

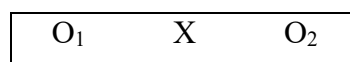
BAB III

METODE PENELITIAN

3.1 Desain penelitian

Metode penelitian merupakan aspek yang sangat penting di dalam sebuah penelitian, karena dengan adanya metode, peneliti bisa menentukan bagaimana suatu masalah dapat dikaji, disimpulkan, serta dianalisis untuk memperoleh hasil yang valid dan memiliki dasar yang kuat untuk dipertanggungjawabkan. Pada penelitian ini menerapkan pendekatan kuantitatif dengan metode eksperimen. Pendekatan penelitian kuantitatif merujuk pada peristiwa objektif yang dikaji secara kuantitatif (Hamdi & Bahrudin, 2015). Artinya data dapat diukur dan dihitung dengan angka serta dianalisis dengan statistik. Sedangkan metode penelitian eksperimen merupakan penelitian yang mengimplementasikan prinsip laboratorium yang menguji pengaruh antara satu variabel dengan variabel lainnya (Hamdi & Bahrudin, 2015). Penelitian eksperimen digunakan untuk mengetahui hubungan sebab-akibat atau pengaruh antara variabel bebas dan variabel terikat (Sugiyono, 2022).

Menurut Tuckman dalam (Sugiyono, 2022), penelitian eksperimen dibagi menjadi empat jenis, salah satunya yaitu *pre-eksperimental design*. Pada penelitian ini menggunakan bentuk *pre-eksperimental design* dengan bentuk *one-group pretest-posttest*. Pada desain ini tidak melibatkan kelompok kontrol, namun terdapat *pretest* dan *posttest* yang digunakan untuk membandingkan nilai atau pengaruh suatu variabel sebelum dan sesudah diberi perlakuan dalam satu kelompok yang sama, dengan begitu, hasil perlakuan terbukti lebih akurat dibandingkan dengan rancangan penelitian *one-shot case study* (Sugiyono, 2022).



Gambar 3. 1: Desain Penelitian *one-group pretest-posttest*

Keterangan:

- O_1 : Hasil tes awal (*pretest*) sebelum penerapan perlakuan (*treatment*)
- O_2 : Hasil tes akhir (*posttest*) setelah penerapan perlakuan (*treatment*)
- X : Perlakuan (*treatment*) model RADEC berbasis portofolio

Pada proses penelitian, sebelum perlakuan model pembelajaran RADEC berbasis portofolio diberikan, peserta didik terlebih dahulu menjalani tes awal (*pretest*) guna mengukur tingkat kemampuan berpikir kritis mereka. Kemudian, setelah data *pretest* diperoleh, peserta didik menerima perlakuan berupa proses pembelajaran menggunakan model pembelajaran RADEC berbasis portofolio pada materi sumber dan perubahan energi sebanyak 4 kali pertemuan. Setelah semua perlakuan diberikan, peserta didik kemudian diminta kembali untuk mengerjakan soal *posttest* untuk mengukur kembali kemampuan berpikir kritisnya setelah perlakuan.

3.2 Tempat dan Waktu Penelitian

Pada penelitian ini dilakukan di SDN Kunci 01, Jl. Raya No.55, Kunci, RT 03 RW 03, Kunci, Kecamatan Sidareja, Kabupaten Cilacap, Provinsi Jawa Tengah. Sekolah tersebut dipilih, karena sekolah tersebut belum mengenal serta belum menerapkan alternatif model pembelajaran RADEC di dalam proses pembelajaran. Selain itu, sekolah tersebut juga sesuai dengan karakteristik model RADEC, di mana peserta didik diharuskan mempunyai kemampuan membaca dasar sebagai fondasi utama dalam proses pembelajaran, khususnya pada tahap *read*. Fokus penelitian pada mata pelajaran Ilmu Pengetahuan Alam dan Sosial (IPAS) Kurikulum Merdeka materi Energi dengan pengimplementasian model pembelajaran *Read-Answer-Discuss-Explain-Create* (RADEC) yang dilakukan di semester genap tahun ajaran 2024/2025. Waktu untuk melaksanakan penelitian ini dilaksanakan pada bulan April 2025.

3.3 Populasi dan Sampel

3.3.1 Populasi

Populasi adalah area umum yang tersusun dari objek atau subjek dengan karakter serta ciri-ciri tertentu yang ditetapkan oleh peneliti untuk dikaji serta disimpulkan (Suriani & Jailani, 2023). Maka dari itu, yang menjadi objek serta perhatian peneliti dari penelitian eksperimen ini ialah seluruh peserta didik kelas VI

SDN Kunci 01 Kabupaten Cilacap tahun ajaran 2024/2025 yang berjumlah 22 orang peserta didik.

3.3.2 Sampel

Sampel merupakan bagian dari populasi yang dipilih untuk merepresentatifkan karakteristik populasi tersebut (Yusuf, 2017: 150). Peneliti menerapkan metode pengambilan sampel jenis *non-probability sampling* dengan teknik sampling jenuh. Teknik ini melibatkan pemilihan seluruh anggota populasi sebagai sampel penelitian, sehingga setiap individu dalam populasi tersebut otomatis menjadi bagian dari sampel (Suriani & Jailani, 2023). Dalam penelitian ini, peserta didik sudah mampu membaca. Oleh karena itu, pada penelitian eksperimen ini sampel ialah peserta didik kelas VI SDN Kunci 01 Kabupaten Cilacap tahun ajaran 2024/2025 berjumlah 22 orang peserta didik yang lancar dalam membaca.

3.4 Teknik Pengumpulan Data

Pengumpulan data merupakan salah satu prosedur krusial dalam pelaksanaan suatu penelitian. Proses ini merujuk pada teknik yang digunakan oleh peneliti untuk memperoleh informasi atau data yang dibutuhkan guna mendukung pelaksanaan dan tujuan penelitian. Dalam penelitian ini menggunakan teknik tes yang meliputi *pretest* dan *posttest* untuk mengumpulkan data tentang kemampuan berpikir kritis siswa sebelum dan sesudah diberi perlakuan. Tes merupakan serangkaian butir soal atau tugas yang harus diselesaikan oleh peserta didik dengan jujur dalam rangka mengevaluasi dimensi tertentu dari peserta didik (Kusumastuti et al., 2020). Dalam penelitian ini, tes dilakukan dengan memberikan soal *pretest* dan *posttest* berbentuk soal pilihan ganda dan lembar penilaian portofolio kemampuan berpikir kritis pada materi energi. *Pretest* bertujuan untuk mengetahui kemampuan awal peserta didik dalam berpikir kritis sebelum diterapkannya model RADEC berbasis portofolio. Sedangkan *posttest*, yaitu tes yang memiliki isi dan bentuk yang sama dengan *pretest*, namun diberikan setelah peserta didik diberi perlakuan untuk mengetahui pengaruh atau efektivitas dari perlakuan yang diberikan serta untuk mengukur kemampuan berpikir kritis peserta didik setelah penerapan perlakuan tertentu.

3.5 Instrumen Penelitian

Pengukuran variabel yang diteliti dalam suatu penelitian memerlukan sebuah alat ukur atau instrumen yang sesuai. Instrumen penelitian merupakan perangkat yang dipergunakan sebagai alat untuk menilai atau menentukan nilai dari variabel-variabel yang diteliti, baik yang berkaitan dengan peristiwa-peristiwa alam atau sosial yang tengah diteliti (Sugiyono, 2022). Dalam penelitian ini yang diukur adalah variabel dependen kemampuan berpikir kritis. Maka dari itu, instrumen penelitian digunakan sebagai alat untuk menghimpun data dan informasi terkait kemampuan berpikir kritis peserta didik. Instrumen penelitian yang digunakan dalam penelitian ini berbentuk instrumen soal pilihan ganda *pretest* dan *posttest*.

3.5.1 Lembar Observasi Keterlaksanaan Pembelajaran

Lembar observasi pelaksanaan pembelajaran dimanfaatkan untuk menilai ketercapaian implementasi model pembelajaran RADEC berbasis portofolio selama proses pembelajaran berlangsung. Instrumen ini berisi lembar observasi untuk aktivitas guru, sedangkan keterlaksanaan aktivitas peserta didik dilihat dari penilaian portofolio. Lembar observasi aktivitas guru mencakup pernyataan-pernyataan yang merepresentasikan seluruh tahapan pembelajaran, mulai dari kegiatan pendahuluan, kegiatan inti, hingga penutup. Penilaian menggunakan kategori “Terlaksana” atau “Tidak Terlaksana”. Setiap tahapan juga dilengkapi dengan kolom keterangan guna observer memberikan catatan, saran, atau komentar tambahan. Observasi dilakukan oleh dua orang, yaitu guru dan seorang observer lainnya, yang menyaksikan langsung jalannya pembelajaran pada setiap sesi.

3.5.2 Kisi-Kisi Instrumen Soal Kemampuan Berpikir Kritis

Instrumen soal disusun dalam bentuk tes tertulis dalam bentuk pilihan ganda. Instrumen dibuat dan dikembangkan berdasarkan pada aspek berpikir kritis menurut Facione meliputi: 1) interpretasi, 2) analisis, dan 3) inferensi. Menggunakan ketiga aspek ini memungkinkan instrumen yang lebih terfokus, efisien, dan mudah digunakan dalam pengukuran kemampuan berpikir kritis tanpa mengurangi kualitas penilaian. Terlalu banyak aspek dapat menyebabkan instrumen menjadi terlalu kompleks dan membingungkan bagi peserta didik. Penyusunan soal

tes kemampuan berpikir kritis dilakukan melalui tahapan penyusunan kisi-kisi soal, perancangan perangkat soal, proses validasi oleh ahli (*expert judgement*), serta pengujian validitas secara empiris.

Berikut ini merupakan tabel 3.1 yang menyajikan kisi-kisi instrumen soal kemampuan berpikir kritis untuk materi IPA topik sumber dan perubahan energi.

Tabel 3. 1 Kisi-Kisi Instrumen Soal Kemampuan Berpikir Kritis

Aspek Berpikir Kritis	Indikator Berpikir Kritis	Indikator Soal	Materi	Nomor Soal
Interpretasi	Memahami sebuah informasi atau permasalahan	Memahami teks dan mengidentifikasi sifat bahan bakar fosil sebagai sumber energi yang tidak dapat diperbarui.	Bentuk-bentuk energi	1
		Menginterpretasikan informasi dalam teks terkait pemanfaatan energi angin.	Energi terbarukan/ energi alternatif	2
	Menafsirkan makna dari tabel, gambar, atau teks yang diberikan	Menafsirkan sumber energi terbarukan yang memiliki potensi terbesar di Indonesia berdasarkan data.	Energi terbarukan/ energi alternatif	3
		Menafsirkan persentase pemanfaatan sumber energi terbarukan di Indonesia berdasarkan data.	Energi terbarukan/ energi alternatif	4

Aspek Berpikir Kritis	Indikator Berpikir Kritis	Indikator Soal	Materi	Nomor Soal
		Mengidentifikasi gambar sumber energi alternatif.	Energi terbarukan/ energi alternatif	5
	Menguraikan informasi yang tersedia	Menguraikan alur perubahan energi dalam pembangkit listrik tenaga uap (PLTU) berdasarkan gambar yang disajikan.	Perubahan bentuk energi	6
		Menguraikan perubahan energi yang terjadi pada berbagai peralatan listrik dalam kehidupan sehari-hari dan menentukan jenis energi yang dihasilkan.	Perubahan bentuk energi	7
Analisis	Mengidenti- fikasi hubungan antar konsep	Mengidentifikasi perilaku yang mendukung penghematan energi listrik dalam kehidupan sehari-hari.	Hemat energi	8
		Mengidentifikasi proses perubahan energi dalam pembangkit listrik tenaga air (PLTA)	Energi terbarukan/ energi alternatif	9

Aspek Berpikir Kritis	Indikator Berpikir Kritis	Indikator Soal	Materi	Nomor Soal
	Memahami apa yang perlu dilakukan untuk menyelesaikan soal	Mengidentifikasi karakteristik bahan bakar fosil berdasarkan pernyataan yang diberikan.	Bentuk-bentuk energi	10
		Mengidentifikasi penggunaan energi cahaya dalam berbagai aktivitas sehari-hari.	Bentuk-bentuk energi	11
		Mengidentifikasi perubahan energi listrik menjadi energi panas/termal dalam berbagai peralatan listrik melalui gambar atau ilustrasi.	Perubahan bentuk energi	12
Inferensi	Menarik kesimpulan berdasarkan informasi yang ada	Menyimpulkan manfaat energi listrik dalam kehidupan sehari-hari berdasarkan informasi dari teks.	Manfaat energi listrik	13
		Menyimpulkan hubungan antara sinar matahari dan proses penguapan dalam kehidupan sehari-hari.	Energi terbarukan/energi alternatif	14

Aspek Berpikir Kritis	Indikator Berpikir Kritis	Indikator Soal	Materi	Nomor Soal
		Menyimpulkan perubahan bentuk energi yang terjadi pada suatu alat berdasarkan informasi yang diberikan.	Perubahan bentuk energi	15
	Memberikan dugaan alternatif	Menduga manfaat energi terbarukan dalam mengurangi dampak negatif bahan bakar fosil.	Energi terbarukan/ energi alternatif	16
		Menduga sumber energi terbarukan yang paling sesuai berdasarkan kondisi alam suatu wilayah.	Energi terbarukan/ energi alternatif	17

Tabel 3.1 menyajikan kisi-kisi dan pemetaan soal berdasarkan aspek dan indikator berpikir kritis dalam mata pelajaran IPA untuk topik sumber dan perubahan energi. Terdapat tiga aspek utama berpikir kritis yang digunakan, yaitu interpretasi, analisis, dan inferensi yang masing-masing diturunkan menjadi indikator berpikir kritis serta nomor soal yang relevan dengan indikator soal dan materi. Pada aspek interpretasi, yang termuat dalam soal nomor 1 sampai 7, indikator yang digunakan meliputi kemampuan memahami informasi atau permasalahan, menafsirkan makna dari teks atau data, serta menguraikan informasi yang tersedia. Contohnya, soal nomor 1 meminta peserta didik memahami teks mengenai bahan bakar fosil, sementara soal nomor 3 dan 4 menuntut siswa menafsirkan data dari grafik atau tabel tentang energi terbarukan. Selain itu, soal nomor 6 dan 7 menilai kemampuan

siswa dalam menguraikan proses perubahan energi melalui gambar atau situasi sehari-hari.

Aspek analisis yang termuat dalam soal nomor 8 sampai 12, terdiri dari indikator mengidentifikasi hubungan antar konsep dan memahami langkah penyelesaian soal. Sebagai contoh, soal nomor 8 meminta siswa mengenali tindakan hemat energi, sedangkan soal nomor 10 dan 11 menilai sejauh mana siswa memahami penggunaan energi dalam kehidupan sehari-hari melalui pengenalan karakteristik bahan atau aktivitas.

Sementara aspek inferensi termuat dalam soal nomor 13 sampai 17, mencakup kemampuan menyimpulkan informasi dan memberikan dugaan alternatif. Misalnya, soal nomor 13 dan 14 meminta siswa menyimpulkan informasi dari teks terkait energi listrik dan proses penguapan. Adapun soal nomor 16 dan 17 menilai kemampuan peserta didik dalam menduga manfaat atau pemilihan sumber energi yang sesuai dengan kondisi lingkungan.

Secara keseluruhan, distribusi soal mencerminkan integrasi antara kemampuan berpikir kritis dan pemahaman konsep IPA, dengan cakupan materi seperti bentuk energi, energi terbarukan, perubahan bentuk energi, manfaat energi listrik, dan hemat energi. Hal ini menunjukkan bahwa soal-soal telah dirancang untuk tidak hanya mengukur pengetahuan faktual saja, tetapi juga mendorong peserta didik berpikir kritis melalui interpretasi, analisis, dan inferensi terhadap berbagai bentuk informasi ilmiah.

3.5.3 Rubrik Penilaian Portofolio

Rubrik penilaian portofolio digunakan untuk menilai tingkat perkembangan kemampuan berpikir kritis peserta didik dari waktu ke waktu melalui aktifitas portofolio yang dibuat berdasarkan pada tahapan pembelajaran RADEC serta aspek kemampuan berpikir kritis. Rubrik penilaian portofolio terdiri dari rubrik penilaian membuat resume, diskusi dan presentasi, serta hasil karya. Berikut ini rubrik penilaian membuat resume disajikan pada tabel 3.2.

Tabel 3. 2 Rubrik Penilaian Membuat Resume

No.	Aspek dan Indikator yang dinilai	Skor			
		3 (Sangat Baik)	2 (Baik)	1 (Cukup)	0 (Kurang)
1.	Isi dan pengetahuan: 1) Akurat dan lengkap meringkas poin-poin penting. 2) Memuat minimal 1 contoh yang relevan. 3) Memberi informasi yang berguna.	Memenuhi semua kriteria yang diharapkan.	Memenuhi 2 kriteria yang diharapkan.	Memenuhi 1 kriteria yang diharapkan.	Seluruh kriteria tidak terpenuhi.
2.	Penggunaan bahasa 1) Tidak ada kesalahan ejaan. 2) Menggunakan kosakata yang paling sesuai untuk menyampaikan makna yang diinginkan. 3) Tidak ada kesalahan tanda baca yang mengganggu pemahaman.	Memenuhi semua kriteria yang diharapkan.	Memenuhi 2 kriteria yang diharapkan.	Memenuhi 1 kriteria yang diharapkan.	Seluruh kriteria tidak terpenuhi.

Tabel 3.2 merupakan rubrik penilaian yang digunakan untuk mengevaluasi hasil kerja peserta didik dalam membuat resume berdasarkan dua aspek utama, yaitu isi dan pengetahuan serta penggunaan. masing-masing aspek memiliki beberapa

indikator yang menjadi tolak ukur dalam menentukan kualitas hasil resume peserta didik dan diberikan skor dengan rentang nilai 0 hingga 3. Dengan adanya tabel ini, proses penilaian menjadi lebih objektif dan terukur karena memberikan acuan yang jelas terhadap sejauh mana peserta didik menguasai materi dan menyampaikannya dengan baik secara bahasa. Rubrik ini juga dapat membantu guru memberikan umpan balik yang spesifik kepada peserta didik untuk meningkatkan kualitas pembelajaran mereka.

Selanjutnya terdapat tabel 3.3 yang menyajikan rubrik penilaian diskusi dan presentasi seperti di bawah ini.

Tabel 3. 3 Rubrik Penilaian Diskusi dan Presentasi

No.	Aspek dan Indikator yang dinilai	Skor			
		3 (Sangat Baik)	2 (Baik)	1 (Cukup)	0 (Kurang)
1.	Partisipasi Aktif	Berpartisipasi memberi ide untuk menjawab 5 soal.	Berpartisipasi memberi ide untuk menjawab 3-4 soal.	Berpartisipasi memberi ide untuk menjawab 1-2 soal.	Tidak memberi ide.
2.	Kedalaman Argumen 1) Mampu menyampaikan argumen dengan jelas, logis, dan terstruktur. 2) Mampu mengevaluasi berbagai sudut pandang terkait topik yang dibahas. 3) Mampu merespon	Memenuhi semua kriteria yang diharapkan.	Memenuhi 2 kriteria yang diharapkan.	Memenuhi 1 kriteria yang diharapkan.	Seluruh kriteria tidak terpenuhi.

	pertanyaan atau sanggahan.				
3.	Kejelasan Komunikasi 1) Penyampaian terstruktur 2) Menggunakan bahasa yang mudah dipahami 3) Melakukan kontak mata dan menggunakan bahasa tubuh yang mendukung komunikasi	Memenuhi semua kriteria yang diharapkan.	Memenuhi 2 kriteria yang diharapkan.	Memenuhi 1 kriteria yang diharapkan.	Seluruh kriteria tidak terpenuhi.

Tabel 3.3 menyajikan rubrik penilaian yang digunakan untuk menilai keterampilan komunikasi dan kontribusi peserta didik dalam diskusi kelompok atau presentasi, yang terdiri dari tiga aspek utama, yaitu partisipasi aktif, kedalaman argumen, dan kejelasan komunikasi. Setiap aspek memiliki indikator tertentu dan diberikan skor dalam skala 0 sampai 3, dengan kategori 3 (Sangat Baik), 2 (Baik), 1 (Cukup), dan 0 (Kurang). Aspek pertama adalah partisipasi aktif, yang menilai sejauh mana peserta didik terlibat dalam memberikan ide atau pendapat saat menjawab soal. Aspek kedua adalah kedalaman argumen, yang menilai kualitas pemikiran peserta didik dalam menyampaikan pendapat. Aspek ketiga adalah kejelasan komunikasi, yang melibatkan penyampaian informasi secara terstruktur, penggunaan bahasa yang mudah dipahami serta penggunaan kontak mata dan bahasa tubuh yang mendukung.

Kemudian rubrik yang terakhir adalah rubrik penilaian hasil karya peserta didik, seperti pada tabel 3.4 berikut ini.

Tabel 3. 4 Rubrik Penilaian Hasil Karya

No.	Aspek dan Indikator yang dinilai	Skor			
		3 (Sangat Baik)	2 (Baik)	1 (Cukup)	0 (Kurang)
1.	Bentuk fisik	Produk atau karya 100% sesuai dengan rancangan dan dapat digunakan.	Produk atau karya hanya 50-80% sesuai dengan rancangan (langkah pembuatan) tetapi dapat digunakan.	Produk atau karya sesuai dengan rancangan tetapi tidak dapat digunakan.	Produk atau karya tidak sesuai rancangan dan tidak dapat digunakan.
2.	Kreativitas dan estetika: 1) memanfaatkan penggunaan bahan yang ada di sekitar; 2) siswa membuat modifikasi atau pengembangan sendiri di luar arahan; 3) tampilan produk menarik, rapi, dan tersusun dengan baik.	Memenuhi semua kriteria yang diharapkan.	Memenuhi 2 kriteria yang diharapkan.	Memenuhi 1 kriteria yang diharapkan.	Seluruh kriteria tidak terpenuhi.

Tabel 3.4 menyajikan rubrik penilaian terhadap produk atau karya peserta didik yang terdiri dari dua aspek utama, yaitu bentuk fisik serta kreativitas dan estetika. Setiap aspek dinilai berdasarkan indikator tertentu dengan rentang skor 0 (Kurang) sampai 3 (Sangat Baik), yang bergantung pada tingkat keterpenuhan indikator.

Aspek pertama yaitu Bentuk Fisik, menilai kesesuaian antara produk akhir dengan rancangan awal serta fungsionalitasnya. Jika produk yang dihasilkan 100% sesuai dengan rancangan dan dapat digunakan dengan baik, maka diberi skor tertinggi yaitu 3. Apabila produk hanya sesuai 50–80% dengan langkah atau rancangan pembuatan tetapi masih berfungsi, diberikan skor 2. Produk yang sesuai dengan rancangan namun tidak dapat digunakan mendapatkan skor 1, sedangkan produk yang tidak sesuai dengan rancangan dan tidak fungsional memperoleh skor 0. Aspek kedua adalah Kreativitas dan Estetika, yang mencakup tiga indikator, yakni: pemanfaatan bahan yang tersedia di lingkungan sekitar, adanya modifikasi atau pengembangan ide secara mandiri di luar arahan guru, serta tampilan produk yang menarik, rapi, dan terstruktur. Secara keseluruhan, rubrik ini bertujuan untuk menilai kualitas hasil karya siswa tidak hanya dari segi fungsi, tetapi juga dari aspek inovasi, kerapian, dan estetika visual yang mencerminkan kreativitas dalam proses pembelajaran berbasis portofolio.

3.5.4 Validitas Instrumen Soal Tes

Instrumen soal tes kemampuan berpikir kritis akan melalui proses validasi yang terdiri dari dua tahap. Tahap pertama melibatkan penilaian validitas instrumen oleh tim ahli (*expert judgment*) untuk mengevaluasi kelayakan isi dan konstruksi instrumen yang akan digunakan. Hasil evaluasi dari tim ahli kemudian akan menjadi dasar pertimbangan untuk melakukan perbaikan dan memastikan kelayakan instrumen tersebut. Berdasarkan hasil validasi oleh ahli menyatakan bahwa instrumen soal kemampuan berpikir kritis layak atau valid untuk digunakan dengan perbaikan untuk proses pengambilan data. Berikut terdapat beberapa rekomendasi oleh ahli untuk perbaikan instrumen yang disajikan dalam tabel 3.5.

Tabel 3. 5 Hasil Validasi Ahli

Nomor soal yang perlu diperbaiki	Saran atau Komentar
1, 3, 4, 6, 11, 16	1. Perbaiki kisi-kisi soal 2. Gunakan gambar dibanding link <i>barcode</i> yang merujuk pada web 3. Perbaiki pernyataan pada soal

Nomor soal yang perlu diperbaiki	Saran atau Komentar
	4. Konsistensi penggunaan kata

Tabel 3.5 di atas menyajikan hasil validasi atau rekomendasi instrumen soal oleh ahli. Berdasarkan tabel tersebut, terdapat daftar nomor soal yang memerlukan perbaikan berdasarkan hasil analisis dan evaluasi instrumen yang dilakukan. Soal-soal yang memerlukan perbaikan diantaranya soal nomor 1, 3, 4, 6, 11, dan 16. Terdapat pula beberapa saran dan komentar yang diberikan untuk meningkatkan kualitas soal-soal tersebut. Pertama, kisi-kisi soal perlu diperbaiki agar lebih sesuai dengan indikator kemampuan yang diukur dan memudahkan penyusunan butir soal yang tepat sasaran. Kedua, penggunaan gambar disarankan menggantikan *barcode* atau tautan yang mengarah ke laman website, untuk mempermudah akses informasi oleh peserta didik dan menghindari ketergantungan pada perangkat teknologi tambahan seperti pemindai kode *QR*. Ketiga, perlu dilakukan perbaikan pada pernyataan soal agar lebih jelas, tidak ambigu, dan mudah dipahami oleh siswa. Terakhir, perlu dijaga konsistensi penggunaan kata dalam soal, termasuk istilah dan pilihan kata yang sesuai dengan bahasa baku serta konteks materi yang diajarkan.

Tahap kedua yaitu melakukan uji coba instrumen soal yang telah divalidasi oleh ahli kepada peserta didik kelas 6 di luar sampel, uji coba dilaksanakan di SDN Dadaha dan SDN 1 Nagrawangi dengan jumlah responden sebanyak 38 peserta didik. Temuan dari pengujian tersebut kemudian akan divalidasi menggunakan pemodelan rasch dengan bantuan perangkat lunak *WINSTEPS* untuk mengidentifikasi item yang valid (*outliers* atau *misfit*). Pemodelan Rasch juga dapat memungkinkan untuk mengidentifikasi *respons errors* atau tebakan serta memprediksi data yang hilang.

Menurut Boone (2014) dalam (Sumintono & Widhiarso, 2015), berikut tabel 3.6 merupakan tolak ukur yang menjadi acuan untuk memeriksa kesesuaian item atau butir soal yang tidak valid (*outliers* atau *misfit*), apabila soal memenuhi salah satu kriteria, maka butir soal tersebut dikatakan *fit* atau valid.

Tabel 3. 6 Kriteria *Item Fit*

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

Tabel 3.6 memuat tiga kriteria utama yang digunakan untuk menilai kualitas atau kesesuaian butir soal dalam analisis model rasch, yaitu *Outfit Mean Square (MNSQ)*, *Outfit Z-Standard (ZSTD)* dan *Point Measure Correlation (Pt Mean Corr)*. setiap kriteria memiliki rentang nilai tertentu yang menunjukkan bahwa butir soal berada dalam kategori layak atau sesuai. Pertama, *nilai Outfit Mean Square (MNSQ)* dianggap sesuai apabila berada dalam rentang lebih dari 0,5 dan kurang dari 1,5. Rentang ini menunjukkan bahwa butir soal tidak terlalu mudah ditebak maupun terlalu acak, serta mampu mengukur kemampuan siswa secara tepat. Kedua, nilai *Outfit Z-Standard (ZSTD)* yang ideal berada antara -0,2 hingga +2,0, yang menandakan bahwa butir tersebut memiliki tingkat penyimpangan yang wajar terhadap model Rasch dan tidak menunjukkan kejanggalan ekstrem. Ketiga, **nilai Point Measure Correlation (Pt Mean Corr)** dinyatakan valid apabila berada dalam rentang 0,4 hingga 0,85, yang berarti butir soal memiliki korelasi positif yang cukup kuat terhadap skor kemampuan siswa secara keseluruhan.

Hasil uji validitas dengan pendekatan pemodelan rasch berbantuan perangkat lunak *WINSTEPS* ditampilkan pada *Output Table 10. Item Fit Order* (Lampiran 2.5), didapatkan hasil validitas soal seperti pada tabel 3.7 berikut.

Tabel 3. 7 Hasil Analisis Validitas Instrumen

Item	Outfit		Pt Measure Corr.	Pemenuhan kriteria	Keterangan
	MNSQ	ZSTD			
1	1,34	1,30	0,20	MNSQ dan ZSTD	<i>Fit</i>
2	0,57	-0,03	0,28	MNSQ dan ZSTD	<i>Fit</i>
3	0,50	-0,36	0,37	3 kriteria terpenuhi	<i>Fit</i>
4	0,59	-1,86	0,71	3 kriteria terpenuhi	<i>Fit</i>

5	1,49	1,24	0,22	MNSQ dan ZSTD	<i>Fit</i>
6	1,65	1,05	0,17	ZSTD	<i>Fit</i>
7	0,85	-0,59	0,54	3 kriteria terpenuhi	<i>Fit</i>
8	1,11	0,39	0,46	3 kriteria terpenuhi	<i>Fit</i>
9	0,70	-1,15	0,62	3 kriteria terpenuhi	<i>Fit</i>
10	0,78	-0,8	0,58	3 kriteria terpenuhi	<i>Fit</i>
11	1,72	1,23	0,17	ZSTD	<i>Fit</i>
12	0,84	-0,36	0,52	3 kriteria terpenuhi	<i>Fit</i>
13	0,57	-0,03	0,28	MNSQ dan ZSTD	<i>Fit</i>
14	0,81	-0,80	0,61	3 kriteria terpenuhi	<i>Fit</i>
15	1,58	1,41	0,24	ZSTD	<i>Fit</i>
16	1,14	0,61	0,48	3 kriteria terpenuhi	<i>Fit</i>
17	0,88	-0,29	0,49	3 kriteria terpenuhi	<i>Fit</i>

Berdasarkan hasil analisis di atas, data tersebut menunjukkan bahwa dari item soal kemampuan berpikir kritis yang telah diujicobakan kepada peserta didik, diantaranya terdapat 3 soal memenuhi 1 kriteria nilai yang diterima, 4 soal memenuhi 2 kriteria, dan 10 soal memenuhi ketiga kriteria yang ditetapkan. Dengan demikian, seluruh item yang ada dapat dikatakan sesuai atau valid, sehingga instrumen soal tersebut layak atau valid digunakan untuk mengukur kemampuan berpikir kritis peserta didik.

3.5.5 Reliabilitas Instrumen Soal Tes

Selain validitas, instrumen soal tes yang diberikan harus reliabel atau ajek, artinya hasil yang didapatkan dari suatu pengukuran melalui tes harus konsisten (Sumintono & Widhiarso, 2015). Uji reliabilitas merupakan proses untuk mengukur sejauh mana suatu instrumen memberikan hasil yang tepat dan konsisten (Rindiasari et al., 2021). Instrumen yang reliabel merupakan instrumen yang mampu memberikan hasil data yang serupa bila diimplementasikan berulang kali dalam waktu berbeda pada objek pengukuran yang sama (Sugiyono, 2022). Untuk mengukur reliabilitas instrumen yang akan digunakan dalam penelitian ini

menggunakan koefisien nilai *Alpha Cronbach*, *Person Reliability* dan *Item Reliability* dari analisis pemodelan Rasch dengan bantuan perangkat lunak *Winstep*. Berikut ini tabel 3.8 yang menyajikan rentang nilai untuk *Alpha Cronbach*.

Tabel 3. 8 Nilai *Alpha Cronbach*

Nilai	Kategori
< 0,5	Buruk
0,5 – 0,6	Jelek
0,6 – 0,7	Cukup
0,7 – 0,8	Baik
> 0,8	Bagus sekali

Tabel 3.8 menunjukkan klasifikasi nilai *Alpha Cronbach* untuk mengukur reliabilitas terhadap respon peserta didik dan soal secara menyeluruh yang dikategorikan ke dalam lima kategori. Jika suatu nilai berada di bawah 0,5, maka dikategorikan sebagai "Buruk". Nilai dalam rentang 0,5 hingga 0,6 dikategorikan sebagai "Jelek". Selanjutnya, nilai antara 0,6 hingga 0,7 masuk dalam kategori "Cukup". Nilai yang berada dalam rentang 0,7 hingga 0,8 termasuk dalam kategori "Baik". Terakhir, nilai yang lebih dari 0,8 dikategorikan sebagai "Bagus sekali". Untuk kategori nilai *Person Reliability* dan *Item Reliability* dapat dilihat pada tabel 3.9 berikut.

Tabel 3. 9 Nilai *Person Reliability* dan *Item Reliability*

Nilai	Kategori
< 0,67	Lemah
0,67 – 0,80	Cukup
0,8 – 0,90	Bagus
0,91 – 0,94	Bagus sekali
> 0,94	Istimewa

Tabel 3.9 menunjukkan kategori dan rentang nilai *Person Reliability* untuk mengukur reliabilitas peserta didik dan *Item Reliability* untuk mengukur reliabilitas

butir soal yang dikategorikan menjadi lima kategori. Nilai yang berada di bawah 0,67 dikategorikan sebagai "Lemah". Nilai dalam rentang 0,67 hingga 0,80 termasuk dalam kategori "Cukup". Sementara itu, nilai antara 0,80 hingga 0,90 dikategorikan sebagai "Bagus". Nilai dalam rentang 0,91 hingga 0,94 masuk dalam kategori "Bagus sekali". Adapun nilai di atas 0,94 termasuk dalam kategori "Istimewa".

Berdasarkan hasil uji reliabilitas terhadap instrumen yang diujicobakan, nilai *Alpha Cronbach*, *Person Reliability*, dan *Item Reliability* disajikan dalam tabel *Sumary Statistic* (Lampiran 2.6). Berikut tabel 3.10 merupakan hasil analisis uji reliabilitas instrumen soal kemampuan berpikir kritis.

Tabel 3. 10 Hasil Analisis Reliabilitas

Analisis Reliabilitas	Nilai	Kategori
<i>Alpha Cronbach</i>	0,71	Baik
<i>Person Reliability</i>	0,62	Lemah
<i>Item Reliability</i>	0,87	Bagus

Berdasarkan hasil analisis tersebut, menunjukkan bahwa nilai *Alpha Cronbach* yang digunakan untuk mengukur reliabilitas atau interaksi antara person dan item soal secara keseluruhan berada dalam kategori bagus dengan nilai sebesar 0.71. Sementara itu, nilai *Person Reliability* berada pada angka 0.62, nilai tersebut mengindikasikan bahwa konsistensi jawaban dari peserta didik masih tergolong lemah. Sedangkan di sisi lain, untuk *Item Reliability* diperoleh nilai sebesar 0.87 yang menunjukkan bahwa kualitas butir soal dalam instrumen tergolong bagus dari aspek reliabilitasnya. Dengan demikian hasil uji reliabilitas terhadap instrumen soal kemampuan berpikir kritis menunjukkan bahwa instrumen tersebut bersifat reliabel dan layak digunakan sebagai alat untuk melakukan pengumpulan data.

3.5.6 Tingkat Kesukaran Instrumen Soal Kemampuan Berpikir Kritis

Tingkat kesukaran menjadi salah satu penentu yang mendukung baiknya tingkat kualitas soal, di mana soal yang berkualitas adalah soal yang memiliki tantangan yang tepat tidak terlalu mudah maupun sukar bagi peserta didik untuk

menjawabnya (Erfan et al., 2020). Tingkat kesukaran butir soal dapat dikelompokkan berdasarkan pada nilai *measure logit* dan nilai deviasi standar (SD) (Sumintono & Widhiarso, 2015), seperti pada tabel 3.11 berikut.

Tabel 3. 11 Kategori Tingkat Kesukaran Item

Nilai <i>Measure (logit)</i>	Kategori Item
$Measure\ logit > SD\ logit$	Sangat Sukar
$0 \leq Measure\ logit \leq SD\ logit$	Sukar
$-SD\ logit \leq Measure\ logit \leq 0$	Mudah
$Measure\ logit < -SD\ logit$	Sangat Mudah

Berdasarkan tabel 3.11, item yang dikategorikan sebagai "Sangat Sukar" apabila nilai *measure logit*-nya lebih besar dari standar deviasi (SD logit), yang menandakan bahwa item tersebut cenderung hanya bisa dijawab oleh peserta didik dengan kemampuan tinggi. Item berada dalam kategori "Sukar" jika nilai *measure logit*-nya berada pada rentang 0 hingga SD logit, yang berarti soal tersebut menantang namun masih dapat dijawab oleh sebagian besar peserta dengan kemampuan rata-rata ke atas. Selanjutnya, item dengan nilai *measure logit* antara -SD logit hingga 0 tergolong "Mudah", artinya soal tersebut dapat dijawab oleh peserta dengan kemampuan sedang. Sedangkan item yang memiliki nilai *measure logit* lebih kecil dari -SD logit termasuk dalam kategori "Sangat Mudah", menunjukkan bahwa soal tersebut bisa dijawab dengan mudah oleh hampir semua peserta didik, termasuk yang berkemampuan rendah.

Pada pemodelan rasch berbantuan program *WINSTEP*, untuk mengetahui tingkat kesukaran item atau soal ditampilkan dalam *Output Table 13. Item Measure* (Lampiran 2.7), tabel tersebut menyajikan tingkat kesukaran soal dari yang tersulit hingga termudah. Nilai yang perlu diperhatikan adalah nilai pada kolom *JMLE Measure* dan nilai SD. Berikut ini tabel 3.12 yang menyajikan hasil analisis tingkat kesukaran terhadap butir soal.

Tabel 3. 12 Hasil Analisis Tingkat Kesukaran

Item	JMLE Measure	SD	Kategori Item
1	0,54	1,35	Sukar
2	-2,25		Sangat Mudah
3	-1,78		Sangat Mudah
4	0,54		Sukar
5	2,52		Sangat Sukar
6	-1.13		Mudah
7	0,69		Sukar
8	-0.66		Mudah
9	0,39		Sukar
10	0,54		Sukar
11	-0,88		Mudah
12	-0,09		Mudah
13	-2,25		Sangat Mudah
14	0,69		Sukar
15	2,52		Sangat Sukar
16	0,54		Sukar
17	0,08		Mudah

Berdasarkan pada hasil analisis di atas, dapat diperoleh bahwa instrumen soal tersebut memiliki nilai deviasi standar (SD) sebesar 1,35, dengan demikian nilai tersebut dapat dijadikan acuan untuk menganalisis lebih lanjut tingkat kesukaran tiap butir soal. Butir soal yang memiliki nilai *measure logit* $> 1,35$ adalah item S5 dan S15, soal tersebut termasuk dalam kategori sangat sukar. Butir soal dengan nilai *measure logit* antara 0,0 s/d 1,35 adalah item S7, S14, S1, S4, S10, S16, dan S9, soal tersebut termasuk dalam kategori soal sukar. Butir soal dengan nilai *measure logit* antara 0,0 s/d -1,35 berkategori mudah adalah soal S17, S12, S8, S11, dan S6. Sedangkan butir soal dengan nilai *measure logit* $< -1,35$ berkategori sangat mudah adalah soal S3, S2, dan S13.

Rekapitulasi analisis tingkat kesukaran soal dalam bentuk persentase disajikan pada tabel 3.13 berikut.

Tabel 3. 13 Rekapitulasi Analisis Tingkat Kesukaran

Kriteria Item	Nomor Soal	Jumlah Soal (Persentase)
Sangat Sukar	S5 dan S15	2 (11,8%)
Sukar	S7, S14, S1, S4, S10, S16, dan S9	7 (41,2%)
Mudah	S17, S12, S8, S11, dan S6	5 (29,4%)
Sangat Mudah	S3, S2, dan S13	3 (17,6%)

Berdasarkan tabel 3.13, dari total 17 soal yang dianalisis, terdapat 2 soal (11,8%) yang dikategorikan sebagai sangat sukar, ini menunjukkan bahwa kedua soal tersebut memerlukan kemampuan tinggi untuk dapat dijawab dengan benar oleh peserta didik. Selanjutnya, terdapat 7 soal (41,2%) yang tergolong sukar, soal-soal ini memiliki tingkat kesulitan yang cukup tinggi namun masih memungkinkan untuk dijawab oleh peserta dengan kemampuan menengah ke atas. Kategori mudah mencakup 5 soal (29,4%), yang relatif dapat dijawab oleh sebagian besar peserta didik dengan tingkat kesulitan yang lebih rendah. Sementara itu, kategori sangat mudah terdiri atas 3 soal (17,6%), yang cenderung dapat dijawab dengan mudah oleh hampir seluruh peserta didik. Distribusi tingkat kesukaran ini menggambarkan bahwa mayoritas soal berada pada kategori sukar, yang menunjukkan bahwa secara umum, instrumen yang digunakan menuntut kemampuan berpikir dan pemahaman yang cukup tinggi dari peserta didik.

3.5.7 Daya Pembeda Instrumen Soal Kemampuan Berpikir Kritis

Daya pembeda butir soal adalah kemampuan butir soal dalam memisahkan tingkat penguasaan peserta didik, antara peserta didik berkemampuan tinggi dan mampu menjawab soal dengan benar, dengan peserta didik yang berkemampuan rendah atau kurang mampu menjawab soal tersebut (Erfan et al., 2020). Dalam pemodelan rasch menggunakan *software WINSTEP*, daya pembeda suatu item dapat ditentukan oleh nilai *PT-Measure Correlation*, yang merujuk pada keterkaitan antara tingkat kesulitan masing-masing butir soal dengan dengan tingkat kesulitan

keseluruhan tes secara umum. Berikut tabel 3.14 kategori nilai *PT-Measure Correlation* (Smiley, 2015; Utari et al., 2021):

Tabel 3. 14 Kategori Daya Pembeda Berdasarkan *PT-Measure Correlation*

Nilai Pt Mean Corr	Kategori Daya Pembeda
$0,40 < ID$	Sangat Baik
$0,30 < ID \leq 0,40$	Baik
$0,20 < ID \leq 0,30$	Kurang Baik
$ID < 0,20$	Buruk

Berdasarkan hasil analisis daya pembeda pada instrumen soal yang diujicobakan, nilai *PT-Measure Correlation* dan kategori daya pembeda dapat dilihat pada tabel 3.15 berikut.

Tabel 3. 15 Hasil Analisis Daya Pembeda

Item	Pt Measure Corr.	Kategori
1	0,20	Kurang baik
2	0,28	Kurang baik
3	0,37	Baik
4	0,71	Sangat baik
5	0,22	Kurang baik
6	0,17	Buruk
7	0,54	Sangat baik
8	0,46	Sangat baik
9	0,62	Sangat baik
10	0,58	Sangat baik
11	0,17	Buruk
12	0,52	Sangat baik
13	0,28	Kurang baik
14	0,61	Sangat baik

15	0,24	Kurang baik
16	0,48	Sangat baik
17	0,49	Sangat baik

Berdasarkan pada hasil analisis tabel 3.15, diperoleh bahwa terdapat 9 (S4, S7, S8, S9, S10, S12, S14, S16, dan S17) dari 17 butir soal yang daya pembedanya termasuk dalam kategori sangat baik dalam mengidentifikasi perbedaan antara peserta didik dengan kemampuan unggul dan yang membutuhkan dukungan lebih. Selanjutnya 1 butir soal (S3) memiliki daya pembeda pada kategori baik, 5 butir soal (S1, S2, S5, S13, dan S15) memiliki daya pembeda pada kategori kurang baik, dan 2 butir soal (S6 dan S11) dalam kategori buruk. Adapun pengelompokkan responden (*person*) dan item (butir soal) dapat dianalisis dari nilai *separation*. Semakin besar nilai *separation* yang diperoleh, maka semakin baik pula kualitas instrumen dalam mengidentifikasi kelompok responden dan butir akan semakin bagus (Sumintono & Widhiarso, 2015).

Pengelompokkan daya pembeda dapat menggunakan rumus pemisah strata, selanjutnya dapat dikategorikan sebarannya seperti pada tabel 3.16.

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

Tabel 3. 16 Kategori Sebaran

Nilai H	Kategori Sebaran	Keterangan
0	Tidak memiliki daya beda	Sangat buruk
1	Mudah	Buruk
2	Sulit dan Mudah	Cukup
3	Sulit, Sedang, dan Mudah	Baik
4	Sangat sulit, sulit, mudah, dan sangat mudah	Sangat baik

(Hasanah & Aini, 2025).

Hasil analisis reliabilitas pada *Output Table 3.1 Summary Statistics* (Lampiran 2.6), menunjukkan bahwa nilai *separation* diperoleh sebesar 2,54, kemudian nilai tersebut dihitung menggunakan rumus pemisah strata (H) sehingga didapatkan nilai

H sebesar 3,72 apabila dibulatkan menjadi 4. Berdasarkan pada nilai H yang didapatkan dan dikategorikan seperti pada tabel 3.16, menyatakan bahwa butir-butir soal dalam instrumen termasuk dalam kategori sangat baik di mana memiliki 4 macam sebaran daya beda yaitu soal sangat sulit, sulit, mudah, dan sangat mudah.

3.6 Prosedur Penelitian

Pada penelitian ini akan melalui 3 (tiga) tahapan, meliputi tahap persiapan, tahap pelaksanaan, serta tahap akhir. Prosedur tahapan tersebut dijabarkan sebagai berikut:

1) Tahap Persiapan

- a. Mengidentifikasi masalah melalui kajian literatur dengan mengkonfirmasi permasalahan tersebut di lapangan.
- b. Melakukan kajian literatur untuk mengeksplorasi solusi alternatif untuk mengatasi permasalahan yang ditemukan.
- c. Menentukan lokasi dan subjek penelitian
- d. Membuat rancangan penelitian yang berisi mengenai rumusan permasalahan, teori yang digunakan dalam kajian pustaka, serta metode yang akan digunakan.
- e. Merancang rencana penerapan model pembelajaran RADEC berbasis portofolio yang berfokus pada peningkatan kemampuan berpikir kritis serta menyusun perangkat pembelajaran
- f. Menyusun instrumen penelitian
- g. Melakukan validasi instrumen oleh ahli
- h. Melakukan perizinan ke sekolah yang akan dijadikan tempat penelitian
- i. Membuat jadwal pelaksanaan penelitian

2) Tahap Pelaksanaan

- a. Melaksanakan *pretest* pada sampel penelitian dengan memberikan soal yang dapat mengasah kemampuan berpikir kritis mengenai materi energi
- b. Melaksanakan analisis data hasil *pretest* untuk mengetahui kemampuan awal sampel penelitian

- c. Memberikan perlakuan (*treatment*) kepada sampel penelitian yang dilakukan sebanyak 4 (empat) kali pertemuan.
- d. Melakukan *posttest* setelah diberikan perlakuan untuk mendapatkan data terkait kemampuan akhir sampel penelitian.

3) Tahap Akhir

- a. Melakukan analisis data yang didapatkan dari hasil tahap pelaksanaan menggunakan pemodelan rasch dengan metode *racking* dan *stacking analysis*.
- b. Hasil analisis data diidentifikasi dan dijabarkan secara mendalam
- c. Hasil penelitian dijadikan bahan untuk menarik kesimpulan yang ditujukan untuk menjawab pertanyaan penelitian serta membuktikan hipotesis penelitian
- d. Hasil penelitian disusun menjadi laporan penelitian.

3.7 Analisis Data

3.7.1 Analisis Keterlaksanaan Pembelajaran

Keterlaksanaan pembelajaran dinilai melalui pengamatan langsung oleh observer terhadap interaksi atau aktivitas guru dan peserta didik selama proses pembelajaran menggunakan model pembelajaran RADEC berbasis portofolio pada materi sumber dan perubahan energi. Observasi tersebut dilakukan untuk menilai kemampuan pendidik dalam mengelola proses pembelajaran serta aktivitas peserta didik selama pembelajaran berlangsung melalui lembar observasi keterlaksanaan pembelajaran selama empat kali pertemuan (Segening et al., 2022). Setiap aktivitas pembelajaran yang diamati diberi skor 0 (menunjukkan “tidak terlaksana”) atau 1 (menunjukkan “terlaksana”). Observasi dalam penelitian ini dilakukan oleh dua orang observer yang telah diberi penjelasan sebelumnya mengenai model pembelajaran RADEC. Peneliti memberikan pemahaman awal kepada kedua observer terkait konsep dasar, tujuan, serta tahapan-tahapan dalam model pembelajaran RADEC, guna memastikan kesamaan persepsi dalam proses pengamatan, sehingga pelaksanaan model pembelajaran dapat diamati secara objektif dan sesuai dengan indikator yang telah ditentukan. Data pengamatan yang

dikumpulkan kemudian diproses untuk menentukan persentase implementasi pembelajaran menggunakan rumus sebagai berikut (Akbar, 2017; Pritasari, 2024):

$$KP_{per\ pertemuan} = \frac{skor\ diperoleh}{total\ skor\ maksimal} \times 100\%$$

Kemudian, hasil perhitungan tersebut diubah menjadi persentase rata-rata observasi keterlaksanaan pembelajaran secara keseluruhan menggunakan rumus berikut:

$$KP_{rata-rata} = \frac{jumlah\ persentase\ tiap\ pertemuan}{jumlah\ pertemuan}$$

Hasil skor observasi keterlaksanaan pembelajaran yang diperoleh kemudian diinterpretasi atau dikategorikan seperti pada tabel 3.17.

Tabel 3. 17 Kategori Keterlaksanaan Pembelajaran

Skor	Kategori
$81,25\% \leq KP \leq 100,00\%$	Sangat terlaksana
$62,50\% \leq KP < 81,25\%$	Terlaksana
$43,75\% \leq KP < 62,50\%$	Kurang terlaksana
$25,00\% \leq KP < 43,75\%$	Tidak terlaksana

Berdasarkan tabel 3.17, dapat dijelaskan bahwa tingkat keterlaksanaan model RADEC berbasis portofolio dikategorikan ke dalam empat rentang skor persentase. Jika nilai persentase keterlaksanaan (KP) berada pada rentang 81,25% hingga 100%, maka kegiatan tersebut dikategorikan "Sangat Terlaksana", yang menunjukkan bahwa hampir seluruh aspek kegiatan berjalan sesuai dengan rencana dan target. Apabila KP berada pada kisaran 62,50% hingga kurang dari 81,25%, maka kegiatan tersebut dianggap "Terlaksana", yang berarti sebagian besar kegiatan telah berjalan dengan baik meskipun mungkin terdapat beberapa kekurangan kecil. Selanjutnya, jika KP berada pada rentang 43,75% hingga kurang dari 62,50%, maka kegiatan masuk dalam kategori "Kurang Terlaksana", yang menunjukkan bahwa masih banyak aspek kegiatan yang belum berjalan sesuai harapan. Terakhir, jika nilai KP berada di antara 25,00% hingga kurang dari 43,75%, maka kegiatan

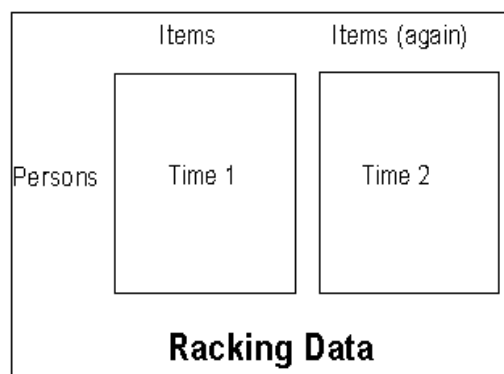
dikategorikan "Tidak Terlaksana", yang mengindikasikan bahwa sebagian besar komponen kegiatan tidak berjalan sebagaimana mestinya dan perlu perbaikan secara menyeluruh.

3.7.2 Analisis Kemampuan Berpikir Kritis

Pada penelitian yang dilakukan ini, analisis mengenai peningkatan kemampuan berpikir kritis dianalisis menggunakan Model Rasch dengan bantuan *software WINSTEP*. Model Rasch memiliki keunggulan dalam memfasilitasi perhitungan perubahan sebelum dan sesudah intervensi dalam mengungkapkan perubahan individu dan item secara lebih rinci (Laliyo, 2021). Dalam penelitian ini menggunakan metode analisis *racking* dan *stacking* yang diperkenalkan oleh Benjamin Wright, metode ini digunakan untuk mengukur perubahan kemampuan peserta didik dan tingkat kesulitan item soal berubah sebelum dan sesudah diberi perlakuan (Laliyo et al., 2022). Analisis *racking* dan *stacking* digunakan untuk melihat perubahan atau peningkatan kemampuan peserta didik berdasarkan nilai logit yang diperoleh peserta didik (Nuraeni & Karlimah, 2023).

3.7.2.1 Racking Analysis

Racking Analysis merupakan metode yang digunakan untuk menganalisis perubahan tingkat kesulitan item soal baik itu soal yang mengalami penurunan atau stagnan sebelum dan setelah diberikan perlakuan. Analisis *racking* digunakan untuk membandingkan tingkat kesulitan butir soal berdasarkan nilai *logit* yang dihasilkan (Hikmah et al., 2022). Pada metode ini, menempatkan data *pretest* dan *posttest* secara sejajar dalam satu baris secara horizontal. Setiap item soal *pretest* dan *posttest* muncul dua kali dalam kumpulan data, sementara peserta didik hanya muncul sekali (Laliyo, 2021). Berikut ini gambar 3.2 yang menunjukkan skema analisis *racking*.

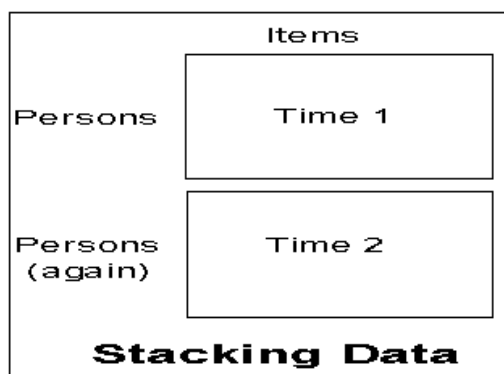


Gambar 3. 2: Skema Analisis *Racking*

Dengan cara ini, data *pretest* dan *posttest* dapat dibandingkan secara langsung, memudahkan analisis perbedaan hasil yang diperoleh sebelum dan setelah perlakuan atau intervensi.

3.7.2.2 Stacking Analysis

Stacking analysis merupakan metode yang digunakan untuk menganalisis perubahan kemampuan pada responden atau peserta didik yang mengalami peningkatan atau stagnan sebelum dan setelah diberikan perlakuan. Data dianalisis dengan membandingkan nilai *logit pretest* dan *posttest* (Hanatan et al., 2023). Pada metode ini, menempatkan data *pretest* dan *posttest* secara sejajar dalam satu kolom secara vertikal. Setiap peserta didik muncul dua kali dalam kumpulan data, pertama untuk hasil *pretest* dan kedua untuk hasil *posttest*, sementara item soal muncul sekali (Laliyo, 2021). Berikut ini gambar 3.3 yang menunjukkan skema analisis *stacking*.



Gambar 3. 3: Skema Analisis *Stacking*

Struktur ini memudahkan perbandingan langsung antara hasil *pretest* dan *posttest* untuk setiap peserta didik, serta memungkinkan analisis yang lebih mendalam terkait perubahan kemampuan mereka setelah mengikuti proses pembelajaran atau intervensi tertentu. Teknik ini memberikan wawasan mengenai bagaimana kemampuan setiap peserta didik berkembang dari waktu ke waktu, sehingga memungkinkan untuk melihat perubahan kemampuan pada tiap individu (Sukarelawan et al., 2024). Pengkategorian kemampuan berpikir kritis peserta didik dapat dikategorikan berdasarkan nilai mean dan standar deviasi (SD) (Sumintono & Widhiarso, 2015). Berikut disajikan kategori kemampuan berpikir kritis pada tabel 3.18.

Tabel 3. 18 Kategori Kemampuan

Nilai Measure (<i>logit</i>)	Kategori Item
$> (Mean + 2 \text{ SD } logit)$	<i>Outliers Atas</i>
$(Mean + 1 \text{ SD } logit) < logit \leq (Mean + 2 \text{ SD } logit)$	Sangat Tinggi
$Mean \leq logit \leq (Mean + 1 \text{ SD } logit)$	Tinggi
$(Mean - 1 \text{ SD } logit) \leq logit \leq Mean$	Rendah
$(Mean - 2 \text{ SD } logit) < logit \leq (Mean - 1 \text{ SD } logit)$	Sangat Rendah
$< (Mean - 2 \text{ SD } logit)$	<i>Outliers Bawah</i>

3.7.3 Analisis Efektivitas Model Pembelajaran RADEC Berbasis Portofolio terhadap Kemampuan Berpikir Kritis

Data yang diperoleh dari hasil penelitian model pembelajaran RADEC berbasis portofolio terhadap kemampuan berpikir kritis peserta didik kemudian dianalisis tingkat efektivitasnya sebagai penarikan kesimpulan akhir penelitian. Efektivitas dianalisis dengan membandingkan hasil *pretest* dan *posttest* peserta didik sebelum dan sesudah peserta didik diberi perlakuan. Efektivitas suatu perlakuan dapat dilihat dari perbedaan hasil belajar yang signifikan secara statistik (Sugiyono, 2022). Sebelum melaksanakan uji efektivitas, terlebih dahulu dilakukan pengujian terhadap asumsi distribusi data untuk menentukan metode uji statistik yang sesuai. Pada penelitian ini ingin melihat apakah terdapat perbedaan signifikan

atau tidak dari hasil *pretest* dan *posttest*, maka uji statistik yang digunakan yaitu statistik inferensial (Sugiyono, 2022). Pada penelitian ini, proses analisis data *pretest* dan *posttest* dilakukan dengan bantuan *software IBM SPSS Statistics*. Terdapat beberapa tahap analisis untuk mengetahui efektivitas dari model pembelajaran RADEC berbasis portofolio terhadap kemampuan berpikir kritis peserta didik, diantaranya sebagai berikut.

3.7.3.1 Uji Normalitas

Uji normalitas merupakan salah satu uji prasyarat yang digunakan untuk menentukan apakah data layak dianalisis dengan menggunakan statistik parametrik atau nonparametrik (Hasan, 2022). Apabila data berdistribusi normal, maka analisis data menggunakan statistik parametrik (*paired sample t-test*). Sebaliknya apabila data tidak berdistribusi normal, maka data dianalisis secara nonparametrik (uji *wilcoxon*). Data dikatakan tidak normal apabila salah satu output analisis data baik *pretest* maupun *posttest* tidak berdistribusi normal. Uji normalitas dilakukan dengan membandingkan nilai signifikansi (*P-value*) hasil uji dengan tingkat signifikansi yang ditetapkan ($\alpha = 0,05$). Jika $\text{Sig.} > 0,05$, maka data dikatakan berasal dari populasi yang berdistribusi normal. Sebaliknya, jika $\text{Sig.} < 0,05$, maka data dikatakan tidak berdistribusi normal. Apabila data yang dianalisis kurang dari 50 (<50) menggunakan teknik *Shapiro-Wilk*, sedangkan bila data lebih dari 50 (>50) menggunakan teknik *Kolmogorov-Smirnov* (Theodorus et al., 2022).

3.7.3.2 Uji Hipotesis

Uji hipotesis merupakan suatu prosedur yang digunakan untuk menghasilkan suatu keputusan, terkait menerima atau menolak hipotesis yang diajukan, di mana dalam pengujian ini sifatnya tidak pasti, sehingga keputusan yang diambil berpotensi benar maupun salah (Hasan, 2022). Setelah mengetahui normalitas dari data, maka selanjutnya dapat melanjutkan analisis data atau pengujian hipotesis. Apabila data berdistribusi normal maka data dianalisis menggunakan uji-t (*paired sample t-test*), sebaliknya apabila data tidak berdistribusi normal maka menggunakan uji *Wilcoxon* untuk melihat apakah ada perubahan kemampuan berpikir kritis yang signifikan setelah peserta didik kelas VI mengikuti proses

pembelajaran menggunakan model pembelajaran RADEC berbasis portofolio. Hasil pengolahan data dalam uji hipotesis diperoleh dengan cara membandingkan nilai signifikansi yang didapatkan dengan taraf signifikansi ($\alpha = 0,05$), kriteria pengujian hipotesis berdasarkan nilai signifikansi sebagai berikut:

1. jika nilai Sig. $< 0,05$, maka H_0 ditolak atau H_a diterima, artinya terdapat perbedaan signifikan.
2. jika nilai Sig. $> 0,05$, maka H_0 diterima atau H_a ditolak, artinya tidak terdapat perbedaan signifikan.