

BAB III

METODE PENELITIAN

Bab ini memaparkan mengenai desain penelitian, partisipan, instrumen penelitian, prosedur penelitian, dan teknik analisis data.

3.1 Desain Penelitian

Metode penelitian adalah suatu pendekatan ilmiah yang digunakan untuk mengumpulkan data untuk tujuan dan kegunaan tertentu (Sugiyono, 2022). Penelitian ini menggunakan metode kuasi eksperimen (*Quasi Experiment*), yaitu jenis penelitian semu yang melibatkan kelompok kontrol namun tidak sepenuhnya mengendalikan variabel eksternal (Sugiyono, 2016). Desain yang digunakan adalah *non-equivalent control group design* (Sugiyono, 2013). Desain ini terdiri dari dua jenis kelompok yaitu kelompok eksperimen dan kelompok kontrol. Penelitian ini akan diawali dengan melakukan tes awal (*pretest*), dilanjutkan dengan pemberian perlakuan atau *treatment* dan diakhiri dengan tes akhir (*posttest*). Kelompok eksperimen pada penelitian ini merupakan kelas yang mendapat pembelajaran model PBL-POE berbasis *Augmented Reality* (AR), sedangkan kelompok kontrolnya merupakan kelas dengan pembelajaran model PBL. Sebelum dan setelah pembelajaran, kedua kelas diberikan tes berpikir tingkat tinggi dan kuesioner motivasi belajar peserta didik. Skor tes berpikir tingkat tinggi dan motivasi belajar peserta didik dari kedua kelompok kemudian dianalisis. Adapun rancangan desain penelitian ini menurut Sugiyono (2013) ditunjukkan pada Tabel 3.1 berikut.

Tabel 3.1 Perbedaan *treatment* kelas eksperimen dan kontrol

<i>Eksperimen Group</i>	Y_1	X_1	Y_2
<i>Control Group</i>	Y_1	X_2	Y_2

(Sugiyono, 2013)

Keterangan:

Y_1 : *Pretest*

Y_2 : *Posttest*

X_1 :Penerapan Model PBL-POE berbantuan *E-Modul* berbasis *Augmented Reality* (AR)

X_2 :Penerapan Model PBL (*Problem Based Learning*)

Berdasarkan Tabel 3.1, dapat melihat perbandingan kondisi peserta didik sebelum dan sesudah pemberian *treatment* (Sugiyono, 2011). Pada tahap awal, peserta didik menjalani *pretest* untuk menentukan kondisi awal mereka. Setelah menerima *treatment*, peserta didik kemudian diminta mengerjakan *posttest* untuk mengetahui perubahan dan mengevaluasi kondisi mereka setelah proses pembelajaran.

3.2 Partisipan

Partisipan adalah individu yang memberikan tanggapan terhadap proses belajar mengajar (Fadliyati, 2015). Dalam penelitian ini, partisipan yang terlibat adalah peserta didik SMA kelas X yang akan mempelajari materi Pemanasan Global, yang kemudian dibagi menjadi dua kelas: kelas eksperimen yang menerapkan model PBL-POE berbantuan *E-Modul* berbasis AR dan kelas kontrol yang menerapkan model PBL.

Penelitian ini dilaksanakan di salah satu SMA negeri di Kota Sukabumi. Terdapat 72 peserta didik kelas X yang berpartisipasi dalam penelitian ini. Dari jumlah tersebut, 36 peserta didik kelas X-1 ditunjuk sebagai partisipan untuk kelas eksperimen, sedangkan 36 peserta didik lainnya kelas X-6 akan menjadi partisipan untuk kelas kontrol.

4.4 Populasi dan Sampel

Populasi merujuk kepada keseluruhan subjek yang menjadi fokus penelitian (Sugiyono 2013: 80). Populasi dalam penelitian ini adalah peserta didik kelas X di salah satu SMA negeri di Kota Sukabumi yang akan mempelajari materi Pemanasan Global. Sedangkan sampel adalah bagian dari populasi yang memiliki karakteristik yang sama (Sugiyono, 2013:81). Teknik pemilihan sampel yang digunakan adalah *purposive sampling*, yang berarti peneliti memilih sampel berdasarkan pertimbangan tertentu. Total sampel dalam penelitian ini adalah 72 peserta didik, terdiri dari 36 peserta didik di kelas eksperimen yaitu kelas X-1 yang mendapat

treatment penerapan model PBL-POE berbantuan *E-Modul* berbasis AR dan 36 peserta didik di kelas kontrol yaitu kelas X-6 yang menerapkan model PBL.

4.5 Variabel Penelitian

Variabel dalam penelitian ini meliputi variabel bebas dan terikat. Model PBL-POE berbantuan *E-Modul* berbasis *Augmented Reality* (AR) pada materi Pemanasan Global adalah variabel bebas (*independent variable*), sementara berpikir tingkat tinggi dan motivasi belajar peserta didik adalah variabel terikat (*dependent variable*).

4.6 Instrumen Penelitian

Instrumen penelitian adalah alat yang digunakan untuk mengukur fenomena yang sedang diamati, baik itu alamiah atau sosial (Sugiyono 2013). Instrumen yang digunakan dalam penelitian ini disusun berdasarkan tujuan penelitian yang ingin dicapai. Jenis instrumen yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.2. berikut.

Tabel 3.2 Jenis Instrumen Penelitian

No	Pertanyaan Penelitian	Data	Bentuk Instrumen	Sumber Data	Waktu
1.	Bagaimana kelayakan <i>E-Modul</i> berbasis <i>Augmented Reality</i> (AR)?	Validasi berbasis <i>Augmented Reality</i> (AR)?	Lembar Validasi <i>E-Modul</i> berbasis <i>Augmented Reality</i> (AR)?	Empat orang validator	Pada tahap persiapan sebelum melakukan <i>treatment</i>
2.	Bagaimana keterlaksanaan penerapan model PBL-POE berbantuan <i>E-Modul</i> berbasis <i>Augmented Reality</i> pada kelas eksperimen dan penerapan model PBL pada kelas kontrol?	Persentase Keterlaksanaan model pembelajaran kedua kelas	Lembar observasi peserta didik dan guru	Dua orang observer	Setiap pertemuan proses pembelajaran
3.	Bagaimana efektivitas model PBL-POE berbantuan <i>E-Modul</i> berbasis <i>Augmented Reality</i> terhadap peningkatan berpikir tingkat tinggi pada materi Pemanasan Global?	Validasi Instrumen dan <i>Pretest-posttest</i> berpikir tingkat tinggi peserta didik kedua kelas	Lembar validasi instrumen dan Soal tipe <i>two tier multiple choice</i> 23 soal materi pemanasan global	Validator dan Peserta didik	Sebelum (<i>pretest</i>) dan setelah (<i>posttest</i>) <i>treatment</i>

No	Pertanyaan Penelitian	Data	Bentuk Instrumen	Sumber Data	Waktu
4.	Bagaimana efektivitas model PBL-POE berbantuan <i>E-Modul</i> berbasis AR terhadap peningkatan motivasi belajar pada materi Pemanasan Global?	<i>Pretest-posttest</i> motivasi belajar peserta didik kedua kelas	Kuesioner motivasi belajar yang terdiri dari 20 soal menggunakan skala likert 1-5	Peserta didik	Sebelum (<i>pre-test</i>) dan setelah (<i>posttest</i>) <i>treatment</i>

Uraian dari setiap jenis instrumen yang digunakan pada penelitian adalah sebagai berikut:

3.5.1 Lembar Angket Validasi *E-Modul* Berbasis *Augmented Reality* (AR)

Media pembelajaran yang digunakan pada penelitian ini merupakan media pembelajaran yang dikembangkan bersama dengan tim pengembang media pembelajaran pada penelitian payung yang diketuai oleh ibu Lina Aviyanti, Ph.D. Media pembelajaran yang dikembangkan berupa *E-Modul* berbasis AR yang terintegrasi dengan LKPD berpikir tingkat tinggi. *E-Modul* yang dibuat terdiri dari tiga materi yang membahas tentang (1) gejala dan dampak pemanasan global, (2) penyebab perubahan lingkungan, dan (3) solusi menanggulangi pemanasan global. *E-Modul* ini menggunakan teknologi *WebGL* dan *JavaScript* yang dilengkapi dengan *Augmented Reality* yang bertujuan untuk menciptakan pengalaman belajar yang lebih interaktif, serta untuk menyederhanakan penjelasan melalui 3D animasi untuk memberikan gambaran menyeluruh terkait beberapa konsep yang kompleks. Selain teknologi AR, *E-modul* ini juga diperkaya dengan berbagai elemen multimedia seperti gambar, video, dan AR. Video dan gambar untuk mendukung pemahaman peserta didik dalam merepresentasikan sebuah fenomena nyata. Di bagian akhir setiap *e-modul* terdapat LKPD yang masing-masing bertujuan untuk melatih berpikir tingkat tinggi peserta didik seperti kemampuan menganalisis (C4), mengevaluasi (C5), dan mencipta (C6).

Sebelum media pembelajaran ini digunakan, akan dilakukan validasi oleh validator untuk mengevaluasi dan melihat bagaimana kelayakan dari media pembelajaran ini menggunakan angket validasi kelayakan *e-modul* berbasis AR. Lembar angket validasi *e-modul* ini terdiri dari 20 pertanyaan yang dikelompokkan menjadi 4 aspek yaitu (1) kelayakan isi, (2) kelayakan penyajian, (3) kelayakan kegrafikan, dan (4) kebahasaan seperti yang ditunjukkan pada Tabel 3.3 berikut.

Tabel 3.3 Kisi-Kisi Angket Validasi *E-Modul* Berbasis AR

No	Aspek yang dinilai	No Soal
1.	Kelayakan Isi	1, 2, 3, 4, 5
2.	Kelayakan Penyajian	6, 7, 8, 9, 10
3.	Kelayakan Kegrafikan	11, 12, 13, 14, 15

Ahmad Rifa'i, 2025

EFEKTIVITAS PBL-POE BERBANTUAN E-MODUL BERBASIS AUGMENTED REALITY TERHADAP PENINGKATAN BERPIKIR TINGKAT TINGGI DAN MOTIVASI BELAJAR PESERTA DIDIK PADA MATERI PEMANASAN GLOBAL
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

No	Aspek yang dinilai	No Soal
4.	Kebahasaan	16, 17, 18, 19, 20

(Dimodifikasi dari Tiara, 2022)

Berikut contoh cuplikan lembar validasi *e-modul* berbasis AR ditunjukkan pada Gambar 3.1 berikut.

LEMBAR PENILAIAN		Nilai Pengamatan				
Aspek	Indikator	1	2	3	4	5
Kelayakan Isi	1. Kesesuaian materi Pemanasan Global yang disajikan dengan Tujuan Pembelajaran.					
	2. Kesesuaian materi Pemanasan Global yang disajikan dengan Keterampilan Berpikir Tingkat Tinggi (C4-C5).					
	3. Kesesuaian materi yang disajikan dalam e-modul IMOGLOW berbasis AR dengan kebenaran teori atau konsep Pemanasan Global.					
	4. E-Modul IMOGLOW berbasis AR membantu penyajian konsep Pemanasan Global yang abstrak atau sulit dipahami tanpa menimbulkan miskonsepsi.					
	5. Penggunaan multimedia (gambar, video, animasi/simulasi dan tampilan AR) yang disajikan mendukung pemahaman materi Pemanasan Global secara jelas.					
Kelayakan Penyajian	6. E-Modul IMOGLOW berbasis AR memfasilitasi pembelajaran lebih interaktif.					
	7. E-Modul IMOGLOW berbasis AR dapat diakses secara mandiri dimana pun dan kapan saja.					
	8. E-Modul IMOGLOW berbasis AR menyediakan panduan penggunaan yang jelas dan mudah diikuti oleh pengguna.					
	9. Penggunaan multimedia (gambar, video, animasi/simulasi dan tampilan AR) dengan kombinasi warna yang disajikan menarik dan terstruktur secara sistematis.					
	10. Penggunaan multimedia (gambar, video, animasi/simulasi dan tampilan AR)					

Gambar 3.1 Cuplikan Lembar Validasi *E-Modul* Berbasis AR

3.5.2 Lembar Observasi Keterlaksanaan Pembelajaran

Lembar observasi keterlaksanaan pembelajaran digunakan untuk mengukur ketercapaian penerapan model PBL-POE berbantuan *E-Modul* berbasis AR pada kelas eksperimen dan penerapan model PBL pada kelas kontrol selama proses pembelajaran berlangsung. Observasi keterlaksanaan pembelajaran menggunakan lembar observasi yang terdiri dari dua jenis lembar observasi yang digunakan yaitu lembar observasi untuk kegiatan peserta didik dan lembar observasi untuk kegiatan guru. Lembar observasi ini terdapat pernyataan dari setiap tahapan pembelajaran dimulai dari pendahuluan, inti, dan penutup disertai dengan kriteria penilaian keterlaksanaan berupa "terlaksana" dan "tidak terlaksana", serta adanya kolom keterangan disetiap tahapan pembelajaran untuk memberikan saran dan komentar. Lembar observasi dibuat setiap pertemuan yang kemudian diisi oleh dua observer (guru dan teman sejawat) yang mengikuti proses pembelajaran pada setiap

pertemuan. Berikut cuplikan lembar observasi keterlaksanaan pembelajaran ditunjukkan pada Gambar 3.2.

**LEMBAR OBSERVASI KETERLAKSANAAN AKTIVITAS GURU PADA
PEMBELAJARAN MODEL PBL-POE BERBANTUAN E-MODUL FLIPBOOK
BERBASIS AR**

Nama Peneliti : Ahmad Rifa'i
 Pertemuan ke- : 1
 Materi : Gejala dan Dampak Pemanasan Global

Petunjuk pengisian:

1. Berdasarkan pendapat Bapak/Ibu, mohon berikan tanda centang (✓) pada kolom "Terlaksana" jika kegiatan dilaksanakan atau kolom "Tidak Terlaksana" jika kegiatan tidak dilaksanakan.
2. Menuliskan keterangan pada setiap kegiatan jika diperlukan.
3. Apabila pada saat pembelajaran diperlukan perbaikan atau masukan, mohon berikan komentar atau masukan secara bebas dan ditulis pada kolom "Catatan"

Kegiatan	Tahapan Pembelajaran	Aktivitas Guru	Terlaksana	Tidak Terlaksana	Keterangan
Pendahuluan	Orientasi	Guru mengucapkan salam dan meminta salah satu perwakilan peserta didik memimpin do'a.			
		Guru memeriksa kehadiran peserta didik.			

Gambar 3.2 Cuplikan Lembar Observasi Keterlaksanaan Pembelajaran

3.5.3 Instrumen Berpikir Tingkat Tinggi *Global Warming*

Tes evaluasi berpikir tingkat tinggi yang digunakan adalah *Two-tier Multiple Choice* dengan jumlah 30 butir soal yang disesuaikan dengan kompetensi berpikir tingkat tinggi. Instrumen tes yang dikembangkan oleh peneliti untuk mengukur berpikir tingkat tinggi yaitu soal pilihan ganda bertingkat dua (*Two-tier Multiple Choice*), tingkat pertama (*first-tier*) terdiri dari pertanyaan dengan lima pilihan jawaban dan tingkat kedua (*second-tier*) terdiri dari lima pilihan alasan yang mengacu pada jawaban tingkat pertama. Pemilihan instrumen *Two-tier Multiple Choice* didasarkan pada hasil penelitian Nanda dkk (2019) yang menunjukkan bahwa instrumen *Two-tier Multiple Choice* membantu peserta didik untuk lebih mendalam memahami konsep fisika dan mengurangi miskonsepsi serta mampu

melihat kemampuan berpikir tingkat tinggi peserta didik. Berikut Tabel 3.4 mengenai kisi-kisi instrumen berpikir tingkat tinggi yang disesuaikan dengan klasifikasi berpikir tingkat tinggi menurut Anderson (2001) dan Brookhart (2010).

Tabel 3.4 Kisi-Kisi Instrumen Berpikir Tingkat Tinggi *Global Warming*

No	Klasifikasi	Indikator	No Item
1	Analisis	<ul style="list-style-type: none"> Menspesifikasikan Elemen-Elemen. Kata Kerja: menganalisis, menghubungkan, Menyusun, menguraikan, memilih	1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 16, 17, 18, 30
2	Evaluasi	<ul style="list-style-type: none"> Mengambil keputusan sendiri. Kata Kerja: memeriksa, mendeteksi, mengevaluasi	4, 14, 15, 19, 20, 21, 22, 23,
3	Mencipta	<ul style="list-style-type: none"> Mengkreasi ide/gagasan sendiri. Kata kerja: membuat, merencanakan, merancang, menciptakan,	5, 24, 25, 26, 27, 28, 29,

(Iskandar, 2015)

Setiap pertanyaan hanya akan diberi skor 1 jika jawaban pilihan dan alasan benar. Perhitungan data dan rubrik penilaian dapat dilihat pada Tabel 3.5 berikut.

Tabel 3.5 Skoring Soal *Two Tier Multiple Choice Berpikir Tingkat Tinggi Global Warming*

No	Distribusi Soal		Penskoran
	Jawab	Pendukung	
1	Benar	Benar	1
2	Salah	Benar	0
3	Benar	Salah	0
4	Salah	Salah	0

3.5.4 Kuesioner Motivasi Belajar Peserta Didik

Kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan atau pernyataan tertulis kepada responden untuk dijawabnya (Sukmadinata, 2005:219). Alat tes untuk mengukur motivasi belajar

yang digunakan mengadaptasi dari penelitian yang dilakukan Aviyanti dkk (2024) berupa kuesioner motivasi belajar yang terdiri dari 20 pertanyaan dan menggunakan skala likert 1-5 yaitu, “Tidak Pernah” (1), “Jarang”(2), “Kadang-kadang” (3), “Sering” (4), dan “Selalu” (5) yang disesuaikan dengan indikator motivasi belajar. Kisi-kisi motivasi belajar peserta didik yang digunakan ditunjukkan pada Tabel 3.6 berikut.

Tabel 3.6 Kisi-Kisi Kuesioner Motivasi Belajar

No.	Indikator	No Item	Jumlah
1.	Ketekunan dalam belajar	1,2,3	3
2.	Ulet dalam menghadapi kesulitan	4,5,6	3
3.	Minat terhadap pelajaran	7,8,9	3
4.	Dapat mempertahankan pendapatnya	10,11,12	3
5.	Tidak mudah melepaskan hal yang diyakini	13,14	2
7.	Senang mencari dan memecahkan masalah soal-soal	15,16,17	3
8.	Berprestasi dalam belajar	18,19,20	3
Jumlah Butir			20

Adapun format kuesioner motivasi belajar peserta didik adalah sebagai berikut

Tabel 3.7 Kuesioner Motivasi Belajar Peserta Didik

No.	Pernyataan	SL	S	KK	J	TP
1.	Saya tetap mengikuti pelajaran, siapapun guru yang mengajarnya					
2.	Saya tidak mengikuti pelajaran, jika pelajaran itu tidak saya sukai					
3.	Saya belajar di luar jam sekolah jika ada tugas dan ulangan saja					
4.	Saya selalu mencoba berulang kali dalam mengerjakan soal fisika yang sulit					
5.	Saya bertanya kepada guru ketika ada penjelasan materi yang tidak saya pahami					
6.	Saya berdiskusi dengan teman jika menemukan kesulitan dalam belajar					
7.	Saya selalu mendengarkan penjelasan guru dengan baik					

No.	Pernyataan	SL	S	KK	J	TP
8.	Saya lebih senang berbicara dengan teman saat pelajaran di kelas					
9.	Saya jarang membaca materi yang akan diajarkan sebelum pembelajaran berlangsung					
10.	Saya yakin atas jawaban tugas yang saya kerjakan					
11.	Saya memilih tidak berdiskusi jika hasil pekerjaan saya berbeda dengan teman					
12.	Saya dapat menjelaskan alasan atau argumen atas jawaban saya					
13.	Saya mengabaikan pendapat teman jika pendapat mereka tidak sesuai dengan pemikiran saya					
14.	Saya yakin dengan rajin berlatih soal-soal akan membuat saya lebih memahami materi pelajaran					
15.	Saya berlatih mengerjakan soal-soal sampai saya bisa					
16.	Saya mencari soal-soal yang menantang (lebih sulit) untuk dikerjakan					
17.	Saya tetap belajar meskipun tidak ada ulangan					
18.	Saya selalu merasa tidak puas dan selalu ingin memperoleh hasil yang lebih baik lagi					
19.	Saya merasa biasa ketika memperoleh nilai yang kurang memuaskan					
20.	Prestasi tinggi dalam belajar, saya peroleh dengan usaha keras saya sendiri					

4.6 Prosedur Penelitian

Prosedur penelitian yang dilakukan terdiri dari tiga tahap sebagai berikut:

3.6.1 Tahap Persiapan

- Melakukan telaah pustaka mengenai model PBL-POE, *e-modul* interaktif, berpikir tingkat tinggi dan motivasi belajar peserta didik.
- Melakukan studi pendahuluan di sekolah yang akan dijadikan tempat penelitian melalui uji coba soal berpikir tingkat tinggi yang terdiri dari 30 butir soal *two tier multiple choice* wawancara dengan guru fisika untuk

mengetahui bagaimana berpikir tingkat tinggi, motivasi belajar, peserta didik, model pembelajaran serta media pembelajaran yang sering digunakan, soal-soal ujian yang digunakan, dan kendala dalam melatih berpikir tingkat tinggi di kelas.

- Merancang perangkat pembelajaran yang meliputi silabus, Rencana Pelaksanaan Pembelajaran (RPP), berupa model PBL-POE berbantuan *E-Modul* berbasis *Augmented Reality* untuk kelas eksperimen dan model PBL untuk kelas kontrol.
- Menyusun instrumen penelitian berupa *tes tipe two tier multiple choice* untuk mengukur berpikir tingkat tinggi pada materi pemanasan global serta membuat kuesioner motivasi belajar peserta didik.
- Melakukan uji coba instrumen tes dan memvalidasi media pembelajaran berupa *E-Modul* berbasis AR.
- Mengolah data hasil uji coba yang meliputi, validitas item soal, tingkat kesukaran, reliabilitas, daya pembeda, serta hasil validasi kelayakan *e-modul*, kemudian menganalisisnya dan menentukan soal yang akan digunakan dalam penelitian.

3.6.2 Tahap Pelaksanaan

Pada tahap ini dilakukan kegiatan sebagai berikut:

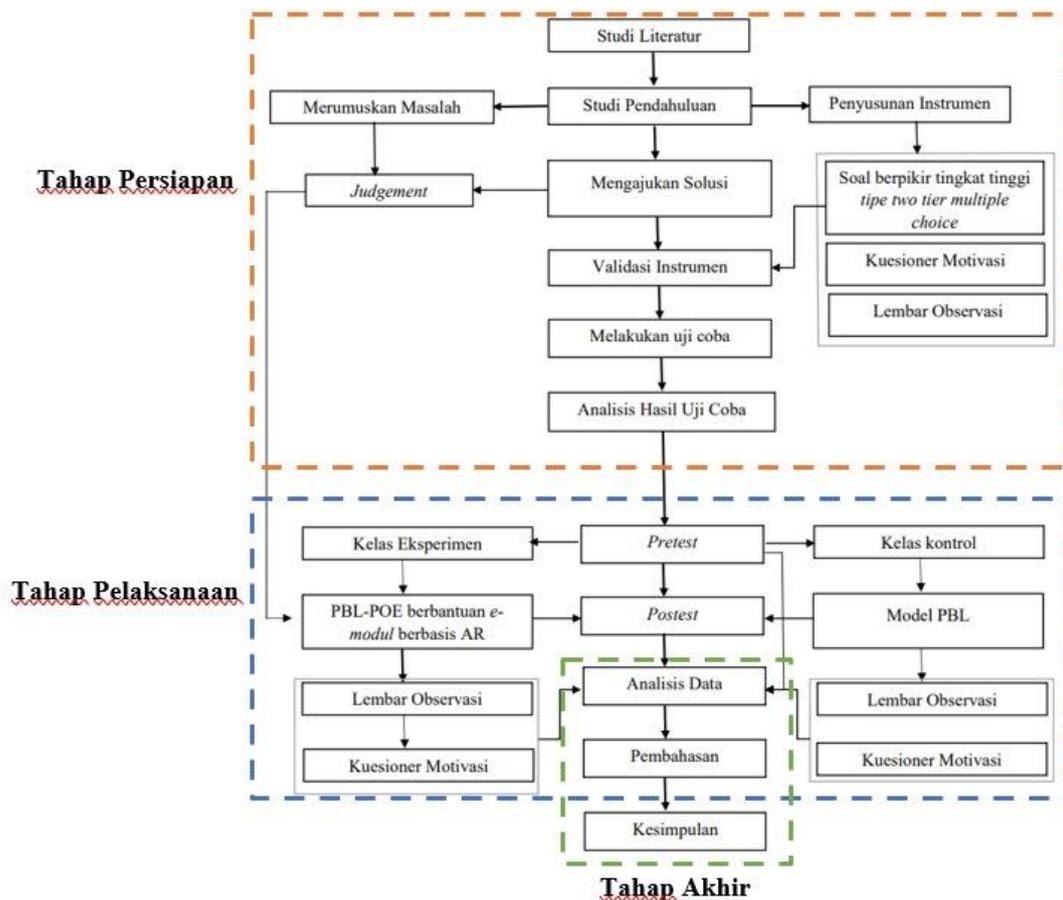
- Memberikan soal *pre-test* berpikir tingkat tinggi dan kuesioner motivasi belajar peserta didik pada kelas eksperimen dan kelas kontrol.
- Mengolah data hasil *pre-test* dan hasil kuesioner pada kelas eksperimen dan kontrol
- Memberikan perlakuan (*treatment*) dengan penggunaan model PBL-POE berbantuan *E-Modul* berbasis *Augmented Reality* pada kelas eksperimen dan model PBL pada kelas kontrol. Selama kegiatan pembelajaran dilakukan observasi keterlaksanaan pembelajaran oleh guru dan teman sejawat.
- Memberikan *post-test* untuk mengetahui perbededaan berpikir tingkat tinggi dan motivasi belajar peserta didik setelah mendapat *treatment*.

3.6.3 Tahap Akhir

Pada tahap ini dilakukan kegiatan sebagai berikut:

- Mengolah data hasil *pretest* dan *post-test*, hasil kuesioner motivasi belajar peserta didik, dan hasil observasi keterlaksanaan pembelajaran dari seluruh pembelajaran yang dilakukan pada kedua kelompok sampel.
- Menganalisis data hasil penelitian dan membahas temuan penelitian.
- Memberikan kesimpulan berdasarkan hasil pengolahan data.
- Memberikan rekomendasi berdasarkan hasil penelitian.

Berikut alur penelitian yang digunakan dalam penelitian digambarkan sebagai berikut:



Gambar 3.3 Prosedur Penelitian

4.7 Teknik Analisis Data

Analisis data pada penelitian ini meliputi analisis lembar kelayakan *E-Modul* berbasis AR, analisis instrumen berpikir tingkat tinggi *global warming*, analisis observasi keterlaksanaan pembelajaran, analisis efektivitas model PBL-POE berbantuan *E-Modul* berbasis *Augmented Reality* (AR) terhadap peningkatan berpikir tingkat tinggi dan motivasi belajar peserta didik pada materi pemanasan global yang dijabarkan sebagai berikut.

3.7.1 Analisis Kelayakan *E-Modul* Berbasis *Augmented Reality* (AR)

Kelayakan *E-Modul* berbasis AR akan divalidasi oleh validator yang terdiri dari dosen dan guru. Lembar validasi *E-Modul* dan LKPD berpikir tingkat tinggi menggunakan skala likert 1-5 yaitu Sangat Sesuai (Skor 5), Sesuai (Skor 4), Cukup Sesuai (Skor 3), Kurang Sesuai (2), dan Tidak Sesuai (Skor 1). Skor tersebut divalidasi pada setiap aspek menggunakan persamaan *V-Aiken* sebagai berikut (Aiken, 1985):

$$V = \frac{\sum s}{n(c - 1)} \quad (3.1)$$

Keterangan:

V = Indeks validitas butir

$\sum s$ = (Skor yang diberikan oleh validator) – (Skor terendah dalam kategori yang dipakai)

n = Banyaknya validator ahli

c = Banyaknya kriteria penskoran yang dapat dipilih validator ahli

Pengambilan interpretasi disesuaikan dengan kategori nilai *V-Aiken* yang ditunjukkan pada Tabel 3.8 berikut.

Tabel 3.8 Kriteria nilai *V-Aiken*

<i>Indeks V-Aiken</i>	Kriteria
$V > 0,8$	Tinggi
$0,4 < V \leq 0,8$	Sedang
$V \leq 0,4$	Rendah

(Retnawati, 2016)

Hasil validasi *E-Modul* berbasis AR yang telah diisi oleh empat validator menggunakan analisis *V-Aiken* ditunjukkan pada Tabel 3.9 berikut:

Tabel 3.9 Hasil Validasi Kelayakan *E-Modul* Berbasis AR

No	Indikator	Nilai <i>V-Aiken</i>	Interpretasi
1	Kelayakan Isi	0,93	Tinggi
2	Kelayakan Penyajian	0,93	Tinggi
3	Kelayakan Kegrafikan	0,96	Tinggi
4	Kebahasaan	0,95	Tinggi
Rata-Rata		0,94	Tinggi

Hasil validasi kelayakan *E-Modul* berbasis *Augmented Reality* (AR) menunjukkan nilai *V-Aiken* yang tinggi untuk semua indikator. Rata-rata nilai *V-Aiken* sebesar 0,94 menunjukkan bahwa *e-modul* ini sangat layak digunakan dalam proses pembelajaran.

3.7.2 Analisis Instrumen Berpikir Tingkat Tinggi *Global Warming*

Analisis instrumen berpikir tingkat tinggi *global warming* terdiri atas uji validitas (validitas isi dan validitas konstruk), uji reliabilitas, tingkat kesukaran *item* soal, dan daya pembeda. Berikut merupakan penjelasan analisis instrumen berpikir tingkat tinggi *global warming* yang digunakan.

3.7.2.1 Uji Validitas

Instrumen tes yang valid adalah instrumen tes yang mampu mengukur apa yang harus diukur (Sumintono & Widhiarso, 2015). Uji validitas merupakan langkah pengujian isi suatu instrumen untuk mengukur keakuratan instrumen yang digunakan dalam penelitian (Sugiyono, 2014; Hakim, 2021). Uji validitas dibedakan menjadi dua jenis, yaitu uji validitas isi dan uji validitas konstruk dengan penjelasan sebagai berikut.

3.7.2.1.1 Uji Validitas Isi

Uji validitas isi dilakukan untuk memastikan bahwa *item* soal pada tes yang digunakan benar-benar mewakili domain yang diukur. Uji validitas ini bukan hanya berdasarkan pada konten soal itu sendiri, melainkan juga mempertimbangkan format, penataan kata, administrasi, dan evaluasi tes (Johnson & Christensen, 2020). Validitas isi dari instrumen tes yang telah dikembangkan diuji oleh para ahli.

Validator untuk instrumen berpikir tingkat tinggi *global warming* melibatkan satu dosen pendidikan fisika dan etiga guru fisika. Aspek yang dinilai dalam validasi isi terdiri dari tiga indikator meliputi 5 pertanyaan aspek materi, 5 pertanyaan aspek kontruksi, dan 4 pertanyaan aspek bahasa seperti yang ditunjukkan pada Tabel 3.10 berikut.

Tabel 3.10 Aspek Penilaian Instrumen Berpikir Tingkat Tinggi *Global*

Warming

Aspek yang dinilai	Pertanyaan
Materi	1. Kesesuaian butir soal dengan indikator soal
	2. Indikator soal menggunakan kalimat sesuai dengan aspek berpikir tingkat tinggi yang diukur.
	3. Kesesuaian butir soal dengan aspek berpikir tingkat tinggi yang diukur.
	4. Pilihan jawaban sudah menunjukkan kesesuaian dengan butir soal dan hanya ada satu jawaban yang benar.
	5. Kesesuaian butir soal dengan konsep atau teori fisika.
Kontruksi	6. Butir soal dirumuskan secara jelas dan benar.
	7. Gambar, grafik, tabel atau bagan diilustrasikan dengan jelas dan berfungsi dengan baik.
	8. Butir soal tidak memberikan langkah penyelesaian atau petunjuk jawaban ke arah jawaban yang benar.
	9. Penyelesaian butir soal membutuhkan pemahaman konsep yang benar.
Bahasa	10. Butir soal menuntut kemandirian peserta didik untuk berpikir tingkat tinggi.
	11. Rumusan kalimat dalam butir soal komunikatif dan jelas.
	12. Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian.
	13. Rumusan kalimat dalam butir soal menggunakan istilah atau kata-kata umum yang sesuai dengan kaidah bahasa Indonesia.
	14. Pilihan jawaban menggunakan bahasa Indonesia yang baku dan mudah dipahami.

Setiap ahli diberikan instrumen tes untuk diisi pada lembar validasi dengan skala penilaian dari 1 (tidak sesuai) hingga 5 (sangat sesuai). Hasil uji validitas isi

oleh ahli diolah menggunakan koefisien validasi *Aiken* (V). Dengan empat validator dengan skala 1-5, nilai validitas harus mencapai 0,88 (Aiken, 1985). Jika aspek nomor 1 tidak valid, soal tersebut tidak digunakan. Jika aspek nomor 2-14 tidak valid, soal diperbaiki dan tetap digunakan. Koefisien validasi *Aiken* dihitung dengan persamaan berikut (Aiken, 1985):

$$V = \frac{\sum s}{n(c - 1)} \quad (3.2)$$

Keterangan:

V = Indeks validitas butir

$\sum s$ = (Skor yang diberikan oleh validator) – (Skor terendah dalam kategori yang dipakai)

n = Banyaknya validator ahli

c = Banyaknya kriteria penskoran yang dapat dipilih validator ahli

Hasil analisis menggunakan *V-Aiken* yang telah disii oleh empat validator ditunjukkan pada Tabel 3.11 berikut.

Tabel 3.11 Validitas Soal menggunakan Validitas Ahli

Aspek Penilaian	Berpikir Tingkat Tinggi <i>Global Warming</i>		
	Nilai V	Rata – Rata V	Kategori
Aspek Materi	0.93	0.93	Tinggi
	0.93		
	0.92		
	0.90		
	0.95		
Aspek Konstruksi	0.90	0.90	Tinggi
	0.89		
	0.89		
	0.91		
	0.92		
Aspek Bahasa	0.91	0.93	Tinggi
	0.90		
	0.95		
	0.95		

Berdasarkan hasil Tabel 3.11, koefisien validitas *Aiken* (V) menunjukkan bahwa dari 30 butir soal yang dinilai, semua soal valid dilihat dari setiap aspek

penilaian. Oleh karena itu, dapat disimpulkan bahwa semua butir soal dapat diuji coba di lapangan.

3.7.2.1.2 Uji Validitas Kontruk

Uji validitas konstruk dilakukan untuk menentukan seberapa jauh *item* soal dan komponen dalam tes sejalan dengan konstruk yang ingin diukur (Sumintono & Widhiarso, 2015). Analisis uji validitas konstruk ini dilakukan dengan menggunakan teknik pemodelan Rasch yang juga dikenal sebagai unidimensionalitas (Sumintono & Widhiarso, 2014). Unidimensionalitas adalah alat pengukuran penting yang digunakan untuk mengevaluasi apakah instrumen tes yang dikembangkan mampu mengukur apa yang seharusnya diukur. Artinya, instrumen penelitian harus dapat mengukur sesuai dengan tujuan yang telah ditetapkan sebelumnya. Dalam *software* WINSTEPS versi 3.73, unidimensionalitas instrumen dapat dilihat dari nilai *raw variance explained by measures* yang dapat ditemukan pada menu *output* tabel 23: *Item dimensionality*. Kriteria unidimensionalitas pada analisis pemodelan Rasch ditunjukkan pada Tabel 3.12.

Tabel 3.12 Kriteria Nilai Unidimensionalitas Instrumen

Nilai <i>Raw Variance Explained by Measures</i> (%)	Kriteria
> 60	Istimewa
> 40	Sesuai
> 20	Terpenuhi

(Sumintono & Widhiarso, 2015)

Validasi kontruk diperoleh dari hasil uji coba instrumen berpikir tingkat tinggi *global warming* yang terdiri dari 30 butir soal *Two-tier Multiple Choice* yang telah diisi oleh 30 orang peserta didik kelas X. Setelah dilakukan penyebaran dan sampel telah mencukupi, data hasil penyebaran instrumen diolah dan dianalisis menggunakan rash model untuk menentukan butir soal yang layak digunakan untuk penelitian dalam bentuk *pre-test* dan *posttest*. Hasil uji unidimensionalitas instrumen berpikir tingkat tinggi *global warming* menggunakan *software*

WINSTEPS versi 3.73 dapat dilihat pada menu *output table 23: Item dimensionality* seperti pada Tabel 3.13 di bawah.

Tabel 3.13 Hasil Uji Dimensionality Instrumen Berpikir Tingkat Tinggi Global Warming

Unidimensionalitas	Nilai indeks	Keterangan
<i>Raw variance explained by measures</i>	41,90 %	Valid

Berdasarkan tabel 3.13 diperoleh nilai *raw variance explained by measures* instrumen berpikir tingkat tinggi *global warming* menggunakan analisis rash model sebesar 41,90%. Menurut Sumintono & Widhiarso (2015), instrumen dapat dikatakan valid apabila nilai *raw variance yang explained by measure* lebih dari 20%. Sehingga, instrumen berpikir tingkat tinggi *global warming* yang dibuat dapat dikatakan valid karena mendapatkan nilai *raw variance explained by measures* lebih dari 20%.

Setelah melakukan uji validitas pada instrumen secara keseluruhan, langkah berikutnya adalah melakukan uji validitas pada setiap *item* soal. Uji ini juga menggunakan analisis pemodelan Rasch untuk mengevaluasi kualitas setiap *item* soal. Informasi untuk uji validitas *item* soal dapat ditemukan di bagian tabel 10: *Item fit order* dari menu *output tables*. Kualitas setiap item soal ditentukan dengan memperhatikan nilai *outfit means square* (MNSQ), *outfit Z-standard* (ZSTD), dan *point measure correlation* (*Pt Measure Corr.*). Masing-masing dari ketiga kriteria ini memiliki standar tertentu yang harus dipenuhi agar *item* soal tersebut dianggap valid. Menurut Sumintono & Widhiarso (2015), kriteria untuk menentukan apakah nilai yang diperoleh berada pada rentang yang diterima atau tidak ditunjukkan pada Tabel 3.14.

Tabel 3.14 Kriteria *Outfit* MNSQ, ZTSD, dan *Pt Measure Corr*

Kriteria	Nilai yang Diterima
<i>Outfit Mean-Square</i> (MNSQ)	$0,5 < \text{MNSQ} < 1,5$
<i>Outfit Z-Standard</i> (ZSTD)	$-2,0 < \text{ZSTD} < 2,0$
<i>Pt Measure Corr</i>	$0,4 < \text{Pt Measure Corr} < 0,85$

(Sumintono & Widhiarso, 2015)

Hasil dari nilai masing-masing kriteria tersebut kemudian diinterpretasikan berdasarkan kriteria nilai *fit-statistic* pada Tabel 3.15 berikut.

Tabel 3.15 Interpretasi Kualitas Butir Soal

Kriteria Nilai <i>Fit-Statistic</i>	Interpretasi
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, 2014)

Interpretasi kualitas setiap butir soal instrumen berpikir tingkat tinggi *global warming* berdasarkan kriteria nilai *fit-statistic* dapat dilihat pada Tabel 3.16 berikut.

Tabel 3.16 Hasil Interpretasi Kualitas Butir Soal Berpikir Tingkat Tinggi

Global Warming

Nomor Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Kriteria Nilai	Interpretasi	Keterangan
	MNSQ	ZSTD				
1	0,60	-1,38	0,45	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
2	0,98	0,31	0,45	Dua dari tiga kriteria nilai terpenuhi	Sangat sesuai	Digunakan
3	1,31	0,76	0,40	Dua dari tiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
4	0,95	-0,07	0,44	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
5	1,04	0,22	0,59	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
6	0,65	-0,89	0,47	Dua dari tiga kriteria	Sangat Sesuai	Digunakan

Nomor Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Kriteria Nilai	Interpretasi	Keterangan
	MNSQ	ZSTD				
				nilai terpenuhi		
7	0,89	-0,24	0,50	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
8	1,12	0,45	0,48	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
9	1,15	0,56	0,40	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
10	1,24	0,69	0,47	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
11	1,19	0,66	0,44	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
12	1,11	0,41	0,59	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
13	1,20	0,60	0,35	Dua dari tiga kriteria nilai terpenuhi	Sesuai	Tidak digunakan
14	1,12	0,39	0,68	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
15	1,23	0,75	0,50	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
16	1,38	0,67	0,18	Dua dari tiga kriteria nilai terpenuhi	Sesuai	Tidak digunakan

Nomor Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Kriteria Nilai	Interpretasi	Keterangan
	MNSQ	ZSTD				
17	0,68	-1,09	0,67	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
18	0,40	-1,45	0,85	Dua dari tiga kriteria nilai terpenuhi	Sesuai	Tidak digunakan
19	1,34	1,06	0,53	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
20	0,82	-0,54	0,65	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
21	0,87	-0,24	0,50	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
22	1,24	0,71	0,32	Dua dari tiga kriteria nilai terpenuhi	Sesuai	Tidak digunakan
23	1,08	0,34	0,47	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
24	1,05	0,26	0,58	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
25	1,00	0,12	0,61	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
26	0,54	-1,53	0,58	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
27	1,79	1,69	0,22	Dua dari tiga kriteria	Sesuai	Tidak digunakan

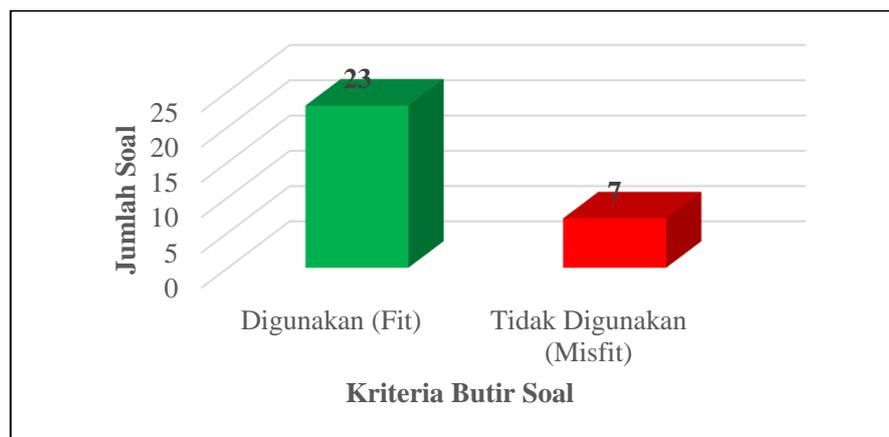
Ahmad Rifa'i, 2025

EFEKTIVITAS PBL-POE BERBANTUAN E-MODUL BERBASIS AUGMENTED REALITY TERHADAP PENINGKATAN BERPIKIR TINGKAT TINGGI DAN MOTIVASI BELAJAR PESERTA DIDIK PADA MATERI PEMANASAN GLOBAL
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Nomor Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Kriteria Nilai	Interpretasi	Keterangan
	MNSQ	ZSTD				
				nilai terpenuhi		
28	0,94	0,03	0,75	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
29	0,75	-0,57	0,79	Ketiga kriteria nilai terpenuhi	Sangat Sesuai	Digunakan
30	0,32	-2,41	0,64	Dua dari tiga kriteria nilai terpenuhi	Sesuai	Tidak digunakan

Tabel 3.16 menunjukkan analisis terhadap validitas tiap butir soal berpikir tingkat tinggi *global warming*. Dari 30 butir soal yang diuji, 23 butir soal memenuhi ketiga kriteria dalam model Rasch, yaitu *outfit Mean-Square* (MNSQ) dengan kisaran nilai antara 0,60 hingga 1,24, *outfit Z-Standard* (ZSTD) yang berada dalam rentang -1,53 hingga 1,06, dan *Point Measure Correlation* (*Pt Measure Corr*) dengan nilai antara 0,40 hingga 0,79. Butir-butir soal ini dinilai sangat sesuai (*fit*) untuk digunakan dalam instrumen pengukuran karena menunjukkan kesesuaian yang baik dengan model Rasch serta korelasi yang kuat dengan skor keseluruhan.

Sebaliknya, terdapat 7 butir soal yang tidak memenuhi semua kriteria (*misfit*). Butir soal nomor 13, 16, 18, 22, 27, dan 30 hanya memenuhi dua dari tiga kriteria validitas dalam Model Rash. Misalnya, butir soal nomor 16 dan 27 memiliki nilai MNSQ yang berada di luar rentang ideal, masing-masing 1,38 dan 1,79, yang menunjukkan adanya ketidaksesuaian dengan model Rasch. Nilai ZSTD yang ekstrem, seperti pada butir soal nomor 30 dengan nilai -2,41, juga menunjukkan ketidaksesuaian yang tinggi. Berdasarkan hasil analisis tersebut dapat disimpulkan bahwa sebanyak 23 butir soal instrumen berpikir tingkat tinggi *global warming* sesuai (*valid*) dan dapat digunakan menjadi instrumen dalam penelitian. Hasil ini dapat dilihat pada Gambar 3.7 berikut.



Gambar 3.4 Hasil Validitas Konstruk Instrumen Berpikir Tingkat Tinggi *Global Warming*

Berdasarkan gambar di atas menunjukkan bahwa dari 30 butir soal yang dibuat dan dianalisis menggunakan rash model terdapat 23 butir soal yang akan digunakan pada penelitian dan 7 soal akan dibuang atau tidak akan digunakan dalam penelitian.

3.8.2.2 Uji Reliabilitas

Reliabilitas instrumen merujuk pada konsistensi atau stabilitas skor dari sebuah tes (Johnson & Christensen, 2020). Artinya, jika instrumen digunakan untuk mengukur kemampuan peserta didik pada waktu tertentu, skor yang diperoleh seharusnya tidak akan berubah secara signifikan jika tes tersebut diulangi pada waktu yang lain (Sumintono & Widhiarso, 2015). Dalam penelitian ini, analisis uji reliabilitas dilakukan dengan menggunakan pemodelan Rasch melalui *software* WINSTEPS versi 3.73. Informasi ini dapat ditemukan pada menu *output tables* bagian tabel 3.1: *summary statistics*. Terdapat tiga indikator yang digunakan dalam penelitian ini, yaitu *person reliability*, *item reliability*, dan *Cronbach Alpha*. *Person reliability* menggambarkan sejauh mana jawaban peserta didik konsisten, sementara *item reliability* mengukur kualitas *item-item* soal dalam instrumen. Sementara itu, nilai *Cronbach Alpha* digunakan untuk mengukur reliabilitas secara keseluruhan, yaitu interaksi antara peserta didik dengan *item* soal. Kriteria untuk ketiga indikator tersebut dapat dilihat pada Tabel 3.17 dan 3.18 sebagai berikut.

Tabel 3.17 Interpretasi Nilai *Person Reliability* dan *Item Reliability*

Rentang Nilai <i>Person Reliability</i> dan <i>Item Reliability</i>	Interpretasi
> 0,94	Istimewa
0,91 – 0,94	Baik sekali
0,81 – 0,90	Baik
0,67 – 0,80	Cukup
< 0,67	Lemah

(Sumintono & Widhiarso, 2015)

Tabel 3.18 Interpretasi Nilai *Cronbach Alpha*

Rentang Nilai <i>Cronbach Alpha</i>	Interpretasi
> 0,80	Sangat baik
0,71 – 0,80	Baik
0,61 – 0,70	Cukup
0,50 – 0,60	Jelek
< 0,50	Buruk

(Sumintono & Widhiarso, 2015)

Interpretasi reliabilitas instrumen berpikir tingkat tinggi *global warming* berdasarkan indikator *person reliability*, *item reliability*, dan *Cronbach Alpha* ditunjukkan pada Tabel 3.19 berikut.

Tabel 3.19 Hasil Uji Reliabilitas Instrumen Berpikir Tingkat Tinggi *Global Warming*

Summary Statistic	Nilai Indeks	Interpretasi
<i>Item reliability</i>	0,88	Baik
<i>Person reliabilty</i>	0,89	Baik
<i>Cronbach's alpha (KR-20)</i>	0,90	Sangat baik

Berdasarkan Tabel 3.17, terlihat bahwa nilai *item reliability* sebesar 0,88 dan *person reliability* sebesar 0,89, keduanya termasuk dalam kategori baik. Hal ini menunjukkan bahwa instrumen berpikir tingkat tinggi *global warming* yang digunakan memberikan hasil pengukuran yang konsisten dan dapat diandalkan. Selain itu, nilai *Cronbach's alpha* sebesar 0,90 menunjukkan bahwa instrumen ini memiliki kualitas yang sangat baik dalam mengidentifikasi hubungan antara peserta didik (*person reliability*) dengan butir soal (*item reliability*). Dengan hasil ini, dapat disimpulkan bahwa instrumen berpikir tingkat tinggi *global warming reliabel* dan

layak digunakan dalam penelitian untuk mengukur kemampuan berpikir tingkat tinggi peserta didik secara akurat dan konsisten dalam konteks pendidikan terkait pemanasan global.

3.8.2.3 Tingkat Kesukaran *Item* Soal

Tingkat kesukaran digunakan untuk menentukan apakah *item* soal termasuk dalam kategori mudah, sedang, atau sulit. Penentuan tingkat kesulitan dilakukan dengan analisis pemodelan Rasch dengan memilih menu *output* tabel 13: *Item Measure* di *software* WINSTEPS versi 3.73. Tingkat kesukaran *item* soal dapat dilihat dari nilai *measure* (ME) dan standar deviasi (SD), dengan membandingkan nilai logit ME pada setiap *item* dan nilai SD. Nilai logit (*log odds unit*) adalah skala yang memiliki interval yang sama dan bersifat linear yang berasal dari data rasio (*odds ratio*) (Sumintono & Widhiarso, 2015). Interpretasi tingkat kesukaran setiap *item* soal dapat ditentukan berdasarkan Tabel 3.20 berikut.

Tabel 3.20 Interpretasi Tingkat Kesukaran Butir Soal

Kriteria	Interpretasi
$Measure\ Logit > SD$	Sangat sulit
$0,00 < Measure\ Logit \leq SD$	Sulit
$-SD < Measure\ Logit \leq 0,00$	Mudah
$Measure\ Logit < -SD$	Sangat mudah

(Sumintono & Widhiarso, 2015)

Tingkat kesulitan setiap butir soal juga dianalisis untuk mengetahui butir soal yang digunakan termasuk kategori mudah, sedang, atau sulit dengan meninjau nilai *measure* (ME) dan standar deviasi (SD) dengan cara membandingkan nilai logit ME pada masing-masing *item* dan nilai SD. Nilai SD yang diperoleh adalah +0,91. Interpretasi tingkat kesulitan butir soal instrumen berpikir tingkat tinggi *global warming* dapat dilihat pada Tabel 3.21 berikut.

Tabel 3.21 Hasil Interpretasi Tingkat Kesulitan Item Berpikir Tingkat Tinggi *Global Warming*

Nomor Butir Soal	<i>Measure</i> (ME)	Standar Deviasi (SD)	Kriteria	Interpretasi
1	0,38	0,91	$-0,91 \leq ME \leq 0,91$	Sedang

Nomor Butir Soal	Measure (ME)	Standar Deviasi (SD)	Kriteria	Interpretasi
2	-0,23	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
3	-0,37	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
4	-0,23	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
5	0,49	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
6	-0,23	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
7	-0,69	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
8	-0,37	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
9	-0,60	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
10	-1,31	0,91	$ME < -0,91$	Mudah
11	-0,45	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
12	0,02	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
13	-1,08	0,91	$ME < -0,91$	Mudah
14	-0,30	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
15	-0,10	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
16	0,66	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
17	0,66	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
18	0,02	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
19	-0,45	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
20	0,55	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
21	-1,08	0,91	$ME < -0,91$	Mudah
22	-0,69	0,91	$-0,91 \leq ME \leq 0,91$	Sedang
23	0,72	0,91	$-0,91 \leq ME \leq 0,91$	Sedang

Tabel 3.21 menunjukkan analisis tingkat kesulitan butir soal berpikir tingkat tinggi *global warming*. Dari 23 butir soal yang diuji, 20 butir soal berada dalam kategori tingkat kesulitan sedang, dengan nilai *Measure* (ME) berkisar antara -0,91 hingga 0,91. Hal ini menunjukkan bahwa mayoritas butir soal memiliki tingkat kesulitan yang moderat dan sesuai untuk mengukur kemampuan berpikir tingkat tinggi peserta didik. Sebanyak 3 butir soal tergolong mudah, dengan nilai ME di bawah -0,91, yang menunjukkan bahwa soal-soal ini lebih mudah dijawab oleh peserta didik. Analisis ini sejalan dengan penelitian yang menyatakan bahwa soal dengan tingkat kesulitan sedang lebih ideal untuk mengukur kemampuan berpikir tingkat tinggi karena dapat menantang peserta didik tanpa membuat mereka frustrasi (Amelia, 2021). Berdasarkan hasil analisis ini, dapat disimpulkan bahwa instrumen berpikir tingkat tinggi *global warming* memiliki distribusi tingkat kesulitan yang baik, dengan proporsi soal sedang yang dominan (87%) dan

beberapa soal mudah (13%) untuk memastikan variasi dalam tingkat kesulitan. Ini menunjukkan bahwa instrumen tersebut efektif dalam mengevaluasi kemampuan berpikir tingkat tinggi peserta didik pada materi pemanasan global. Hasil lebih lanjut mengenai tingkat kesulitan butir soal berpikir tingkat tinggi *global warming* disajikan pada Tabel 3.22 berikut.

Tabel 3.22 Hasil Analisis Kesulitan Butir Soal Berpikir Tingkat Tinggi
Global Warming

Kriteria	Frekuensi	Persentase (%)
Mudah	3	13
Sedang	20	87

Tabel 3.20 menunjukkan bahwa dari 23 butir soal berpikir tingkat tinggi *global warming* yang diuji, sebanyak 20 butir soal (87%) memiliki tingkat kesulitan sedang, sedangkan 3 butir soal (13%) tergolong mudah. Proporsi besar soal dengan tingkat kesulitan sedang memastikan bahwa instrumen ini mampu mendukung efektivitas instrumen dalam mengukur kemampuan berpikir tingkat tinggi.

3.8.2.4 Daya Pembeda

Daya pembeda digunakan untuk menunjukkan sejauh mana suatu *item* soal mampu membedakan antara individu yang memiliki kemampuan tinggi dan rendah (Sumintono & Widhiarso, 2015). Dalam penelitian ini, daya pembeda dianalisis dengan metode pemodelan Rasch menggunakan *software* WINSTEPS versi 3.37, melibatkan nilai *Point Measure Correlation (Pt Mean Corr)*, yang mengacu pada hubungan antara tingkat kesulitan tiap *item* soal dengan tingkat kesulitan tes secara umum. Nilai satu menunjukkan adanya korelasi sempurna antara *respons item* soal dengan kemampuan peserta tes. Nilai nol menunjukkan tidak ada hubungan antara *respons item* soal tertentu dengan *item* soal secara keseluruhan. Sementara nilai negatif menunjukkan *item* soal memiliki kekurangan pada beberapa aspek fundamental dan perlu ditinjau kembali atau dihapus dari tes (Smiley, 2015). Interpretasi daya pembeda untuk setiap *item* soal ditunjukkan pada Tabel 3.21 berikut.

Tabel 3.21 Interpretasi Daya Pembeda Butir Soal

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

(Brown, 2005; Smiley, 2015)

Hasil analisis daya pembeda instrumen berpikir tingkat tinggi *global warming* menggunakan *software* WINSTEPS versi 3.37 ditunjukkan pada Tabel 3.22 berikut.

**Tabel 3.22 Hasil Interpretasi Daya Pembeda Item Berpikir Tingkat Tinggi
*Global Warming***

Nomor soal	<i>Point measure correlation</i>	Keterangan	Level Kognitif
1	0,45	Sangat baik	C4
2	0,35	Baik	C4
3	0,30	Kurang baik	C4
4	0,44	Sangat baik	C5
5	0,59	Sangat baik	C6
6	0,37	Baik	C4
7	0,50	Sangat baik	C4
8	0,48	Sangat baik	C4
9	0,40	Sangat baik	C4
10	0,47	Sangat baik	C4
11	0,44	Sangat baik	C4
12	0,59	Sangat baik	C4
13	0,35	Baik	C4
14	0,68	Sangat baik	C5
15	0,50	Sangat baik	C5
16	0,18	Buruk	C4
17	0,67	Sangat baik	C4
18	0,85	Sangat baik	C4
19	0,53	Sangat baik	C5
20	0,65	Sangat baik	C5
21	0,50	Sangat baik	C5
22	0,32	Baik	C5
23	0,47	Sangat baik	C5
24	0,58	Sangat baik	C6

Nomor soal	<i>Point measure correlation</i>	Keterangan	Level Kognitif
25	0,61	Sangat baik	C6
26	0,58	Sangat baik	C6
27	0,22	Kurang baik	C6
28	0,75	Sangat baik	C6
29	0,79	Sangat baik	C6
30	0,64	Sangat baik	C4

Tabel 3.22 menyajikan hasil interpretasi daya pembeda butir soal berpikir tingkat tinggi *global warming* berdasarkan *point measure correlation* dan level kognitif. Dari tabel tersebut, terlihat bahwa mayoritas butir soal (23 dari 30) memiliki daya pembeda yang sangat baik, dengan *point measure correlation* berkisar antara 0,40 hingga 0,85. Butir soal dengan *point measure correlation* di atas 0,40 menunjukkan kemampuan yang baik dalam membedakan antara peserta didik dengan kemampuan tinggi dan rendah, sehingga dianggap sesuai. Sebaliknya, 7 butir soal dinilai tidak sesuai karena memiliki *point measure correlation* di bawah 0,40, yang menunjukkan bahwa soal-soal ini kurang efektif dalam membedakan kemampuan peserta didik. Daya pembeda yang baik sangat penting untuk memastikan bahwa butir soal dapat mengukur kemampuan berpikir tingkat tinggi dengan akurat pada berbagai level kognitif, dari C4 (Analisis), C5 (Evaluasi), hingga C6 (Mencipta). Hasil ini menunjukkan bahwa sebagian besar butir soal berpikir tingkat tinggi *global warming* dapat diandalkan untuk mengukur kemampuan berpikir tingkat tinggi peserta didik, meskipun beberapa butir soal perlu direvisi untuk meningkatkan daya pembedanya.

3.7.3 Analisis Keterlaksanaan Pembelajaran

Data analisis keterlaksanaan pembelajaran diperoleh melalui pengamatan seorang observer terhadap aktivitas guru dan peserta didik selama proses pembelajaran model PBL-POE berbantuan *E-Modul* berbasis *Augmented Reality* (AR) dikelas eksperimen dan model PBL di kelas kontrol pada materi pemanasan global. Skor penilaian untuk keterlaksanaan pembelajaran dimulai dari 0 untuk "tidak terlaksana", 1 untuk "terlaksana". Pengolahan hasil observasi dilakukan

dengan cara menghitung persentase keterlaksanaan pembelajaran dengan rumus berikut.

$$\%KP = \frac{\sum \text{jumlah skor yang diperoleh}}{\sum \text{jumlah skor maksimal}} \times 100\% \quad (3.3)$$

(Sudaryono, 2017)

Kategori keterlaksanaan pembelajaran dan interpretasi berdasarkan persentase yang diperoleh disajikan pada Tabel 3.23 berikut.

Tabel 3.23 Interpretasi Persentase Keterlaksanaan Pembelajaran

Keterlaksanaan Pembelajaran (KP) (%)	Kategori
$KP > 80$	Sangat baik
$60 < KP \leq 80$	Baik
$40 < KP \leq 60$	Cukup baik
$20 < KP \leq 40$	Lemah
$KP \leq 20$	Sangat lemah

(Avianti & Yonata, 2015)

3.7.4 Analisis Efektivitas Model PBL-POE Berbantuan *E-Modul* Berbasis *Augmented Reality* (AR) Terhadap Peningkatan Berpikir Tingkat Tinggi

Untuk menilai efektivitas sebuah intervensi dalam penelitian, penting untuk menggunakan beberapa kriteria yang komprehensif. Menurut Kyriakides & Leonidas (2015), beberapa kriteria yang digunakan untuk mengukur efektivitas mencakup peningkatan variabel terikat, perbedaan variabel terikat antara kelompok eksperimen dan kontrol, serta ukuran efek (*effect size*) atau uji dampak. Cohen (1988) juga menjelaskan bahwa ukuran efek sangat penting untuk menentukan sejauh mana intervensi memberikan dampak yang signifikan. Dengan menerapkan kriteria-kriteria ini, peneliti dapat memperoleh gambaran yang lebih jelas dan menyeluruh mengenai sejauh mana intervensi yang dilakukan memberikan dampak yang signifikan dan relevan pada variabel yang diteliti. Berikut penjelasan kriteria untuk menilai efektivitas model PBL-POE berbantuan *E-Modul* berbasis AR terhadap berpikir tingkat tinggi.

3.7 4.1 Peningkatan Berpikir Tingkat Tinggi

Peningkatan berpikir tingkat tinggi peserta didik diperoleh dari data kuantitatif nilai *pre-test* dan *posttest* peserta didik di kelas eksperimen dengan menggunakan pemodelan Rasch berbantuan *software* WINSTEPS versi 3.73 teknik *stacking*. Pemodelan Rasch dipilih karena mampu memberikan analisis pada tingkat individu (*individual-centered statistics*). Hasil *pre-test* dan *posttest* diolah menjadi skala pengukuran dengan interval yang sama. Pemodelan Rasch merubah data ordinal menjadi data rasio untuk analisis statistik. Perbandingan rasio ini menunjukkan hubungan antara tingkat kemampuan peserta didik (*person ability*) dan tingkat kesulitan soal (*item difficulty*). Fungsi logaritma digunakan untuk mendapatkan pengukuran dengan interval yang sama dan hasilnya menciptakan satuan baru yang disebut logit (*log odds unit*). Menurut Englehard, keberhasilan peserta didik dalam mengerjakan soal sangat ditentukan oleh tingkat kemampuan peserta didik dan tingkat kesulitan soal (Sumintono, 2021). Nilai probabilitas *odd ratio* individu dapat ditentukan melalui persamaan berikut.

$$Odd\ ratio = \frac{P}{N - P} \quad (3.4)$$

(Sumintono & Widhiarso, 2015)

Keterangan:

P: jumlah soal yang dikerjakan peserta didik

N: total soal yang diujikan

Nilai probabilitistik *odd ratio* diubah menjadi bentuk logaritma untuk menciptakan skala dengan interval yang sama. Formula untuk logit (*unit log odds*) yang dirumuskan oleh Sumintono dan Widhiarso (2015) adalah sebagai berikut:

$$Logit = \log \frac{P}{N - P} \quad (3.5)$$

Pada penelitian ini, interpretasi berpikir tingkat tinggi peserta didik diukur menggunakan nilai logit dan perubahan posisi pada *vertical ruler*. Berikut adalah penjelasan lebih lanjut tentang teknik analisis *stacking-racking*.

3.7.4.1.1 Analisis Stacking

Analisis *stacking* digunakan dalam studi longitudinal untuk membandingkan kemampuan peserta didik sebelum dan sesudah intervensi (Sukarelawan dkk., 2024). Teknik ini melibatkan pengumpulan data *pre-test* dan *posttest* secara bersamaan, yang kemudian diubah menjadi skala interval yang sama (*logit*). Perubahan dalam kemampuan peserta didik dilihat berdasarkan perubahan dalam posisi level kemampuan mereka pada *vertical ruler* dan perubahan nilai logit pada *software* WINSTEPS versi 3.73. Analisis *stacking* memberikan pandangan yang lebih rinci tentang bagaimana kemampuan Peserta didik berubah seiring waktu, terutama setelah intervensi (Sukarelawan dkk., 2024). Teknik ini dipilih karena dapat mengidentifikasi perubahan pada tingkat individu, pengukuran efek intervensi, pemahaman yang lebih mendalam, identifikasi reaksi terhadap intervensi, serta memberikan pengukuran yang akurat dan tepat (Sukarelawan dkk., 2024).

3.7.4.1.2 Analisis Racking

Analisis *racking* digunakan untuk membandingkan tingkat kesulitan soal sebelum dan sesudah intervensi (Laliyo, 2021). Seperti analisis *stacking*, data *pre-test* dan *posttest* pada kelas eksperimen dikumpulkan secara bersamaan dan diubah menjadi skala interval yang sama (*logit*). Perubahan dalam tingkat kesulitan soal dilihat berdasarkan perubahan dalam posisi *item* soal pada *vertical ruler* dan perubahan nilai logit pada *software* WINSTEPS versi 3.73

3.7 4.2 Uji Dampak (Effect Size)

Untuk melihat seberapa besar dampak model PBL-POE berbantuan *E-Modul* berbasis AR terhadap peningkatan berpikir tingkat tinggi peserta didik diperoleh melalui perhitungan *effect size*. *Effect size* menunjukkan seberapa besar pengaruh dari suatu variabel terhadap variabel lain dalam suatu penelitian. *Effect size* pada

penelitian ini diperoleh melalui pengolahan hasil nilai *posttest* dari kelas eksperimen dan kelas kontrol. Persamaan yang digunakan untuk menghitung *effect size* adalah menggunakan persamaan *Cohen's d* (Thalheimer & Cook, 2002) yang ditunjukkan oleh persamaan (3.6) berikut.

$$d = \frac{\bar{x}_t - \bar{x}_c}{S_{pooled}} \quad (3.6)$$

Dengan

- d : Cohen's d *Effect Size*
 \bar{x}_t : Nilai rata-rata *posttest* peserta didik kelas eksperimen
 \bar{x}_c : Nilai rata-rata *posttest* peserta didik kelas kontrol
 S_{pooled} : Standar deviasi gabungan

Perhitungan *Cohen's d* menggunakan bantuan *software* SPSS versi 26 dan hasilnya akan diinterpretasikan berdasarkan Tabel 3.24 berikut.

Tabel 3.24 Interpretasi nilai *Cohen's d*

Nilai <i>Cohen's d</i>	Interpretasi
$0,80 \leq d \leq 2,00$	Besar
$0,50 \leq d < 0,80$	Sedang
$0,20 \leq d < 0,50$	Kecil
$0,00 \leq d < 0,20$	Kurang

(Lalongo, 2016)

Ada beberapa tahap analisis data sebelum kita melakukan perhitungan *effect size*, diantaranya sebagai berikut:

3.7.4.1.3 Uji Normalitas

Uji normalitas bertujuan untuk mengetahui sejauh mana variabel acak terdistribusi secara normal (Susetyo, 2017). Data yang terdistribusi normal dapat menggunakan teknik analisis statistik parametrik, sedangkan data yang tidak terdistribusi normal dapat menggunakan teknik analisis statistik nonparametrik. Terdapat beberapa metode uji normalitas yang dapat digunakan yaitu uji Chi-Square, Kolmogorov-Smirnov, Liliefors, Shapiro-Wilk yang disesuaikan dengan jumlah sampel. Uji normalitas dilakukan dengan membandingkan nilai *P-value* signifikansi hasil pengujian dengan taraf signifikansi ($\alpha = 0,05$). Jika *Sig.* > 0,05,

maka sampel dianggap berasal dari populasi dengan distribusi normal. Pengujian normalitas dilakukan dengan menggunakan *software* SPSS versi 26.

3.7.4.1.4 Uji Homogenitas

Uji homogenitas digunakan untuk memastikan apakah data berasal dari varians yang sama atau tidak. Uji ini penting dalam penelitian kuasi untuk memastikan apakah kelompok eksperimen dan kontrol berasal dari varians yang sama. Pada penelitian ini, uji homogenitas yang digunakan adalah uji homogenitas *Levene test* yang diolah dengan menggunakan *software* SPSS versi 26. Jika *Sig.* \geq 0,05, maka sampel dianggap berasal dari varians yang sama (homogen).

3.7.4.1.5 Uji Hipotesis

Uji hipotesis dilakukan untuk memverifikasi kebenaran parameter berdasarkan data yang diperoleh dari sampel penelitian. Ada dua metode yang dapat digunakan dalam uji hipotesis, yaitu menggunakan uji statistik parametrik dan uji statistik nonparametrik. Untuk menentukan metode uji hipotesis yang tepat, diperlukan terlebih dahulu uji normalitas dan uji homogenitas terhadap data yang diperoleh. Jika data menunjukkan sifat normal tetapi tidak homogen, maka dilakukan uji nonparametrik seperti Uji Mann-Whitney U. Namun, jika data menunjukkan sifat normal dan homogen, maka dilakukan uji parametrik seperti Uji *Independent t-test*. uji *Independent Sample t-test* digunakan untuk mengetahui perbedaan rata-rata kemampuan berpikir tingkat tinggi pada nilai posttest dari kelas eksperimen dan kelas kontrol. Berikut hipotesis yang diajukan pada penelitian ini.

H_0 = Tidak terdapat perbedaan yang signifikan dalam kemampuan berpikir tingkat tinggi peserta didik antara model pembelajaran PBL-POE berbantuan *E-Modul* berbasis AR dan model pembelajaran PBL pada materi pemanasan global.

H_a = Terdapat perbedaan yang signifikan dalam kemampuan berpikir tingkat tinggi peserta didik antara model pembelajaran PBL-POE berbantuan *E-Modul* berbasis AR dan model pembelajaran PBL pada materi pemanasan global.

Hasil uji hipotesis dapat diperoleh dengan cara membandingkan nilai signifikansi yang didapatkan dengan taraf signifikansi ($\alpha = 0,05$) atau dengan membandingkan nilai t_{hitung} dengan t_{tabel} . Adapun kriteria pengujian dari uji hipotesis berdasarkan nilai signifikansi dan nilai t adalah sebagai berikut.

- Jika $t_{hitung} > t_{tabel}$ dan Jika nilai *Sig. (2-tailed)* $< 0,05$, Artinya H_0 ditolak
- Jika $t_{hitung} < t_{tabel}$ dan Jika nilai *Sig. (2-tailed)* $> 0,05$, Artinya H_0 diterima

Jika kedua kriteria peningkatan berpikir tingkat tinggi dan uji dampak terpenuhi maka dapat dikatakan bahwa model PBL-POE berbantuan *E-Modul* berbasis AR efektif dalam meningkatkan berpikir tingkat tinggi.

3.7.5 Efektivitas Model PBL-POE Berbantuan *E-Modul* Berbasis *Augmented Reality* (AR) Terhadap Peningkatan Motivasi Belajar

Ada beberapa kriteria yang digunakan untuk mengukur efektivitas intervensi dalam pembelajaran yaitu peningkatan variabel terikat, perbedaan variabel terikat antara kelompok eksperimen dan kontrol, serta ukuran efek (*effect size*) atau uji dampak (Kyriakides & Leonidas, 2015). Berikut penjelasan kriteria untuk menilai efektivitas model PBL-POE berbantuan *E-Modul* berbasis AR terhadap berpikir tingkat tinggi.

3.7.5.1 Peningkatan Motivasi Belajar

Kuesioner motivasi belajar dilakukan untuk mengetahui peningkatan motivasi belajar peserta didik setelah diterapkan model PBL-POE berbantuan *E-Modul Flipbook* berbasis AR. Peningkatan motivasi belajar peserta didik diperoleh dari data kuantitatif nilai kuesioner *pre-test* dan *posttest* peserta didik di kelas eksperimen dengan menggunakan pemodelan Rasch berbantuan *software* WINSTEPS versi 3.73 teknik *stacking*. Teknik *stacking* melibatkan pengumpulan data *pre-test* dan *posttest* secara bersamaan, yang kemudian diubah menjadi skala interval yang sama (*logit*). Teknik ini dipilih karena dapat mengidentifikasi perubahan pada tingkat individu, pengukuran efek intervensi, pemahaman yang lebih mendalam, identifikasi reaksi terhadap intervensi, serta memberikan pengukuran yang akurat dan tepat (Sukarelawan dkk., 2024). Pada penelitian ini,

interpretasi peningkatan motivasi belajar peserta didik dilihat berdasarkan perubahan dalam posisi level kemampuan mereka pada *vertical ruler* dan perubahan nilai logit pada *software* WINSTEPS versi 3.73

3.7.5.2 Uji Dampak (*Effect Size*)

Sebelum melakukan uji dampak ada beberapa tahapan analisis yang harus dipenuhi yaitu uji prasyarat (normalitas dan homogenitas) serta uji hipotesis. Berikut hipotesis yang diajukan pada penelitian ini.

H_0 = Tidak terdapat perbedaan yang signifikan dalam motivasi belajar peserta didik antara model pembelajaran PBL-POE berbantuan *E-Modul* berbasis AR dan model pembelajaran PBL pada materi pemanasan global.

H_a = Terdapat perbedaan yang signifikan dalam motivasi belajar peserta didik antara model pembelajaran PBL-POE berbantuan *E-Modul* berbasis AR dan model pembelajaran PBL pada materi pemanasan global.

Hasil uji hipotesis dapat diperoleh dengan cara membandingkan nilai signifikansi yang didapatkan dengan taraf signifikansi ($\alpha = 0,05$) atau dengan membandingkan nilai t_{hitung} dengan t_{tabel} . Adapun kriteria pengujian dari uji hipotesis berdasarkan nilai signifikansi dan nilai t adalah sebagai berikut.

- Jika $t_{hitung} > t_{tabel}$ dan Jika nilai *Sig. (2-tailed)* $< 0,05$, Artinya H_0 ditolak
- Jika $t_{hitung} < t_{tabel}$ dan Jika nilai *Sig. (2-tailed)* $> 0,05$, Artinya H_0 diterima

Setelah uji prasyarat dan uji hipotesis terpenuhi, maka dapat dilakukan perhitungan ukuran efek (*effect size*) dengan menggunakan hasil nilai *posttest* dari kelas eksperimen dan kelas kontrol. Persamaan yang digunakan untuk menghitung ukuran efek adalah Cohen's d (Thalheimer & Cook, 2002), ditunjukkan oleh persamaan berikut:

$$d = \frac{\bar{x}_t - \bar{x}_c}{S_{pooled}} \quad (3.6)$$

Dengan

- d : Cohen's *d Effect Size*
 \bar{x}_t : Nilai rata-rata posttest peserta didik kelas eskperimen
 \bar{x}_c : Nilai rata-rata posttest peserta didik kelas kontrol
 S_{pooled} : Standar deviasi gabungan

Perhitungan *Cohen's d* menggunakan bantuan *software* SPSS versi 26 dan hasilnya akan diinterpretasikan berdasarkan Tabel 3.30 yang telah dijelaskan sebelumnya.