

BAB III

METODE PENELITIAN

Bagian ini menjelaskan metode penelitian yang digunakan mencakup desain penelitian, partisipan, populasi dan sampel, instrumen penelitian, prosedur penelitian, dan teknik analisis data.

3.1 Desain Penelitian

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, 2011) yang melibatkan dua kelompok: kelompok eksperimen yang menggunakan model pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*), dan kelompok kontrol yang menggunakan model pembelajaran Learning Cycle 5E. Sebelum dan setelah pembelajaran, semua kelompok diberikan tes keterampilan berpikir kritis. Skor keterampilan berpikir kritis dari kedua kelompok kemudian dianalisis. Berikut tabel 3.1 mengenai desain penelitian yang akan digunakan.

Tabel 3. 1 Desain Penelitian

Kelompok	Pre-test	Treatment	Post-test
Kelas Eksperimen	O_1	X_1	O_2
Kelas Kontrol	O_1	X_2	O_2

(Sugiyono, 2011)

Keterangan :

O_1 : *Pre-test* keterampilan berpikir kritis

O_2 : *Post-test* keterampilan berpikir kritis

X_1 : Pembelajaran menggunakan model LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*)

X_2 : Pembelajaran menggunakan model *Learning Cycle 5E*

Berdasarkan tabel 3.1, kita dapat melihat perbandingan sebelum dan sesudah pemberian *treatment* (Sugiyono, 2011). Pada tahap awal, siswa menjalani *pretest* untuk menentukan kondisi awal mereka. Setelah menerima *treatment*, siswa kemudian diminta mengerjakan *posttest* untuk mengevaluasi kondisi mereka

setelah proses pembelajaran. Penelitian ini dilaksanakan dalam waktu dua pertemuan, dengan durasi masing-masing pertemuan selama 2 x 45 menit dan 3x 45 menit.

3.2 Partisipan

Partisipan dalam penelitian merujuk pada individu atau kelompok orang yang berpartisipasi dan berkontribusi dalam proses penelitian. Menurut Fadliyati (dalam Cilviani, 2022), partisipan adalah sekelompok orang yang terlibat secara fisik, mental, dan emosional sebagai responden atau peserta yang memberikan respons terhadap proses belajar-mengajar. Keterlibatan partisipan tidak hanya berarti memberikan dukungan moral dan material, tetapi juga membawa tanggung jawab atas partisipasinya

Penelitian ini akan dilaksanakan di salah satu SMA Negeri 15 Bandung. Terdapat 55 siswa kelas XI IPA yang aktif dan terdaftar sebagai peserta didik di SMA Negeri 15 Bandung untuk tahun ajaran 2024/2025 yang berpartisipasi dalam penelitian ini. Dari jumlah tersebut, 27 siswa kelas XI IPA dari SMA Negeri 15 Bandung ditunjuk sebagai partisipan untuk kelas eksperimen, sedangkan 28 siswa lainnya akan menjadi partisipan untuk kelas kontrol.

3.3 Populasi dan Sampel

Sugiyono (2020) mendefinisikan populasi sebagai wilayah yang lebih luas yang berisi objek atau individu dengan susunan dan sifat tertentu yang ditetapkan oleh peneliti. Dalam konteks penelitian ini, populasi adalah semua siswa kelas XI IPA yang aktif dan terdaftar di SMA Negeri 15 Kota Bandung untuk tahun ajaran 2024/2025.

Sugiyono (2017) menyatakan bahwa sampel merupakan sebagian dari populasi yang diteliti. Populasi tersebut memiliki sejumlah ciri yang ditetapkan oleh peneliti. Metode pengambilan sampel, menurut Sugiyono (2016), adalah metode yang menentukan sampel yang akan digunakan. Dalam penelitian ini, sampel dipilih berdasarkan kelas yang bersedia berpartisipasi dan dengan tujuan tertentu (*purposive sampling*). Sampel penelitian terdiri dari 55 siswa kelas XI di SMA Negeri 15 Bandung. Dari jumlah tersebut, 27 siswa berada di kelompok eksperimen yang menerima perlakuan menggunakan model pembelajaran LCV-Lab (*Learning*

Cycle 9E berbantuan *Virtual Laboratory*), sedangkan 28 siswa lainnya berada di kelompok kontrol yang menerima perlakuan menggunakan model pembelajaran LC 5E (*Learning Cycle 5E*).

3.4 Instrumen Penelitian

Untuk mendapatkan data dan informasi yang diperlukan dalam penelitian tentang “Pengaruh Pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) terhadap Peningkatan Keterampilan berpikir kritis Siswa pada Materi Fluida Dinamis”, peneliti memerlukan bantuan instrumen penelitian. Instrumen ini berfungsi sebagai alat pendukung peneliti dan menjadi panduan tentang apa yang harus dilakukan untuk mengumpulkan data dan informasi yang diperlukan. Instrumen penelitian yang digunakan dalam studi ini dibagi menjadi dua kategori, yaitu instrumen perangkat pembelajaran dan instrumen pengumpulan data. Instrumen perangkat pembelajaran terdiri dari rencana pelaksanaan pembelajaran (RPP) dan lembar kerja peserta didik (LKPD). Sementara itu, instrumen pengumpulan data terdiri dari lembar observasi keterlaksanaan, dan tes keterampilan berpikir kritis, seperti yang ditunjukkan dalam tabel 3.2 berikut:

Tabel 3. 2 Jenis Instrumen Penelitian

Pertanyaan	Data	Instrumen
Bagaimana peningkatan keterampilan berpikir kritis siswa setelah diterapkan model pembelajaran LCV-Lab (<i>Learning Cycle 9E</i> berbantuan <i>Virtual Laboratory</i>) di kelas eksperimen dengan model pembelajaran LC (<i>Learning Cycle 5E</i>) di kelas kontrol pada materi fluida dinamis?	Nilai <i>pre-test</i> dan <i>post-test</i> keterampilan berpikir kritis siswa dari kedua kelompok (LCV-Lab dan LC 5E)	Tes keterampilan berpikir kritis
Bagaimana perbedaan keterampilan berpikir kritis antara siswa yang menggunakan model pembelajaran LCV-Lab (<i>Learning Cycle 9E</i> berbantuan <i>Virtual Laboratory</i>) dengan siswa yang menggunakan model pembelajaran LC	Nilai <i>post-test</i> keterampilan berpikir kritis siswa dari kedua kelompok (LCV-Lab dan LC 5E)	Tes keterampilan berpikir kritis

Pertanyaan	Data	Instrumen
(Learning Cycle 5E) pada materi fluida dinamis?	Bagaimana ukuran dampak pembelajaran menggunakan model LCV-Lab (Learning Cycle 9E berbantuan Virtual Laboratory) terhadap peningkatan keterampilan berpikir kritis siswa pada materi fluida dinamis?	Nilai <i>post-test</i> dari dua kelompok (LCV-Lab dan LC 5E) Tes keterampilan berpikir kritis

3.4.1 Rencana Pelaksanaan Pembelajaran

Rencana Pelaksanaan Pembelajaran (RPP) adalah panduan yang digunakan oleh guru saat mengajar di kelas. RPP dirancang untuk mendukung pengajaran agar sesuai dengan kompetensi dasar yang diperlukan dalam pertemuan tersebut. Dalam RPP ini, terdapat dua pertemuan yang disesuaikan dengan kurikulum yang berlaku di sekolah, yaitu Kurikulum Merdeka. RPP mencakup model pembelajaran, tujuan pembelajaran, media pembelajaran, materi pembelajaran, kegiatan pembelajaran, dan lembar kerja siswa. Untuk pembahasan sub materi pada setiap pertemuan seperti yang tercantum pada tabel 3.3 berikut:

Tabel 3. 3 Pembahasan Sub Materi pada setiap pertemuan

Pertemuan ke-	Sub Materi
1	Asas Kontinuitas dan Prinsip Bernoulli
2	Penerapan Prinsip Bernoulli

3.4.2 Lembar Observasi Model Pembelajaran

Selama proses pembelajaran, lembar observasi digunakan untuk mengukur ketercapaian pembelajaran LCV-Lab (Learning Cycle 9E berbantuan Virtual Laboratory). Lembar observasi terdiri dari penilaian keterlaksanaan yang didasarkan pada kegiatan pendahuluan, inti, dan penutup. Lembar observasi dibuat setiap pertemuan, dan diisi oleh *observer* dengan memberikan tanda checklist (\checkmark) pada kolom “YA” jika aktivitas guru dan siswa terlaksana atau pada kolom “TIDAK” jika aktivitas guru dan siswa tidak terlaksana sesuai dengan kolom indikator yang sedang diobservasi. Berikut Cuplikan lembar observasi disajikan pada gambar 3.1 berikut.

LEMBAR OBSERVASI KETERLAKSANAAN AKTIVITAS GURU DAN SISWA PADA PEMBELAJARAN LCV-LAB			
Nama Peneliti	: Gani Febriani		
Nama Observer	:		
Pertemuan ke-	: 1		
Materi	: Asas Kontinuitas dan Prinsip Bernoulli		
Hari/Tanggal	:		
Kelas yang diamati	:		
Petunjuk pengisian:			
<ol style="list-style-type: none"> Berdasarkan pendapat Bapak/Ibu, mohon berikan tanda centang (✓) pada kolom "Ya" jika kegiatan dilaksanakan atau kolom "Tidak" jika kegiatan tidak dilaksanakan. Apabila pada saat pembelajaran diperlukan perbaikan atau masukan, mohon berikan komentar atau masukan secara bebas dan ditulis pada kolom yang telah disediakan. 			
Tahapan Pembelajaran	Aspek yang diamati	Keterlaksanaan	
		YA	TIDAK
Pendahuluan			
Orientasi	<ul style="list-style-type: none"> Guru mengucapkan salam Guru memeriksa kehadiran siswa 		
Apersepsi	<ul style="list-style-type: none"> Guru mengulas kembali materi sebelumnya dan menghubungkannya dengan materi yang akan dibahas. Guru memberikan apersepsi kepada siswa berkaitan dengan materi yang telah dipelajari. 		

Gambar 3.1 Cuplikan Lembar Obervasi Keterlaksanaan

3.4.3 Lembar Kerja Peserta Didik (LKPD)

Tujuan Lembar Kerja Peserta Didik (LKPD) adalah untuk membantu siswa dalam melakukan kegiatan percobaan atau eksperimen. LKPD dibuat mengikuti tahapan model *Learning Cycle 9E*, yang mencakup tugas-tugas yang harus dilakukan siswa untuk menemukan dan menerapkan ide. Dalam penelitian ini, terdapat dua LKPD: persamaan Kontinuitas, prinsip Bernoulli dan penerapannya. Berikut Cuplkan LKPD disajikan pada Gambar 3.2.



Gambar 3.2 Cuplikan LKPD pada Pertemuan 1

3.4.4 Tes Keterampilan Berpikir Kritis

Tes keterampilan berpikir kritis adalah alat evaluasi yang digunakan untuk menilai peningkatan keterampilan berpikir kritis siswa pada topik fluida dinamis. Instrumen tes keterampilan berpikir kritis terdiri dari soal *pre-test* dan *post-test* berjumlah 20 soal dalam bentuk *two-tier multiple choice*. Setiap butir soal terdiri dari dua komponen yaitu tier pertama berupa pilihan ganda, tier kedua berupa pilihan alasan terhadap tier pertama. Soal dalam tes keterampilan berpikir kritis mengacu pada indikator menurut Ennis (1985) yaitu kemampuan *elementary clarification* (memberikan penjelasan sederhana), *the basic support* (membangun keterampilan dasar), *inference* (menyimpulkan), *advanced clarification* (memberikan klarifikasi lanjutan), dan *strategy and tactics* (menyusun strategi dan

taktik). Matrik instrument tes keterampilan berpikir kritis ditunjukkan pada tabel 3.4 berikut.

Tabel 3. 4 Matrik Instrumen Tes Berpikir Kritis

Nomor Soal	Sub Materi	Aspek Berpikir Kritis	Indikator Soal
1	Persamaan Kontinuitas	<i>Elementary Clarification</i>	Disajikan gambar fenomena sedang menyiram tanaman dengan selang, peserta didik dapat merumuskan pertanyaan yang menjelaskan fenomena tersebut
2	Persamaan Kontinuitas	<i>Elementary Clarification</i>	Disajikan beberapa argumen tentang aliran darah, peserta didik dapat menganalisis argumen yang benar menjelaskan atas kontinuitas.
3	Persamaan Kontinuitas	<i>The Basic Support</i>	Disajikan hasil observasi aliran air pada saluran irigasi yang menyempit, peserta didik dapat menentukan pernyataan yang paling tepat menggambarkan aliran lebih cepat di bagian menyempit.
4	Persamaan Kontinuitas	<i>Inference</i>	Disajikan ilustrasi gambar pipa dengan luas penampang berbeda, peserta didik dapat menyimpulkan kecepatan aliran air pada masing-masing pipa.
5	Persamaan Kontinuitas	<i>Advanced Clarification</i>	Disajikan informasi mengenai penggunaan selang air dan fenomena yang terjadi, peserta didik dapat mendefinisikan istilah yang tepat berdasarkan fenomena tersebut.
6	Persamaan Kontinuitas	<i>Strategy and Tactics</i>	Disajikan permasalahan pipa pembuangan air hujan, peserta didik dapat menentukan strategi yang paling efektif untuk mengatasi permasalahan tersebut.
7	Persamaan Bernoulli	<i>Elementary Clarification</i>	Disajikan gambar fenomena pesawat terbang sedang lepas landas, peserta didik dapat merumuskan pertanyaan yang menjelaskan fenomena tersebut.
8	Persamaan Bernoulli	<i>Elementary Clarification</i>	Disajikan beberapa argumen tentang konsep dasar dari prinsip

Nomor Soal	Sub Materi	Aspek Berpikir Kritis	Indikator Soal
			Bernoulli, peserta didik dapat menganalisis argumen yang benar menjelaskan prinsip Bernoulli.
9	Persamaan Bernoulli	<i>The Basic Support</i>	Disajikan hasil observasi perahu yang melaju berdekatan dengan kecepatan tinggi, peserta didik dapat menentukan pernyataan yang menggambarkan penurunan tekanan yang menyebabkan benturan.
10	Persamaan Bernoulli	<i>Inference</i>	Disajikan gambar ilustrasi pipa dengan jenis cairan dan ketinggian berbeda, peserta didik dapat menyimpulkan kedua pipa tersebut.
11	Persamaan Bernoulli	<i>Inference</i>	Disajikan gambar ilustrasi desain sayap pesawat dan mobil balap, peserta didik dapat menyimpulkan prinsip kerja yang digunakan pada pesawat terbang dan mobil balap.
12	Persamaan Bernoulli	<i>Inference</i>	Disajikan tabel dan dua gambar pipa dengan kondisi yang berbeda, peserta didik dapat menyimpulkan hubungan tekanan terhadap ketinggian, dan luas penampang pada pipa tersebut.
13	Persamaan Bernoulli	<i>Advanced Clarification</i>	Disajikan informasi mengenai penggunaan venturimeter dalam mengukur laju aliran air, peserta didik dapat mendefinisikan istilah yang tepat berdasarkan fenomena tersebut.
14	Persamaan Bernoulli	<i>Strategy and Tactics</i>	Disajikan gambar pipa PDAM, peserta didik dapat menentukan strategi pemilihan diameter yang tepat agar dihasilkan aliran yang efisien.
15	Teorema Toricelli	<i>Elementary Clarification</i>	Disajikan gambar ilustrasi sebuah tangki dengan air yang mengalir keluar dari lubang disampingnya, peserta didik dapat merumuskan pertanyaan yang menjelaskan fenomena tersebut
15	Teorema Toricelli	<i>Elementary Clarification</i>	Disajikan beberapa argumen tentang hasil pola aliran air dari

Nomor Soal	Sub Materi	Aspek Berpikir Kritis	Indikator Soal
			beberapa lubang dengan ketinggian berbeda, peserta didik dapat menganalisis argumen yang benar menjelaskan teorema Toricelli
17	Teorema Toricelli	<i>The Basic Support</i>	Disajikan data hasil observasi hubungan selisih kedalaman air dengan jarak jatuhnya air, peserta didik dapat mempertimbangkan hasil observasi untuk menentukan grafik yang tepat
18	Teorema Toricelli	<i>Inference</i>	Disajikan data hasil percobaan botol air berlubang dengan diameter sama tetapi ketinggian berbeda, peserta didik dapat menyimpulkan kecepatan aliran air dari lubang-lubang tersebut.
19	Teorema Toricelli	<i>Advanced Clarification</i>	Disajikan gambar ilustrasi tangka air untuk mengairi sawah, peserta didik dapat mendefinisikan istilah yang tepat berdasarkan fenomena tersebut.
20	Teorema Toricelli	<i>Strategy and Tactics</i>	Disajikan permasalahan dan strategi penyelesaiannya, peserta didik dapat menentukan strategi yang paling efektif dan efisien untuk mengatasi permasalahan tersebut

Setiap pertanyaan diberi skor 3 jika jawaban dan alasan benar. Jika jawaban salah tetapi alasan benar, mendapat skor 2. Jika jawaban benar tetapi alasan salah, mendapat skor 1. Jika jawaban dan alasan keduanya salah, mendapat skor 0. Rumus penghitungan data dan rubrik penilaian dapat dilihat pada tabel 3.5 berikut.

Tabel 3. 5 Skoring Soal Two Tier Multiple Choice

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

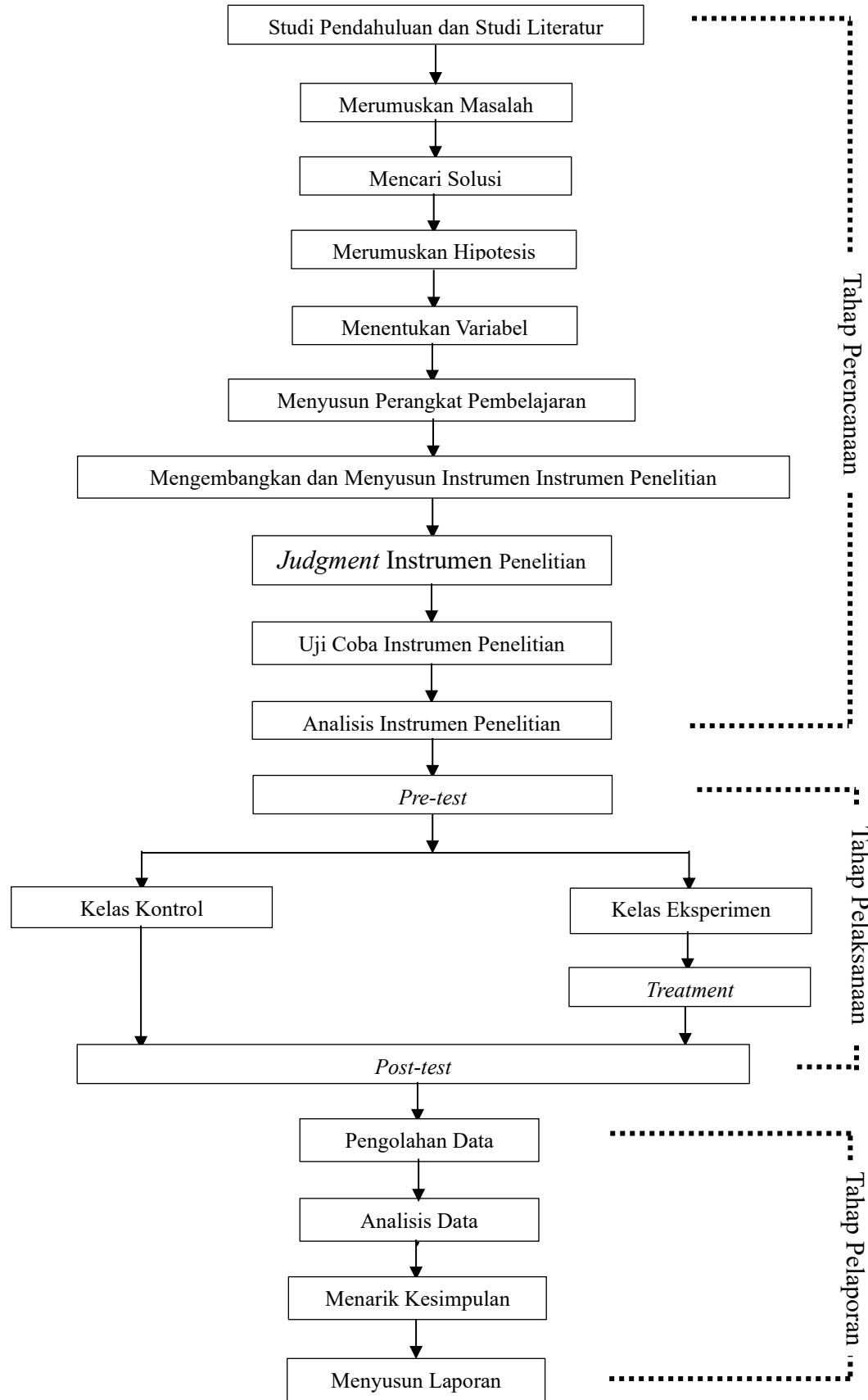
3.5 Prosedur Penelitian

Pada penelitian ini, prosedur penelitian dibagi menjadi tiga tahap, yaitu tahap persiapan, tahap pelaksanaan, dan tahap pelaporan. Berikut masing-masing prosedur dari setiap tahapnya.

- 1) Tahap Persiapan
 - a) Studi pendahuluan dan studi literatur dari jurnal-jurnal, buku, skripsi, tesis, dan sumber lainnya untuk mencari informasi dari permasalahan.
 - b) Merumuskan masalah yang akan dikaji untuk penelitian.
 - c) Mencari alternatif solusi dari permasalahan yang telah dirumuskan.
 - d) Merumuskan Hipotesis
 - e) Menentukan variabel penelitian
 - f) Menentukan metode penelitian dan teknik pengolahan data untuk mendapatkan jawaban dari rumusan masalah yang telah dibuat.
 - g) Menentukan dan membuat instrumen penelitian sebagai teknik pengumpulan data.
 - h) Menyusun RPP mengenai model pembelajaran *Learning Cycle (9E)* dan *Learning Cycle (5E)*.
 - i) Validasi dan *judgment* instrumen oleh ahli.
 - j) Pengujian instrumen ke sekolah sebelum digunakan oleh sampel.
 - k) Pengolahan, revisi, dan seleksi soal dari data hasil uji coba instrumen.
 - l) Analisis hasil uji coba instrument
- 2) Tahap Pelaksanaan
 - a) Melakukan *pre-test* pada kedua kelas penelitian yaitu kelas eksperimen dan kelas kontrol menggunakan instrumen tes keterampilan berpikir kritis sebelum diberi perlakuan pada kelas eksperimen.
 - b) Melaksanakan pembelajaran pada kelas eksperimen dan kelas kontrol. Pada kelas eksperimen diberi perlakuan dengan kegiatan belajar menggunakan model pembelajaran *Learning Cycle 9E* berbantuan *Virtual Laboratory*. sedangkan pada kelas kontrol diberi perlakuan model LC 5E (*Learning Cycle 5E*)

- c) Melaksanakan observasi terkait keterlaksanaan kegiatan pembelajaran di kelas yang dilakukan oleh *observer*.
 - d) Melakukan *post-test* pada kedua kelas penelitian yaitu kelas eksperimen dan kelas kontrol menggunakan instrumen tes keterampilan berpikir kritis setelah diberi perlakuan pada kelas eksperimen.
- 3) Tahap Pelaporan
- a) Melakukan pengolahan instrumen dari hasil penelitian berupa *pretest*, dan *posttest*.
 - b) Menganalisis data hasil pengolahan data
 - c) Menarik kesimpulan
 - d) Menyusun laporan hasil penelitian
- Berikut ini merupakan skema penelitian yang digunakan sebagai acuan dalam pelaksanaan penelitian.

Gambar 3.3 Skema Prosedur Penelitian



3.6 Teknik Analisis Data

3.6.1 Uji Instrumen Penelitian

Uji Instrumen penelitian dilakukan terhadap instrumen soal tes keterampilan berpikir kritis. Beberapa uji instrumen yang dilakukan antara lain sebagai berikut:

3.6.1.1 Validitas

Validitas adalah ukuran yang menunjukkan seberapa valid atau sah suatu instrumen. Valid berarti instrumen tersebut dapat digunakan untuk mengukur apa yang seharusnya diukur (Sumintono & Widhiarso, 2015). Validitas tes ini diuji secara isi dan konstruks.

3.6.1.1.1 Validitas isi

Uji validitas isi bertujuan untuk menilai sejauh mana butir soal dalam tes mewakili domain yang diukur. Validitas ini tidak hanya berdasarkan pada isi soal, tetapi juga mempertimbangkan format, susunan kata, administrasi, dan penilaian tes (Johnson & Christensen, 2020). Dalam penelitian ini, uji validitas isi instrumen dilakukan oleh lima ahli, termasuk tiga dosen dan dua guru. Lima ahli tersebut menilai relevansi butir soal dengan keterampilan berpikir kritis serta kualitas kontennya. Uji validasi dilakukan terhadap setiap soal yang mengacu pada aspek dilembar validasi yang terdiri dari tiga pernyataan aspek materi, tiga pernyataan kontruksi, dan 3 pertanyaan aspek bahasa. Aspek penilaian instrument tes keterampilan berpikir kritis dapat dilihat pada tabel 3.6 berikut.

Tabel 3. 6 Aspek Penilaian Instrumen Tes Keterampilan Berpikir Kritis

No	Aspe yang dinilai	
1	Butir soal dapat mengukur KBK	
2	Butir soal sesuai dengan aspek KBK	Materi
3	Butir soal sesuai dengan indikator soal	
4	Butir soal dirumuskan dengan jelas dan benar	
5	Pilihan jawaban benar dan hanya ada satu jawaban yang tepat	Konstruksi
6	Jika terdapat gambar/grafik/tabel dalam soal dapat terlihat dengan jelas dan berfungsi	
7	Rumusan kalimat dalam butir soal komunikatif dan jelas.	
8	Butir soal menggunakan bahasa indonesia yang baik dan benar	Bahasa
9	Kalimat soal tidak mengandung penafsiran ganda	

Setiap ahli diberikan instrumen tes untuk diisi pada lembar validasi dengan skala penilaian dari 0 (tidak sesuai) hingga 3 (sangat sesuai). Hasil uji validitas isi

oleh ahli diolah menggunakan koefisien validasi Aiken (V). Dengan 5 orang ahli dan 4 kategori, nilai validitas harus mencapai 0,87. Jika aspek nomor 1 tidak valid, soal tersebut tidak digunakan. Jika aspek nomor 2-9 tidak valid, soal diperbaiki dan tetap digunakan. Koefisien validasi Aiken dihitung dengan persamaan berikut (Aiken, 1985):

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

(Aiken, 1985)

Dengan $s = \sum(r - l_0)$

Keterangan :

V = koefisien validitas Aiken

r = skor yang diberikan validator

l_0 = Skor terendah pada kategori

n = jumlah validator

c = jumlah kategori penilaian

Setelah dinilai oleh para ahli, validitas logis diuji menggunakan validasi Aiken (lihat Tabel 3.7) dan masukan para ahli dicatat pada Tabel 3.8 berikut

Tabel 3. 7 Validitas Butir Soal Menggunakan Validitas Ahli

No	Aspek yang dinilai									Rata-Rata	Ket
	1	2	3	4	5	6	7	8	9		
1	0,93	0,93	0,93	0,93	0,93	0,93	0,87	0,93	0,93	0,93	V
2	0,93	0,93	0,93	0,93	0,87	0,93	0,93	0,93	0,93	0,93	V
3	1,00	1,00	0,93	0,93	1,00	0,93	0,93	0,93	0,93	0,96	V
4	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	V
5	0,93	0,93	0,93	0,93	0,93	0,93	0,93	0,93	0,93	0,93	V
6	0,93	0,93	0,93	0,93	1,00	0,93	0,93	0,93	0,93	0,94	V
7	0,93	0,93	0,93	0,93	1,00	0,93	0,93	0,93	0,93	0,94	V
8	0,93	0,93	0,93	0,93	1,00	0,93	0,93	0,93	0,93	0,94	V
9	0,93	0,93	0,93	1,00	1,00	0,93	0,93	0,93	0,93	0,95	V
10	0,93	0,93	0,93	1,00	1,00	0,93	0,93	0,93	0,93	0,95	V
11	0,93	0,93	0,93	1,00	1,00	0,93	0,93	0,93	0,93	0,95	V
12	0,93	0,93	0,93	1,00	1,00	0,93	0,93	0,93	0,93	0,95	V
13	0,93	0,93	0,93	1,00	1,00	1,00	0,93	0,93	0,93	0,96	V
14	0,93	0,93	0,93	1,00	1,00	1,00	0,93	0,93	0,93	0,96	V
15	1,00	1,00	1,00	1,00	1,00	1,00	0,93	0,93	0,93	0,98	V
16	0,93	0,93	0,93	1,00	1,00	1,00	0,93	0,93	0,93	0,96	V
17	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	V

No	Aspek yang dinilai									Rata-Rata	Ket
	1	2	3	4	5	6	7	8	9		
18	0,93	0,93	0,93	1,00	1,00	0,93	0,87	0,93	0,93	0,94	V
19	0,93	0,93	0,93	1,00	1,00	1,00	0,87	0,93	0,93	0,95	V
20	0,93	0,93	0,93	1,00	1,00	1,00	0,87	0,93	0,93	0,95	V
21	0,93	0,93	0,93	1,00	1,00	0,93	0,93	0,93	0,93	0,95	V
22	1,00	1,00	1,00	1,00	1,00	0,93	0,93	0,93	0,93	0,97	V
23	0,93	0,93	0,93	1,00	1,00	1,00	0,87	0,93	0,93	0,95	V

Berdasarkan hasil Tabel 3.7, koefisien validitas Aiken (V) menunjukkan bahwa dari 23 butir soal yang dinilai, semua soal valid. Oleh karena itu, dapat disimpulkan bahwa semua butir soal dapat diuji coba di lapangan.

3.6.1.1.2 Validitas Konstruk

Uji validitas konstruk bertujuan untuk menilai sejauh mana butir soal dan komponen tes sesuai dengan konsep yang diukur (Sumintono & Widhiarso, 2015). Dalam penelitian ini, interpretasi validitas konstruk instrumen tes keterampilan berpikir kritis dilakukan dengan menggunakan model *Rasch* melalui *software Winstep* versi 3.73. Untuk memastikan bahwa instrumen tes secara keseluruhan dapat mengukur keterampilan berpikir kritis sesuai dengan tujuan yang telah ditetapkan, maka dianalisis undimensionalitas instrumen berdasarkan nilai *raw variance explained by measures* yang dapat ditemukan pada menu *output* tabel 23. Selain itu, analisis menggunakan *Principal Component Analysis* (PCA) dari residual digunakan untuk mengukur sejauh mana instrumen mengukur apa yang seharusnya diukur, dengan memperhatikan nilai *unexplained variance in contrasts* 1-5 (Fisher, 2007; Sumintono & Widhiarso, 2015).

Tabel 3. 8 Kriteria Nilai Unidimensionalitas Instrumen

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

(Sumintono & Widhiarso, 2015)

Tabel 3. 9 Kriteria Nilai Unexplained Variance in Contrasts

Nilai Unexplained Variance in Contrasts 1-5 (%)	Kriteria
< 3	Istimewa
3 – 5	Sangat baik
5 – 10	Baik
10 – 15	Cukup baik

Nilai Unexplained Variance in Contrasts 1-5 (%)	Kriteria
> 15	Buruk
	(Fisher, 2007)

Hasil *unidimensionalitas* instrumen dari *output tables* bagian tabel 23: *item dimensionality* pada software Winstep ditunjukkan pada Gambar 3.6 berikut.

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)			
	-- Empirical --	Modeled	
Total raw variance in observations =	31.0	100.0%	100.0%
Raw variance explained by measures =	8.0	25.8%	25.8%
Raw variance explained by persons =	2.5	7.9%	7.9%
Raw Variance explained by items =	5.6	17.9%	17.9%
Raw unexplained variance (total) =	23.0	74.2% 100.0%	74.2%
Unexplned variance in 1st contrast =	4.2	13.6%	18.4%
Unexplned variance in 2nd contrast =	3.3	10.7%	14.5%
Unexplned variance in 3rd contrast =	2.6	8.3%	11.2%
Unexplned variance in 4th contrast =	1.8	5.7%	7.6%
Unexplned variance in 5th contrast =	1.7	5.4%	7.3%

Gambar 3. 6 Hasil Uji *Unidimensionalitas* Instrumen KBK

Gambar 3.4 menunjukkan bahwa nilai *raw variance explained by measures* dari uji coba lapangan instrumen dengan 23 butir soal adalah 25,8%. Berdasarkan Tabel 3.8, nilai ini memenuhi kriteria “terpenuhi”, yang berarti instrumen dapat mengukur satu variabel tanpa dipengaruhi oleh variabel lain. Selain nilai *raw variance explained by measures*, *unidimensionalitas* instrumen juga dapat dilihat dari nilai *unexplained variance in 1st contrast*. Jika nilainya kurang dari 15%, instrumen dianggap memiliki *unidimensionalitas* yang baik. Hasil analisis menunjukkan bahwa nilai *unexplained variance in 1st contrast* adalah 13,6%, yang berarti instrumen memiliki *unidimensionalitas* yang cukup baik.

Analisis validitas untuk setiap butir soal menggunakan model Rasch diperoleh dengan melihat nilai *Mean-Square Outfit* (MNSQ), *Z-Standard Outfit* (ZSTD), dan *Point Measure Correlation* (*Pt Measure Corr*). Hasil analisis dapat dilihat pada menu *output* bagian tabel 10: *Item fit order*. Menurut Sumintono & Widhiarso (2015), kriteria untuk menilai apakah nilai yang diperoleh berada dalam rentang yang diterima atau tidak dapat dilihat pada tabel 3.10 berikut.

Tabel 3. 10 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)	$-2,0 < \text{ZSTD} < 2,0$

Kriteria	Nilai yang Diterima
Pt Measure Corr	0,4 < Pt Measure Corr < 0,85 (Sumintono & Widhiarso 2015)

Hasil dari nilai kriteria tersebut kemudian diinterpretasikan berdasarkan kriteria nilai *fit-statistic* pada Tabel 3.11 berikut.

Tabel 3. 11 Interpretasi Kualitas Butir Soal

Kriteria Nilai <i>Fit-Statistic</i>	Interpetasi
Ketiga kriteria nilai terpenuhi	Digunakan
Dua dari tiga kriteria nilai terpenuhi	Digunakan
Satu dari tiga kriteria nilai terpenuhi	Tidak Digunakan
Semua kriteria nilai tidak terpenuhi	Tidak Digunakan

(Sumintono & Widhiarso, 2014)

Kategori kualitas butir soal instrumen tes keterampilan berpikir kritis berdasarkan hasil uji coba lapangan dengan 47 responden ditunjukkan pada Gambar 3.4 dan Tabel 3.12 melalui nilai *fit-statistic*.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	OUTFIT MNSQ	PT-MEASURE CORR.	EXACT EXP.	MATCH OBS%	ITEM EXP%
9	18	47	.03	.32	1.76	5.3	2.67	5.6 A-.45	.37	40.4 68.7
7	5	47	1.89	.49	1.28	.9	2.24	1.7 B-.18	.24	89.4 89.3
8	3	47	2.48	.61	1.11	.4	1.81	1.1 C-.01	.19	93.6 93.6
5	21	47	-.28	.32	1.32	2.7	1.28	1.5 D .10	.38	46.8 66.4
3	33	47	-1.55	.35	1.15	.9	1.06	.3 E .28	.39	70.2 74.6
2	23	47	-.48	.32	1.13	1.2	1.08	.6 F .28	.39	51.1 65.8
4	26	47	-.78	.32	1.07	.7	1.00	.0 G .35	.40	53.2 66.7
14	18	47	.03	.32	1.07	.6	.96	-.1 H .33	.37	53.2 68.7
6	35	46	-1.92	.38	1.03	.2	.89	-.3 I .38	.38	73.9 78.6
19	14	47	.47	.34	1.02	.2	.95	-.1 J .34	.35	68.1 72.7
17	20	47	-.18	.32	1.00	.1	.93	-.3 K .39	.38	61.7 67.0
23	17	47	.14	.33	.98	-.1	.87	-.5 L .41	.37	61.7 69.7
15	18	47	.03	.32	.96	-.3	.87	-.6 k .43	.37	70.2 68.7
20	8	47	1.29	.41	.92	-.3	.70	-.6 j .40	.29	80.9 83.2
22	7	47	1.47	.43	.91	-.2	.67	-.6 i .39	.27	83.0 85.1
1	35	47	-1.80	.36	.90	-.5	.78	-.7 h .49	.39	78.7 77.5
16	17	47	.14	.33	.89	-.9	.80	-.9 g .48	.37	74.5 69.7
11	20	47	-.18	.32	.86	-1.3	.78	-1.2 f .53	.38	66.0 67.0
13	26	47	-.78	.32	.83	-1.6	.77	-1.4 e .56	.40	74.5 66.7
10	20	47	-.18	.32	.81	-1.7	.76	-1.4 d .56	.38	83.0 67.0
18	13	47	.59	.35	.73	-1.8	.60	-1.5 c .61	.34	83.0 74.1
12	22	47	-.38	.32	.72	-2.9	.65	-2.3 b .66	.39	78.7 65.9
21	19	47	-.07	.32	.69	-3.0	.62	-2.2 a .67	.38	80.9 67.8
MEAN	19.0	47.0	.00	.36	1.01	-.1	1.03	-.2		70.3 72.8
S.D.	8.4	.2	1.05	.07	.23	1.7	.51	1.6		13.8 7.9

Gambar 3.4 Hasil Uji Validitas Instrumen KBK

Tabel 3. 12 Validitas butir soal

Nomor Butir Soal	Nilai Outfit		Pt. Measure Corr.	Kriteria Nilai	Ket
	MNSQ	ZSTD			
1	0,78	-0,70	0,49	Ketiga kriteria nilai terpenuhi	Digunakan
2	1,08	0,60	0,28	Dua dari tiga kriteria nilai terpenuhi	Digunakan
3	1,06	0,30	0,28	Dua dari tiga kriteria nilai terpenuhi	Digunakan
4	1,00	00,0	0,35	Dua dari tiga kriteria nilai terpenuhi	Digunakan
5	1,28	1,5	0,10	Dua dari tiga kriteria nilai terpenuhi	Digunakan
6	0,89	-0,30	0,38	Dua dari tiga kriteria nilai terpenuhi	Digunakan
7	2,24	1,70	-0,18	Satu dari tiga kriteria nilai terpenuhi	Tidak digunakan
8	1,81	1,10	-0,01	Satu dari tiga kriteria nilai terpenuhi	Tidak digunakan
9	2,67	5,60	-0,45	Ketiga kriteria nilai tidak terpenuhi	Tidak digunakan
10	0,76	-1,40	0,56	Ketiga kriteria nilai terpenuhi	Digunakan
11	0,78	-1,20	0,53	Ketiga kriteria nilai terpenuhi	Digunakan
12	0,65	-2,30	0,66	Dua dari tiga kriteria nilai terpenuhi	Digunakan
13	0,60	-1,50	0,61	Ketiga kriteria nilai terpenuhi	Digunakan
14	0,96	-0,10	0,33	Dua dari tiga kriteria nilai terpenuhi	Digunakan
15	0,87	-,60	0,43	Ketiga kriteria nilai terpenuhi	Digunakan
16	0,80	-0,90	0,48	Ketiga kriteria nilai terpenuhi	Digunakan
17	0,93	-0,30	0,39	Dua dari tiga kriteria nilai terpenuhi	Digunakan
18	0,65	-2,30	0,66	Ketiga kriteria nilai terpenuhi	Digunakan
19	0,95	-0,10	0,34	Dua dari tiga kriteria nilai terpenuhi	Digunakan
20	0,78	-1,20	0,53	Ketiga kriteria nilai terpenuhi	Digunakan

Nomor Butir Soal	Nilai Outfit		Pt. Measure Corr.	Kriteria Nilai	Ket
	MNSQ	ZSTD			
21	0,62	-2,20	0,67	Satu dari tiga kriteria nilai terpenuhi	Digunakan
22	0,67	-0,60	0,39	Dua dari tiga kriteria nilai terpenuhi	Digunakan
23	0,87	-0,50	0,41	Ketiga kriteria nilai terpenuhi	Digunakan

Tabel 3.12 menunjukkan hasil validasi instrumen tes keterampilan berpikir kritis. Dari 23 butir soal yang dianalisis, 20 butir soal dinyatakan valid dan digunakan, sedangkan 3 butir soal tidak digunakan. Dengan demikian, dapat disimpulkan bahwa 20 butir soal valid dan dapat digunakan sebagai instrumen penelitian.

3.6.1.2 Uji Reliabilitas

Reliabilitas instrumen mengacu pada konsistensi atau kestabilan skor dari suatu tes (Johnson & Christensen, 2020). Dalam konteks penelitian, reliabilitas menunjukkan sejauh mana instrumen pengukur dapat menghasilkan hasil yang seragam ketika diulang pada waktu yang berbeda. Misalnya, ketika instrumen digunakan untuk mengukur kemampuan peserta didik, skor yang diperoleh seharusnya tidak berbeda secara signifikan saat diberikan kembali pada waktu yang berbeda. Analisis uji reliabilitas sering dilakukan menggunakan metode pemodelan Rasch dengan *software* seperti Winstep versi 3.73. Hasil uji reliabilitas instrumen keterampilan berpikir kritis menggunakan *software* Winstep versi 3.73 dapat dilihat pada menu *output tables* bagian tabel 3.1: *summary statistics*. Hasil uji reliabilitas mencakup beberapa indikator, termasuk *person reliability*, *item reliability*, dan nilai *Cronbach Alpha*. *Person reliability* mengukur konsistensi jawaban peserta didik, sedangkan *item reliability* menilai kualitas butir-butir soal dalam instrumen. Nilai *Cronbach Alpha* menggambarkan reliabilitas secara keseluruhan, mengukur interaksi antara person dan item. Kriteria reliabilitas ini dapat dilihat dalam tabel 3.13 dan 3.14 berikut.

Tabel 3. 13 Interpretasi Nilai Person Reliability dan Item Reliability

Rentang Nilai Person Reliability dan Item Reliability	Interpretasi
$r > 0,94$	Baik Sekali
$0,90 < r \leq 0,94$	Sangat Baik

Rentang Nilai Person Reliability dan Item Reliability	Interpretasi
$0,80 < r \leq 0,90$	Baik
$0,67 < r \leq 0,80$	Cukup
$< 0,67$	Rendah

(Sumintono & Widhiarso, 2015)

Tabel 3. 14 Interpretasi Nilai *Cronbach Alpha*

Rentang Nilai <i>Cronbach Alpha</i>	Interpretasi
$KR - 20 > 0,80$	Sangat Tinggi
$0,70 \leq KR - 20 < 0,80$	Tinggi
$0,60 \leq KR - 20 < 0,70$	Baik
$0,5 \leq KR - 20 < 0,60$	Sedang
$< 0,50$	Rendah

(Sumintono & Widhiarso, 2015)

Nilai *separation* merupakan faktor penting dalam menilai kualitas instrumen. Separation mengacu pada kemampuan instrumen untuk membedakan antara peserta didik dan butir soal. Semakin tinggi nilai separation, semakin baik instrumen dalam mengukur keseluruhan respons peserta didik dan kualitas butir-butir soal (Sumintono & Widhiarso, 2015). Analisis lebih lanjut mengenai pengelompokan peserta didik dan butir soal dapat ditemukan melalui persamaan pemisahan strata seperti pada persamaan 3.1 berikut.

$$H = \frac{[(4 \times SEPARATION) + 1]}{3} \quad (3.1)$$

(Sumintono & Widhiarso, 2015)

Hasil uji reliabilitas instrumen dari output tabel *summary statistic* pada software Winstep ditunjukkan pada Gambar 3.5 berikut.

SUMMARY OF 47 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	9.3	23.0	-.55	.51	.99	.0	1.03	.1
S.D.	3.8	.1	.97	.10	.17	.8	.47	1.0
MAX.	18.0	23.0	1.53	1.04	1.41	1.8	3.03	3.5
MIN.	1.0	22.0	-3.53	.46	.61	-1.4	.31	-1.2
REAL RMSE	.53	TRUE SD	.81	SEPARATION	1.52	Person	RELIABILITY	.70
MODEL RMSE	.52	TRUE SD	.82	SEPARATION	1.59	Person	RELIABILITY	.72
S.E. OF Person MEAN	=	.14						
Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .70								
SUMMARY OF 23 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	19.0	47.0	.00	.36	1.01	-.1	1.03	-.2
S.D.	8.4	.2	1.05	.07	.23	1.7	.51	1.6
MAX.	35.0	47.0	2.48	.61	1.76	5.3	2.67	5.6
MIN.	3.0	46.0	-1.92	.32	.69	-3.0	.60	-2.3
REAL RMSE	.38	TRUE SD	.98	SEPARATION	2.57	Item	RELIABILITY	.87
MODEL RMSE	.36	TRUE SD	.98	SEPARATION	2.70	Item	RELIABILITY	.88
S.E. OF Item MEAN	=	.22						

Gambar 3.5 Hasil Uji Reliabilitas *Output Summary Statistic*

Berdasarkan Gambar 3.13 *person reliability* bernilai 0,70 dengan kategori “Cukup”, sementara item reliability sebesar 0,87 dengan kategori “Baik”. Ini menunjukkan bahwa instrumen tes keterampilan berpikir kritis konsisten dalam bobot soal dan pilihan jawabannya. Nilai *Cronbach's Alpha* (KR-20) sebesar 0,70 dengan kategori “Tinggi” menunjukkan kualitas instrumen yang tinggi karena dapat mengidentifikasi hubungan antara siswa (*person reliability*) dengan butir soal (*item reliability*). Berdasarkan analisis ini, dapat disimpulkan bahwa instrumen tes keterampilan berpikir kritis reliabel untuk digunakan dalam penelitian ini.

Pengelompokan peserta didik dan butir soal berdasarkan nilai *person separation*, *item separation*, dan pemisahan strata (nilai H) ditunjukkan pada Tabel 3.15 berikut.

Tabel 3.15 Nilai *Separation* dan Pemisahan Strata Instrumen KBK

	Nilai <i>Separation</i>	Nilai H
Person	1,52	2,36

	Nilai Separation	Nilai H
Item	2,57	3,76
Berdasarkan tabel 3.15 nilai person strata (H) sebesar 2,36, yang dibulatkan menjadi 3, menunjukkan bahwa sebaran kemampuan peserta didik terbagi menjadi tiga kelompok. Nilai item strata (H) sebesar 3,76, yang dibulatkan menjadi 4, menunjukkan bahwa tingkat kesulitan butir soal terbagi menjadi empat kelompok.		

3.6.1.3 Taraf Kesukaran

Tingkat kesulitan digunakan untuk menggambarkan sejauh mana peserta didik memahami suatu butir soal dibandingkan dengan item yang digunakan. Dalam penelitian ini, tingkat kesulitan (TK) dianalisis menggunakan pemodelan *Rasch* dengan bantuan *software* Winstep 3.73. Proses analisis melibatkan pengurutan nilai *logarithm odd unit* (logit) dari setiap butir soal. Data logit memiliki interval yang seragam, memungkinkan kita untuk memahami perbedaan kesulitan antar butir soal. Nilai TK diperoleh dari pengukuran (ME) dan standar deviasi (SD) dengan membandingkan logit ME pada setiap item dan nilai SD. Hasil penentuan tingkat kesulitan butir soal dapat ditemukan dalam *output* tabel 13: *item measure*. Interpretasi lebih lanjut mengenai tingkat kesulitan butir soal dapat ditemukan pada Tabel 3.16 berikut.

Tabel 3. 16 Interpretasi Taraf Kesukaran

Kriteria	Interpretasi
<i>Measure Logit</i> > +SD	Sulit
$-1SD < Measure Logit \leq +SD$	Sedang
<i>Measure Logit</i> < -SD	Mudah

(Sumintono & Widhiarso, 2015)

Item STATISTICS: MEASURE ORDER																
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	OUTFIT ZSTD	MNSQ	ZSTD	PT-MEASURE CORR.	EXACT EXP.	OBS% EXP%	Item				
8	3	47	2.48	.61	1.11	.4	1.81	1.1	-.01	.19	93.6	93.6	B8			
7	5	47	1.89	.49	1.28	.9	2.24	1.7	-.18	.24	89.4	89.3	B7			
22	7	47	1.47	.43	.91	-.2	.67	-.6	.39	.27	83.0	85.1	B22			
20	8	47	1.29	.41	.92	-.3	.70	-.6	.40	.29	80.9	83.2	B20			
18	13	47	.59	.35	.73	-1.8	.60	-1.5	.61	.34	83.0	74.1	B18			
19	14	47	.47	.34	1.02	.2	.95	-.1	.34	.35	68.1	72.7	B19			
16	17	47	.14	.33	.89	-.9	.80	-.9	.48	.37	74.5	69.7	B16			
23	17	47	.14	.33	.98	-.1	.87	-.5	.41	.37	61.7	69.7	B23			
9	18	47	.03	.32	1.76	5.3	2.67	5.6	-.45	.37	40.4	68.7	B9			
14	18	47	.03	.32	1.07	.6	.96	-.1	.33	.37	53.2	68.7	B14			
15	18	47	.03	.32	.96	-.3	.87	-.6	.43	.37	70.2	68.7	B15			
21	19	47	-.07	.32	.69	-3.0	.62	-2.2	.67	.38	80.9	67.8	B21			
10	20	47	-.18	.32	.81	-1.7	.76	-1.4	.56	.38	83.0	67.0	B10			
11	20	47	-.18	.32	.86	-1.3	.78	-1.2	.53	.38	66.0	67.0	B11			
17	20	47	-.18	.32	1.00	.1	.93	-.3	.39	.38	61.7	67.0	B17			
5	21	47	-.28	.32	1.32	2.7	1.28	1.5	.10	.38	46.8	66.4	B5			
12	22	47	-.38	.32	.72	-2.9	.65	-2.3	.66	.39	78.7	65.9	B12			
2	23	47	-.48	.32	1.13	1.2	1.08	.6	.28	.39	51.1	65.8	B2			
4	26	47	-.78	.32	1.07	.7	1.00	.0	.35	.40	53.2	66.7	B4			
13	26	47	-.78	.32	.83	-1.6	.77	-1.4	.56	.40	74.5	66.7	B13			
3	33	47	-1.55	.35	1.15	.9	1.06	.3	.28	.39	70.2	74.6	B3			
1	35	47	-1.80	.36	.90	-.5	.78	-.7	.49	.39	78.7	77.5	B1			
6	35	46	-1.92	.38	1.03	.2	.89	-.3	.38	.38	73.9	78.6	B6			
MEAN	19.0	47.0	.00	.36	1.01	-.1	1.03	-.2			70.3	72.8				
S.D.	8.4	.2	1.05	.07	.23	1.7	.51	1.6			13.8	7.9				

Gambar 3.6 Hasil taraf kesukaran *output item measure*

Tabel 3. 17 Hasil Interpretasi Daya Pembeda Item KBK

Nomor Butir Soal	Measure (ME)	Standar Deviasi (SD)	Kriteria	Interpretasi
1	-1,80	1,05	$-1,80 < -1,05$	Mudah
2	-0,48	1,05	$-1,05 < -0,48 \leq 1,05$	Sedang
3	-1,55	1,05	$-1,55 > -1,05$	Mudah
4	-0,78	1,05	$-1,05 < -0,78 \leq 1,05$	Sedang
5	-0,28	1,05	$-1,05 < -0,28 \leq 1,05$	Sedang
6	-1,92	1,05	$-1,92 < -1,05$	Mudah
10	-0,18	1,05	$-1,05 < -0,18 \leq 1,05$	Sedang
11	-0,18	1,05	$-1,05 < -0,18 \leq 1,05$	Sedang
12	-0,38	1,05	$-1,05 < -0,38 \leq 1,05$	Sedang
13	-0,78	1,05	$-1,05 < -0,78 \leq 1,05$	Sedang
14	0,03	1,05	$-1,05 < 0,03 \leq 1,05$	Sedang
15	0,03	1,05	$-1,05 < 0,03 \leq 1,05$	Sedang
16	0,14	1,05	$-1,05 < 0,14 \leq 1,05$	Sedang
17	-0,18	1,05	$-1,05 < -0,18 \leq 1,05$	Sedang
18	0,59	1,05	$-1,05 < 0,59 \leq 1,05$	Sedang
19	0,47	1,05	$-1,05 < 0,47 \leq 1,05$	Sedang

Nomor Butir Soal	Measure (ME)	Standar Deviasi (SD)	Kriteria	Interpretasi
20	1,29	1,05	$1,29 > 1,05$	Sulit
21	-0,07	1,05	$-1,05 < -0,07 \leq 1,05$	Sedang
22	1,47	1,05	$1,47 > 1,05$	Sulit
23	0,14	1,05	$-1,05 < 0,14 \leq 1,05$	Sedang

Berdasarkan Tabel 3.15, 2 butir soal termasuk tingkatan mudah, 16 butir soal termasuk tingkatan sedang, dan 2 butir soal termasuk tingkatan sulit. Hal ini menunjukkan bahwa taraf kesukaran terdistribusi dengan baik.

3.6.1.4 Daya Pembeda

Daya pembeda mengukur sejauh mana suatu butir soal mampu membedakan antara individu yang memiliki kemampuan tinggi dan rendah (Sumintono & Widhiarso, 2015). Dalam penelitian ini, kami menganalisis daya pembeda menggunakan pemodelan *Rasch* dengan memperhatikan nilai *Point Measure Correlation (Pt Mean Corr)*. *Pt Mean Corr* mengacu pada korelasi antara tingkat kesulitan masing-masing butir soal dengan tingkat kesulitan tes secara keseluruhan. Nilai *Pt Mean Corr* sebesar satu menunjukkan adanya korelasi sempurna antara respons butir soal dan kemampuan peserta tes. Nilai nol menandakan bahwa tidak ada hubungan antara respons butir soal tertentu dengan keseluruhan butir soal. Sedangkan nilai negatif mengindikasikan bahwa butir soal memiliki kelemahan pada beberapa aspek mendasar dan perlu diperiksa ulang atau bahkan dihapus dari pengujian (Smiley, 2015). Interpretasi lebih lanjut mengenai daya pembeda butir soal dapat ditemukan dalam Tabel 3.18.

Tabel 3. 18 Interpretasi Daya Pembeda

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

(Brown, 2005; Smiley, 2015)

Hasil analisis daya pembeda instrumen tes keterampilan berpikir kritis dengan software Winstep versi 3.73 ditampilkan pada Gambar 3.6 dan Tabel 3.19 berikut

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Item
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	
9	18	47	.03	.32	1.76	5.3	2.67	5.6	A .45	.37	40.4	68.7	B9
7	5	47	1.89	.49	1.28	.9	2.24	1.7	B -.18	.24	89.4	89.3	B7
8	3	47	2.48	.61	1.11	.4	1.81	1.1	C -.01	.19	93.6	93.6	B8
5	21	47	-.28	.32	1.32	2.7	1.28	1.5	D .10	.38	46.8	66.4	B5
3	33	47	-1.55	.35	1.15	.9	1.06	.3	E .28	.39	70.2	74.6	B3
2	23	47	-.48	.32	1.13	1.2	1.08	.6	F .28	.39	51.1	65.8	B2
4	26	47	-.78	.32	1.07	.7	1.00	.0	G .35	.40	53.2	66.7	B4
14	18	47	.03	.32	1.07	.6	.96	-.1	H .33	.37	53.2	68.7	B14
6	35	46	-1.92	.38	1.03	.2	.89	-.3	I .38	.38	73.9	78.6	B6
19	14	47	.47	.34	1.02	.2	.95	-.1	J .34	.35	68.1	72.7	B19
17	20	47	-.18	.32	1.00	.1	.93	-.3	K .39	.38	61.7	67.0	B17
23	17	47	.14	.33	.98	-.1	.87	-.5	L .41	.37	61.7	69.7	B23
15	18	47	.03	.32	.96	-.3	.87	-.6	k .43	.37	70.2	68.7	B15
20	8	47	1.29	.41	.92	-.3	.70	-.6	j .40	.29	80.9	83.2	B20
22	7	47	1.47	.43	.91	-.2	.67	-.6	i .39	.27	83.0	85.1	B22
1	35	47	-1.80	.36	.90	-.5	.78	-.7	h .49	.39	78.7	77.5	B1
16	17	47	.14	.33	.89	-.9	.80	-.9	g .48	.37	74.5	69.7	B16
11	20	47	-.18	.32	.86	-1.3	.78	-1.2	f .53	.38	66.0	67.0	B11
13	26	47	-.78	.32	.83	-1.6	.77	-1.4	e .56	.40	74.5	66.7	B13
10	20	47	-.18	.32	.81	-1.7	.76	-1.4	d .56	.38	83.0	67.0	B10
18	13	47	.59	.35	.73	-1.8	.60	-1.5	c .61	.34	83.0	74.1	B18
12	22	47	-.38	.32	.72	-2.9	.65	-2.3	b .66	.39	78.7	65.9	B12
21	19	47	-.07	.32	.69	-3.0	.62	-2.2	a .67	.38	80.9	67.8	B21
MEAN	19.0	47.0	.00	.36	1.01	-.1	1.03	-.2			70.3	72.8	
S.D.	8.4	.2	1.05	.07	.23	1.7	.51	1.6			13.8	7.9	

Gambar 3.6 Hasil Daya Pembeda *Output Item fit order***Tabel 3. 19 Hasil Interpretasi Daya Pembeda Item KBK**

Nomor Butir Soal	Pt Mean Corr	Interpretasi
1	0,49	Sangat baik
2	0,28	Sangat baik
3	0,28	Sangat baik
4	0,35	Baik
5	0,10	Buruk
6	0,38	Baik
10	0,56	Sangat Baik
11	0,53	Sangat Baik
12	0,66	Sangat Baik
13	0,61	Sangat Baik
14	0,33	Baik
15	0,43	Sangat Baik
16	0,48	Sangat Baik

Nomor Butir Soal	Pt Mean Corr	Interpretasi
17	0,39	Baik
18	0,66	Sangat Baik
19	0,34	Baik
20	0,53	Sangat Baik
21	0,67	Sangat Baik
22	0,39	Baik
23	0,41	Sangat Baik

Berdasarkan hasil interpretasi daya pembeda *item* KBK, sebagian besar butir soal pada item KBK memiliki daya pembeda yang sangat baik, artinya soal-soal tersebut efektif untuk membedakan kemampuan siswa. Beberapa soal seperti nomor 1, 6, 10, 12, 13 dan lainnya memiliki korelasi rata-rata titik yang tinggi, menunjukkan kualitas dalam mengukur keterampilan berpikir kritis. Namun, soal nomor 5 memiliki daya pembeda yang buruk dan perlu direvisi. Beberapa soal lain, seperti nomor 14, 17, 19, dan 22, cukup baik, tetapi masih bisa ditingkatkan. Dengan memperbaiki soal yang kurang efektif dan menjaga kualitas yang ada, sehingga evaluasi keterampilan berpikir siswa dapat menjadi lebih akurat dan efektif.

3.6.2 Analisis Keterlaksanaan Pembelajaran

Analisis berdasarkan lembar observasi dari keterlaksanaan pembelajaran yang digunakan untuk mengukur sejauh mana proses pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) berlangsung, Skor keterlaksanaan adalah 1 jika berhasil terlaksana dan 0 jika tidak terlaksana, Persentase keterlaksanaan dihitung menggunakan persamaan 3.1 sebagai berikut.

$$P = \frac{f}{N} \times 100\% \quad (3,1)$$

Keterangan :

P = persentase ketuntasan belajar siswa

f = jumlah skor yan diperoleh

N = Skor Maksimum

Tabel 3. 20 Interpretasi 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)

3.6.3 Analisis Pengaruh Keterampilan berpikir kritis

Tujuan penelitian adalah untuk melihat pengaruh model pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) terhadap peningkatan keterampilan berpikir kritis siswa pada materi fluida dinamis, Teknik analisis data yang digunakan adalah *stacking-racking*, uji hipotesis dan *effect size*.

3.6.3.1 Analisis Peningkatan Keterampilan Berpikir Kritis

Peningkatan keterampilan berpikir kritis diperoleh dari data kuantitatif nilai pretest dan posttest peserta didik menggunakan pemodelan Rasch dengan teknik *stacking* dan *racking*. Pemodelan Rasch dipilih karena mampu memberikan analisis individu (*individual-centered statistics*). Data mentah dari *pretest* dan *posttest* diolah untuk menghasilkan skala pengukuran dengan interval yang sama. Dalam pemodelan Rasch, data ordinal diubah menjadi data rasio untuk keperluan analisis statistik, menunjukkan hubungan antara kemampuan peserta didik (*person ability*) dan kesulitan butir soal (*item difficult*). Fungsi logaritma digunakan untuk mendapatkan pengukuran dengan interval yang sama, dan hasilnya menciptakan satuan baru yang disebut logit (*log odds unit*). Englehard menyatakan bahwa keberhasilan peserta didik dalam menjawab soal sangat bergantung pada kemampuan dan tingkat kesulitan soal (Sumintono, 2021). Menurut Sumintono dan Widhiarso (2015), probabilitas *odd ratio* individu dapat ditentukan dengan persamaan:

$$Odd\ ratio = \frac{P}{N - P} \quad (3.3)$$

Keterangan:

P : jumlah soal yang dikerjakan peserta didik

N : total soal yang diujikan

Probabilitas *odd ratio* kemudian diubah menjadi bentuk logaritma untuk mendapatkan skala dengan interval yang seragam. Menurut Sumintono dan Widhiarso (2015), rumus *logarithm odds unit* (logit) adalah sebagai berikut.

$$\text{Logit} = \log \frac{P}{N - P} \quad (3.4)$$

Dalam penelitian ini, peningkatan kemampuan keterampilan berpikir kritis siswa diinterpretasikan dengan menggunakan nilai logit (*log odds unit*) dan perubahan posisi pada *vertical ruler*. Berikut adalah penjelasan mengenai teknik analisis *stacking* dan *racking*.

3.6.3.2 Teknik Analisis *Stacking*

Analisis *stacking* adalah teknik analisis longitudinal yang digunakan untuk membandingkan kemampuan siswa sebelum dan sesudah intervensi (Sukarelawan dkk., 2024). Data *pretest* dan *posttest* dikumpulkan bersamaan, lalu diubah menjadi skala interval yang sama (logit). Perubahan kemampuan siswa diukur berdasarkan perubahan posisi pada *vertical ruler* dan nilai logit pada software Winstep versi 3.73. Teknik ini memungkinkan peneliti melihat perubahan kemampuan siswa dari waktu ke waktu dengan lebih mendalam, terutama setelah intervensi (Sukarelawan dkk., 2024). Teknik ini digunakan karena dapat mengidentifikasi perubahan individu, mengukur dampak intervensi, memberikan pemahaman yang lebih mendalam, mengenali respons terhadap intervensi, serta memberikan pengukuran yang tepat dan akurat (Sukarelawan dkk., 2024).

3.6.3.3 Teknik Analisis *Racking*

Analisis *racking* adalah teknik yang digunakan untuk membandingkan tingkat kesulitan butir soal sebelum dan sesudah intervensi (Laliyo, 2021). Mirip dengan teknik *stacking*, data *pretest* dan *posttest* siswa dikumpulkan bersamaan, lalu diubah menjadi skala interval yang sama (logit). Perubahan kesulitan butir soal diukur berdasarkan perubahan posisi pada *vertical ruler* dan nilai logit di software Winstep versi 3.73.0.

3.6.3.4 Uji Prasyarat

3.6.3.4.1 Uji Normalitas

Uji normalitas digunakan untuk menentukan apakah data berdistribusi normal atau tidak. Salah satu cara untuk melakukan ini adalah dengan Uji Kolmogorov

Smirnov dan Shapiro Wilk. Uji Kolmogorov Smirnov digunakan dalam data ini karena sampel berjumlah 55, sedangkan uji Shapiro Wilk lebih cocok untuk sampel kurang dari 50 (Razali dan Wah, 2011). Uji Kolmogorov Smirnov sering digunakan untuk pengujian normalitas, dengan membandingkan distribusi data dengan distribusi normal baku. Uji normalitas ini diolah menggunakan *SPSS for Windows* 26.0. Ketetapan dari uji normalitas ialah nilai signifikansi $\alpha = 0,05$ dengan dasar pengambilan keputusan sebagai berikut:

- Jika nilai *Asymp. Sig (2-tailed)* $> 0,05$ maka data terdistribusi normal
- Jika nilai *Asymp. Sig (2-tailed)* $< 0,05$ maka data tidak terdistribusi normal

3.6.3.4.2 Uji homogenitas

Uji Homogenitas adalah metode untuk mengetahui apakah dua atau lebih sampel memiliki variansi populasi yang sama. Salah satu cara untuk melakukan pengujian homogenitas adalah dengan uji Levene. Uji Levene digunakan untuk menguji variansi dari dua atau lebih kelompok data dengan mencari selisih antara data dan rata-rata kelompoknya. Setelah dilakukan uji normalitas, juga dilakukan uji homogenitas untuk memastikan apakah dua sampel memiliki variasi populasi yang homogen. Pengujian ini diolah menggunakan *SPSS for Windows* 26.0. Dasar pengambilan keputusan menggunakan SPSS versi 27 dengan $\alpha = 5\%$ sebagai berikut:

- Jika nilai *Asymp. Sig.* $> 0,05$ maka data homogen
- Jika nilai *Asymp. Sig.* $< 0,05$ maka data tidak homogen

3.6.3.5 Uji Hipotesis

Uji Hipotesis adalah metode statistik yang digunakan untuk menentukan kebenaran/kepalsuan hipotesis nol, yang menyatakan bahwa tidak ada perbedaan signifikan antara dua rata-rata sampel acak dari populasi yang sama (Sugiyono, 2016), Adapun hipotesis pada penelitian ini yaitu:

Ho : Tidak terdapat perbedaan dalam keterampilan berpikir kritis antara siswa yang menggunakan model pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) dengan siswa yang menggunakan model pembelajaran LC 5E (*Learning Cycle 5E*) pada materi fluida dinamis

Ha : Terdapat perbedaan dalam keterampilan berpikir kritis antara siswa yang menggunakan model pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) dengan siswa yang menggunakan model pembelajaran LC 5E (*Learning Cycle 5E*) pada materi fluida dinamis

Uji hipotesis yang digunakan dalam penelitian ini adalah uji *Mann Whitney*. Uji *Mann Whitney* adalah metode nonparametrik yang digunakan untuk mengidentifikasi perbedaan median antara dua kelompok independen ketika skala data variabel terikatnya adalah ordinal atau interval/rasio tetapi tidak berdistribusi normal (Sotakova dkk., 2020). Dasar pengambilan keputusan menggunakan SPSS versi 26 dengan $\alpha = 5\%$ sebagai berikut:

- Jika nilai *Asymp Sig (2-tailed)* < 0,05 Artinya H_0 ditolak
- Jika nilai *Asymp Sig. (2-tailed)* > 0,05 Artinya H_0 diterima

3.6.3.6 Effect Size

Tujuan penelitian ini adalah untuk melihat pengaruh pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) terhadap keterampilan berpikir kritis pada materi fluia dinamis, Untuk melihat besarnya pengaruh dianalisis menggunakan perhitungan *Effect size*, *Effect Size* merupakan ukuran besarnya kekuatan hubungan antara sebuah variabel bebas dengan variabel terikat (Dunst, dkk, 2004), Yang dimaksud hubungan dalam penelitian ini adalah kuat lemahnya peningkatan keterampilan berpikir kritis, Kuat lemahnya peningkatan keterampilan berpikir kritis tersebut menggambarkan besar kecilnya kontribusi model pembelajaran LCV-Lab (*Learning Cycle 9E* berbantuan *Virtual Laboratory*) terhadap keterampilan berpikir kritis, *Effect size* dihitung menggunakan rumus Cohen (dalam Becker, 2000) sebagai berikut:

$$d = \frac{|M_E - M_K|}{SD_{pool}} \quad (3,3)$$

$$SD_{pool} = \sqrt{\frac{SD_E^2 + SD_K^2}{2}} \quad (3,4)$$

Keterangan :

d = effect size

M_E = Nilai rata – rata kelas eksperimen

M_K = Nilai rata – rata kelas kontrol

SD_{pool} = Standar deviasi untuk kedua kelas partisipan

SD_E = Standar deviasi kelas eksperimen

SD_K = Standar deviasi kelas kontrol

Nilai *effect size d* yang diperoleh kemudian diinterpretasi dengan menggunakan kriteria Cohen (dalam Becker, 2000) di bawah ini

Tabel 3. 21 Interpretasi *Effect Size*

<i>Effect Size</i>	Interpretasi
$0,8 < d < 2,0$	Tinggi
$0,5 \leq d < 0,8$	Sedang
$0,2 \leq d < 0,5$	Rendah

Cohen (dalam Becker, 2000)