

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Metode penelitian yang digunakan adalah metode kuantitatif. Penelitian dengan metode kuantitatif merupakan penelitian untuk menguji teori objektif dengan menguji hubungan antar variabel (Creswell, 2014). Penelitian ini dilakukan dengan eksperimen menggunakan desain penelitian *One Group Pretest-Posttest Design*. Dalam pelaksanaannya, sampel penelitian diberikan *pretest* (O_1) terlebih dahulu untuk mengidentifikasi kemampuan berpikir kritis peserta didik. Kemudian sampel penelitian tersebut diberikan perlakuan (X) berupa penerapan model *Problem Based Learning* dalam pembelajaran materi suhu dan kalor. Tahap terakhir, peserta didik diberikan *posttest* (O_2) untuk melihat peningkatan kemampuan berpikir kritis peserta didik setelah diberikan perlakuan tersebut. Adapun desain pada penelitian ini dapat dilihat pada Tabel 3.1.

Tabel 3.1 Desain Penelitian

O_1	X	O_2
<i>Pretest</i>	Penerapan PBL dalam pembelajaran materi Suhu dan Kalor	<i>Posttest</i>

(Sugiyono, 2019)

Keterangan:

O_1 : *Pretest* (Tes awal kemampuan berpikir kritis dilakukan sebelum penerapan pembelajaran dengan model *Problem Based Learning* (PBL)).

O_2 : *Posttest* (Tes akhir kemampuan berpikir kritis dilakukan setelah penerapan pembelajaran dengan model *Problem Based Learning* (PBL)).

X : Penerapan model *Problem Based Learning* (PBL) dalam pembelajaran materi Suhu dan Kalor.

3.2 Populasi dan Sampel

Populasi dalam penelitian ini adalah siswa kelas XI pada SMA Pasundan 8 Bandung. Teknik pengambilan sampel penelitian menggunakan teknik *purposive sampling* yaitu teknik penentuan sampel dengan pertimbangan tertentu (Sugiyono, 2019). Penerapan metode *convenience sampling* digunakan agar sampel yang diinginkan sesuai berdasarkan kepentingan peneliti. Sampel dibutuhkan dari satu kelas yang sama dan representatif, selain itu sampel dipilih berdasarkan hasil pertimbangan pihak sekolah berdasarkan kelas dengan jadwal yang memungkinkan untuk dilakukan penelitian. Sampel pada penelitian ini adalah kelas XI Mipa 3 berjumlah 33 peserta didik dipilih berdasarkan hasil pertimbangan guru mata pelajaran fisika yaitu kelas yang belum dan akan mempelajari materi “Suhu dan Kalor”.

3.3 Instrumen Penelitian

Instrumen penelitian terdiri dari instrumen tes dan instrumen *non-test*. Instrumen tes merupakan instrumen pengumpulan data berupa tes kemampuan berpikir kritis yang disajikan berbentuk soal uraian. Sedangkan instrumen *non-test* merupakan pendukung instrumen tes berupa lembar observasi keterlaksanaan model pembelajaran dan angket respon peserta didik.

3.3.1 Tes Kemampuan Berpikir Kritis

Pada penelitian ini, tes kemampuan berpikir kritis digunakan untuk *pretest* dan *posttest* berupa 10 butir soal uraian untuk mengukur kemampuan berpikir kritis peserta didik. Tes kemampuan berpikir kritis mengacu pada indikator kemampuan berpikir kritis Ennis (2011) dan sesuai dengan materi yang diambil yaitu suhu dan kalor. Berikut matriks instrumen tes kemampuan berpikir kritis ditunjukkan pada Tabel 3.2.

Tabel 3.2 Matriks instrumen kemampuan berpikir kritis

Materi	Sub Materi	Nomor Soal	Indikator Kemampuan Berpikir Kritis (Ennis, 2011)				
			<i>Elementary Clarification</i>	<i>Strategy and Tactics</i>	<i>Advance Clarification</i>	<i>Basic Support</i>	<i>Inference</i>
Suhu dan Pemuain	Suhu	1	√				
	Skala Termometer	2		√			
	Pemuain	3			√		
		4			√		
Kalor	Kalor	5					√
	Perubahan Wujud Zat	6	√				
	Azas Black	7				√	
Perpindahan Kalor	Konduksi	8		√			
	Konveksi	9					√
	Radiasi	10				√	

3.3.2 Lembar Observasi Keterlaksanaan Model Pembelajaran

Lembar observasi digunakan untuk mengetahui keterlaksanaan dari model pembelajaran berbasis masalah yang digunakan dan terdapat indikator kemampuan berpikir kritis dari beberapa sintaks *problem based learning*. Lembar observasi berisi pernyataan yang berkaitan dengan aktivitas guru dan peserta didik selama proses pembelajaran berbentuk daftar *checklist* yang diisi oleh *observer* yaitu salah satu guru mata pelajaran fisika di SMA Pasundan 8 Bandung. *Observer* dapat memberi *checklist* (√) pada kolom “Ya” jika aktivitas guru/peserta didik terlaksana atau kolom “Tidak” jika aktivitas guru/peserta didik tidak terlaksana. Terlaksana atau tidaknya suatu aktivitas disesuaikan berdasarkan rubrik penilaian lembar observasi keterlaksanaan model pembelajaran yang telah peneliti susun dan disajikan dalam lampiran. Skor keterlaksanaannya yaitu 1 jika terlaksana dan 0 jika tidak terlaksana. Hasil data lembar observasi keterlaksanaan model pembelajaran ini kemudian akan dilihat persentase keterlaksanaannya.

3.3.3 Angket Respon Peserta Didik

Peneliti menyebarkan angket kepada peserta didik selaku responden pada pertemuan akhir. Angket ini diisi responden sebagai respon peserta didik apakah hasil belajar dirasa efektif menggunakan model pembelajaran yang diterapkan.

Angket respon peserta didik yang digunakan dalam penelitian ini, terdiri dari 4 aspek yang ditunjukkan pada Tabel 3.3.

Tabel 3.3 Aspek angket respon peserta didik

No.	Aspek	Deskripsi	Nomor Pernyataan	
			Positif	Negatif
1.	Minat	Adanya ketertarikan peserta didik terhadap pembelajaran model <i>problem based learning</i>	1,3	2,4
2.	Motivasi	Peningkatan motivasi peserta didik dalam memecahkan suatu permasalahan pada pembelajaran model <i>problem based learning</i>	5,7	6,8
3.	Kepuasan	Menilai seberapa jauh pembelajaran model <i>problem based learning</i> membantu peserta didik dalam memahami materi	9,11	10,12
4.	Penilaian	Peningkatan kemampuan peserta didik pada pembelajaran model <i>problem based learning</i>	13,15	14,16

(Hairina et al., 2021)

Pernyataan pada angket terdiri atas 8 pernyataan positif dan 8 pernyataan negatif. Kategori pernyataan angket respon peserta didik ditunjukkan pada Tabel 3.4.

Tabel 3.4 Kategori butir pernyataan

Nomor Butir Pernyataan	Kategori
1,3,5,7,9,11,13,15	Positif
2,4,6,8,10,12,14,16	Negatif

Hasil pengukuran angket respon peserta didik digunakan skala *likert*. Skala *likert* mengacu pada pedoman Sugiyono (2019) disajikan pada Tabel 3.5.

Tabel 3.5 Skala likert angket respon peserta didik

Kriteria	Skala <i>Likert</i>	
	Pernyataan Positif	Pernyataan Negatif
Sangat Setuju	4	1
Setuju	3	2
Tidak Setuju	2	3
Sangat Tidak Setuju	1	4

(Sugiyono, 2019)

3.4 Prosedur Penelitian

Pada prosedur penelitian terdapat tiga tahapan, diantaranya yaitu, tahap perencanaan, tahap pelaksanaan, dan tahap akhir. Tahapan tersebut dijelaskan sebagai berikut:

3.4.1 Tahap Perencanaan

- 1) Melakukan studi pendahuluan
 - Studi literatur terhadap jurnal, buku, dan penelitian-penelitian terdahulu untuk mengetahui rangkaian penelitian yang akan dikembangkan.
 - Studi lapangan, dilakukan untuk menentukan masalah yang akan dikaji dalam penelitian.
- 2) Merumuskan masalah dalam bentuk pertanyaan penelitian
- 3) Menyusun instrumen penelitian berupa RPP, 10 soal uraian sebagai tes kemampuan berpikir kritis peserta didik beserta rubrik penilaiannya, serta lembar observasi keterlaksanaan pembelajaran.
- 4) Melakukan validasi dan uji coba instrumen,
- 5) Berdasarkan hasil uji coba yang dilakukan, soal dikelompokkan menjadi tiga kelompok, yaitu soal yang dapat langsung digunakan, soal yang direvisi sebelum langsung digunakan dan soal yang tidak bisa digunakan/dibuang. Soal yang tidak layak digunakan akan dibuang sebelum diaplikasikan.
- 6) Melakukan survei untuk menentukan populasi dan sampel penelitian serta tempat dan waktu penelitian.

3.4.2 Tahap Pelaksanaan

- 1) Memberikan *pretest* pada kelas eksperimen untuk mengetahui kemampuan berpikir kritis awal peserta didik yang menjadi sampel penelitian sebelum diberikan perlakuan pada kelas tersebut.
- 2) Memberikan perlakuan melalui kegiatan pembelajaran pada kelas eksperimen, yaitu menerapkan PBL dalam pembelajaran materi suhu dan kalor.
- 3) Memberikan *posttest* untuk mengetahui kemampuan akhir siswa terkait dengan kemampuan berpikir kritis peserta didik setelah perlakuan selesai diberikan.

3.4.3 Tahap Akhir

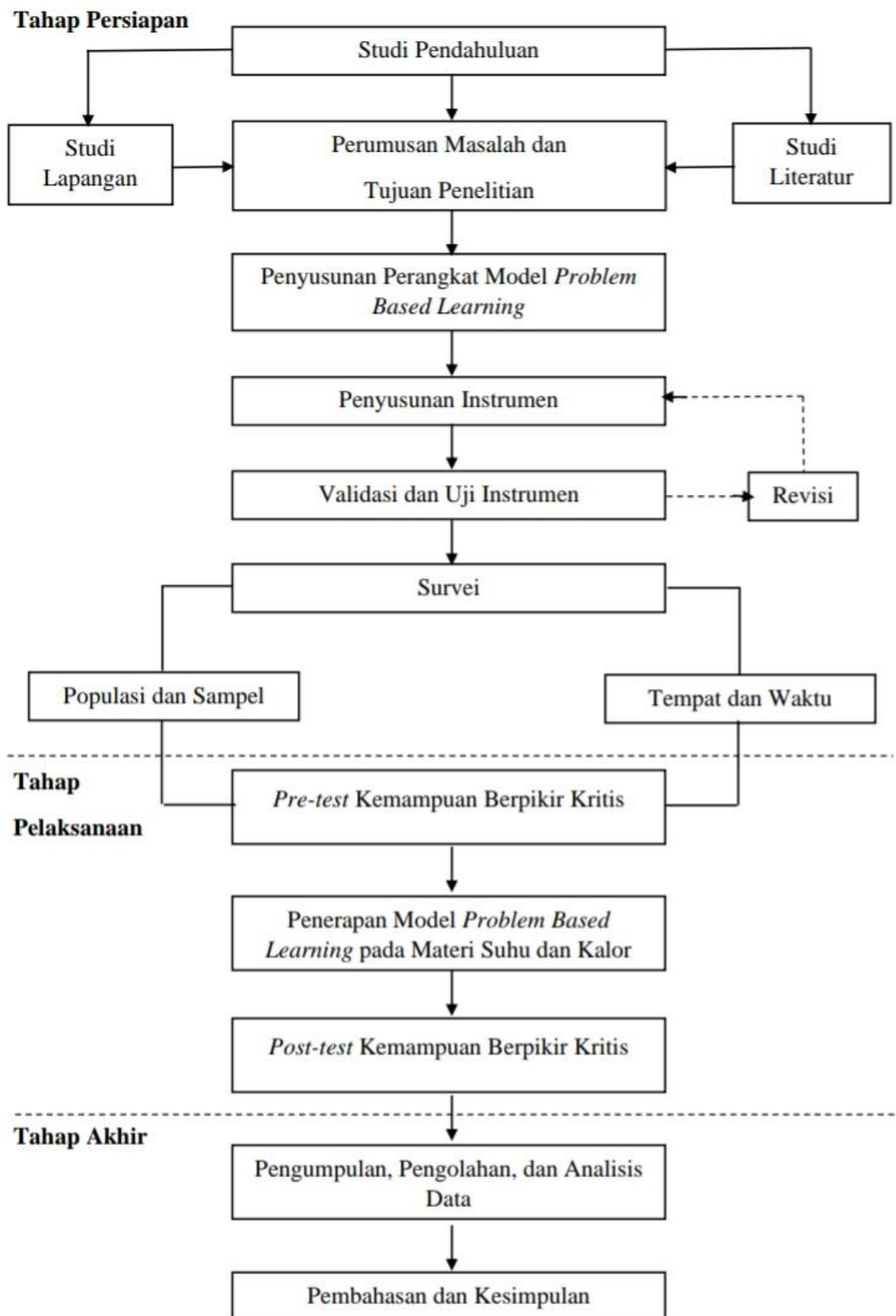
Tahap akhir dalam penelitian merupakan proses pengolahan dan analisis seluruh data yang telah dikumpulkan hingga diperoleh kesimpulan terhadap hasil penelitian yang telah dilaksanakan.

3.5 Alur Penelitian

Alur penelitian yang digunakan pada penelitian ini disajikan pada Gambar 3.1.

3.6 Analisis Tes Instrumen Penelitian

Pada penelitian ini, tes instrumen harus diuji terlebih dahulu sebelum digunakan pada sampel. Uji instrumen oleh validasi ahli menggunakan validitas Aiken. Sedangkan uji instrumen penelitian secara empiris dilakukan menggunakan pemodelan Rasch, untuk mendapatkan hasil analisis uji validitas, reliabilitas, dan tingkat kesukaran yang dibantu dengan *software ministep*. Adapun penjelasan dan hasil analisis pengujian instrumennya adalah sebagai berikut.



Gambar 3.1 Alur Penelitian

3.6.1 Validasi Ahli (*Expert Judgement*)

Salah satu langkah awal yang dilakukan setelah penyusunan tes instrumen yaitu melakukan uji instrumen validitas isi. Validitas isi tes instrumen dilakukan untuk mengetahui apakah butir soal dalam tes layak mewakili indikator yang diukur (Azwar, 2018). Layak atau tidaknya suatu *item* dalam instrumen dapat ditentukan oleh ahli berdasarkan hasil penilaian (*judgement*). Metode untuk menganalisis validitas isi dari setiap butir soal instrumen tes pada penelitian ini digunakan indeks Aiken's V. Validasi tes instrumen kemampuan berpikir kritis dilakukan oleh 3 validator yaitu dua dosen ahli pembelajaran fisika dan satu guru fisika SMA.

Indikator validasi yang akan digunakan untuk tiap nomor soal diantaranya yaitu kemampuan butir soal dalam mengukur indikator kemampuan berpikir kritis, kebenaran konsep materi, kejelasan pertanyaan, bahasa pada butir soal sesuai dengan Ejaan Bahasa Indonesia, dan rubrik penilaian dapat mengukur target yang diukur. Setelah ahli melakukan pengecekan instrumen, para ahli memberikan penilaian terhadap setiap butir soal dengan rentang skala penilaian 1-3. Selanjutnya, peneliti mengolah hasil nilai yang telah diberikan oleh ahli menggunakan formula Aiken (1985) dengan rumus sebagai berikut.

$$V = \sum \frac{S}{[n(c - 1)]}$$

Keterangan:

$$S = r - l_0$$

r : Angka yang diberikan penilai

l_0 : Angka penilaian validitas yang terendah (misalnya 1)

c : Angka penilaian validitas tertinggi (misalnya 3)

n : Jumlah penilai

Setelah didapatkan indeks V berdasarkan perhitungan di atas, Aiken (1985) juga memberikan panduan untuk menentukan *item* tersebut diterima atau tidak. Nilai indeks V tersebut ditampilkan pada Gambar 3.2.

No. of Items (m) or Raters (n)	Number of Rating Categories (c)													
	2		3		4		5		6		7			
	V	p	V	p	V	p	V	p	V	p	V	p		
2							1.00	.040	1.00	.028	1.00	.020		
3							1.00	.008	1.00	.005	1.00	.003		
3			1.00	.037	1.00	.016	.92	.032	.87	.046	.89	.029		
4					1.00	.004	.94	.008	.95	.004	.92	.006		
4			1.00	.012	.92	.020	.88	.024	.85	.027	.83	.029		
5			1.00	.004	.93	.006	.90	.007	.88	.007	.87	.007		
5	1.00	.031	.90	.025	.87	.021	.80	.040	.80	.032	.77	.047		
6			.92	.010	.89	.007	.88	.005	.83	.010	.83	.008		
6	1.00	.016	.83	.038	.78	.050	.79	.029	.77	.036	.75	.041		
7			.93	.004	.86	.007	.82	.010	.83	.006	.81	.008		
7	1.00	.008	.86	.016	.76	.045	.75	.041	.74	.038	.74	.036		
8	1.00	.004	.88	.007	.83	.007	.81	.008	.80	.007	.79	.007		
8	.88	.035	.81	.024	.75	.040	.75	.030	.72	.039	.71	.047		
9	1.00	.002	.89	.003	.81	.007	.81	.006	.78	.009	.78	.007		
9	.89	.020	.78	.032	.74	.036	.72	.038	.71	.039	.70	.040		
10	1.00	.001	.85	.005	.80	.007	.78	.008	.76	.009	.75	.010		
10	.90	.001	.75	.040	.73	.032	.70	.047	.70	.039	.68	.048		
11	.91	.006	.82	.007	.79	.007	.77	.006	.75	.010	.74	.009		
11	.82	.033	.73	.048	.73	.029	.70	.035	.69	.038	.68	.041		
12	.92	.003	.79	.010	.78	.006	.75	.009	.73	.010	.74	.008		
12	.83	.019	.75	.025	.69	.046	.69	.041	.68	.038	.67	.049		
13	.92	.002	.81	.005	.77	.006	.75	.006	.74	.007	.72	.010		
13	.77	.046	.73	.030	.69	.041	.67	.048	.68	.037	.67	.041		
14	.86	.006	.79	.006	.76	.005	.73	.008	.73	.007	.71	.009		
14	.79	.029	.71	.035	.69	.036	.68	.036	.66	.050	.66	.047		
15	.87	.004	.77	.008	.73	.010	.73	.006	.72	.007	.71	.008		
15	.80	.018	.70	.040	.69	.032	.67	.041	.65	.048	.66	.041		
16	.88	.002	.75	.010	.73	.009	.72	.008	.71	.007	.70	.010		
16	.75	.038	.69	.046	.67	.047	.66	.046	.65	.046	.65	.046		
17	.82	.006	.76	.005	.73	.008	.71	.010	.71	.007	.70	.009		
17	.76	.025	.71	.026	.67	.041	.66	.036	.65	.044	.65	.039		
18	.83	.004	.75	.006	.72	.007	.71	.007	.70	.007	.69	.010		
18	.72	.048	.69	.030	.67	.036	.65	.040	.64	.042	.64	.044		
19	.79	.010	.74	.008	.72	.006	.70	.009	.70	.007	.68	.009		
19	.74	.032	.68	.033	.65	.050	.64	.044	.64	.040	.63	.048		
20	.80	.006	.72	.009	.70	.010	.69	.010	.68	.010	.68	.008		
20	.75	.021	.68	.037	.65	.044	.64	.048	.64	.038	.63	.041		
21	.81	.004	.74	.005	.70	.010	.69	.008	.68	.010	.68	.009		
21	.71	.039	.67	.041	.65	.039	.64	.038	.63	.048	.63	.045		
22	.77	.008	.73	.006	.70	.008	.68	.009	.67	.010	.67	.008		
22	.73	.026	.66	.044	.65	.035	.64	.041	.63	.046	.62	.049		
23	.78	.005	.72	.007	.70	.007	.68	.007	.67	.010	.67	.009		
23	.70	.047	.65	.048	.64	.046	.63	.045	.63	.044	.62	.043		
24	.79	.003	.71	.008	.69	.006	.68	.008	.67	.010	.66	.010		
24	.71	.032	.67	.030	.64	.041	.64	.035	.62	.041	.62	.046		
25	.76	.007	.70	.009	.68	.010	.67	.009	.66	.009	.66	.009		
25	.72	.022	.66	.033	.64	.037	.63	.038	.62	.039	.61	.049		

Gambar 3.2 Indeks V Aiken

Penelitian ini menetapkan peluang error sebesar 5%, maka dapat dilihat baris kedua dengan jumlah rater 3 pada Gambar 3.5, maka nilai minimum indeks V pada tabel V adalah 1.00, artinya apabila nilai V_{hitung} lebih dari sama dengan nilai V_{tabel}

maka butir tersebut dinyatakan valid. Berikut telah dilakukan analisis menggunakan formula Aiken dan diperoleh hasil data yang disajikan pada Tabel 3.6.

Tabel 3.6 Hasil analisis data validasi pre-post test kemampuan berpikir kritis

Butir Soal	Kemampuan butir soal dalam mengukur indikator kemampuan berpikir kritis		Kebenaran konsep materi		Kejelasan Pertanyaan		Bahasa pada butir soal sesuai dengan Ejaan Bahasa Indonesia		Rubrik penilaian dapat mengukur target yang diukur	
	N	%	N	%	N	%	N	%	N	%
1	0,67	67	1,00	100	0,83	83	0,83	83	0,67	67
2	0,83	83	1,00	100	0,83	83	0,83	83	0,67	67
3	1,00	100	1,00	100	0,83	83	1,00	100	0,67	67
4	1,00	100	1,00	100	0,83	83	1,00	100	0,83	83
5	1,00	100	1,00	100	0,83	83	0,83	83	0,83	83
6	1,00	100	1,00	100	1,00	100	1,00	100	0,83	83
7	1,00	100	1,00	100	1,00	100	1,00	100	1,00	100
8	1,00	100	1,00	100	0,83	83	1,00	100	1,00	100
9	1,00	100	1,00	100	0,83	83	0,83	83	0,83	83
10	1,00	100	1,00	100	0,83	83	1,00	100	0,83	83

Berdasarkan Gambar 3.5, V_{tabel} pada penelitian ini yaitu 1,00 dan pada Tabel 3.6 terdapat nilai $V_{\text{hitung}} \leq V_{\text{tabel}}$. Oleh sebab itu, menurut Nudin & Hidayatullah (2023) untuk pengambilan keputusan tiap butir soal diambil dari persentase nilai rata-rata setiap indikator dan dapat dikategorikan valid tidaknya berdasarkan acuan skala likert pada Tabel 3.7.

Tabel 3.7 Skala Likert

Persentase	Kategori
0% – 20%	Sangat Tidak Valid
21% – 40%	Tidak Valid
41% – 60%	Cukup Valid
61% – 80%	Valid
81% – 100%	Sangat Valid

(Riduwan, 2015)

Dari nilai rata-rata persentase pada Tabel 3.6 yang disesuaikan dengan skala likert pada Tabel 3.7 maka dapat disimpulkan bahwa setiap butir soal pada tes instrumen kemampuan berpikir kritis dikategorikan valid dan dapat digunakan untuk pengujian lebih lanjut dengan mempertimbangkan beberapa saran/masukan validator yang terdapat pada lembar hasil validasi ahli dalam lampiran. Berikut rangkuman saran/masukan validator untuk perbaikan soal tes instrumen kemampuan berpikir kritis, disajikan pada Tabel 3.8.

Tabel 3.8 Rangkuman masukan validator

Indikator Validasi	Masukan Validator
Kemampuan butir soal dalam mengukur indikator kemampuan berpikir kritis	<ul style="list-style-type: none"> • Pada soal nomor 2, belum nampak unsur untuk <i>strategy and tactics</i>. • Pada soal nomor 6, kurang nampak unsur untuk <i>inference</i>.
Kebenaran konsep materi	<ul style="list-style-type: none"> • Soal yang disajikan harus berupa fenomena secara faktual.
Kejelasan pertanyaan	<ul style="list-style-type: none"> • Tambahkan gambar pada beberapa soal tertentu agar peserta didik dapat membayangkannya. • Fokus pada konsep. • Tidak mengulang-ulang kalimat.
Bahasa pada butir soal sesuai dengan Ejaan Bahasa Indonesia	<ul style="list-style-type: none"> • Lebih efektifkan kalimat pada soal yang disajikan. • Kalimat “terdapat kalor” diganti dengan “terjadi perpindahan kalor”.
Rubrik penilaian dapat mengukur target yang diukur	<ul style="list-style-type: none"> • Opsi pada rubrik harus ada karakteristiknya. • Rubrik penilaian dibuat lebih detail untuk memperjelas skor pada saat penilaian.

Dalam hal ini validator menyarankan untuk dilakukannya perbaikan terkait butir soal sebelum diuji cobakan agar butir soal tersebut berguna dalam pengujian. Perbaikan butir soal dalam penelitian ini, mempertimbangkan saran dan masukan dari para validator yang terlampir pada Tabel 3.8. Setelah kegiatan perbaikan butir soal selesai, kegiatan pengembangan instrumen dilanjutkan dengan uji instrumen penelitian secara empiris menggunakan pemodelan Rasch, untuk mendapatkan hasil analisis uji validitas, reliabilitas, dan tingkat kesukaran yang dibantu dengan *software winsteps*.

3.6.2 Uji Validitas Konstruk

Uji validitas perlu dilakukan pada soal yang akan digunakan untuk mengukur kemampuan peserta didik. Uji validitas dalam analisis pemodelan rasch dikenal dengan nama unidimensionalitas instrumen. Unidimensionalitas instrumen merupakan ukuran yang penting untuk mengevaluasi apakah instrumen yang dikembangkan mampu mengukur apa yang seharusnya diukur sehingga dapat dikatakan valid (Sumintono & Widhiarso, 2015).

Soal tes kemampuan berpikir kritis diuji cobakan kepada peserta didik kelas XII MIPA di SMA Pasundan 8 Bandung dengan jumlah sampel uji coba 36 orang, kemudian data tersebut dimasukkan ke *Microsoft Excel* lalu diolah dengan *software winsteps*. Adapun hasil dari pengolahan uji validitas menggunakan *software winsteps* dapat dilihat pada Gambar 3.3.

Table of RAW RESIDUAL variance in Eigenvalue units			
	Eigenvalue	Observed	Expected
Total raw variance in observations =	28.6252	100.0%	100.0%
Raw variance explained by measures =	18.6252	65.1%	64.7%
Raw variance explained by persons =	13.3960	46.8%	46.5%
Raw Variance explained by items =	5.2292	18.3%	18.2%
Raw unexplained variance (total) =	10.0000	34.9%	100.0% 35.3%
Unexplned variance in 1st contrast =	2.5282	8.8%	25.3%
Unexplned variance in 2nd contrast =	1.7602	6.1%	17.6%
Unexplned variance in 3rd contrast =	1.2500	4.4%	12.5%
Unexplned variance in 4th contrast =	1.1891	4.2%	11.9%
Unexplned variance in 5th contrast =	1.0687	3.7%	10.7%

Gambar 3.3 Hasil analisis uji validitas

Unidimensionalitas dapat ditunjukkan jika tidak ada butir soal yang bermasalah maka nilai *raw variance* > 20% dan nilai *observed* dalam *unexplained* < 15% untuk menunjukkan butir soal yang sesuai (Sumintono & Widhiarso, 2015). Berdasarkan hasil analisis unidimensionalitas pada Gambar 3.5, diperoleh *raw variance* sebesar 65.1% dan nilai *observed* dalam *unexplained variance 1 st contrast* menunjukkan nilai 8.8%, sehingga dapat disimpulkan bahwa tidak ada butir soal yang bermasalah dan setiap butir soal mampu mengukur rentang kemampuan (*ability*) peserta didik secara komprehensif.

3.6.3 Uji Validitas Instrumen

Validitas instrumen merupakan suatu argumen untuk mengetahui apakah tiap butir soal dapat mengukur kemampuan peserta didik. Butir soal yang cocok (*fit*) berarti soal tersebut berperilaku secara konsisten dengan apa yang diharapkan oleh model (Benyamin, 1998). Beberapa indeks fit dalam analisis model rasch diantaranya, memiliki nilai *outfit mean square* $0.5 < MNSQ < 1.5$, nilai *outfit Z Standardized* $-2.0 < ZSTD < 2.0$, dan nilai *Point Measure Correlation (Pt Mean Corr)* tidak negatif (Sumintono & Widhiarso, 2015). Berikut hasil analisis data fit *pre-posttest* kemampuan berpikir kritis dengan model rasch dapat dilihat pada Gambar 3.4 dan Tabel 3.9.

TABLE 3.1 C:\Users\Admin\Downloads\Data Mentah (ZOU640WS.TXT Nov 11 2023 08:4
INPUT: 36 Person 10 Item REPORTED: 36 Person 10 Item 4 CATS MINISTEP 5.6.3.

SUMMARY OF 36 MEASURED Person

	TOTAL		MEASURE	MODEL S.E.	INFIT		OUTFIT	
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD
MEAN	26.7	10.0	.57	.59	.97	-.22	.97	-.17
SEM	1.1	.0	.35	.01	.14	.23	.14	.23
P.SD	6.3	.0	2.08	.04	.82	1.35	.83	1.35
S.SD	6.4	.0	2.11	.04	.83	1.37	.84	1.37
MAX.	37.0	10.0	4.26	.72	5.18	5.34	5.15	5.29
MIN.	14.0	10.0	-3.78	.55	.22	-2.47	.22	-2.48
REAL RMSE	.64	TRUE SD	1.98	SEPARATION	3.08	Person	RELIABILITY	.90
MODEL RMSE	.59	TRUE SD	2.00	SEPARATION	3.39	Person	RELIABILITY	.92
S.E. OF Person MEAN = .35								

Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .92 SEM = 1.81
STANDARDIZED (50 ITEM) RELIABILITY = .98

SUMMARY OF 10 MEASURED Item

	TOTAL		MEASURE	MODEL S.E.	INFIT		OUTFIT	
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD
MEAN	96.0	36.0	.00	.31	.98	-.16	.97	-.15
SEM	4.4	.0	.42	.00	.11	.49	.10	.46
P.SD	13.3	.0	1.25	.01	.32	1.46	.31	1.38
S.SD	14.0	.0	1.32	.01	.33	1.54	.33	1.45
MAX.	121.0	36.0	1.86	.34	1.53	2.06	1.45	1.78
MIN.	76.0	36.0	-2.45	.30	.50	-2.73	.51	-2.69
REAL RMSE	.33	TRUE SD	1.21	SEPARATION	3.70	Item	RELIABILITY	.93
MODEL RMSE	.31	TRUE SD	1.22	SEPARATION	3.93	Item	RELIABILITY	.94
S.E. OF Item MEAN = .42								

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

Gambar 3.4 Hasil analisis data fit

Tabel 3.9 Hasil analisis data fit

	<i>Outfit MNSQ</i>	<i>Outfit ZSTD</i>
Person	0.97	-0.17
Item	0.97	-0.15

Berdasarkan hasil analisis data fit pada Tabel 3.9, diketahui nilai rata-rata *outfit Mean Squared (outfit MNSQ)* sebesar 0.97 pada kolom *person* dan *item*, menunjukkan nilai tersebut tergolong kategori fit yaitu terletak di antara $0.5 < \text{MNSQ} < 1.5$. Hal ini membuktikan bahwa instrumen tes yang digunakan sudah sesuai dengan model untuk mengukur kemampuan berpikir kritis peserta didik. Selain itu, diketahui nilai rata-rata *outfit Z Standardized (outfit ZSTD)* sebesar -0.17 untuk *person* dan nilai -0.15 untuk *item*. Kedua nilai tersebut berada pada rentang $-2 < \text{ZSTD} < 2,0$. Artinya secara keseluruhan butir soal telah sesuai dengan model rasch dan dapat dijadikan instrumen untuk mengukur kemampuan berpikir kritis peserta didik pada materi suhu dan kalor.

3.6.4 Uji Reliabilitas

Reliabilitas dalam model rasch terdapat *reliability person* dan *reliability item*, dan terdapat nilai *Cronbach Alpha* yang dapat mengukur reliabilitas dari interaksi individu dan item secara keseluruhan (Sumintono & Widhiarso, 2015). Berikut Tabel 3.10 dan Tabel 3.11 merupakan tabel kriteria *Person Reliability*, *Item Reliability*, dan *Cronbach Alpha* menurut Sumintono dan Widhiarso (2015).

Tabel 3.10 Kriteria *Person Reliability* dan *Item Reliability*

Nilai <i>r</i>	Kriteria
$r > 0,94$	Istimewa
0,91 - 0,94	Bagus Sekali
0,81 - 0,90	Bagus
0,67 - 0,80	Cukup
$r < 0,67$	Lemah

(Sumintono & Widhiarso, 2015)

Tabel 3.11 Kriteria Nilai *Cronbach Alpha*

Nilai α	Kriteria
$\alpha \geq 0,8$	Bagus Sekali
$0,70 \leq \alpha < 0,80$	Bagus
$0,60 \leq \alpha < 0,70$	Cukup
$0,50 \leq \alpha < 0,60$	Jelak
$\alpha < 0,50$	Buruk

(Sumintono & Widhiarso, 2015)

Berikut merupakan hasil analisis data menggunakan pemodelan rasch pada *software winsteps table 3.1 summary statistic*, diperoleh informasi penting yang dapat dilihat pada Gambar 3.5 dan Tabel 3.12.

TABLE 3.1 C:\Users\Admin\Downloads\Data Mentah (ZOU640WS.TXT Nov 11 2023 08:4
INPUT: 36 Person 10 Item REPORTED: 36 Person 10 Item 4 CATS MINISTEP 5.6.3.

SUMMARY OF 36 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	26.7	10.0	.57	.59	.97	-.22	.97	-.17
SEM	1.1	.0	.35	.01	.14	.23	.14	.23
P.SD	6.3	.0	2.08	.04	.82	1.35	.83	1.35
S.SD	6.4	.0	2.11	.04	.83	1.37	.84	1.37
MAX.	37.0	10.0	4.26	.72	5.18	5.34	5.15	5.29
MIN.	14.0	10.0	-3.78	.55	.22	-2.47	.22	-2.48

REAL RMSE .64 TRUE SD 1.98 SEPARATION 3.08 Person RELIABILITY .90
MODEL RMSE .59 TRUE SD 2.00 SEPARATION 3.39 Person RELIABILITY .92
S.E. OF Person MEAN = .35

Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .92 SEM = 1.81
STANDARDIZED (50 ITEM) RELIABILITY = .98

SUMMARY OF 10 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	96.0	36.0	.00	.31	.98	-.16	.97	-.15
SEM	4.4	.0	.42	.00	.11	.49	.10	.46
P.SD	13.3	.0	1.25	.01	.32	1.46	.31	1.38
S.SD	14.0	.0	1.32	.01	.33	1.54	.33	1.45
MAX.	121.0	36.0	1.86	.34	1.53	2.06	1.45	1.78
MIN.	76.0	36.0	-2.45	.30	.50	-2.73	.51	-2.69

REAL RMSE .33 TRUE SD 1.21 SEPARATION 3.70 Item RELIABILITY .93
MODEL RMSE .31 TRUE SD 1.22 SEPARATION 3.93 Item RELIABILITY .94
S.E. OF Item MEAN = .42

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

Gambar 3.5 Hasil analisis uji reliabilitas untuk nilai *cronbach alpha*, *item reliability*, dan *person reliability*

Tabel 3.12 Hasil analisis uji reliabilitas untuk nilai *cronbach alpha*, *item reliability*, dan *person reliability*

	Rata-rata Logit	Separation	Reliability	Alpha Cronbach
Person	0.57 (2.08)	3.08	0.90	0.92
Item	0.0 (1.25)	3.70	0.93	

Berdasarkan hasil analisis instrumen pada Tabel 3.12, diketahui rata-rata logit *person* 0.57, sedangkan rata-rata logit *item* 0.00. Hal ini menunjukkan bahwa rata-rata logit *person* lebih besar dibandingkan dengan rata-rata logit *item*, sehingga membuktikan bahwa kemampuan peserta didik pada umumnya lebih besar daripada kesukaran butir soal instrumen. Selanjutnya, mengenai *separation* atau pengelompokan *person* dan *item*. *Separation* menunjukkan seberapa bagus seperangkat butir pada tes instrumen kemampuan berpikir kritis dalam penyebaran rentang kemampuan logit. Semakin tinggi nilai *separation*, maka semakin bagus instrumen yang dibuat. Hal ini menunjukkan butir soal didalamnya dapat menjangkau individu dengan kemampuan level tinggi hingga rendah.

Konsistensi jawaban dari peserta didik (*person reliability*) memiliki nilai 0.90 dan kualitas butir soal instrumen (*item reliability*) dengan nilai 0.93. Selain itu, nilai *Cronbach Alpha* untuk tes kemampuan berpikir kritis memiliki nilai 0.92. Hal ini menunjukkan bahwa terdapat kesesuaian antara butir soal dengan *person* (peserta didik), sehingga dapat disimpulkan bahwa tes kemampuan berpikir kritis materi suhu dan kalor dapat dikatakan reliabel.

3.6.5 Tingkat Kesukaran

Tingkat kesukaran menunjukkan sukar mudahnya butir soal dalam suatu *item* dalam mengukur kemampuan peserta didik. Berikut hasil analisis data berdasarkan *measure order* pada *software winsteps* disajikan pada Gambar 3.6.

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S. E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
5	76	36	1.86	.31	.99	.03	1.05	.28	.68	.75	63.9	65.7	S5
6	85	36	1.02	.30	1.02	.17	1.00	.09	.80	.76	69.4	62.5	S6
7	86	36	.93	.30	.84	-.71	.83	-.73	.81	.76	55.6	62.1	S7
9	89	36	.66	.30	1.38	1.63	1.42	1.78	.60	.76	44.4	61.2	S9
3	92	36	.39	.30	1.09	.45	1.11	.56	.67	.76	66.7	60.5	S3
2	95	36	.12	.30	.70	-1.43	.69	-1.53	.83	.76	75.0	60.7	S2
4	95	36	.12	.30	.50	-2.73	.51	-2.69	.80	.76	80.6	60.7	S4
1	107	36	-.98	.31	1.53	2.06	1.45	1.75	.87	.76	52.8	65.3	S1
10	114	36	-1.68	.32	1.20	.87	1.10	.46	.75	.76	66.7	66.4	S10
8	121	36	-2.45	.34	.59	-1.93	.54	-1.49	.84	.74	77.8	69.2	S8
MEAN	96.0	36.0	.00	.31	.98	-.16	.97	-.15			65.3	63.4	
P. SD	13.3	.0	1.25	.01	.32	1.46	.31	1.38			10.9	2.9	

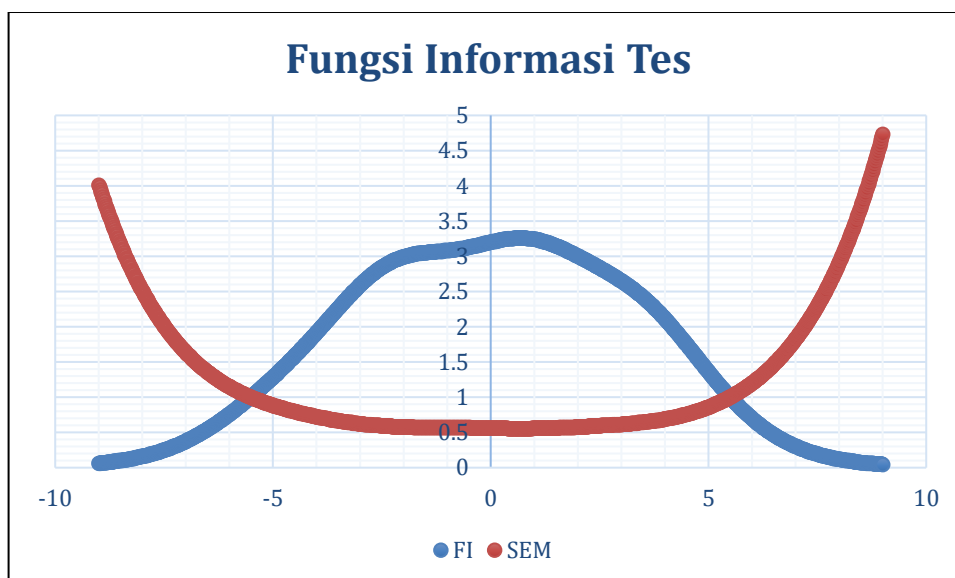
Gambar 3.6 Hasil analisis tingkat kesukaran dan SEM *item*

Tingkat ketelitian *item* mengukur atau *standard error measurement* (SEM) menunjukkan nilai diskriminasi baik jika $SEM < 0.5 \text{ logit}$. Berdasarkan Gambar 3.6 pada kolom 5 menunjukkan bahwa secara keseluruhan butir soal memiliki $SEM < 0.5 \text{ logit}$ dengan rata-rata 0.31 sehingga dapat dikatakan bahwa butir soal “baik” dalam ketelitian pengukuran. Selain itu, ditunjukkan *measure* pada Tabel 3 kolom 2. Menurut Sumintono dan Widhiarso (2015) tingkat kesukaran butir soal memiliki rentang $-2 < \text{measure} < 2$. Butir soal dikategorikan mudah apabila memiliki nilai mendekati -2.00 logit , butir soal dikategorikan sedang apabila $-1.00 \text{ logit} < \text{Measure} < +1.00 \text{ logit}$, dan butir soal dikategorikan sukar apabila nilai mendekati $+2.00 \text{ logit}$ (Hambleton et al., 1991).

Berdasarkan hasil analisis logit instrumen kemampuan berpikir kritis, maka diketahui terdapat 20% butir soal yang termasuk kategori tingkat kesukaran sulit, 50% butir soal termasuk kategori sedang, 20% butir soal termasuk kategori mudah, dan 10% butir soal termasuk kategori sangat mudah. Rata-rata tingkat kesukaran keseluruhan butir soal sebesar 0.00 logit dengan standar deviasi 1.25 logit. *Measure* pada butir soal nomor 8 memiliki logit -2.45 yang menunjukkan bahwa butir soal nomor 8 termasuk kategori soal sangat mudah dan butir soal nomor 5 memiliki logit paling tinggi di antara butir soal lainnya, yaitu sebesar 1.86 yang menunjukkan bahwa butir soal nomor 5 termasuk kategori sulit.

3.6.6 Fungsi Informasi

Terkait fungsi informasi yang berkaitan dalam menyatakan kekuatan suatu butir soal pada suatu instrumen tes. Fungsi informasi dan *Standard Error of Measurement* (SEM) memiliki hubungan yang berbanding terbalik secara kuadratik, yaitu semakin besar fungsi informasi, maka semakin kecil SEM yang dimilikinya. Potongan antara kurva fungsi informasi dengan SEM dipandang sebagai kelayakan instrumen dalam mengukur kemampuan peserta didik (Retnawati, 2017). Di luar rentang tersebut, tes kemampuan berpikir kritis materi suhu dan kalor tidak akan cocok dengan peserta didik dikarenakan nilai SEM lebih besar daripada nilai fungsi informasi. Berikut grafik fungsi informasi instrumen tes kemampuan berpikir kritis, disajikan pada Gambar 3.7.



Gambar 3.7 Fungsi informasi instrumen kemampuan berpikir kritis

Berdasarkan Gambar 3.7, kurva *Test Information Function* (TIF) memiliki titik puncak atau informasi paling tinggi pada teta sekitar 0.70 logit. Hal ini menunjukkan bahwa informasi tes instrumen paling besar ketika digunakan atau diuji cobakan kepada peserta didik yang memiliki kemampuan sekitar 0.70 logit. Titik potong kedua kurva yaitu pada teta -5.51 dan $+5.53$. Pada rentang teta ini, nilai fungsi informasi tes lebih besar daripada standar error nya. Suatu tes dapat dipercaya ketika nilai informasi yang disampaikan lebih dominan dibandingkan dengan error

(Arlinwibowo et al., 2021). Sehingga dapat disimpulkan bahwa instrumen reliabel untuk peserta didik yang memiliki kemampuan -5.51 logit sampai dengan $+5.53$ logit

3.6.7 Pengambilan Keputusan Butir Soal

Butir soal dikatakan fit dengan model apabila butir soal yang dikatakan cocok adalah butir yang memiliki nilai *outfit mean square* $0.5 < \text{MNSQ} < 1.5$, nilai *outfit Z Standardized* $-2.0 < \text{ZSTD} < 2.0$, dan nilai *Point Measure Correlation (Pt Mean Corr)* tidak negatif (Sumintono & Widhiarso, 2015). Apabila butir soal belum memenuhi ketiga syarat fit model, maka instrumen tetap dapat diterima dengan mempertimbangkan syarat fit model lainnya. Berikut hasil analisis parameter butir soal dapat dilihat pada Tabel 3.13.

Tabel 3.13 Hasil analisis parameter butir soal

Kode Item	Measure	SEM	Outfit MNSQ	Outfit ZSTD	Point Measure Correlation	Keputusan
S1	-0.98	0.31	1.45	1.75	0.87	Diterima
S2	0.12	0.30	0.69	-1.53	0.83	Diterima
S3	0.39	0.30	1.11	0.56	0.67	Diterima
S4	0.12	0.30	0.51	-2.69	0.80	Diterima
S5	1.86	0.31	1.05	0.28	0.68	Diterima
S6	1.02	0.30	1.00	0.09	0.80	Diterima
S7	0.93	0.30	0.83	-0.73	0.81	Diterima
S8	-2.45	0.34	0.54	-1.49	0.84	Diterima
S9	0.66	0.30	1.42	1.78	0.60	Diterima
S10	-1.68	0.32	1.10	0.46	0.75	Diterima

Berdasarkan Tabel 3.13, ada satu butir soal yang tidak memenuhi ketiga syarat fit model. *Outfit ZSTD* pada soal nomor 4 menunjukkan nilai -2.69 (berada diluar rentang syarat *outfit ZSTD*) yang berarti soal nomor 4 tidak kompatibel dengan model yang diharapkan. Tetapi, soal nomor 4 memenuhi syarat *outfit MNSQ* dan *Pt Measure Corr* yaitu dengan nilai *outfit MNSQ* sebesar 0.51 dan nilai *Pt Measure Corr* sebesar 0.80 . Dengan mempertimbangkan syarat item fit lainnya, maka soal nomor 4 akan tetap digunakan dan juga secara keseluruhan item pada validitas instrumen memiliki nilai rata-rata *infit outfit ZSTD* yang sudah sesuai dengan syarat fit. Maka dari itu, seluruh soal uji coba dalam instrumen tes kemampuan berpikir kritis akan diterima dan tetap digunakan.

Butir soal dengan nilai *Pt Measure Corr* negatif mengindikasikan butir soal yang menyesatkan karena peserta tes dengan kemampuan rendah mampu menjawab butir dengan benar dan peserta tes dengan kemampuan tinggi justru menjawab salah (Sumintono & Widhiarso, 2015). Soal-soal dengan nilai korelasi negatif harus diperiksa untuk melihat apakah kunci jawaban salah, perlu direvisi, atau dihapus. Dikarenakan *Pt Measure Corr* seluruh butir soal instrumen tes kemampuan berpikir kritis tidak bernilai negatif, maka peneliti akan menggunakan seluruh butir soal untuk menguji kemampuan berpikir kritis peserta didik.

3.7 Teknik Pengumpulan Data

Teknik pengumpulan data secara lengkap dijelaskan pada Tabel 3.14.

Tabel 3.14 Teknik pengumpulan data

Jenis Instrumen		Sumber Data	Tujuan	Waktu
<i>Test</i>	Tes kemampuan berpikir kritis	Peserta didik	Menunjukkan data kemampuan berpikir kritis yang dimiliki peserta didik	Awal dan akhir pembelajaran
	Lembar observasi keterlaksanaan pembelajaran	Peserta didik dan guru	Menunjukkan keterlaksanaan model pembelajaran yang diterapkan	Selama proses pembelajaran
<i>Non-test</i>	Angket respon peserta didik	Peserta didik	Menunjukkan hasil respon peserta didik untuk mengetahui apakah hasil belajar dirasa efektif menggunakan model pembelajaran yang diterapkan	Akhir pembelajaran

3.8 Analisis Data Penelitian

3.8.1 Analisis Keterlaksanaan Model Pembelajaran

Lembar observasi digunakan sebagai alat ukur untuk melihat keterlaksanaan proses pembelajaran dengan menerapkan model *Problem Based Learning* (PBL) selama kegiatan pembelajaran berlangsung. Persentase terlaksananya dirumuskan sebagai berikut menurut Fakhrunnisa dan Mahmudi (2016).

$$\text{Persentase keterlaksanaan} = \frac{\sum \text{kegiatan yang terlaksana}}{\sum \text{kegiatan keseluruhan}} \times 100\%$$

Dengan klasifikasi keterlaksanaan model pembelajaran seperti disajikan pada Tabel 3.15.

Tabel 3.15 Klasifikasi keterlaksanaan model pembelajaran

Persentase (%)	Kategori
$x \leq 25.00$	Sangat Kurang
$25.00 < x \leq 37.60$	Kurang
$37.60 < x \leq 62.60$	Sedang
$62.60 < x \leq 87.60$	Baik
$87.60 < x \leq 100.00$	Sangat Baik

(Koswara , 2015)

3.8.2 Analisis Peningkatan Kemampuan Berpikir Kritis Peserta Didik

3.7.2.1 Uji Normalitas

Uji normalitas dilakukan untuk mengetahui apakah sampel yang diambil dalam penelitian berdistribusi normal atau tidak. Uji normalitas populasi harus dipenuhi sebagai syarat untuk menentukan perhitungan yang akan dilakukan pada hipotesis berikutnya. Untuk menguji normalitas dapat menggunakan *software* IBM SPSS 23 dengan analisis uji Shapiro Wilk. Uji Shapiro Wilk adalah sebuah metode atau rumus perhitungan sebaran data yang dibuat oleh Shapiro dan Wilk. Menurut Razali, N.M & Wah, Y.B. (2011) menyatakan bahwa uji Shapiro dan Wilk pada umumnya digunakan untuk sampel yang jumlahnya kecil atau kurang dari 50 data. Dikarenakan sampel pada penelitian ini berjumlah 33 peserta didik atau < 50 , maka uji Shapiro Wilk merupakan metode uji normalitas yang efektif untuk penelitian ini. Dasar pengambilan keputusan jika nilai signifikan $> 0,05$, maka data penelitian berdistribusi normal. Jika nilai signifikan $< 0,05$, maka data penelitian tidak berdistribusi normal.

3.7.2.2 Uji Hipotesis

Uji hipotesis ini menggunakan uji beda dua rata-rata untuk melihat adakah perbedaan signifikan antara rata-rata nilai *pre-test* dan *post-test* dari kelompok yang diuji kemampuan berpikir kritis setelah dilakukan pembelajaran dengan model PBL. Untuk data hasil kemampuan berpikir kritis, uji hipotesis menggunakan uji Wilcoxon jika data tidak berdistribusi normal dan menggunakan uji paired sample T test jika data berdistribusi normal. Hasil analisis output yang telah didapatkan, diinterpretasikan sebagai berikut:

- a) H_0 = Tidak terdapat perbedaan yang signifikan terhadap penerapan model PBL dalam meningkatkan kemampuan berpikir kritis peserta didik.
- b) H_1 = Terdapat perbedaan yang signifikan terhadap penerapan model PBL dalam meningkatkan kemampuan berpikir kritis peserta didik.

Dari hasil uji hipotesis untuk kemampuan berpikir kritis, jika sudah diketahui bahwa terdapat perbedaan rata-rata yang signifikan antara *pre-test* dan *post-test*, maka dapat dilanjutkan dengan uji N-Gain.

3.7.2.3 Uji N-gain

Dilakukannya uji N-Gain bertujuan untuk mengetahui peningkatan dari kemampuan berpikir kritis peserta didik setelah penerapan model pembelajaran (Sundayana, 2018). Peningkatan Kemampuan Memahami diperoleh berdasarkan hasil *pre-test* dan *post-test* yang telah dikerjakan oleh peserta didik. Nilai skor n-gain dapat dinyatakan menggunakan rumus berikut (Hake, 2002).

$$\text{N-Gain} = \frac{\text{Skor Posttest} - \text{Skor Pretest}}{\text{Skor Ideal} - \text{Skor Pretest}} \times 100 \%$$

Keterangan:

Skor ideal = Skor maksimum yang dapat diperoleh

Berikut adalah tingkatan skor n-gain yang telah dibagi menjadi 3 kategori berdasarkan Hake (2002).

Tabel 3.16 Kriteria skor N-Gain

N-Gain	Kategori
$g < 0,3$	Rendah
$0,3 \leq g < 0,7$	Sedang
$g \geq 0,7$	Tinggi

Keterangan: $\langle g \rangle$ = skor N-Gain

3.8.3 Analisis Efektivitas Penerapan Model *Problem Based Learning*

Efektivitas model pembelajaran yang diterapkan dapat diketahui dengan digunakan uji *effect size*. *Effect size* adalah ukuran untuk mengetahui besarnya efek suatu variabel pada variabel lain, besarnya perbedaan ataupun hubungan yang bebas dari pengaruh jumlah sampel (Aldila & Mukhaiyar, 2020). Berikut rumus yang digunakan dalam uji *effect size* menurut Cohen (1988).

$$d = \frac{M_{posttest} - M_{pretest}}{\sqrt{\frac{SD_{posttest}^2 + SD_{pretest}^2}{2}}}$$

Keterangan:

d = *Effect size*

M = Rata-rata

SD = Standar deviasi

Hasil perhitungan *effect size* yang telah didapatkan kemudian dapat diinterpretasikan berdasarkan kategori berikut.

Tabel 3.17 Kategori *d value*

Ukuran Efek (d)	Kategori
$0,8 \leq d < 2,0$	Tinggi
$0,5 \leq d < 0,8$	Sedang
$d < 0,5$	Rendah

(Becker, 2000)

3.8.4 Analisis Hasil Angket Respon Peserta Didik

Pada penelitian ini angket respon peserta didik dilakukan untuk mengetahui apakah hasil belajar dirasa efektif menggunakan model pembelajaran yang

diterapkan dengan menggunakan kuesioner tertutup yang diisi oleh peserta didik yang telah mendapatkan perlakuan, yaitu pembelajaran materi suhu dan kalor menggunakan model *problem based learning*. Pernyataan pada angket terdiri atas 8 pernyataan positif dan 8 pernyataan negatif yang diukur menggunakan skala likert. Skala likert digunakan untuk mengukur pendapat dan persepsi seseorang atau sekelompok tentang suatu kejadian (Asnawi, 2018). Berikut skala likert untuk angket respon peserta didik pada penelitian ini, disajikan pada Tabel 3.18.

Tabel 3.18 Skala likert angket respon

Kriteria	Skala	
	Pernyataan Positif	Pernyataan Negatif
Sangat Setuju	4	1
Setuju	3	2
Tidak Setuju	2	3
Sangat Tidak Setuju	1	4