

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Pendekatan penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen. Metode eksperimen dipilih karena memungkinkan peneliti untuk mengkaji pengaruh suatu variabel independen terhadap variabel dependen dalam situasi yang terkendali dan terstruktur (Sugiyono, 2019). Pemilihan metode ini dilatarbelakangi oleh kesesuaian antara pendekatan tersebut dengan fokus permasalahan yang dikaji, yakni untuk mengidentifikasi efektivitas penerapan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga terhadap literasi sains siswa sekolah dasar.

Desain eksperimen yang digunakan dalam penelitian ini adalah *pre-eksperimental* dengan jenis *one-group pretest-posttest design*. Pada desain ini, tidak terdapat kelompok kontrol sebagai pembanding, melainkan hanya satu kelompok yang diberikan tes awal (*pretest*), perlakuan (*treatment*), dan kemudian tes akhir (*posttest*) untuk mengukur perubahan setelah perlakuan diberikan. Keberadaan tes awal (*pretest*) dan tes akhir (*posttest*) memungkinkan peneliti memperoleh hasil yang lebih akurat dari perlakuan yang diberikan, karena memungkinkan perbandingan antara kondisi sebelum dan setelah perlakuan diberikan (Sugiyono, 2019), sehingga desain ini dianggap lebih informatif dibandingkan dengan *one-shot case study* yang hanya melakukan pengukuran setelah perlakuan tanpa adanya pembanding kondisi awal.

Sebelum model pembelajaran RADEC berbasis etnosains budaya Kampung Naga diterapkan, siswa terlebih dahulu mengikuti tes awal (*pretest*) sebagai bagian dari proses pelaksanaan. Melalui tes awal, peneliti dapat memetakan kemampuan dasar siswa sebagai acuan sebelum intervensi pembelajaran diberikan. Setelah data *pretest* diperoleh, siswa kemudian diberikan perlakuan berupa penggunaan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga selama empat kali pertemuan. Setelah seluruh rangkaian perlakuan selesai dilaksanakan, siswa diminta untuk mengerjakan *posttest* atau tes akhir.

Pemilihan desain ini didasarkan pada kesesuaiannya dengan fokus penelitian, yaitu untuk menganalisis perubahan yang terjadi sebelum dan sesudah penerapan model RADEC berbasis etnosains budaya Kampung Naga terhadap literasi sains. Rancangan desain penelitian secara lengkap disajikan sebagai berikut.

$O_1 \times O_2$

Gambar 3. 1 *Design One-Group Pretest-Posttest*

Sumber: Sugiyono (2019)

Keterangan:

O_1 : Nilai *pretest* (sebelum menggunakan model RADEC berbasis etnosains)

X : Perlakuan (*Treatment*) dengan menggunakan model RADEC berbasis etnosains

O_2 : Nilai *posttest* (setelah menggunakan model RADEC berbasis etnosains)

3.2 Tempat, dan Waktu Penelitian

Lokasi penelitian dilakukan di SD Al Falah yang terletak di Kecamatan Tawang, Kota Tasikmalaya, dengan waktu pelaksanaan penelitian yang dijadwalkan pada semester kedua tahun ajaran 2024/2025. Sekolah yang menjadi lokasi penelitian, sebelumnya telah menjadi bagian dari program Kampus Mengajar, yang bertujuan dalam meningkatkan kemampuan literasi dan kemampuan numerasi siswa. Keikutsertaan ini menunjukkan bahwa sekolah memiliki komitmen terhadap perbaikan mutu pembelajaran dan terbuka terhadap berbagai inovasi pendidikan. Meskipun telah mendapatkan intervensi peningkatan literasi dan numerasi secara umum, pembelajaran IPA di sekolah ini belum secara khusus dikembangkan melalui pendekatan kontekstual yang berbasis budaya lokal. Oleh karena itu, penerapan model pembelajaran RADEC berbasis etnosains dalam penelitian ini diharapkan dapat menjadi inovasi lanjutan yang tidak hanya mendukung literasi sains siswa, tetapi juga memperkuat keterkaitan antara pembelajaran dan kearifan lokal yang relevan dengan kehidupan siswa.

Pelaksanaan pembelajaran menggunakan model pembelajaran RADEC berbasis etnosains adalah sebanyak empat kali pertemuan pada saat jam pelajaran dengan jumlah alokasi waktu yang sama setiap pertemuannya yaitu 2 JP. Waktu

penelitian yang telah ditentukan memfokuskan pada penerapan model tersebut selama beberapa pertemuan, dengan pembelajaran yang terintegrasi dengan konten etnosains lokal, untuk menilai efektivitas dan dampaknya terhadap kemampuan literasi sains siswa.

3.3 Populasi dan Sampel

3.3.1 Populasi

Populasi merupakan keseluruhan kelompok subjek yang memiliki karakteristik tertentu dan berada dalam rentang yang telah ditetapkan oleh peneliti, yang dijadikan sumber untuk menarik kesimpulan penelitian (Sugiyono, 2019). Dalam konteks penelitian ini, populasi yang dimaksud adalah seluruh siswa kelas V di SD Al Falah, Desa Cikalang, Kecamatan Tawang, Kota Tasikmalaya. Jumlah populasi yang terdaftar sebanyak 25 orang.

3.3.2 Sampel

Sampel merupakan bagian dari populasi yang secara sengaja dipilih untuk mewakili karakteristik yang ada dalam populasi secara keseluruhan (Cooper & Schindler, 2014). Artinya, sampel mencerminkan sejumlah individu yang memiliki ciri-ciri tertentu yang serupa dengan populasi. Dalam penelitian ini, teknik pengambilan sampel yang digunakan adalah sampling jenuh. Menurut Sugiyono (2018) sampling jenuh adalah teknik pengambilan sampel di mana keseluruhan anggota populasi digunakan sebagai sampel. Teknik ini umumnya digunakan ketika jumlah populasi kecil, yakni di bawah 30 orang. Berdasarkan pertimbangan tersebut, seluruh siswa kelas V dengan dipilih sebagai sampel dalam penelitian ini.

Pemilihan fase C dengan jenjang kelas V sebagai sampel penelitian didasarkan pada pertimbangan bahwa fase ini dianggap sebagai tahap penting dalam perkembangan kognitif siswa. Pada fase ini siswa mulai mampu memahami konsep-konsep abstrak dan menerapkannya dalam situasi yang lebih kompleks. Siswa sebelumnya telah mendapatkan pembiasaan membaca, berhitung, dan berpikir logis melalui kegiatan dari program Kampus Mengajar. Hal tersebut menjadikan siswa memiliki dasar yang cukup untuk diarahkan pada literasi tingkat lanjut, termasuk literasi sains. Pemilihan ini juga didasarkan pada relevansi materi penelitian dengan kurikulum yang diajarkan pada jenjang tersebut.

3.4 Teknik Pengumpulan Data

Teknik pengumpulan data diperlukan untuk memperoleh informasi yang relevan dengan tujuan penelitian (Sugiyono, 2019). Penelitian ini menggunakan tes (*pretest* dan *posttest*) sebagai teknik pengumpulan data dan observasi sebagai teknik pelengkap. Tes adalah alat ukur berupa pertanyaan untuk menilai kemampuan kognitif siswa sebelum dan setelah pembelajaran, dengan variasi bentuk seperti pilihan ganda atau esai (Sani, 2022). Tes berfungsi untuk mengevaluasi hasil belajar dalam kurun waktu tertentu (Sukmadinata, 2016: 223). Tes digunakan untuk menilai kemampuan siswa yang dilakukan sebelum dan setelah perlakuan diberikan.

Dalam penelitian ini, instrumen tes yang digunakan berbentuk pilihan ganda dikerjakan secara individu oleh siswa. Terdapat dua jenis tes yang diberikan, yaitu:

1. *Pretest*, yang bertujuan untuk mengukur kemampuan awal literasi sains siswa sebelum penerapan model RADEC berbasis etnosains budaya Kampung Naga dalam pembelajaran.
2. *Posttest*, yang memiliki isi dan bentuk yang sama dengan *pre-test*, namun diberikan setelah siswa mengikuti pembelajaran tersebut. Tes diberikan untuk mengidentifikasi perubahan yang ditimbulkan model pembelajaran yang diterapkan. Penggunaan instrumen tes untuk mengukur variabel dependen utama dalam penelitian ini, yaitu kemampuan literasi sains siswa.

Data yang diperoleh dari hasil tes literasi sains kemudian dianalisis menggunakan Model Rasch. Analisis ini digunakan untuk menjawab pertanyaan penelitian dengan memberikan gambaran yang lebih objektif mengenai tingkat kemampuan literasi sains siswa berdasarkan *logit* skor yang terstandarisasi.

Teknik observasi digunakan sebagai teknik pelengkap untuk memperoleh data mengenai keterlaksanaan model pembelajaran RADEC. Observasi dilakukan oleh observer dengan menggunakan lembar observasi yang telah disusun sesuai dengan sintaks pembelajaran.

3.5 Instrumen Penelitian

Instrumen penelitian didefinisikan sebagai alat bantu ilmiah yang digunakan untuk mengobservasi dan mengidentifikasi variabel-variabel penelitian, sehingga menjamin keteraturan dalam pelaksanaan penelitian (Sugiyono, 2019). Dalam penelitian ini, variabel yang akan diamati yaitu kemampuan literasi sains. Instrumen dalam penelitian ini berupa seperangkat soal tes literasi sains yang diberikan kepada siswa dalam dua tahap pengumpulan data, yaitu pada tahap awal (*pretest*) dan tahap akhir (*posttest*).

3.5.1 Lembar Observasi Keterlaksanaan Pembelajaran

Ketercapaian penerapan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga diukur selama proses pembelajaran menggunakan lembar observasi keterlaksanaan pembelajaran. Instrumen ini memuat dua format utama, yaitu lembar observasi aktivitas siswa dan guru. Instrumen ini disusun sebagai alat bantu evaluasi proses pembelajaran, baik dari sisi guru maupun siswa, berdasarkan sintaks-sintaks dalam model RADEC (*Read, Answer, Discuss, Explain, Create*) yang telah diadaptasi dengan konteks etnosains. Masing-masing lembar mencakup pernyataan-pernyataan yang mewakili seluruh tahapan pembelajaran, mulai dari kegiatan pendahuluan, inti, hingga penutup, dengan indikator penilaian berupa kategori “terlaksana” dan “tidak terlaksana”. Selain itu, disediakan pula kolom keterangan pada setiap tahap pembelajaran untuk memberikan catatan, saran, atau komentar tambahan. Dua orang observer, yakni guru dan rekan sejawat berperan dalam pengisian lembar observasi, yang secara langsung mengamati seluruh rangkaian pembelajaran pada setiap pertemuan.

3.5.2 Soal Tes Kemampuan Literasi Sains

Instrumen soal tes literasi sains yang diterapkan dalam penelitian ini disusun mengacu pada tiga indikator utama kompetensi literasi sains sebagaimana ditetapkan oleh OECD (2019), yang meliputi: 1) Menjelaskan fenomena secara ilmiah, 2) Mengevaluasi dan merancang penyelidikan ilmiah dan 3) Menginterpretasikan data dan bukti ilmiah. Pengembangan instrumen tes dilakukan melalui tahapan menyusun kisi-kisi instrumen soal, perancangan butir soal, validasi

oleh tim ahli (*expert judgement*), dan uji validitas empiris yang dijelaskan secara rinci sebagai berikut.

3.5.2.1 Kisi-Kisi Instrumen Soal Literasi Sains

Instrumen tes literasi sains dirancang menjadi bentuk tes tertulis yaitu pilihan ganda. Terdapat 20 soal dalam instrumen ini, untuk *pretest* dan *posttest*. Kisi-kisi tes literasi sains yang dikembangkan dalam penelitian ini dapat dilihat pada Tabel 3.1 berikut:

Tabel 3. 1 Kisi-Kisi Instrumen Soal Tes Literasi Sains

No	Aspek	Indikator	Sebaran Soal	Jumlah Soal
1.	Menjelaskan fenomena secara ilmiah (<i>Explaining phenomena scientifically</i>)	1. Menghubungkan fenomena ilmiah dengan konsep sains.	1, 2	2
		2. Menentukan hipotesis atau dugaan sementara berdasarkan fenomena ilmiah.	3, 4, 5	3
		3. Menghubungkan fenomena ilmiah dengan dampaknya terhadap kehidupan sehari-hari.	6,7	2
2.	Mengevaluasi dan merancang penyelidikan ilmiah (<i>Evaluating and Designing Enquiry</i>)	1. Menyebutkan langkah-langkah dalam melakukan percobaan ilmiah sederhana.	8,9	2
		2. Mengidentifikasi variabel dalam suatu percobaan ilmiah sederhana.	10,11	2
		3. Menyusun pertanyaan ilmiah yang dapat diuji dengan percobaan sederhana.	12, 13	2
3.	Menginterpretasikan data dan bukti ilmiah	1. Menafsirkan data ilmiah yang disajikan	14, 15, 16	3

No	Aspek	Indikator	Sebaran Soal	Jumlah Soal
	<i>(Interpreting data and Evidence Scientifically)</i>	dalam tabel, grafik, atau gambar.		
		2. Menilai klaim atau argumen tentang fenomena ilmiah didukung oleh data atau bukti yang valid.	17, 18	2
		3. Menarik kesimpulan yang tepat berdasarkan data ilmiah tertentu.	19, 20	2

Tabel 3.1 memuat perincian kisi-kisi instrumen yang disusun untuk mengukur kemampuan literasi sains siswa. Instrumen tes literasi sains disusun berdasarkan tiga aspek utama dalam kerangka PISA 2018, yaitu: (1) menjelaskan fenomena secara ilmiah, (2) mengevaluasi dan merancang penyelidikan ilmiah, serta (3) menginterpretasikan data dan bukti ilmiah.

Aspek pertama terdiri dari tiga indikator dengan total 7 soal, meliputi kemampuan menghubungkan fenomena dengan konsep sains, menyusun hipotesis, dan mengaitkan fenomena dengan kehidupan sehari-hari. Aspek kedua juga mencakup tiga indikator dengan 6 soal, yang mengukur kemampuan siswa dalam merancang dan mengevaluasi eksperimen sederhana. Aspek ketiga terdiri dari tiga indikator dengan 7 soal, mengukur keterampilan menafsirkan data dan menilai argumen berdasarkan bukti ilmiah. Secara keseluruhan, instrumen terdiri dari 20 soal yang tersebar merata sesuai indikator untuk mencakup seluruh aspek literasi sains yang diukur.

3.5.2.2 Validitas Instrumen Soal Literasi Sains

Proses validasi instrumen tes literasi sains dilakukan melalui dua tahapan. Tahap pertama merupakan validasi isi oleh tim ahli (*expert judgement*), yang terdiri dari dosen ahli di bidang evaluasi pendidikan. Dosen ahli bertugas mengevaluasi kelayakan isi dan konstruk instrumen yang disusun. Masukan dan hasil penilaian dari tim ahli digunakan sebagai dasar untuk melakukan revisi instrumen agar sesuai

dengan tujuan pengukuran. Berdasarkan hasil penilaian, instrumen dinyatakan valid dan layak digunakan dalam pengumpulan data penelitian.

Rekomendasi dari tim ahli terkait perbaikan instrumen disajikan dalam Tabel 3.2 sebagai berikut.

Tabel 3. 2 Rekomendasi Dosen Ahli

No.	Rekomendasi	Revisi
1.	Perbaiki istilah indikator dan sub indikator	Memperbaiki indikator menjadi lebih spesifik atau terukur.
2.	Perbaiki penulisan/typo yang masih keliru	Memperbaiki kesalahan penulisan (<i>typo</i>) dan memperbaiki ejaan sesuai kaidah bahasa yang benar.
3.	Perbaiki kalimat yang masih ambigu serta mudah ditebak jawabannya oleh siswa	Menyederhanakan kalimat pada soal dan pilihan jawaban agar lebih jelas serta menghindari jawaban yang terlalu mudah ditebak.

Hasil validasi dari tim ahli sesuai Tabel 3.2 menghasilkan tiga saran utama untuk penyempurnaan instrumen literasi sains. Pertama, istilah pada indikator dan subindikator disarankan untuk dirumuskan secara lebih jelas, spesifik, dan terukur. Kedua, ditemukan sejumlah kesalahan penulisan yang perlu disesuaikan dengan kaidah bahasa Indonesia yang baku. Ketiga, terdapat butir soal yang dianggap memiliki kalimat kurang jelas atau terlalu mudah ditebak jawabannya. Oleh karena itu, perbaikan melalui penyederhanaan redaksi soal dan pilihan jawaban agar lebih jelas dan tidak memicu jawaban yang terlalu mudah dikenali oleh siswa.

Tahap kedua dilakukan dengan melakukan uji coba soal yang telah melalui proses validasi oleh tim ahli dan diperbaiki kepada siswa. Siswa yang terlibat dalam kegiatan uji coba adalah siswa kelas 5 di dua sekolah berbeda, yaitu SD Dadaha dan SD Nagarawangi 1 yang dipilih berdasarkan kesamaan karakteristik demografis tempat penelitian dengan jumlah responden keseluruhan adalah 33 siswa.

Validasi butir soal hasil uji coba dilakukan dengan pendekatan model Rasch menggunakan perangkat lunak Winsteps versi 5.7.3.0. Salah satu aspek penting

dalam validasi instrumen menggunakan model Rasch adalah pengujian *unidimensionalitas*, yaitu sejauh mana instrumen secara konsisten mengukur satu konstruk yang dimaksud (Sumintono & Widhiarso, 2015). Pengujian unidimensionalitas dilihat melalui nilai *raw variance explained by measures* dan *eigenvalue*, yang interpretasinya disajikan dalam Tabel 3.3. Adapun kriteria unidimensionalitas yang baik ditunjukkan apabila nilai *unexplained variance in the first contrast* memiliki *eigenvalue* kurang dari 3 dan nilai *observed* tidak melebihi 15% (Andrich & Marais, 2019).

Tabel 3. 3 Interpretasi Unidimensionalitas

Interpretasi	<i>Raw Variance Explained by Measure</i>
Terpenuhi	> 20%
Bagus	> 40%
Istimewa	> 60%

(Linacre, 2002)

Tabel 3.3 menyajikan pedoman interpretasi untuk menilai tingkat unidimensionalitas suatu instrumen berdasarkan nilai *Raw Variance Explained by Measure*. Nilai ini menunjukkan seberapa besar persentase varian data yang dapat dijelaskan oleh konstruk utama dalam model Rasch (Sumintono & Widhiarso, 2015). Kategori "Terpenuhi" diberikan apabila nilai *Raw Variance Explained by Measure* lebih dari 20%. Ini menunjukkan bahwa instrumen telah memenuhi batas minimum unidimensionalitas dan dapat digunakan untuk pengukuran meskipun kualitasnya belum optimal. Kategori "Bagus" diberikan jika nilai yang diperoleh lebih dari 40%. Hal ini menunjukkan bahwa instrumen tidak hanya memenuhi unidimensionalitas, tetapi juga memiliki kualitas yang baik karena konstruk utama menjelaskan sebagian besar variansi data. Kategori "Istimewa" diberikan apabila nilai yang diperoleh melebihi 60%. Ini menunjukkan bahwa instrumen memiliki unidimensionalitas yang sangat kuat, di mana hampir seluruh variansi data dijelaskan oleh satu dimensi utama.

Hasil uji coba instrumen dianalisis menggunakan perangkat lunak Winsteps versi 5.7.3.0 yang ditunjukkan pada Gambar 3.2 berikut:

INPUT: 33 Person 20 Item REPORTED: 33 Person 20 Item 2 CATS MINISTEP 5.4.2.0				
Table of RAW RESIDUAL variance in Eigenvalue units				
	Eigenvalue	Observed	Expected	
Total raw variance in observations =	29.7092	100.0%	100.0%	
Raw variance explained by measures =	9.7092	32.7%	33.3%	
Raw variance explained by persons =	5.0130	16.9%	17.2%	
Raw Variance explained by items =	4.6962	15.8%	16.1%	
Raw unexplained variance (total) =	20.0000	67.3%	100.0%	66.7%
<u>Unexplned</u> variance in 1st contrast =	2.6367	10.3%	15.3%	
<u>Unexplned</u> variance in 2nd contrast =	2.2147	8.9%	13.2%	
<u>Unexplned</u> variance in 3rd contrast =	1.7819	7.4%	11.0%	
<u>Unexplned</u> variance in 4th contrast =	1.7461	7.3%	10.8%	
<u>Unexplned</u> variance in 5th contrast =	1.1829	5.4%	8.0%	

Gambar 3. 2 Unidimensionalitas Hasil Uji Coba Instrumen Tes Literasi Sains

Berdasarkan hasil uji coba instrumen tes menggunakan model Rasch, nilai *raw variance explained by measures* yang diperoleh sebesar 32,7% dan *eigenvalue* pada *unexplained variance in 1st contrast* yaitu sebesar 2.6367 dan nilai *observerd* 10,3% yaitu kurang dari 15%. Interpretasi unidimensionalitas yang ditunjukkan pada Tabel 3.3 menunjukkan bahwa instrumen tes yang dikembangkan dapat dikatakan valid dengan kriteria terpenuhi.

Model Rasch juga dapat digunakan untuk mengidentifikasi responden yang tidak menunjukkan pola jawaban yang sesuai (disebut *outlier* atau *misfit*). Responden seperti ini mungkin menjawab secara acak atau mengalami kesulitan dalam memahami pertanyaan, sehingga respons yang diberikan menjadi tidak konsisten dengan model yang diharapkan. Untuk menentukan apakah data dari responden tergolong sesuai (*fit*) atau tidak sesuai (*misfit*), digunakan dua indikator, yaitu nilai *Infit* dan *Outfit*. *Infit* (*inlier-sensitive* atau *information-weighted fit*) menggambarkan sejauh mana pola respons konsisten terhadap item yang sesuai dengan tingkat kemampuan responden, atau sebaliknya (Sumintono & Widhiarso, 2015). Sementara itu, *Outfit* (*outlier-sensitive fit*) mengukur konsistensi pola respons terhadap item yang memiliki tingkat kesulitan tertentu, khususnya terhadap respons yang tidak biasa atau mencolok (Linacre, 2002).

Nilai *Infit* dalam model Rasch dilaporkan dalam bentuk *mean-square* yang diperoleh dari pembagian nilai chi-kuadrat dengan total pembobotannya.

Sementara itu, nilai *outfit* juga dinyatakan sebagai *mean-square*, namun dihitung dengan membagi chi-kuadrat dengan derajat kebebasan. Chi-kuadrat sendiri merupakan ukuran yang umum digunakan dalam statistik. Dalam konteks Rasch, *outliers* bisa muncul akibat tebakan yang kebetulan benar (*lucky guesses*) atau kesalahan yang terjadi karena ketidaksengajaan (*careless mistakes*), yang dapat menyebabkan item yang sebenarnya valid tampak bermasalah (Sumintono & Widhiarso, 2015). Untuk mengatasi hal ini, statistik *Infit* dirancang agar dapat meminimalkan pengaruh *outliers* dengan lebih menekankan pada pola respons butir-butir item yang memiliki tingkat kesukaran atau kemampuan dari responden. Namun demikian, nilai *Outfit* dianggap lebih praktis untuk dilaporkan dalam analisis pengukuran karena lebih mudah diinterpretasikan (Linacre, 2002).

Penentuan item yang tidak sesuai (*misfit*) dalam model Rasch dilakukan dengan mempertimbangkan tiga indikator utama, yaitu nilai *outfit* MNSQ, *outfit* ZSTD, dan *Point Measure Correlation* (PTMEA Corr). Statistik MNSQ (*mean-square*) digunakan untuk melihat sejauh mana jawaban peserta tes sesuai atau menyimpang dari yang diharapkan oleh model pengukuran. MNSQ menunjukkan tingkat ketidakteraturan dalam pola jawaban. Nilai MNSQ idealnya adalah 1, yang mencerminkan sistem pengukuran yang ideal. Jika nilai MNSQ mendekati 1, berarti penyimpangan dalam pengukuran relatif kecil. Nilai MNSQ yang lebih tinggi dari 1 menunjukkan bahwa hasil respons sulit diprediksi, Sementara itu, nilai statistik MNSQ yang berada di bawah 1 mengindikasikan respons yang terlalu konsisten atau terlalu mudah diprediksi, yang dapat menunjukkan bahwa data kurang memberikan informasi baru. Terdapat parameter tertentu yang digunakan dalam model Rasch untuk mengidentifikasi item maupun responden yang tidak sesuai (*misfit*).

ZSTD (*Z-standardized*) merupakan bentuk uji statistik t untuk menguji apakah data sesuai (*fit*) dengan model Rasch. Indikator ini menyajikan signifikansi statistik dari nilai chi-kuadrat berdasarkan asumsi bahwa data mengikuti model. Nilai ZSTD dilaporkan dalam bentuk skor z, dengan nilai ideal sebesar 0,0. Jika nilai ZSTD melebihi 0,0, hal ini mengindikasikan bahwa respons siswa sulit diprediksi oleh model, sedangkan nilai yang lebih rendah dari 0,0 mengisyaratkan

bahwa respons terlalu mudah diprediksi. *Point Measure Correlation* pada analisis Rasch, digunakan sebagai pemeriksaan langsung untuk mengetahui tingkat penilaian respon *fit* sehingga jika korelasi bernilai negatif maka kemungkinan ada sesuatu yang salah. *Point Measure Correlation* mengukur kekuatan dan arah hubungan antara jawaban peserta dan estimasi kemampuan (Sumintono & Widhiarso, 2015).

Tabel 3. 4 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)	$-2 < \text{ZSTD} < 2$.
Nilai <i>Point Measure Correlation</i> (Pt Measure Corr)	$0.4 < \text{Pt Measure Corr} < 0.85$

(Sumintono & Widhiarso, 2015)

Butir tes keterampilan literasi sains yang memenuhi setidaknya salah satu kriteria pada Tabel 3.4, maka butir tersebut dapat digunakan, dengan kata lain butir soal tersebut valid dan layak digunakan dalam pengukuran (Sumintono & Widhiarso, 2015). Berdasarkan hasil uji *Item Statistics: Misfit Order Rasch Model* (lihat lampiran 2.4), diperoleh validitas butir soal literasi sains dengan analisis Model Rasch yang disajikan pada Tabel 3.5 sebagai berikut.

Tabel 3. 5 Hasil Analisis Validitas Instrumen

Kriteria Validitas yang Diterima	Jumlah Item	Item
MNSQ dan ZSTD	8	S5, S8, S10, S11, S14, S17, S18, S19
ZSTD dan PT Measure Corr	1	S16
Ketiga Kriteria	11	S1, S2, S3, S4, S6, S7, S9, S12, S13, S15, S20
Total	20	

Hasil analisis berdasarkan kriteria validitas yang telah ditetapkan sebelumnya, menunjukkan bahwa item soal literasi sains yang diujikan kepada siswa diantaranya 9 item soal memenuhi dua kriteria dan 11 item soal memenuhi

ketiga kriteria. Oleh karena itu, seluruh item soal dapat dikatakan *fit* atau sesuai sehingga dapat dinyatakan valid yang berarti dapat mengukur apa yang hendak diukur.

3.5.2.3 Reliabilitas Instrumen Soal Literasi Sains

Uji reliabilitas bertujuan dalam mengukur sejauh mana instrumen tes mampu memberikan hasil yang konsisten dan dapat dipercaya (Sumintono & Widhiarso, 2015). Suatu alat ukur dinyatakan reliabel jika memberikan hasil konsisten meskipun pengukuran dilakukan pada waktu yang berbeda (Sugiyono et al., 2020). Koefisien reliabilitas yang tinggi mengindikasikan tingkat kesalahan pengukuran yang rendah, sehingga instrumen dapat dikatakan reliabel. Sebaliknya, rendahnya nilai koefisien reliabilitas menunjukkan tingginya potensi kesalahan pengukuran, yang berimplikasi pada rendahnya tingkat keandalan alat ukur tersebut (Suharman, 2018).

Uji reliabilitas dalam penelitian ini, menggunakan analisis pemodelan Rasch dengan bantuan *software* Winsteps versi 5.7.3.0. Data reliabilitas diperoleh melalui menu *output tables*, khususnya pada Tabel 3.1: *Summary Statistics* (lihat lampiran 2.). Terdapat tiga indikator utama digunakan dalam pengukuran ini, yakni *person reliability*, *item reliability*, dan *Cronbach Alpha*. Indikator *Person reliability* menggambarkan tingkat konsistensi respon yang diberikan siswa, sedangkan *item reliability* menunjukkan sejauh mana kualitas butir-butir item soal yang digunakan dalam instrumen. Adapun nilai *Cronbach's Alpha* digunakan untuk menilai reliabilitas secara keseluruhan, yakni berdasarkan interaksi antara siswa dan item. Kriteria interpretasi ketiga indikator tersebut (Sumintono & Widhiarso, 2015) ditunjukkan pada Tabel 3.6 dan Tabel 3.7 berikut.

Tabel 3. 6 Interpretasi Nilai *Person Reliability* dan *Item Reliability*

Rentang Nilai <i>Person</i> dan <i>Item Reliability</i>	Interpretasi
> 0.94	Istimewa
0.91-0.94	Bagus Sekali
0.81-0.90	Bagus
0.67-0.80	Cukup

Rentang Nilai <i>Person</i> dan <i>Item Reliability</i>	Interpretasi
< 0.67	Lemah

Tabel 3. 7 Interpretasi Nilai *Cronbach Alpha*

Rentang Nilai <i>Cronbach Alpha</i>	Interpretasi
>0.8	Bagus Sekali
0.7-0.8	Bagus
0.6-0.7	Cukup
0.5-0.6	Jelek
<0.5	Buruk

Interpretasi terhadap reliabilitas instrumen tes literasi sains berbasis etnosains budaya Kampung Naga dilakukan dengan mengacu pada tiga indikator utama, yaitu *person reliability*, *item reliability*, dan *Cronbach Alpha*. Hasil analisis interpretasi terhadap ketiga indikator reliabilitas tersebut disajikan pada Tabel 3.8 sebagai berikut.

Tabel 3. 8 Hasil Analisis Uji Reliabilitas Instrumen Literasi Sains

Summary Statistic	Nilai Indeks	Interpretasi
<i>Person Reliability</i>	0,75	Cukup
<i>Item Reliability</i>	0,84	Bagus
<i>Cronbach's Alpha.</i>	0,79	Bagus

Tabel 3.8 memuat hasil analisis *output tabel summary statistic* (lihat lampiran 2.5) yang menunjukkan reliabilitas keseluruhan tes literasi sains dengan nilai *Cronbach alpha* 0,79 menunjukkan interaksi yang bagus antara *person* dan butir-butir soal. *Person reliability* dengan nilai 0,75 menunjukkan tingkat konsistensi jawaban siswa yang berada dalam kategori cukup dan *item reliability* dengan nilai 0,84 yang menunjukkan kualitas butir-butir soal bagus. Dengan demikian nilai hasil analisis reliabilitas menunjukkan bahwa instrumen yang dikembangkan reliabel

sehingga layak digunakan sebagai alat pengumpul data dalam penelitian ini (Sumintono & Widhiarso, 2015).

3.5.2.4 Daya Pembeda Instrumen Soal Literasi Sains

Fungsi daya pembeda untuk mengindikasikan sejauh mana butir soal dapat membedakan peserta tes yang berkemampuan tinggi dan rendah (Sumintono & Widhiarso, 2015). Dalam penelitian ini, analisis daya pembeda dilakukan melalui pendekatan pemodelan Rasch melalui perangkat lunak Winsteps 5.7.3.0. Indikator yang digunakan sebagai dasar analisis adalah *Point Measure Correlation (Pt Mean Corr)* sebagai acuan yang menggambarkan korelasi antara tingkat kesulitan suatu item soal dan tingkat kesulitan tes. Hasil interpretasi daya pembeda untuk tiap item soal dapat dilihat pada Tabel 3.9 berikut.

Tabel 3. 9 Interpretasi Daya Pembeda Item Soal

<i>PT Measure Corr</i>	Interpretasi
$ID \geq 0,40$	Bagus Sekali
$0,30 \leq ID < 0,40$	Bagus
$0,20 \leq ID < 0,30$	Cukup
$ID < 0,20$	Buruk

(Smiley, 2015 dalam Utari et al., 2021)

Nilai 1 menunjukkan adanya korelasi sempurna antara respons peserta terhadap suatu item dengan tingkat kemampuan mereka. Sebaliknya, nilai 0 menandakan tidak adanya hubungan antara respons peserta pada item tertentu dengan keseluruhan tes. Apabila nilai yang diperoleh bersifat negatif, hal ini mengindikasikan bahwa item tersebut mengandung kelemahan mendasar dan sebaiknya dievaluasi kembali atau dikeluarkan dari instrumen tes (Smiley, 2015 dalam (Utari et al., 2021).

Hasil analisis daya pembeda instrumen tes kemampuan literasi sains yang diperoleh melalui *software* Winsteps versi 5.7.3.0 , disajikan pada Tabel 3.10.

Tabel 3. 10 Hasil Analisis Daya Pembeda Item Tes Kemampuan Literasi Sains

<i>Item</i>	<i>PT Measure Corr</i>	Interpretasi
1	0,47	Bagus Sekali
2	0,58	Bagus Sekali

<i>Item</i>	<i>PT Measure Corr</i>	Interpretasi
3	0,53	Bagus Sekali
4	0,53	Bagus Sekali
5	0,32	Bagus
6	0,65	Bagus Sekali
7	0,51	Bagus Sekali
8	0,27	Cukup
9	0,41	Bagus Sekali
10	0,39	Bagus
11	0,30	Bagus
12	0,64	Bagus Sekali
13	0,40	Bagus Sekali
14	0,33	Bagus
15	0,51	Bagus Sekali
16	0,66	Bagus Sekali
17	0,38	Bagus
18	0,37	Bagus
19	0,34	Bagus
20	0,49	Bagus Sekali

Tabel 3.10 menyajikan hasil interpretasi daya pembeda butir soal literasi sains, yang dianalisis berdasarkan nilai *Point Measure Correlation*. Berdasarkan tabel tersebut, diketahui bahwa sebagian besar butir soal (12 dari 20) memiliki daya pembeda yang sangat baik, ditunjukkan oleh nilai *Point Measure Correlation* antara 0,40 hingga 0,66. Nilai *Point Measure Correlation* (*PT Measure Corr*) yang melebihi 0,40 menunjukkan bahwa butir soal tersebut efektif dalam membedakan siswa dengan kemampuan tinggi dan rendah, sehingga dapat dinyatakan layak digunakan dalam pengukuran. Selain itu, terdapat 7 butir soal yang tergolong memiliki daya pembeda yang cukup baik, ditunjukkan dengan rentang nilai *PT Measure Corr* antara 0,30 hingga 0,39. Sementara itu, terdapat satu soal dinilai kurang sesuai karena memiliki nilai *PT Measure Corr* di bawah 0,30 yang menunjukkan kelemahan dalam membedakan kemampuan siswa. Temuan ini menunjukkan bahwa mayoritas butir soal yang dikembangkan telah memenuhi kriteria untuk mengukur kemampuan literasi sains, meskipun beberapa item perlu direvisi guna meningkatkan kualitas daya pembedanya.

Kemampuan butir soal dalam membedakan tingkat kemampuan juga dapat dianalisis melalui indeks separasi responden. Menurut Sumintono & Widhiarso (2015) semakin tinggi nilai indeks pemisahan, maka semakin baik pula kualitas instrumen, karena mampu mengelompokkan siswa berdasarkan tingkat kemampuannya secara lebih akurat. Untuk menganalisis pemisahan kelompok responden secara lebih rinci, dapat digunakan rumus pemisahan strata (H) berikut.

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

Keterangan:

H : Nilai strata (pengelompokan)

Separation : Nilai *separation* untuk responden yang dihasilkan

Berdasarkan *output summary measure item* (lihat lampiran 2.5) diketahui bahwa nilai separasi butir soal adalah 2,26 maka pengelompokkan butir soal dapat ditentukan dengan perhitungan separasi sebagai berikut.

$$H = \frac{[(4 \times 2,26) + 1]}{3} = \frac{10,04}{3} = 3,347$$

Berdasarkan hasil perhitungan indeks separasi, diperoleh nilai strata pemisahan (H) sebesar 3,347 yang jika dibulatkan menunjukkan adanya tiga tingkatan atau kategori butir soal, yaitu butir soal dengan tingkat kesulitan tinggi, sedang, dan rendah. Hal ini menunjukkan bahwa daya pembeda instrumen tergolong baik karena mampu mengelompokkan pemahaman siswa berdasarkan tingkat kemampuannya, mulai dari siswa berkemampuan tinggi hingga rendah.

3.5.2.5 Tingkat Kesukaran Instrumen Soal Literasi Sains

Pada teori tes modern Rasch menggunakan bantuan program Winsteps versi 5.7.3.0 tingkat kesukaran suatu butir soal dapat dianalisis melalui perbandingan antara nilai *logit measure* dan standar deviasi (SD) yang diperoleh dari *output* Tabel 13 bagian *Item Fit Statistic* (Boone et al., 2014). Nilai *measure* (ME) menunjukkan tingkat kesulitan butir soal dalam satuan *logit*, yang merupakan skala interval linear berbasis rasio (*log odds unit*) (Sumintono & Widhiarso, 2015). Interpretasi terhadap tingkat kesukaran butir soal dilakukan dengan membandingkan nilai *logit* masing-

masing item terhadap nilai standar deviasinya. Panduan pengelompokan kategori kesukaran tersebut ditampilkan pada Tabel 3.11.

Tabel 3. 11 Kriteria Tingkat Kesukaran Butir Soal

Nilai Measure (<i>Logit</i>)	Interpretasi
$Logit > +1SD$	Sangat Sukar
$0,0 \leq Logit \leq +1SD$	Sukar
$-1SD \leq Logit \leq Mean$	Mudah
$Logit < -SD$	Sangat Mudah

(Sumintono & Widhiarso, 2015)

Nilai *logit* yang tinggi mengindikasikan bahwa suatu butir soal memiliki tingkat kesukaran yang tinggi. Tingkat kesukaran tersebut ditunjukkan melalui nilai *measure*, yaitu hasil konversi dari data ordinal (skor jawaban siswa) ke dalam skala *logit* yang bersifat *equal interval* (Boone, 2016). Berdasarkan *output summary measure item* (lihat lampiran 2.5) diketahui bahwa nilai rata-rata *logit* item adalah 0,00 dan nilai standar deviasi (S.D) adalah 1,12. Nilai tersebut dikombinasikan untuk mengklasifikasikan interpretasi kesukaran butir soal yang dapat dilihat pada Tabel 3.12.

Tabel 3. 12 Hasil Kriteria Tingkat Kesukaran Butir Soal

Nilai Measure (<i>Logit</i>)	Interpretasi
$> 1,12$	Sangat Sukar
$0,0 - 1,12$	Sukar
$0,0$	Sedang
$(-1,12) - 0,0$	Mudah
$< (-1,12)$	Sangat Mudah

Tingkat kesukaran butir soal ke dalam lima kategori, sebagaimana ditunjukkan pada Tabel 3.12 menunjukkan butir soal dengan nilai *logit* di atas 1,12 dikategorikan sebagai sangat sukar, sementara yang berada pada rentang 0,0 hingga 1,12 termasuk dalam kategori sukar. Nilai *logit* tepat 0,0 diklasifikasikan sebagai

sedang, sedangkan nilai *logit* antara -1,12 hingga 0,0 dianggap mudah. Adapun butir soal dengan nilai *logit* di bawah -1,12 masuk dalam kategori sangat mudah.

Setelah ditentukan kriteria klasifikasi tingkat kesukaran soal, tahap selanjutnya adalah melakukan analisis terhadap *Output Tabel 13. Item Measured* (lihat lampiran 2.6) yang dihasilkan oleh perangkat lunak Winsteps dengan pendekatan Rasch Model. Untuk hasil yang lebih jelas dapat dilihat pada tabel 3.7 sebagai berikut.

Tabel 3. 13 Rekapitulasi Hasil Analisis *Item Measured*

Kategori	No. Butir Soal	Jumlah	Persentase
Sangat Sukar	11, 15, 16	3	15%
Sukar	2, 6, 9, 10, 12, 13, 14	7	35%
Mudah	1, 7, 17, 18, 19, 20	6	30%
Sangat Mudah	3, 4, 5, 8	4	20%

Hasil analisis pada Tabel 3.13 menunjukkan tingkat kesukaran 20 butir soal tes kemampuan literasi sains (S adalah soal) tersebut dapat diketahui bahwa 3 butir soal (S11, S15, S16,) termasuk kategori sangat sukar dengan persentase 15%, 7 butir soal (S2, S6, S9, S10, S12, S13, S14) termasuk kategori sukar dengan persentase 35 %, 6 butir soal (S1, S7, S17, S18, S19, S20) termasuk kategori mudah dengan persentase 30%, dan 4 butir soal (S3, S4, S5, S8) termasuk kategori sangat mudah dengan persentase 20%.

Dengan demikian, dapat disimpulkan bahwa soal tes kemampuan literasi sains berdasarkan perbandingan butir soal dengan kategori sangat sukar, sukar, mudah dan sangat mudah sudah seimbang yaitu butir soal memuat semua kategori tingkat kesulitan soal. Sejalan dengan hal tersebut Mukhlis (2023) menyatakan bahwa karakteristik soal yang baik ditandai dengan tingkat kesulitan yang seimbang, artinya tidak didominasi oleh soal yang terlalu mudah maupun terlalu sulit.

3.6 Prosedur Penelitian

Prosedur penelitian ini terdiri atas tiga tahap utama, yakni tahap persiapan, tahap pelaksanaan, dan tahap akhir yang mencakup proses pengumpulan data, analisis data, serta penyusunan laporan hasil penelitian. Setiap tahapan dijelaskan lebih lanjut sebagai berikut.

3.6.1 Tahap Persiapan

1. Mengidentifikasi permasalahan melalui kegiatan observasi langsung dilapangan dan kajian literatur untuk mengeksplorasi alternatif solusi yang potensial untuk diterapkan dalam menyelesaikan permasalahan yang ditemukan.
2. Melakukan kajian literatur terkait variabel-variabel penelitian yang akan dikaji, berdasarkan teori-teori yang relevan.
3. Menentukan subjek dan lokasi penelitian.
4. Menyusun rancangan desain dan perangkat pembelajaran yang mengadopsi model RADEC berbasis etnosains budaya Kampung Naga, dengan fokus pada kemampuan literasi sains.
5. Menyusun instrumen pengumpulan data berupa tes literasi sains yang dirancang khusus untuk mengukur tingkat pencapaian literasi sains siswa.
6. Melakukan validasi (*expert judgment*) rancangan instrumen literasi sains oleh tim ahli.
7. Mengajukan surat izin penelitian kepada pihak Kampus UPI Tasikmalaya, kepada kepala sekolah tempat penelitian dilaksanakan.
8. Berkonsultasi dengan kepala sekolah dan guru kelas untuk menentukan waktu pelaksanaan, subjek, dan sampel penelitian yang akan terlibat.
9. Melaksanakan uji coba instrumen dan menganalisis hasilnya untuk menentukan kelayakan serta perbaikan instrumen sebelum digunakan dalam penelitian.

3.6.2 Tahap Pelaksanaan

1. Melaksanakan *pretest* dengan memberikan soal literasi sains kepada siswa sebelum perlakuan dilakukan.

2. Memberikan perlakuan yang berlangsung selama empat kali pertemuan, dengan tahapan kegiatan sebagai berikut.
 - a. Pertemuan 1–3
Tahapan *Read, Answer, Discuss, Explain* dilaksanakan melalui kegiatan membaca referensi budaya dan IPAS, menjawab soal pra-pembelajaran, diskusi kelompok, dan presentasi.
 - b. Pertemuan 4
Tahapan *Create* dilaksanakan melalui kegiatan merancang dan menyajikan karya yang berkaitan dengan pelestarian SDA dan praktik budaya lokal.
3. Melaksanakan *posttest* untuk mengetahui perubahan atau peningkatan kemampuan literasi sains siswa setelah mendapat perlakuan.

3.6.3 Tahap Akhir

1. Data hasil *pretest* dan *posttest* dikumpulkan lalu dikelompokkan sesuai dengan aspek-aspek literasi sains yang menjadi fokus penelitian.
2. Data tersebut kemudian diproses dan dianalisis menggunakan Model Rasch dengan metode *Racking* dan *Stacking Analysis*.
3. Hasil analisis dibahas secara mendalam dan rinci.
4. Analisis yang telah dilakukan digunakan untuk menarik kesimpulan dan menjawab rumusan pertanyaan penelitian.
5. Hasil dan temuan penelitian ini kemudian disusun secara sistematis dan dijabarkan dalam bentuk laporan penelitian.

3.7 Analisis Data

Penelitian ini menggunakan analisis data berupa analisis lembar observasi keterlaksanaan pembelajaran, analisis kemampuan literasi sains siswa sebelum dan setelah dilakukannya perlakuan (*treatment*), analisis aspek literasi sains yang mengalami perubahan dan analisis efektivitas model pembelajaran RADEC berbasis etnosains budaya Kampung Naga terhadap literasi sains siswa.

3.7.1 Analisis Keterlaksanaan Pembelajaran

Data analisis lembar observasi keterlaksanaan pembelajaran diperoleh melalui pengamatan oleh observer terhadap aktivitas guru dan siswa selama penerapan pembelajaran model RADEC berbasis etnosains budaya Kampung

Naga. Skor penilaian untuk keterlaksanaan pembelajaran dimulai dari 0 untuk “tidak terlaksana”, 1 untuk “terlaksana”. Pengolahan hasil observasi dilakukan dengan menghitung persentase keterlaksanaan pembelajaran (KP) menggunakan rumus berikut.

$$\%KP = \frac{\sum \text{skor keterlaksanaan}}{\sum \text{skor maksimal}} \times 100\%$$

(Sudaryono, 2016)

Rumus persentase keterlaksanaan di atas, digunakan untuk mengetahui hasil keterlaksanaan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga setiap pertemuan. Selanjutnya, dihitung rata-rata persentase keterlaksanaannya dengan rumus berikut.

$$\overline{KP} = \frac{\sum \%KP \text{ Setiap Pertemuan}}{\sum \text{Pertemuan}}$$

Modifikasi dari Vilosofah et al (2025)

Keterlaksanaan pembelajaran dikatakan telah terlaksana dengan benar apabila persentase rata-rata yang diperoleh $\geq 61\%$. Kategori keterlaksanaan pembelajaran dan interpretasi berdasarkan persentase yang diperoleh disajikan pada Tabel 3.14.

Tabel 3. 14 Interpretasi Persentase Keterlaksanaan Pembelajaran

Keterlaksanaan Pembelajaran (KP) (%)	Interpretasi
81% - 100%	Sangat Baik
61% - 80%	Baik
41% - 60%	Cukup Baik
21% - 40%	Buruk
0% - 20%	Sangat Buruk

Riduwan (dalam Firdausichuuriyah et al., 2017)

Tabel 3.14 menunjukkan kategori yang digunakan sebagai acuan untuk menilai sejauh mana pelaksanaan kegiatan pembelajaran telah sesuai dengan perencanaan dan prosedur yang ditetapkan.

3.7.2 Analisis Uji Kemampuan Literasi Sains Siswa

Proses menganalisis hasil asesmen dalam dunia pendidikan dapat dilakukan menggunakan dua pendekatan utama (Susdelina et al., 2018). Salah satu metode pengukuran konvensional dan masih banyak dipakai adalah *Classical Test Theory* (CTT) atau Teori Tes Klasik. Dikembangkan pertama kali oleh Charles Spearman pada awal abad ke-20 yaitu tahun 1904, pendekatan ini telah menjadi dasar dalam memprediksi performansi peserta ujian (Sumintono & Widhiarso, 2015). CTT mengevaluasi kualitas soal melalui dua indikator utama yaitu tingkat kesulitan dan kemampuan soal dalam membedakan peserta atau daya pembeda. Namun, pendekatan ini memiliki keterbatasan signifikan. Salah satunya adalah sifat soal yang tidak stabil, karena sangat dipengaruhi oleh kemampuan peserta tes. Selain itu, CTT hanya mampu menghitung *margin of error* secara kelompok, bukan per individu, sehingga kurang akurat untuk penilaian personal (Perdana, 2018).

Pendekatan kedua sebagai solusi atas keterbatasan CTT, berupa pemodelan Rasch yang dikembangkan oleh Georg Rasch, seorang ahli statistik asal Denmark pada tahun 1960-an. Berbeda dengan CTT, model ini mengonversi perolehan skor mentah menjadi skala pengukuran interval linier yang memfasilitasi penilaian lebih akurat baik terhadap kemampuan peserta tes maupun kualitas butir soal. Pemodelan Rasch mengubah data ordinal menjadi data rasio untuk keperluan analisis statistik. Perbandingan rasio ini menunjukkan korelasi antara tingkat kemampuan siswa (*person ability*) dan tingkat kesulitan butir soal (*item difficulty*). Untuk menghasilkan pengukuran yang berskala interval sama, digunakan fungsi logaritma, yang kemudian menghasilkan satuan pengukuran baru yang dikenal sebagai *logit* (*log odds unit*). Menurut Englehard, keberhasilan siswa dalam mengerjakan soal sangat ditentukan oleh tingkat kemampuan siswa dan tingkat kesulitan soal (Sumintono, 2021).

Nilai probabilitas *odd ratio* individu dapat ditentukan melalui persamaan berikut:

$$Odd Ratio = \frac{P}{(N - P)}$$

(Sumintono & Widhiarso, 2015)

Keterangan:

P: jumlah soal yang dikerjakan dengan benar oleh siswa

N: jumlah total soal

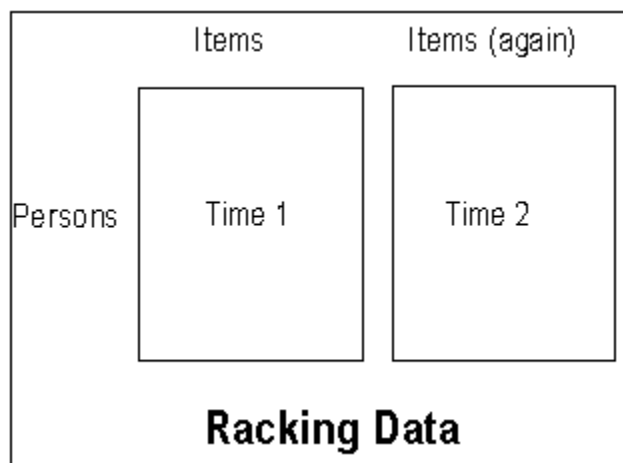
Nilai probabilistik *odd ratio* diubah menjadi bentuk logaritma untuk menciptakan skala pengukuran yang memiliki interval yang sama. Formula untuk *logit* (unit *log odds*) yang dirumuskan oleh Sumintono dan Widhiarso (2015) adalah sebagai berikut.

$$Logit = \text{Log} \left(\frac{P}{(N - P)} \right)$$

Keunggulan utama analisis Rasch terletak pada kemampuannya menyajikan karakteristik butir soal dan siswa dalam satu metrik terpadu, menjadikannya alat evaluasi yang lebih andal dalam konteks pendidikan (Sumintono & Widhiarso, 2015). Berdasarkan penjelasan tersebut, dalam penelitian ini pendekatan yang digunakan untuk menganalisis perubahan kemampuan literasi sains siswa adalah pendekatan teori tes modern yaitu pemodelan Rasch dengan teknik *Racking* dan *Stacking*. Kemudian interpretasi pengukurannya menggunakan nilai *logit* dan perubahan lokasi pada *Wright Maps*.

3.7.2.1 *Racking Analysis*

Metode *racking* digunakan untuk menganalisis perubahan tingkat kesulitan butir soal sebelum dan sesudah intervensi dengan cara membandingkan data *pre-test* dan *posttest*. Dalam metode ini, data *pretest* dan *posttest* ditempatkan secara horizontal, data diletakkan dalam baris yang sejajar atau horizontal (Laliyo, 2021).



Gambar 3. 3 Skema Data *Racking*

Dengan menempatkan data *pre-test* dan *posttest* dalam satu baris atau kolom, dapat dilihat apakah ada perubahan dalam tingkat kesulitan soal antara kedua tes.

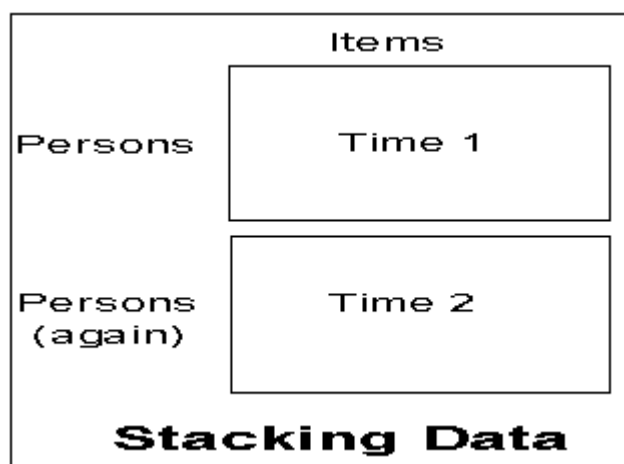
Setiap item soal akan muncul dua kali dalam data ini, satu kali untuk *pretest* dan satu kali untuk *posttest*, meskipun siswa hanya muncul satu kali dalam data (karena mereka hanya diuji satu kali), setiap soal yang diujikan akan muncul dua kali dalam dataset yang terpisah untuk *pretest* dan *posttest*. Dengan adanya data *pretest* dan *posttest* yang ditempatkan secara terpisah namun mengacu pada soal yang sama, dapat dilihat tingkat kesulitan soal mengalami perubahan. Misalnya, jika soal yang sama dianggap sulit oleh siswa di *pretest* namun menjadi lebih mudah di *posttest*, itu menunjukkan adanya perubahan tingkat kesulitan soal. Sebaliknya, jika soal yang sama tetap sulit di kedua tes, maka tingkat kesulitan item tersebut tidak mengalami perubahan.

Tingkat kesulitan masing-masing item dalam analisis *racking* ditentukan melalui perbandingan skor *pretest* dan *posttest*. Tingkat kesulitan item dalam penelitian ini diuraikan berdasarkan aspek-aspek literasi sains yang terdapat dalam soal. Hal ini memungkinkan peneliti untuk mendapatkan informasi tentang aspek-aspek literasi sains mana saja yang mengalami perubahan, baik itu peningkatan atau penurunan, setelah intervensi dilakukan. Dengan demikian, metode *racking* memberikan gambaran yang jelas mengenai efektivitas intervensi terhadap kemampuan literasi sains siswa.

3.7.2.2 *Stacking Analysis*

Analisis *Stacking* dalam pemodelan Rasch, merupakan teknik analisis longitudinal yang membandingkan kemampuan siswa sebelum dan setelah perlakuan atau intervensi tertentu dalam lingkungan belajar (Sukarelawan et al., 2024). Teknik ini memungkinkan peneliti atau guru untuk melacak perkembangan keterampilan siswa secara individual dari waktu ke waktu. Metode ini memerlukan pengumpulan informasi dari sekelompok siswa yang mengikuti tes awal (*pretest* yang dilakukan sebelum intervensi) dan tes akhir (*posttest* yang dilakukan setelah intervensi) untuk menilai perubahan dalam kemampuan siswa. Model Rasch kemudian digunakan untuk mengevaluasi data ini, mengubah hasil tes menjadi pengukuran kemampuan yang serupa (*logit*) (Sukarelawan et al., 2024).

Analisis teknik *Stacking* menempatkan data *pretest* dan *posttest* secara vertikal yang berarti data dari kedua tes tersebut diorganisir dalam bentuk kolom yang saling berdampingan, satu untuk *pretest* dan satu untuk *posttest*. Hal tersebut memungkinkan peneliti untuk membandingkan hasil individu (dalam hal ini, siswa) sebelum dan sesudah intervensi dilakukan (Laliyo, 2021) dan mendapatkan informasi tentang siswa kategori mana saja yang mengalami perubahan, apakah ada peningkatan atau penurunan dalam kemampuan kemampuan literasi sains atau jawaban siswa setelah mereka mendapatkan intervensi.



Gambar 3. 4 Skema Data *Stacking*

Metode ini memberikan gambaran yang jelas mengenai dampak intervensi pada setiap individu berdasarkan perbandingan hasil *pre-posttest* yang telah dikerjakan oleh siswa.

Pengelompokan tingkat kemampuan atau *ability* siswa dilakukan dengan menggunakan nilai rata-rata (*mean*) dan standar deviasi (SD) dari nilai *logit* siswa pada masing-masing tes dengan menjumlahkan nilai *logit* rata-rata dengan S.D (Sumintono & Widhiarso, 2015). Berdasarkan hasil analisis nilai *measure* (*non-extreme*), nilai separasi dihitung melalui selisih antara nilai *logit* rata-rata dan S.D, yang kemudian digunakan untuk menentukan jarak dari median ke satu deviasi standar (S.D) (Ramadhan & Hidayatullah, 2023). Kategori pengelompokan tingkat kemampuan literasi sains siswa disajikan pada Tabel 3.14 sebagai berikut.

Tabel 3. 15 Kategori Kemampuan Siswa

Nilai <i>Measure</i> (<i>Logit</i>)	Interpretasi
$Logit > (Mean + 2SD)$	<i>Outlier Atas</i>
$(Mean + 1SD) < Logit \leq (Mean + 2SD)$	Sangat Tinggi
$Mean \leq Logit \leq (Mean + 1SD)$	Tinggi
$(Mean - 1SD) \leq Logit \leq Mean$	Rendah
$(Mean - 2SD) \leq Logit < (Mean - 1SD)$	Sangat Rendah
$Logit < (Mean - 2SD)$	<i>Outlier Bawah</i>

(Sumintono & Widhiarso, 2015)

Langkah-langkah analisis perubahan kemampuan siswa menggunakan analisis *stacking* yaitu sebagai berikut.

- 1) Mengumpulkan data *pretest* sebelum intervensi dan data *posttest* setelah intervensi dari siswa yang sama;
- 2) Menggunakan model Rasch untuk mentransformasi data mentah menjadi skala *logit* yang mengukur kemampuan secara komparatif;
- 3) Menyusun data dalam format *stacking* untuk membandingkan nilai kemampuan *pretest* dan *posttest* setiap individu;

- 4) Membandingkan nilai *logit pretest* dan *posttest* untuk mengidentifikasi perubahan kemampuan;
- 5) Melakukan analisis statistik untuk mengevaluasi signifikansi perbedaan antara data *pretest* dan *posttest*; dan
- 6) Meninjau hasil perbandingan untuk menilai efektivitas strategi pembelajaran. Jika terdapat peningkatan signifikan, hal ini menunjukkan keberhasilan intervensi.

3.7.3 Analisis Efektivitas Model Pembelajaran RADEC Berbasis Etnosains Budaya Kampung Naga

Analisis efektivitas bertujuan untuk mengkaji keefektifan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga terhadap peningkatan kemampuan literasi sains siswa sekolah dasar. Menurut Sugiyono (2019), efektivitas diukur dengan membandingkan nilai awal (*pretest*) dan nilai akhir (*posttest*) siswa setelah mengikuti pembelajaran menggunakan model tersebut. Sebelum dilakukan uji efektivitas, terlebih dahulu dilakukan pengujian terhadap asumsi distribusi data untuk menentukan jenis uji statistik yang tepat. Teknik analisis yang digunakan mencakup uji normalitas terhadap data hasil uji *stacking pre-posttest*, kemudian dilanjutkan dengan uji hipotesis nonparametrik apabila data tidak berdistribusi normal.

3.7.3.1 Uji Asumsi Normalitas

Uji asumsi normalitas pada penelitian ini dilakukan terhadap nilai *logit* yang diperoleh dari hasil analisis pemodelan Rasch dengan teknik *stacking* (Sukarelawan et al., 2024). Teknik *stacking* digunakan untuk menggabungkan data *pretest* dan *posttest* dalam satu file analisis, di mana setiap peserta direpresentasikan oleh dua entri, yaitu nilai *logit* saat *pretest* dan *posttest*. Nilai *logit* ini merupakan bentuk transformasi dari data mentah (skor) ke skala interval logaritmik, sehingga lebih representatif untuk analisis statistik.

Uji normalitas bertujuan untuk melihat apakah sebaran nilai *logit* gabungan tersebut mengikuti distribusi normal. Hal ini penting untuk menentukan jenis uji statistik yang tepat, apakah menggunakan uji parametris (seperti *paired t-test*) atau non-parametris (seperti *Wilcoxon signed-rank test*). Pengujian normalitas dilakukan

dengan menggunakan uji *Shapiro-Wilk* penelitian yang dilakukan oleh Shapiro, Wilk, dan Chen, uji Shapiro-Wilk dirancang dan terbukti efektif digunakan pada jumlah sampel yang relatif kecil, yakni berkisar antara 10 hingga tidak lebih dari 50 responden. Dengan mempertimbangkan bahwa jumlah sampel dalam penelitian ini sebanyak 25 orang, maka uji *Shapiro-Wilk* dipandang sebagai metode yang paling tepat untuk menguji normalitas data (Field et al., 2012).

Proses analisis normalitas dalam penelitian ini dilakukan menggunakan perangkat lunak IBM SPSS *Statistics*. Jika nilai signifikansi (p) $> 0,05$ data dianggap berdistribusi normal, sehingga dapat digunakan uji parametrik. Jika $p \leq 0,05$, artinya data tidak berdistribusi normal, sehingga uji non-parametrik lebih tepat digunakan. Seperti dijelaskan oleh Sukarelawan et al. (2024), penting untuk memastikan bahwa nilai *logit* hasil *stacking* memenuhi asumsi distribusi normal sebelum digunakan dalam analisis lebih lanjut. Uji ini membantu menjaga validitas interpretasi perubahan kemampuan literasi sains siswa sebelum dan setelah perlakuan pembelajaran.

3.7.3.2 Uji Hipotesis

Pengujian hipotesis dalam penelitian ini dilakukan dengan menggunakan uji-*t* berpasangan (*Paired Sample t-test*) apabila data terdistribusi normal, atau uji *Wilcoxon* jika data tidak berdistribusi normal. Kedua uji ini digunakan untuk membandingkan rata-rata dua pengukuran yang berasal dari kelompok yang sama, yaitu membandingkan skor *pretest* dan *posttest* dari siswa yang sama. Tujuannya adalah untuk mengetahui apakah terdapat perbedaan yang signifikan dalam kemampuan literasi sains siswa sebelum dan sesudah diterapkannya model pembelajaran RADEC berbasis etnosains budaya Kampung Naga.

Analisis data statistik dalam penelitian kuantitatif, bertujuan untuk menguji kebenaran hipotesis yang diajukan (Sugiyono, 2019). Hipotesis dalam konteks ini diartikan sebagai dugaan sementara terhadap pertanyaan penelitian. Proses pengujian hipotesis dilaksanakan dengan melakukan perbandingan nilai signifikansi (sig.) terhadap tingkat alpha 0,05. Kriteria pengambilan keputusannya adalah:

1. Jika signifikansi bernilai $< 0,05$, maka hipotesis alternatif (H_a) diterima yang berarti terdapat perbedaan skor *logit* kemampuan literasi sains siswa yang signifikan sebelum dan setelah diterapkan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga.
2. Jika signifikansi bernilai $> 0,05$, maka hipotesis alternatif ditolak dan hipotesis nol (H_0) yang diterima, menunjukkan tidak adanya perbedaan skor *logit* kemampuan literasi sains siswa yang signifikan sebelum dan setelah diterapkan model pembelajaran RADEC berbasis etnosains budaya Kampung Naga