ditentukan. Berikut desain instrumen *five-tier test* dapat dilihat pada Tabel 3.2.

Tabel 3.2 Desain Instrumen Five-Tier Test

Pertanyaan : Mendiagnosis konsepsi siswa
Jawaban Pilihan (Tier-1) A. Alternatif jawaban pertama B. Alternatif jawaban kedua C. Alternatif jawaban ketiga D. Alternatif jawaban keempat E. Alternatif jawaban kelima
Tingkat keyakinan (Tier-2) A. Alternatif jika siswa “Yakin” terhadap jawaban yang dipilih B. Alternatif jika siswa “Tidak Yakin” terhadap jawaban yang dipilih
Alasan (Tier-3) A. Alternatif jawaban pertama B. Alternatif jawaban kedua C. Alternatif jawaban ketiga D. Alternatif jawaban keempat E. Alternatif jawaban kelima
Tingkat keyakinan kedua (Tier-4) A. Alternatif jika siswa “Yakin” terhadap jawaban yang dipilih B. Alternatif jika siswa “Tidak Yakin” terhadap jawaban yang dipilih
Sumber pilihan jawaban pada tier-1 dan tier-3 (tier-5) A. Alternatif jawaban pertama B. Alternatif jawaban kedua C. Alternatif jawaban ketiga D. Alternatif jawaban keempat E. Alternatif jawaban kelima

Sebaran distribusi soal tentang Kalor dan Termodinamika yang telah dirancang pada instrumen *five tier test*, ditunjukkan pada Tabel 3.3.

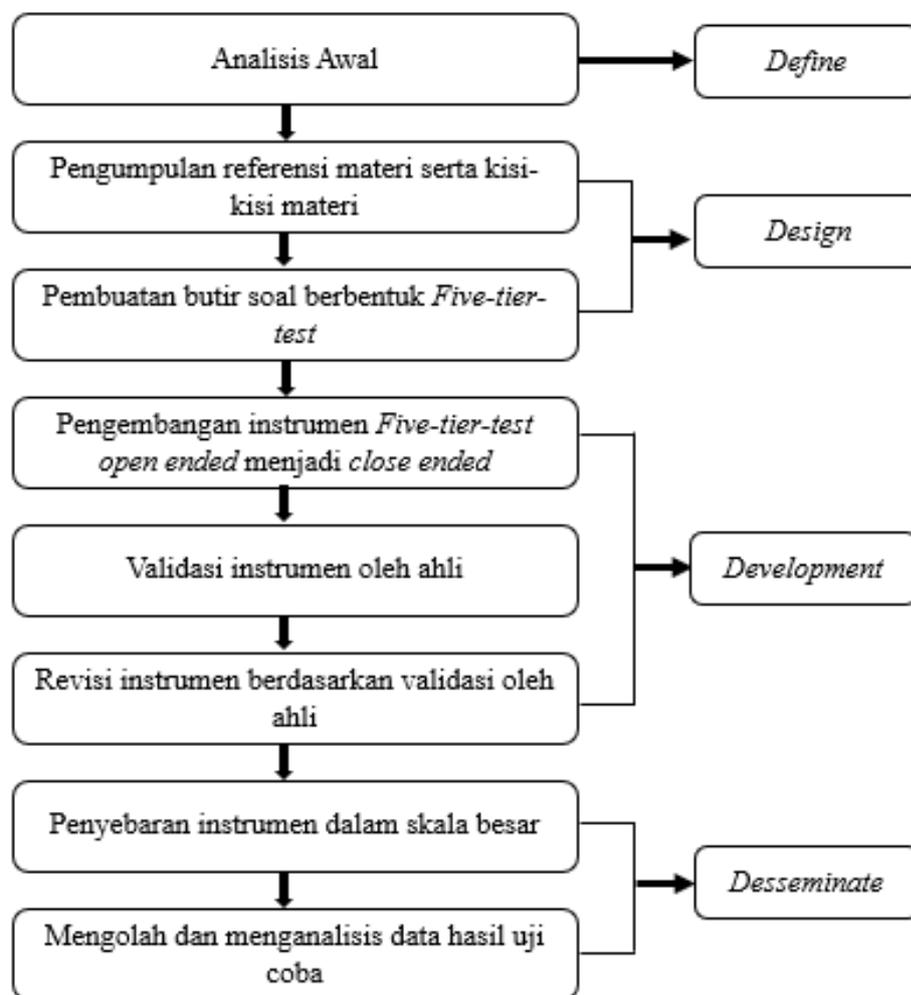
Tabel 3.3 Distribusi Soal Five-Tier Test

Sub Konsep	Jumlah Soal	Nomor Soal
Kalor Jenis	1	1
Perpindahan Kalor	3	2, 3, 4
Keseimbangan Termal	2	5, 6
Perubahan Wujud Zat	2	7, 8
Perubahan Suhu	2	9, 10
Hukum I Termodinamika	2	11, 12
Hukum II Termodinamika	1	13
Mesin Carnot	2	14, 15

3.5 Prosedur Penelitian

3.5.1 Model Pengembangan 4D

Prosedur dan pengembangan merupakan tahap penelitian yang akan dilakukan untuk pengembangan produk. Model pengembangan 4D terdiri dari empat langkah yang meliputi *define* (pendefinisian), *design* (perancangan), *development* (pengembangan), dan *dissemination* (penyebarluasan). Berdasarkan model pengembangan tersebut, berikut adalah bagan diagram alir prosedur pengembangan media pembelajaran:



Gambar 3.1 Prosedur Pengembangan 4D

Model pengembangan 4D terdiri dari 4 tahap yaitu sebagai berikut :

1. Tahap *Define* (Mendefinisikan)

Pada tahap *define*, penulis melakukan analisis awal mengenai miskonsepsi yang biasanya terjadi pada materi fisika. Berdasarkan analisis yang telah dilakukan, penulis mengidentifikasi bahwa materi yang sering terjadi miskonsepsi yakni pada materi kalor dan termodinamika yang merupakan konsep fisika yang abstrak namun erat kaitannya dengan kehidupan sehari-hari. Penulis juga melakukan analisis awal mengenai instrumen yang dapat digunakan untuk mengidentifikasi miskonsepsi pada siswa. Instrumen yang biasa digunakan untuk mengidentifikasi miskonsepsi yaitu tes diagnostik *one-tier*, *two-tier*, *three-tier*, dan *four-tier*. Dari keempat tes diagnostik tersebut, tidak dapat mengidentifikasi penyebab miskonsepsi yang terjadi pada siswa. Sedangkan untuk mengatasi miskonsepsi, lebih baik jika penyebab miskonsepsi juga diidentifikasi. Oleh karena itu, pada penelitian ini dilakukan pengembangan instrumen *Five-tier test* untuk mengidentifikasi miskonsepsi dan penyebab miskonsepsi yang terjadi pada siswa.

2. Tahap *Design* (Merancang)

Pada tahap *design*, dilakukan dengan penyusunan kisi-kisi instrumen penelitian sehingga butir soal tersebar secara merata untuk setiap konsep. Sebaran kisi-kisi tersebut terdiri dari miskonsepsi, konsepsi ilmiah, dan konstruksi soal. Penyusunan kisi-kisi instrumen penelitian berdasarkan pada informasi tentang miskonsepsi siswa yang diperoleh dari studi literatur. Kemudian disusunlah instrumen *five tier test* sesuai dengan sebaran kisi-kisi yang telah dibuat. Desain instrumen *five tier open ended test* pada materi kalor dan termodinamika dapat dilihat pada Tabel 3.4.

Tabel 3.4 Desain instrumen *five tier open ended test* pada materi kalor dan termodinamika

Pertanyaan :

Mendiagnosis konsepsi siswa

Jawaban Pilihan (*Tier-1*)

- A. Alternatif jawaban pertama
- B. Alternatif jawaban kedua
- C. Alternatif jawaban ketiga
- D. Alternatif jawaban keempat

E. Alternatif jawaban kelima

Tingkat keyakinan (*Tier-2*)

A. Alternatif jika siswa “Yakin” terhadap jawaban yang dipilih

B. Alternatif jika siswa “Tidak Yakin” terhadap jawaban yang dipilih

Alasan (*Tier-3*)

.....

.....

.....

.....

.....

Tingkat keyakinan kedua (*Tier-4*)

A. Alternatif jika siswa “Yakin” terhadap jawaban yang dipilih

B. Alternatif jika siswa “Tidak Yakin” terhadap jawaban yang dipilih

Sumber pilihan jawaban pada *tier-1* dan *tier-3* (*tier-5*)

A. Alternatif jawaban pertama

B. Alternatif jawaban kedua

C. Alternatif jawaban ketiga

D. Alternatif jawaban keempat

E. Alternatif jawaban kelima

3. Tahap *Development* (Mengembangkan)

Pada tahap *development*, instrumen *Five-tier test open ended* yang sudah dibuat oleh peneliti berdasarkan studi literatur kemudian dikembangkan menjadi instrumen *Five-tier test close ended*. Berikut merupakan contoh pengembangan instrumen *Five-tier test open ended* menjadi *close ended* yang ditunjukkan pada Gambar 3.2.

FIVE TIER DIAGNOSTIC TEST OPEN ENDED
PADA MATERI KALOR DAN TERMODINAMIKA

1.1 Raihian melakukan percobaan dengan menggunakan batang logam, korek api kayu dan lilin seperti pada Gambar 1.



Gambar 1. Raihian memanaskan salah satu ujung batang logam diatas lilin dan menempatkan korek api pada batang logam

Pada percobaan tersebut, mula-mula Raihian memanaskan ujung logam A di atas lilin sedangkan ujung B tidak dipanaskan. Korek api C diletakkan pada batang dekat dengan ujung batang A dan korek api D diletakkan dekat dengan ujung batang B. Setelah beberapa lama, korek api manakah yang lebih cepat menyala?

A. Korek api C karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konveksi
 B. Korek api D karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konveksi
 C. Korek api C karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konduksi
 D. Korek api D karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konduksi
 E. Korek api C karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses radiasi

1.2 Tingkat keyakinan terhadap pilihan jawabannya!
 A. Yakin
 B. Tidak yakin

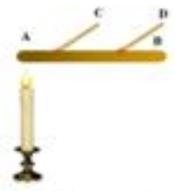
1.3 Alasan dari jawaban tersebut adalah ...

1.4 Tingkat keyakinan terhadap pilihan jawaban!
 A. Yakin
 B. Tidak yakin

1.5 Dimana sumber yang Anda gunakan untuk menjawab pertanyaan tersebut?
 A. Buku
 B. Penjelasan guru
 C. Pendapat pribadi
 D. Penjelasan Teman
 E. Lainnya

FIVE TIER DIAGNOSTIC TEST CLOSE ENDED
PADA MATERI KALOR DAN TERMODINAMIKA

1.1 Raihian melakukan percobaan dengan menggunakan batang logam, korek api kayu dan lilin seperti pada Gambar 1.



Gambar 1. Raihian memanaskan salah satu ujung batang logam diatas lilin dan menempatkan korek api pada batang logam

Pada percobaan tersebut, mula-mula Raihian memanaskan ujung logam A di atas lilin sedangkan ujung B tidak dipanaskan. Korek api C diletakkan pada batang dekat dengan ujung batang A dan korek api D diletakkan dekat dengan ujung batang B. Setelah beberapa lama, korek api manakah yang lebih cepat menyala?

A. Korek api C karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konveksi
 B. Korek api D karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konveksi
 C. Korek api C karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konduksi
 D. Korek api D karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses konduksi
 E. Korek api C karena ujung batang A lebih cepat menerima panas dibanding ujung batang B karena proses radiasi

1.2 Tingkat keyakinan terhadap pilihan jawabannya!
 A. Yakin
 B. Tidak yakin

1.3 Alasan dari jawaban tersebut adalah ...
 A. Kalor berpindah dari suhu tinggi ke suhu rendah disertai dengan perpindahan partikel-partikel tersebut
 B. Kalor berpindah dari suhu tinggi ke suhu rendah tanpa disertai dengan perpindahan partikel-partikel tersebut
 C. Kalor berpindah dari suhu tinggi ke suhu rendah disertai dengan perubahan wujud pada benda tersebut
 D. Kalor berpindah dari suhu tinggi ke suhu rendah tanpa disertai dengan perubahan wujud pada benda tersebut
 E. Kalor berpindah dari suhu tinggi ke suhu rendah disertai dengan perubahan warna pada benda tersebut

1.4 Tingkat keyakinan terhadap pilihan jawaban!
 A. Yakin
 B. Tidak yakin

1.5 Dimana sumber yang Anda gunakan untuk menjawab pertanyaan tersebut?
 A. Buku
 B. Penjelasan guru
 C. Pendapat pribadi
 D. Penjelasan Teman
 E. Lainnya

(a)

(b)

Gambar 3.2 (a) FTDT open ended (b) FTDT close ended

Gambar 3.2 menunjukkan pengembangan instrumen *five tier test open ended* menjadi *five tier test close ended* yang dibuat berdasarkan hasil studi

Aulia Ayuni Novita, 2025

PENGEMBANGAN INSTRUMEN FIVE-TIER TEST UNTUK MENGIDENTIFIKASI MISKONSEPSI PESERTA DIDIK PADA MATERI KALOR DAN TERMODINAMIKA

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

pendahuluan alasan siswa pada *tier* ketiga *five tier test open ended* yang telah disebarakan kepada 25 siswa kelas XI di salah satu SMA Negeri Kabupaten Karawang yang kemudian dikembangkan menjadi *five tier test close ended*. Instrumen tersebut kemudian divalidasi oleh lima orang validator yaitu 2 dosen dan 3 guru dengan menggunakan lembar validasi instrumen *five tier test* pada materi kalor dan termodinamika. Pada lembar validasi terdapat aspek materi, konstruk dan bahasa yang diberikan penilaian dari rentang 1 – 5 oleh validator. Selain itu, validator dapat memberikan catatan berupa saran terhadap instrumen yang telah dibuat sehingga instrumen dapat direvisi agar lebih layak untuk digunakan.

4. Tahap *Desseminate*

Pada tahap *desseminate*, instrumen *five-tier test* yang telah dikembangkan dan dinyatakan layak untuk digunakan kemudian akan disebarakan dan diujicobakan kepada siswa yang berasal dari ketiga SMA Negeri di Kabupaten Karawang. Instrumen *five tier test* pada materi kalor dan termodinamika yang disebarakan terdiri dari 15 butir soal dan telah diisi oleh 195 siswa kelas XI dari ketiga SMA Negeri di Kabupaten Karawang. Responden pada saat uji coba skala besar berbeda dengan responden pada saat uji coba studi pendahuluan. Pengisian instrumen *five tier test* pada materi kalor dan termodinamika dikerjakan secara daring dengan bantuan *Google formulir*. Instrumen dianalisis menggunakan pemodelan Rasch dengan bantuan aplikasi *Winstep* versi 5.1.7.0. Analisis instrumen yang digunakan adalah uji validitas (validitas isi dan konstruk), uji reliabilitas, tingkat kesulitan butir soal, dan daya pembeda.

3.5.2 Lembar Validasi Instrumen

Instrumen yang digunakan dalam penelitian ini harus memenuhi persyaratan yang ada. Sebelum instrumen *five-tier test* diberikan kepada siswa, instrumen tersebut harus melalui tahap validasi terlebih dahulu oleh beberapa validator yakni dua dosen ahli serta dua guru Fisika di SMA. Pada proses validasi tersebut mencakup aspek materi, konstruk dan bahasa yang diberikan penilaian dari rentang 1 – 5 oleh validator, seperti yang ditunjukkan pada Tabel 3.5.

Tabel 3.5 Lembar Validasi Instrumen Five-Tier Test

No	Aspek Penilaian	Butir Soal Nomor										
		1	2	3	4	5	6	7	8	9	10	dst
Materi												
1	Item sesuai dengan kisi-kisi											
2	Pertanyaan dan rubrik terumuskan dengan benar											
3	Materi sesuai dengan tujuan untuk pengukuran kemampuan mengidentifikasi miskonsepsi											
4	Materi sesuai dengan tingkat kognisi peserta didik											
Konstruksi												
1	Pengecoh berfungsi (ada beberapa pilihan yang hampir benar)											
2	Hanya ada satu jawaban yang benar											
3	Pokok soal dirumuskan dengan jelas dan tegas											
4	Pokok soal TIDAK mengarah ke jawaban yang benar											
5	Pokok soal TIDAK mengandung pernyataan negatif ganda											
6	Pilihan homogen dan logis ditinjau dari segi materi											
7	Panjang rumusan pilihan relatif sama. Jika pilihan tidak sama panjang, telah diurutkan dari yang terpendek ke yang terpanjang atau sebaliknya											
8	Pilihan tidak mengandung pernyataan "Semua jawaban benar / salah"											
9	Pilihan berbentuk angka telah diurutkan dari terkecil ke terbesar, atau sebaliknya											
10	Gambar, grafik dan stimulus dinyatakan dengan jelas dan berfungsi dengan baik											

No	Aspek Penilaian	Butir Soal Nomor										
		1	2	3	4	5	6	7	8	9	10	dst
11	Butir soal tidak bergantung kepada jawaban soal sebelumnya											
12	Opsi jawaban tidak mengarah kepada alasan yang harus dituliskan											
Bahasa												
1	Rumusan kalimat soal komunikatif											
2	Kesesuaian bahasa yang digunakan dengan bahasa Indonesia yang baku											
3	Ketiadaan makna ganda pada kalimat yang digunakan											
4	Rumusan pokok soal tidak mengandung ungkapan yang bermakna tidak pasti, misal sebaliknya, pada umumnya, kadang-kadang											

3.6 Teknik Analisis Data

Setelah instrumen tes *five-tier test* materi kalor dan termodinamika di uji coba lapangan, maka hasil penelitian tersebut perlu diolah menggunakan analisis Model Rasch berbantuan *software* Winstep 5.1.7.0. Analisis Rasch merupakan bagian dari teori respon butir yang dapat mengelompokkan perhitungan *person* (subyek) dan *item* (soal) dalam suatu peta distribusi, sehingga model tersebut mampu menggambarkan keterkaitan antara tingkat kemampuan subyek dengan item soal (Priyayi, 2022). Sebelum dilakukan analisis data untuk menjawab rumusan masalah penelitian, dilakukan pengkodean dan penskoran terlebih dahulu pada jawaban siswa terhadap instrumen *Five-tier test* berdasarkan rubrik kategori konsepsi yang di adaptasi dari Aminudin dkk. (2019) dan Rosita dkk. (2020). Rubrik kategori konsepsi tersebut ditunjukkan pada Tabel 3.6.

Tabel 3.6 Kategori Konsepsi dan Skor Kombinasi Jawaban Five-Tier-Test

Tier I	Tier II	Tier III	Tier IV	Tier V	Kategori	Skor
				Pemikiran pribadi	SU-PT	
Benar	Y	Benar	Y	Guru	SU-T	4
				Buku	SU-B	
				Teman	SU-PE	
				Internet	SU-I	
Benar	Y	Benar	TY	Pemikiran pribadi	PP-PT	3
Benar	TY	Benar	Y	Guru	PP-T	
				Buku	PP-B	
Benar	TY	Benar	TY	Teman	PP-PE	
				Internet	PP-I	
Benar	Y	Salah	Y	Pemikiran pribadi	PN-PT	1
Benar	TY	Salah	Y			
Benar	Y	Salah	TY	Guru	PN-T	
Benar	TY	Salah	TY			
Salah	Y	Benar	Y	Buku	PN-B	
Salah	TY	Benar	Y			
Salah	Y	Benar	TY	Teman	PN-PE	
Salah	TY	Benar	TY	Internet	PN-I	
				Pemikiran pribadi	MC-PT	0
Salah	Y	Salah	Y	Guru	MC-T	
				Buku	MC-B	
				Teman	MC-PE	
				Internet	MC-I	
Salah	Y	Salah	TY	Pemikiran pribadi	NU-PT	0
Salah	TY	Salah	Y	Guru	NU-T	
				Buku	NU-B	
Salah	TY	Salah	TY	Teman	NU-PE	
				Internet	NU-I	

Keterangan :

Y = Yakin

TY = Tidak Yakin

SU = *Sound Understanding*PT = *Personal Thoughts*PP = *Partial Positif*T = *Teacher*PN = *Partial Negatif*B = *Book*MC = *Misconception*PE = *People's Explanation*NU = *No Understanding*I = *Internet*

Aulia Ayuni Novita, 2025

PENGEMBANGAN INSTRUMEN FIVE-TIER TEST UNTUK MENGIDENTIFIKASI MISKONSEPSI PESERTA DIDIK PADA MATERI KALOR DAN TERMODINAMIKA

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

Ada beberapa model yang bisa digunakan dalam analisis butir politomi, salah satunya dengan *Partial Credit Model* (PCM). Pada penelitian ini, maka akan menggunakan analisis butir menggunakan *partial credit model* (PCM). Pada Tabel 3.4 menunjukkan bahwa konsepsi siswa dikelompokkan ke dalam beberapa kategori, yakni SU (*Sound Understanding*) menunjukkan konsepsi pada siswa telah utuh karena siswa dapat menjawab pertanyaan dengan benar serta alasan yang benar juga yakin dengan jawabannya, PP (*Partial Positif*) menunjukkan pemahaman siswa terhadap suatu konsep belum utuh karena keraguan dalam menjawab soal, PN (*Partial Negatif*) menunjukkan pemahaman terhadap suatu konsep belum utuh dikarenakan hanya benar pada salah satu tier dan tingkat keyakinannya berbeda-beda, MC (*Misconception*) menunjukkan pemahaman konsep siswa tidak sesuai dengan konsep ilmiah tetapi siswa yakin atas pemahamannya tersebut, sedangkan NU (*No Understanding*) menunjukkan keadaan siswa yang tidak paham mengenai konsep ilmiah. Selain itu, dikelompokkan juga berdasarkan sumber penyebab miskonsepsinya, yaitu *Personal Thoughts* (PT)/pemikiran pribadi, *Teacher/guru*, *Book (B)/Buku*, *People's Explanation* (PE)/penjelasan orang lain, dan *Internet (I)/internet*.

3.6.1 Karakteristik Instrumen *Five-Tier Test*

3.6.1.1 Validitas Instrumen *Five-Tier Test* Pada Materi Kalor dan Termodinamika

Menurut Sugiyono (dalam Maharrani, 2023), mengemukakan bahwa validitas merupakan suatu ukuran yang menunjukkan tingkat kevalidan atau kesahihan suatu instrumen. Suatu instrumen dapat dikatakan valid jika alat ukur yang digunakan dapat mengukur yang hendak diukur. Validitas instrumen terdiri dari validitas konstruk dan isi, serta validitas empiris.

A. Validitas Konstruk dan Isi

Untuk menguji validitas konstruk dan isi, maka dapat digunakan pendapat para ahli (*judgement experts*) (Sugiyono, Statistika Untuk Penelitian, 2019). Hasil data validasi oleh ahli kemudian akan dianalisis menggunakan Formula Aiken. Formula Aiken digunakan untuk menentukan koefisien V dari

data berskala rating secara statistik (Aiken, 1985). Formula Aiken dirumuskan sebagai berikut :

$$V = \frac{\sum s}{n(c - 1)}$$

Keterangan :

V = indeks validitas item

$s = r - l_0$ = skor yang diberi penilai; l_0 = skor terendah

c = jumlah kategori penilaian

n = jumlah validator

Koefisien validitas Aiken (Aiken's V) yang diperoleh berdasarkan perhitungan menggunakan formula Aiken kemudian dibandingkan dengan nilai nilai standar Aiken's V pada Tabel 3.7.

Tabel 3.7 Standar Koefisien Validitas (V) Aiken

Jumlah Item (m) atau Validator (n)	Jumlah Kategori Rating (c)							
	2		3		4		5	
	V	P	V	P	V	P	V	P
2							1,00	0,40
3							1,00	0,008
3			1,00	0,037	1,00	0,016	0,92	0,032
4					1,00	0,004	0,94	0,008
4			1,00	0,012	0,92	0,020	0,88	0,024
5			1,00	0,004	0,93	0,006	0,90	0,007
5	1,00	0,031	0,90	0,025	0,87	0,021	0,80	0,040
6			0,92	0,010	0,89	0,007	0,88	0,005
6	1,00	0,016	0,83	0,038	0,78	0,050	0,79	0,029
7			0,93	0,004	0,86	0,007	0,82	0,010
7	1,00	0,008	0,86	0,016	0,76	0,045	0,75	0,041
8	1,00	0,004	0,88	0,007	0,83	0,007	0,81	0,008
8	0,88	0,035	0,81	0,024	0,75	0,040	0,75	0,030
9	1,00	0,002	0,89	0,003	0,81	0,007	0,81	0,006
9	0,89	0,020	0,78	0,032	0,74	0,036	0,72	0,038
10	1,00	0,001	0,85	0,005	0,80	0,007	0,78	0,008
10	0,90	0,001	0,75	0,040	0,73	0,032	0,70	0,047
11	0,91	0,006	0,82	0,007	0,79	0,007	0,77	0,006
11	0,82	0,033	0,73	0,048	0,73	0,029	0,70	0,035

Aulia Ayuni Novita, 2025

PENGEMBANGAN INSTRUMEN FIVE-TIER TEST UNTUK MENGIDENTIFIKASI MISKONSEPSI
PESERTA DIDIK PADA MATERI KALOR DAN TERMODINAMIKA

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

12	0,92	0,003	0,79	0,010	0,78	0,006	0,75	0,009
12	0,83	0,019	0,75	0,025	0,69	0,046	0,69	0,041
13	0,92	0,002	0,81	0,005	0,77	0,006	0,75	0,006
13	0,77	0,046	0,73	0,030	0,69	0,041	0,67	0,048
14	0,86	0,006	0,79	0,006	0,76	0,005	0,73	0,008
14	0,79	0,029	0,71	0,035	0,69	0,036	0,68	0,036
15	0,87	0,004	0,77	0,008	0,73	0,010	0,73	0,006
15	0,80	0,018	0,70	0,040	0,69	0,032	0,67	0,041
16	0,88	0,002	0,75	0,010	0,73	0,009	0,72	0,008
16	0,75	0,038	0,69	0,046	0,67	0,047	0,66	0,046
17	0,82	0,006	0,76	0,005	0,73	0,008	0,71	0,010
17	0,76	0,025	0,71	0,026	0,67	0,041	0,66	0,036
18	0,83	0,004	0,75	0,006	0,72	0,007	0,71	0,007
18	0,72	0,048	0,69	0,030	0,67	0,036	0,65	0,040
19	0,79	0,010	0,74	0,008	0,72	0,006	0,70	0,009
19	0,74	0,032	0,68	0,033	0,65	0,050	0,64	0,044
20	0,80	0,006	0,72	0,009	0,70	0,010	0,69	0,010
20	0,75	0,021	0,68	0,037	0,65	0,044	0,64	0,048

(Aiken, 1985)

Ketika nilai Aiken's V hitung lebih besar dari nilai standar Aiken's V pada tabel, maka item tersebut dapat dikatakan "valid".

Hasil validasi ahli (*judgement expert*) pada instrumen *Five Tier Test* pada materi kalor dan termodinamika yang dilakukan kepada 2 orang dosen pendidikan fisika dan 3 orang guru fisika dianalisis menggunakan Formula Aiken V dalam menentukan koefisien validitas Aiken (Aiken's V) berbantuan *Microsoft Excel*. Koefisien validitas Aiken (Aiken's V) yang telah diperoleh berdasarkan hasil perhitungan menggunakan formula Aiken's V akan dibandingkan dengan nilai standar minimal Aiken's V (nilai Aiken's V Tabel). Hal ini dilakukan untuk menentukan validitas tiap butir soal dalam instrumen penelitian yang digunakan. Dikarenakan instrumen penelitian divalidasi oleh 5 orang ahli dan 20 kategori penilaian, maka nilai standar koefisien Aiken's V minimal yang diperoleh yaitu 0,69 dengan probabilitas 0,010. Berdasarkan hasil perhitungan validasi ahli pada instrumen *five-tier test* pada materi kalor dan termodinamika dapat dilihat pada Tabel 3.8.

Tabel 3.8 Hasil Uji Validitas Ahli

Butir	Penilai					$V_{rata-rata}$	Kategori
	I	II	III	IV	V		
1	3	5	5	4	4	0,8	Valid
2	3	4	4	4	4	0,7	Valid
3	4	4	5	4	4	0,8	Valid
4	3	4	4	4	4	0,7	Valid
5	4	3	4	4	4	0,7	Valid
6	3	4	4	4	4	0,7	Valid
7	4	5	5	4	4	0,85	Valid
8	3	5	5	4	4	0,8	Valid
9	2	3	2	2	2	0,3	Tidak Valid
10	5	5	5	4	4	0,9	Valid
11	5	5	5	4	4	0,9	Valid
12	3	2	3	2	2	0,35	Tidak Valid
13	4	4	3	4	4	0,7	Valid
14	4	4	4	4	4	0,75	Valid
15	5	5	5	4	4	0,9	Valid
16	5	4	4	4	4	0,8	Valid
17	3	4	4	4	4	0,7	Valid
18	4	4	4	4	4	0,75	Valid
19	5	5	5	4	4	0,9	Valid
20	5	5	4	4	4	0,85	Valid

Dari Tabel 4.1 diatas dapat dilihat bahwa terdapat 18 butir soal dengan kategori “Valid” yang memiliki nilai lebih besar dari nilai standar minimal Aiken’s V tabel dan terdapat 2 butir soal dengan kategori “Tidak Valid” yang memiliki nilai lebih kecil dari nilai standar minimal Aiken’s V tabel. Sehingga dapat disimpulkan bahwa dari hasil validasi ahli terdapat 18 butir soal valid yang dapat digunakan sebagai instrumen dalam penelitian ini.

B. Validitas Empiris *Five Tier Test* Pada Materi Kalor dan Termodinamika

1. Kualitas Butir Soal (Validitas)

Dalam penelitian ini uji validitas empiris akan dilakukan dengan analisis model Rasch menggunakan *software* dengan memasukkan skor konsepsi siswa berdasarkan hasil implementasi instrumen *Five-Tier Test*. Uji validitas ini dianalisis berdasarkan keseluruhan butir soal disebut dengan *Unidimensionality* serta untuk setiap butir soal (Sumintono, 2015). Uji validitas berdasarkan keseluruhan butir soal diperoleh dari menu *output tables* pada bagian *item : dimensionality*. Nilai dari *raw variance explained by measures* merupakan nilai validitas berdasarkan keseluruhan butir soal yang kemudian dapat diinterpretasikan berdasarkan Tabel 3.9.

Tabel 3.9 Interpretasi Unidimensionalitas Instrumen

<i>Raw variance explained by measure</i>	Interpretasi
$21\% \leq \text{value} \leq 40\%$	Terpenuhi
$41\% \leq \text{value} \leq 60\%$	Sesuai
$\text{value} \geq 61\%$	Istimewa

(Sumintono & Widhiarso, 2015)

Setelah uji validitas untuk keseluruhan instrumen dilakukan, selanjutnya adalah uji kualitas untuk setiap butir soal. Kualitas setiap butir soal dapat dilihat berdasarkan nilai *outfit mean square* (MNSQ) yang menunjukkan keacakan berupa jumlah penyimpangan dalam sistem pengukuran, nilai *outfit Z-standard* (ZSTD) yang menunjukkan seberapa besar kemungkinan terjadi penyimpangan, serta nilai *Pt measure corr* yang menunjukkan hubungan antara kesulitan setiap item dan kesulitan instrumen keseluruhan (Sumintono, 2015). Kriteria untuk memeriksa kualitas setiap butir soal dapat dilihat pada Tabel 3.10 berikut.

Tabel 3.10 Kriteria Nilai *Outfit MNSQ*, *Outfit ZSTD*, dan *Pt Measure Corr*

Kriteria	Rentang yang diterima	Keterangan
<i>Outfit Mean Square</i> (MNSQ)	$0,5 < \text{MNSQ} < 1,5$	<i>Fit</i>
<i>Outfit Z Standard</i> (ZSTD)	$-2,0 < \text{ZSTD} < + 2,0$	<i>Fit</i>
<i>Point Measure Correlation</i> (<i>Pt Measure Corr</i>)	$0,4 < \text{Pt Measure Corr} < 0,85$	<i>Fit</i>

(Sumintono & Widhiarso, 2015)

Menurut (Boone, 2014), menyatakan bahwa hasil nilai *outfit man square*, *outfit Z-standard*, dan *Pt measure correlation* merupakan kriteria yang digunakan untuk melihat tingkat kesesuaian butir (*item fit*). Sehingga jika pada butir soal ketiga kriteria tersebut tidak terpenuhi, maka dapat dipastikan bahwa butir soal yang dibuat kurang bagus sehingga perlu diperbaiki atau diganti. Hasil dari nilai masing-masing kriteria tersebut diinterpretasikan berdasarkan nilai *fit-statistic* pada Tabel 3.11 berikut.

Tabel 3.11 Interpretasi kualitas butir soal

Kriteria Nilai <i>Fit-Statistic</i>	Rentang yang diterima
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

(Sumintono & Widhiarso, 2015)

Hasil uji validitas empiris pada instrumen *Five Tier Test* pada materi kalor dan termodinamika dianalisis berdasarkan model Rasch menggunakan bantuan *software Winstep* versi 5.1.7.0. Instrumen *Five Tier Test* pada materi kalor dan termodinamika diberikan sebanyak 18 butir soal menggunakan media *google formulir* kepada 25 siswa yang terdiri dari 10 siswa laki-laki dan 15 siswa perempuan.

Uji validitas secara keseluruhan menggunakan model Rasch disebut dengan unidimensionalitas. Hasil dari unidimensionalitas instrumen Five-Tier Diagnostic Test pada materi kalor dan termodinamika yang didapatkan dari *output tabel item dimensionality* pada *software Winstep* versi 5.7.1 ditunjukkan pada Gambar 3.3 berikut.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units			
	Eigenvalue	Observed	Expected
Total raw variance in observations =	22.0835	100.0%	100.0%
Raw variance explained by measures =	7.0835	32.1%	32.6%
Raw variance explained by persons =	1.7416	7.9%	8.0%
Raw Variance explained by items =	5.3419	24.2%	24.6%
Raw unexplained variance (total) =	15.0000	67.9%	100.0%
Unexplned variance in 1st contrast =	2.4973	11.3%	16.6%
Unexplned variance in 2nd contrast =	2.2198	10.1%	14.8%
Unexplned variance in 3rd contrast =	2.0419	9.2%	13.6%
Unexplned variance in 4th contrast =	1.5266	6.9%	10.2%
Unexplned variance in 5th contrast =	1.4229	6.4%	9.5%

Gambar 3.3 Output tabel item dimensionality

Aulia Ayuni Novita, 2025

PENGEMBANGAN INSTRUMEN FIVE-TIER TEST UNTUK MENGIDENTIFIKASI MISKONSEPSI PESERTA DIDIK PADA MATERI KALOR DAN TERMODINAMIKA

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

Pada Gambar 4.1 menunjukkan bahwa nilai *raw variance explained by measures* yang diperoleh dari uji coba lapangan instrumen *Five-Tier Diagnostic Test* pada materi kalor dan termodinamika sebanyak 18 butir soal adalah sebesar 32,1%. Berdasarkan Tabel 3.8, nilai tersebut memenuhi kriteria “terpenuhi”, artinya instrumen tersebut dapat mengukur satu variabel tanpa dipengaruhi oleh variabel-variabel yang lainnya. Selain itu, unidimensionalitas instrumen juga dapat dilihat dari nilai *unexplained variance in 1st contrast* jika nilainya kurang dari 15%, maka instrumen tersebut memiliki kuantitas *unidimensionalitas* yang baik (Fisher, 2007). Dapat dilihat dari *output* tabel diatas, nilai *unexplained variance in 1st contrast* sebesar 11,3% sehingga dapat dikatakan bahwa kuantitas *unidimensionalitas* instrumen dapat dikatakan baik.

Setelah dilakukan uji validitas secara keseluruhan, langkah selanjutnya adalah uji validitas untuk setiap butir soal. Berdasarkan nilai *fit-statistic* dijelaskan mengenai uji validitas untuk setiap butir soal pada Tabel 3.12 berikut.

Tabel 3.11 Hasil Interpretasi Kualitas Butir Soal

Nomor Butir Soal	MNSQ		ZSTD		Pt Measure Corr.		Kriteria Nilai	Interp.	Keterangan
	Nilai	Ket	Nilai	Ket	Nilai	Ket			
1	0,89	<i>Fit</i>	-0,01	<i>Fit</i>	0,44	<i>Fit</i>	Ketiga kriteria terpenuhi	Sangat sesuai	Digunakan
2	1,14	<i>Fit</i>	0,51	<i>Misfit</i>	0,34	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
3	0,74	<i>Fit</i>	-0,63	<i>Fit</i>	0,48	<i>Fit</i>	Ketiga kriteria terpenuhi	Sangat sesuai	Digunakan
4	0,86	<i>Fit</i>	-0,40	<i>Fit</i>	0,55	<i>Fit</i>	Ketiga kriteria terpenuhi	Sangat sesuai	Digunakan
5	0,79	<i>Fit</i>	-0,36	<i>Misfit</i>	0,39	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
6	1,33	<i>Fit</i>	1,22	<i>Fit</i>	0,29	<i>Fit</i>	Ketigaa kriteria terpenuhi	Sangat sesuai	Digunakan

7	1,06	<i>Fit</i>	0,29	<i>Misfit</i>	0,30	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
8	1,03	<i>Fit</i>	0,21	<i>Misfit</i>	0,31	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
9	1,63	<i>Misfit</i>	1,61	<i>Fit</i>	0,27	<i>Misfit</i>	Satu kriteria terpenuhi	Kurang sesuai	Tidak digunakan
10	0,85	<i>Fit</i>	-0,44	<i>Fit</i>	0,47	<i>Fit</i>	Ketiga kriteria terpenuhi	Sangat sesuai	Digunakan
11	1,27	<i>Fit</i>	0,65	<i>Fit</i>	0,25	<i>Misfit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
12	0,47	<i>Fit</i>	-0,32	<i>Misfit</i>	0,34	<i>Fit</i>	Satu kriteria terpenuhi	Kurang sesuai	Tidak digunakan
13	1,02	<i>Fit</i>	0,25	<i>Misfit</i>	0,20	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
14	0,55	<i>Fit</i>	-0,32	<i>Misfit</i>	0,34	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
15	0,63	<i>Fit</i>	-0,98	<i>Fit</i>	0,56	<i>Fit</i>	Ketiga kriteria terpenuhi	Sangat sesuai	Digunakan
16	1,72	<i>Misfit</i>	1,73	<i>Fit</i>	0,28	<i>Misfit</i>	Satu kriteria terpenuhi	Kurang sesuai	Tidak digunakan
17	0,62	<i>Fit</i>	0,16	<i>Fit</i>	0,28	<i>Misfit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan
18	1,05	<i>Fit</i>	0,26	<i>Misfit</i>	0,32	<i>Fit</i>	Dua kriteria terpenuhi	Sesuai	Digunakan

Tabel 3.12 menunjukkan hasil validitas setiap butir soal instrumen *Five-Tier Test* pada materi kalor dan termodinamika. Hasilnya menunjukkan bahwa pada nomor 1, 3, 4, 6, 10 dan 15 memenuhi ketiga kriteria dengan interpretasi “sangat sesuai”. Pada nomor 2, 5, 7, 8, 11, 13, 14, 17 dan 18 memenuhi dua kriteria dengan interpretasi “sesuai”. Sedangkan pada nomor 9, 12 dan 16 hanya memenuhi satu kriteria dengan interpretasi “tidak sesuai”,

Aulia Ayuni Novita, 2025

PENGEMBANGAN INSTRUMEN FIVE-TIER TEST UNTUK MENGIDENTIFIKASI MISKONSEPSI PESERTA DIDIK PADA MATERI KALOR DAN TERMODINAMIKA

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

sehingga butir soal tersebut tidak bisa digunakan. Berdasarkan hasil analisis tersebut dapat disimpulkan bahwa sebanyak 15 butir soal pada instrumen *Five-Tier Test* pada materi kalor dan termodinamika dikatakan valid dan dapat digunakan menjadi instrumen dalam penelitian.

2. Reliabilitas Instrumen

Reliabilitas instrumen merupakan ketetapan instrumen dalam pengukuran yang dilakukan dalam sebuah penelitian atau ketetapan responden dalam menjawab instrumen soal. Dengan demikian, pengukuran yang dilakukan secara berulang kali akan menghasilkan informasi yang konsisten. Kesetaraan tersebut menunjukkan bahwa instrumen yang digunakan yaitu reliabel (Sumintono & Widhiarso, 2015). Uji reliabilitas pada penelitian ini dilakukan dengan analisis pemodelan *Rasch* dan menggunakan *software Winstep* versi 5.7.1. Data yang akan didapatkan untuk realibilitas ini antara lain *person reability* (p), *item reability* (r), dan *alpha cronbach* ($KR-20$) yang menunjukkan interaksi yang terjadi antara *person reability* dan *item reability* (Sumintono & Widhiarso, 2015). Interpretasi ketiganya dapat dilihat pada Tabel 3.13 berikut :

Tabel 3.13 Interpretasi item and person reliability dan cronbach alpha

Statistik	Nilai Indeks	Interpretasi
<i>Person and Item Reability</i>	$r \leq 0,67$	Lemah
	$0,67 < r \leq 0,80$	Cukup
	$0,81 < r \leq 0,90$	Bagus
	$0,91 < r \leq 0,94$	Bagus Sekali
	$>0,94$	Istimewa
<i>Alpha Cronbach (KR-20)</i>	$KR-20 < 0,50$	Buruk
	$0,50 < KR-20 < 0,6$	Jelek
	$0,60 < KR-20 < 0,70$	Cukup
	$0,70 < KR-20 < 0,80$	Bagus
	$KR-20 < 0,80$	Bagus Sekali

(Sumintono & Widhiarso, 2015)

Hasil reliabilitas instrumen pada penelitian ini dianalisis menggunakan *software* Winstep versi 5.1.7.0 pada menu utama *Output Tables* kemudian memilih menu *Summary Statistics*. Data yang akan didapatkan untuk realibilitas ini antara lain *person reability* (p), *item reability* (r), dan *alpha cronbach* (KR-20) yang menunjukkan interaksi yang terjadi antara *person reability* dan *item reability*. Hasil uji reliabilitas instrumen *five tier test* pada materi kalor dan termodinamika ditunjukkan pada Gambar 4.2.

SUMMARY OF 25 MEASURED (EXTREME AND NON-EXTREME) PERSON									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	4.8	18.0	-1.23	.74					
SEM	.5	.0	.25	.05					
P.SD	2.3	.0	1.21	.26					
S.SD	2.4	.0	1.23	.26					
MAX.	9.0	18.0	.55	1.86					
MIN.	.0	18.0	-4.80	.60					
REAL RMSE	.81	TRUE SD	.89	SEPARATION	1.10		PERSON RELIABILITY	.55	
MODEL RMSE	.78	TRUE SD	.92	SEPARATION	1.18		PERSON RELIABILITY	.58	
S.E. OF PERSON MEAN = .25									
PERSON RAW SCORE-TO-MEASURE CORRELATTON = .97									
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .54 SEM = 1.56									
STANDARDIZED (50 ITEM) RELIABILITY = .79									
SUMMARY OF 15 MEASURED (NON-EXTREME) ITEM									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	7.9	25.0	.00	.58	1.01	.12	.92	.03	
SEM	1.4	.0	.38	.05	.04	.17	.06	.14	
P.SD	5.3	.0	1.44	.19	.14	.64	.24	.54	
S.SD	5.5	.0	1.49	.20	.14	.67	.24	.56	
MAX.	20.0	25.0	2.44	1.04	1.25	1.25	1.33	1.22	
MIN.	1.0	25.0	-2.97	.46	.73	-1.34	.47	-.98	
REAL RMSE	.63	TRUE SD	1.29	SEPARATION	2.04		ITEM RELIABILITY	.81	
MODEL RMSE	.62	TRUE SD	1.30	SEPARATION	2.11		ITEM RELIABILITY	.82	
S.E. OF ITEM MEAN = .38									

Gambar 3.4 Output table summary statistic

Berdasarkan Gambar 3.4 menunjukkan bahwa *person reability* bernilai 0,55 dengan interpretasi “lemah”. Hasil interpretasi pada *person reability* yang lemah menunjukkan kemungkinan jumlah sampel dan butir soal yang sedikit, untuk mendapatkan koefisien *person reability* yang tinggi maka diperlukan sampel dengan rentang kemampuan yang besar serta instrumen yang digunakan memiliki jumlah item yang banyak. Untuk nilai *item reability* sebesar 0,81 dengan interpretasi “bagus”. Hasil interpretasi ini menunjukkan bahwa instrumen *Five-Tier Test* pada materi kalor dan

termodinamika memiliki kualitas dinyatakan baik dalam hal konsistensi bobot soal dengan pilihan jawabannya. Kemudian untuk nilai *cronbach's alpha* (KR-20) sebesar 0,54 dengan interpretasi “rendah”. Berdasarkan hasil analisis tersebut dapat disimpulkan bahwa instrumen *Five-Tier Test* pada materi kalor dan termodinamika reliabel untuk digunakan sebagai instrumen dalam penelitian ini.

3. Tingkat Kesukaran Butir Soal

Tingkat kesukaran butir soal (*item measure*) dapat diurutkan dari tingkat kesukaran tinggi hingga rendah, dari yang sulit sampai mudah. Rasch Model juga dapat mengetahui setiap butir soal yang diolah menggunakan *software winstep* pada menu *output tables* pada bagian *item measure* dengan data nilai *Measure* (M) dan Standar Deviasi (SD) (Sumintono, 2015). Untuk mengetahui tingkat kesulitan butir soal dapat dilihat dari nilai logit tiap butir soal yang dapat dilihat pada kolom *measure*. Nilai logit yang tinggi menunjukkan tingkat kesulitan soal yang paling tinggi. Berikut adalah interpretasi tingkat kesulitan setiap butir soal berdasarkan Tabel 3.14.

Tabel 3.14 Interpretasi tingkat kesulitan butir soal

Kriteria	Interpretasi
$ME < -1SD$	Mudah
$-1SD \leq ME \leq +1SD$	Sedang
$ME > +1SD$	Sulit

(Sumintono & Widhiarso, 2015)

Hasil analisis tingkat kesukaran instrumen *five tier test* pada materi kalor dan termodinamika dianalisis menggunakan *software winstep* versi 5.1.7.0 pada menu *item : measure*. Berdasarkan *output* tabel *item measure* diperoleh dari nilai standar deviasi (SD) dari hasil uji coba instrumen *Five-Tier Test* pada materi kalor dan termodinamika sebesar 1,89. Hasil uji tingkat kesukaran instrumen *Five-Tier Test* pada materi kalor dan termodinamika ditunjukkan pada Tabel 4.3.

Tabel 2.15 Hasil Interpretasi Tingkat Kesulitan Butir Soal

Nomor Butir Soal	Measure (ME)	Standar Deviasi (SD)	Kriteria	Interpretasi
1	-2,97	1,89	$ME < -1,89$	Mudah
2	-0,23	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
3	0,00	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
4	-1,88	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
5	0,24	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
6	-1,66	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
7	-0,44	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
8	-0,23	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
9	-0,44	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
10	2,44	1,89	$ME > 1,89$	Sulit
11	0,81	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
12	1,67	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
13	0,00	1,89	$-1,89 \leq ME \leq 1,89$	Sedang
14	2,44	1,89	$ME > 1,89$	Sulit
15	0,24	1,89	$-1,89 \leq ME \leq 1,89$	Sedang

Berdasarkan Tabel 3.15 menunjukkan bahwa butir soal nomor 1 termasuk tingkatan yang “mudah” dan butir soal nomor 10 dan 14 termasuk tingkatan yang “sulit”. Sedangkan butir soal nomor 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13 dan 15 termasuk tingkatan yang “sedang”. Analisis lebih lanjut mengenai tingkat kesulitan dapat dilakukan dengan menghitung frekuensi dan persentase untuk setiap interpretasi tingkat kesulitan butir soal yang disajikan pada Tabel 3.16 berikut.

Tabel 3.16 Frekuensi dan persentase tingkat kesulitan butir soal

Interpretasi	Frekuensi	Persentase (%)
Sulit	2	13
Sedang	12	80
Mudah	1	7

Berdasarkan Tabel 4.4 didapatkan bahwa frekuensi terbesar berada pada tingkat kesulitan dengan interpretasi “sedang” yaitu sebanyak 12 butir soal dengan persentase sebesar 80%. Untuk interpretasi “mudah” yaitu sebanyak 1 butir soal dengan persentase 7%. Sedangkan interpretasi “sulit” sebanyak 2 butir soal dengan persentase 13%. Hal tersebut menunjukkan bahwa tingkat kesulitan instrumen *Five-Tier-Test* pada materi kalor dan termodinamika terdistribusi cukup baik.

4. Daya Pembeda Butir Soal

Daya pembeda butir soal digunakan untuk menunjukkan seberapa jauh sebuah butir soal dapat membedakan peserta didik yang memiliki kemampuan tinggi dan peserta didik yang memiliki kemampuan rendah (Sumintono & Widhiarso, 2015). Daya pembeda dapat dianalisis menggunakan pemodelan Rasch berbantuan *software winstep* versi 5.1.7.0 yang melibatkan nilai *Point Measure Correlation (Pt Mean Corr)* dengan merujuk pada tingkat kesulitan masing-masing butir soal dengan tingkat kesulitan keseluruhan butir soal. Interpretasi daya pembeda untuk butir soal dapat ditunjukkan pada Tabel 3.17 berikut.

Tabel 3.17 Interpretasi daya pembeda butir soal

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

Hasil nalisis daya pembeda dapat dilakukan dengan menggunakan pemodelan Rasch berbantuan *software winstep* versi 5.1.7.0 untuk membedakan peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah dalam menjawab butir soal. Hasil uji daya pembeda instrumen *Five Tier Test* pada materi kalor dan termodinamika ditunjukkan pada Tabel 3.18 berikut.

Tabel 3.18 Hasil Interpretasi Daya Pembeda Butir Soal

Nomor Butir Soal	<i>Pt Measure Corr.</i>	Interpretasi
1	0,44	Sangat Baik
2	0,34	Baik
3	0,48	Sangat Baik
4	0,55	Sangat Baik
5	0,39	Baik
6	0,29	Kurang Baik
7	0,30	Kurang Baik
8	0,31	Baik
9	0,27	Kurang Baik
10	0,47	Sangat Baik
11	0,25	Kurang Baik
12	0,34	Baik
13	0,20	Kurang Baik
14	0,34	Baik
15	0,56	Sangat Baik
16	0,28	Kurang Baik
17	0,28	Kurang Baik
18	0,32	Baik

Berdasarkan Tabel 3.18 dapat diketahui bahwa butir soal nomor 1, 3, 4, 10, dan 15 termasuk ke dalam daya pembeda yang “sangat baik”.

Butir soal nomor 2, 5, 8, 12, 14, dan 18 termasuk ke dalam daya pembeda yang “baik”. Sedangkan pada butir soal nomor 6, 7, 9, 11, 13, 16 dan 17 termasuk ke dalam daya pembeda yang “kurang baik”. Analisis lebih lanjut mengenai daya pembeda instrumen *Five Tier Test* pada materi kalor dan termodinamika dapat dilakukan dengan menghitung frekuensi dan persentase untuk setiap interpretasi daya pembeda setiap butir soal yang disajikan pada Tabel 3.19 berikut.

Tabel 3.19 Frekuensi dan Persentase Daya Pembeda Butir Soal

Interpretasi	Frekuensi	Persentase (%)
Sangat Baik	5	28
Baik	6	33
Kurang Baik	7	39

Berdasarkan Tabel 3.19 dapat dikatakan bahwa frekuensi terbesar terdapat pada interpretasi “kurang baik” yaitu sebanyak 7 butir soal dengan persentase sebesar 39%. Untuk interpretasi “baik” yaitu sebanyak 6 butir soal dengan persentase sebesar 33%. Sedangkan untuk interpretasi “sangat baik” yaitu sebanyak 5 butir soal dengan persentase sebesar 28%. Hal tersebut menunjukkan bahwa daya pembeda instrumen *Five-Tier Test* pada materi kalor dan termodinamika terdistribusi cukup baik.

3.6.2 Analisis Profil Miskonsepsi dan Penyebab Miskonsepsi Siswa

Setelah data yang dianalisis mengenai konsepsi siswa telah dianalisis, pada siswa yang teridentifikasi miskonsepsi perlu dilakukan analisis lebih mendalam untuk mengetahui miskonsepsi yang terjadi pada siswa serta untuk mengetahui sumber penyebab siswa mengalami miskonsepsi berdasarkan pada konsep-konsep yang telah diujikan pada instrumen *five-tier-test*.