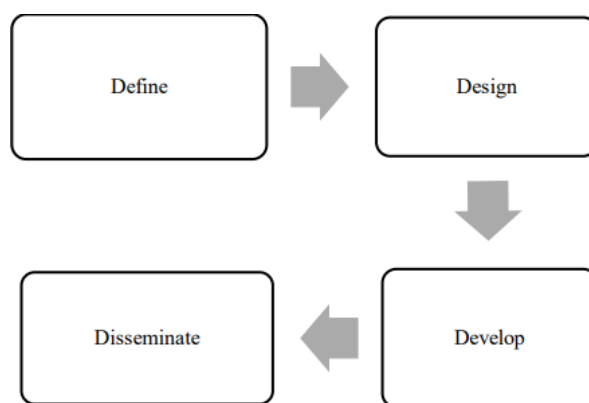


BAB III

METODE PENELITIAN

3.1 Metode dan Desain Penelitian

Penelitian ini merupakan *Research and Development* (RnD) yaitu metode penelitian untuk mengembangkan dan menguji produk berupa instrument tes. Terdapat berbagai jenis metode RnD salah satunya adalah model 4D (*Define, Design, Develop, Disseminate*) dengan tahapan sebagai berikut.



Gambar 3.1 Tahapan Model 4D

Menurut Thiagarajan, S., Semmel, D. S., Semmel (1974) ke empat tahapan tersebut diuraikan menjadi:

1. Pertama, *Define*. Tahap ini merupakan tahap analisis kebutuhan. Pada tahap ini peneliti melakukan analisis kebutuhan instrumen *technology engineering literacy* bagi pendidik.
2. Kedua, *Design*. Tahap ini merupakan tahap perancangan. Pada tahap ini peneliti merancang instrument yang akan dikembangkan melalui tahap awal yaitu analisis indicator TEL yang relevan dengan materi energi terbarukan dan kurikulum yang berlaku di Indonesia.
3. Ketiga, *Develop*. Tahap ini merupakan tahap untuk menghasilkan produk pengembangan atau model. Pada tahap ini peneliti mengembangkan butir soal sesuai dengan indicator yang telah dipilih. Selain itu pada tahap ini peneliti mengembangkan website yang akan dijadikan sarana untuk TEL. Setelah itu peneliti melakukan uji coba terbatas.

4. Keempat, *Disseminate*. Tahap ini merupakan tahap penyebarluasan produk atau model melalui pada individu, kelompok, atau sistem. Pada tahap ini peneliti melakukan uji coba dalam skala luas.

3.2 Partisipan Penelitian

Partisipan merupakan elemen penting dalam penelitian pendidikan karena karakteristiknya dapat memengaruhi kualitas data dan generalisasi temuan (Creswell, J.W., & Creswell, 2018). Dalam penelitian ini, partisipan didefinisikan sebagai individu atau kelompok yang menjadi sumber data utama dalam proses pengumpulan informasi yang relevan dengan tujuan penelitian. Pemilihan partisipan harus mempertimbangkan keterkaitan antara populasi sasaran, desain penelitian, dan kriteria inklusi yang telah ditetapkan, sehingga data yang diperoleh dapat merepresentasikan fenomena yang dikaji secara valid. Dalam konteks pengembangan dan pengujian instrumen asesmen, keterlibatan partisipan dengan latar belakang pengetahuan yang sesuai dengan topik penelitian menjadi krusial agar instrumen dapat dievaluasi dari segi validitas, reliabilitas, tingkat kesukaran, dan daya pembeda secara optimal (Bond, T. G., & Fox, 2015).

Partisipan dalam penelitian ini yaitu 263 orang peserta didik kelas X di beberapa Sekolah Menengah Atas (SMA) yang terletak di kota Bandung dan Tasikmalaya dengan rincian 90 orang peserta didik terlibat pada tahap uji coba terbatas dan 173 peserta didik penyebaran skala luas. Adapun demografi partisipan adalah sebagai berikut:

Tabel 3.1 Demografi Partisipan Penelitian

Tahap	STEM	Non-STEM	Jumlah
Uji Coba	- 60 orang peserta didik kota Bandung - 30 orang peserta didik Kabupaten Tasikmalaya		90 orang
Diseminasi	- 32 Orang peserta didik kota Bandung	- 22 orang peserta didik Kabupaten Tasikmalaya	173 orang

Tahap	STEM	Non-STEM	Jumlah
	- 31 orang peserta didik Kab. Tasikmalaya	- 26 orang peserta didik Kab. Bandung Barat	
	- 62 orang peserta didik Kota Tasikmalaya		
Total			263 orang

Untuk menghindari faktor bias daerah kota atau desa, maka distribusi partisipan pada tahap ujicoba diambil dari 2 sekolah yaitu 1 sekolah di Kabupaten Tasikmalaya dan 1 sekolah di Kota Bandung. Sedangkan pada tahap diseminasi partisipan dipilih beragam dengan pertimbangan kriteria tertentu. Lebih lengkapnya mengenai karakteristik partisipan disajikan dalam Tabel 3.1.

Tabel 3.2 Karakteristik Partisipan Pada Tahap Diseminasi

Jenis Pendekatan	Aktivitas Pembelajaran
STEM	Pembelajaran pada materi energi terbarukan dilakukan menggunakan model PBL dan PjBL dengan mengintegrasikan pendekatan STEM, peserta didik terbagi menjadi beberapa kelompok yang ditugaskan membuat <i>prototype</i> sebagai solusi dari permasalahan yang diberikan.
Non-STEM	Pembelajaran pada materi energi terbarukan dilakukan dengan model direct instruction dan PBL dengan memberikan penugasan berupa presentasi kelompok menganalisis isu energi terbarukan.

Distribusi strategis ini memperkuat ketelitian dan penerapan hasil penelitian, memberikan gambaran representatif dari berbagai pengalaman pendidikan di antara peserta didik di Jawa Barat. Namun kelompok STEM dan non-STEM mungkin tidak secara inheren setara, hal ini karena pemilihan sampel bersifat *purposive* dan berdasarkan keterlibatan dalam pendekatan pengajaran yang sangat berbeda. Untuk memastikan kesetaraan dua kelompok sejak awal peneliti berupaya untuk mengendalikan potensi variabel pengganggu seperti kemampuan akademik awal peserta didik, latar belakang sosial ekonomi, dan kualitas guru. Salah satunya dengan memilih peserta didik berdasarkan prestasi akademik tinggi di setiap sekolah yang berpartisipasi.

3.3 Instrumen Penelitian

Instrumen penelitian merupakan alat yang digunakan untuk mengumpulkan data sesuai dengan tujuan penelitian, sehingga kualitas instrumen akan sangat menentukan validitas dan reliabilitas temuan (Sugiyono, 2015). Dalam penelitian kuantitatif, instrumen berfungsi sebagai sarana untuk mengukur variabel secara objektif dan sistematis, baik dalam bentuk tes, kuesioner, lembar observasi, maupun pedoman wawancara. Penyusunan instrumen harus melalui proses yang terstruktur, mulai dari perumusan indikator berdasarkan kajian teori, penyusunan butir-butir pertanyaan atau soal, hingga pengujian validitas dan reliabilitasnya (Creswell, J.W., & Creswell, 2018; Creswell, 2014). Pada penelitian pengembangan, instrumen tidak hanya berperan dalam pengumpulan data, tetapi juga menjadi objek utama yang diuji kualitasnya melalui analisis statistik, seperti uji validitas isi, validitas konstruk, reliabilitas, tingkat kesukaran, daya pembeda, dan keberfungsian distractor. Instrumen yang digunakan dalam penelitian ini yaitu sebagai berikut.

3.4.1. Lembar Validasi Ahli (*Expert Judgement*)

Lembar validasi ahli digunakan untuk mengetahui kualitas instrumen dengan cara meminta hasil pandangan dari validator. Lembar validasi digunakan oleh validator untuk menilai kualitas butir soal diantaranya dengan menentukan kesesuaian antara indikator TEL dengan butir soal yang telah dibuat. Lembar validasi ini terdiri atas rubrik penskoran secara kuantitatif dan uraian pendapat validator secara kualitatif. Adapun lembar validasi ahli yang digunakan yaitu sebagai berikut.

No.	Indikator TEL		Soal	Aspek yang dinilai	Skor (*berikan tanda V)					Catatan
	Assessment Target	Assessment Area			1	2	3	4	5	
			A.	Soal sesuai dengan indikator TEL						
				Konsep yang disajikan dalam soal bebas dari miskonsepsi						
				Materi yang disajikan sesuai dengan level peserta didik						
				Soal dirumuskan dengan singkat, jelas, dan tegas						
				Stimulus yang disajikan jelas dan berfungsi						
				Bahasa yang digunakan pada soal sesuai dengan kaidah EYD						
				Bahasa yang digunakan bebas dari ambiguitas						

Gambar 3.2 Format Lembar Validasi Ahli

3.4.2. Lembar Instrumen Tes

Lembar instrumen tes terdiri atas butir soal yang harus dikerjakan oleh peserta didik. Tes yang digunakan berupa pilihan ganda yang terdiri atas lima opsi jawaban disajikan menggunakan media nearpod. Tampilan halaman tes pada nearpod adalah sebagaimana **Gambar 3.3** berikut ini.



Gambar 3.3 Halaman Tes TEL dalam Nearpod

3.4 Prosedur Penelitian

Berdasarkan tahapan 4D maka prosedur dalam penelitian ini yaitu sebagai berikut:

a. Tahap Pendefinisian (*Define*)

Tahap ini bertujuan untuk menetapkan dan merumuskan kebutuhan pengembangan instrumen *Technology and Engineering Literacy* (TEL) pada materi energi terbarukan. Langkah-langkah yang dilakukan meliputi:

1) Analisis Awal

Peneliti mengidentifikasi permasalahan dan kebutuhan di lapangan melalui observasi modul ajar energi terbarukan yang digunakan di SMA. Fokus analisis adalah pada kesesuaian materi, bentuk pembelajaran, dan asesmen yang digunakan.

2) Analisis Kebutuhan

Peneliti melakukan wawancara semi-terstruktur dengan salah satu guru SMA untuk menggali informasi terkait pengalaman mengajar materi energi terbarukan, hambatan dalam penilaian ketercapaian kompetensi peserta didik, serta ekspektasi guru terhadap instrumen yang akan dikembangkan.

3) Analisis Kurikulum

Peneliti menelaah Capaian Pembelajaran dan Alur Tujuan Pembelajaran pada Kurikulum Merdeka, khususnya fase E untuk mata pelajaran yang memuat materi energi terbarukan. Analisis ini digunakan untuk menentukan indikator TEL yang relevan dengan konteks pembelajaran di SMA.

4) Perumusan Tujuan Pembelajaran dan Indikator

Berdasarkan hasil observasi, wawancara, dan analisis dokumen, peneliti merumuskan tujuan pembelajaran instrumen TEL yang dapat mengukur kemampuan peserta didik pada materi energi terbarukan secara valid, reliabel, dan sesuai dengan kerangka *Technology and Engineering Literacy* NAEP.

b. Tahap Perancangan (*Design*)

Pada tahap ini peneliti mendesain atau merancang instrumen *technology engineering literacy* (TEL) melalui beberapa tahapan sebagai berikut.

- 1) Melakukan analisis indikator kemampuan *literacy technology engineering* mengacu pada *framework* TEL.
- 2) Melakukan analisis kesesuaian indikator dengan materi energi terbarukan, pada proses ini peneliti memilih beberapa indikator yang relevan untuk digunakan.
- 3) Merumuskan indikator pencapaian kompetensi atau tujuan pembelajaran yang sesuai untuk mencapai target pembelajaran pada materi energi terbarukan.
- 4) Membuat kartu soal dan kisi-kisi instrumen tes berdasarkan indikator yang telah dirancang sebelumnya.

c. Tahap Pengembangan (*Develop*)

Tahap ini bertujuan untuk menghasilkan instrumen *Technology and Engineering Literacy* (TEL) yang siap diimplementasikan dan telah teruji secara psikometrik. Adapun langkah-langkah yang dilakukan pada tahap ini adalah sebagai berikut:

1) Penyusunan Instrumen Tes

Butir soal yang telah dirumuskan pada tahap perancangan dituangkan ke dalam lembar instrumen tes TEL sesuai format yang telah ditetapkan. Setiap butir soal dilengkapi dengan petunjuk pengerjaan serta penentuan kunci jawaban. Penyusunan dilakukan dengan mempertimbangkan kaidah penulisan soal pilihan ganda yang baik, seperti kesesuaian materi, konstruksi, dan bahasa (Nitko & Brookhart, 2014).

2) Validasi Ahli

Instrumen yang telah disusun kemudian divalidasi oleh para ahli yang memiliki kompetensi di bidang pendidikan sains,

pengukuran dan evaluasi pendidikan, serta literasi teknologi. Validasi meliputi aspek isi, kesesuaian indikator, tingkat kesulitan, dan kejelasan bahasa. Hasil validasi dianalisis menggunakan koefisien Aiken's V untuk menentukan tingkat kesepakatan antar ahli terhadap kelayakan butir soal (Aiken, 1985).

3) Uji Coba Terbatas (Uji Keterbacaan)

Instrumen yang telah direvisi berdasarkan masukan ahli diuji cobakan pada kelompok kecil (uji keterbacaan) untuk memastikan kejelasan redaksi, kesesuaian konteks, dan waktu pengerjaan. Data hasil uji coba terbatas dianalisis untuk mengidentifikasi potensi masalah sebelum pengujian skala luas.

4) Revisi Instrumen

Berdasarkan hasil uji keterbacaan dan validasi ahli, butir soal yang tidak memenuhi kriteria perbaikan dihapus atau direvisi. Proses ini bertujuan menghasilkan instrumen final yang valid, reliabel, dan layak untuk diimplementasikan pada uji coba skala luas.

d. Tahap Penyebarluasan (*Disseminate*)

Tahap penyebaran bertujuan untuk mengimplementasikan instrumen *Technology and Engineering Literacy* (TEL) yang telah dikembangkan pada populasi sasaran secara luas, sehingga dapat mengukur profil TEL peserta didik secara komprehensif. Adapun Langkah-langkahnya sebagai berikut:

1) Penentuan Subjek Uji Coba Skala Luas

Instrumen final disebarkan kepada peserta didik kelas X yang telah mempelajari materi energi terbarukan, baik melalui pendekatan pembelajaran STEM maupun non-STEM. Pemilihan subjek dilakukan secara purposive untuk

memastikan keterwakilan kedua kelompok tersebut, sehingga memungkinkan perbandingan capaian TEL berdasarkan perbedaan pendekatan pembelajaran.

2) Pelaksanaan Tes

Uji coba skala luas dilakukan di beberapa SMA di Kota Bandung. Tes dilaksanakan pada jam pelajaran yang relevan, dengan pengawasan langsung dari guru mata pelajaran dan peneliti untuk memastikan kepatuhan terhadap prosedur pelaksanaan tes. Peserta didik diberikan penjelasan singkat mengenai tujuan penelitian, tata cara pengerjaan soal, serta ketentuan waktu pengerjaan.

3) Pengumpulan dan Pengolahan Data

Jawaban peserta dikumpulkan dan diinput ke dalam perangkat lunak analisis, dalam hal ini *Winsteps* untuk pemodelan Rasch, guna mengevaluasi validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta keberfungsian distraktor pada konteks implementasi luas.

4) Analisis Hasil Profil TEL

Data hasil tes dianalisis lebih lanjut untuk membandingkan profil TEL peserta didik yang belajar menggunakan pendekatan STEM dan non-STEM. Analisis ini mencakup perbedaan capaian berdasarkan aspek TEL, area TEL, dan indikator TEL sesuai kerangka *NAEP Technology and Engineering Literacy Framework*.

3.5 Teknik Analisis Data

Teknik analisis data merupakan prosedur sistematis yang digunakan untuk mengolah dan menginterpretasikan data penelitian sehingga dapat menjawab pertanyaan penelitian atau menguji hipotesis yang telah dirumuskan (Creswell, J.W., & Creswell, 2018). Pemilihan teknik analisis harus disesuaikan dengan tujuan penelitian, jenis data yang dikumpulkan, serta rancangan penelitian yang digunakan. Dalam penelitian pengembangan instrumen, analisis data tidak hanya berfokus pada pengujian hipotesis, tetapi juga pada evaluasi kualitas instrumen, meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan keberfungsian distraktor.

Pada penelitian ini, analisis data dilakukan melalui dua tahap utama. Pertama, hasil validasi ahli terhadap instrumen dianalisis menggunakan **koefisien Aiken's V** untuk menentukan tingkat kesepakatan antar validator terhadap relevansi, kejelasan, dan kesesuaian setiap butir soal. Kedua, analisis kuantitatif terhadap hasil uji coba instrumen dilakukan dengan menggunakan **Rasch Model** melalui perangkat lunak *Winsteps*. Pendekatan ini dipilih karena memiliki keunggulan dalam mengukur kualitas butir soal dan kemampuan responden secara simultan berdasarkan skala logit, serta memberikan informasi yang komprehensif terkait unidimensionalitas, *item fit*, reliabilitas butir dan responden, tingkat kesukaran, daya pembeda, dan analisis *Differential Item Functioning* (DIF). Selain itu, untuk membandingkan profil *Technology and Engineering Literacy* (TEL) antara kelompok peserta didik yang mengikuti pembelajaran berbasis STEM dan non-STEM, digunakan analisis komparatif seperti uji *t-test* atau *Mann-Whitney* sesuai dengan distribusi data. Dengan demikian, teknik analisis yang diterapkan mampu menghasilkan temuan yang valid, reliabel, dan dapat dipertanggungjawabkan secara ilmiah.

1. Uji Validitas

Validitas instrumen didefinisikan sebagai seberapa jauh pengukuran oleh instrumen dapat mengukur atribut yang seharusnya diukur (Sumintono & Widhiarso, 2014). Pada umumnya terdapat dua jenis

validitas yang sering digunakan yaitu validitas isi (*content validity*) dan validitas konstruk (*construct validity*).

a. Validitas Isi (*Content Validity*)

Untuk mengetahui valid tidaknya instrumen tes dianalisis dengan validitas isi (*content validity*). Untuk memvalidasi instrumen dalam penelitian ini dilakukan dengan meminta pertimbangan ahli/*judgment* kepada tiga orang dosen ahli dan dua orang guru fisika. Hal ini bertujuan untuk mengukur kesesuaian antara soal yang dibuat dengan indikator kemampuan berpikir kritis yang diukur. Hasil *judgment* dari lima orang ahli tersebut selanjutnya dihitung dan dianalisis menggunakan Aiken V untuk mendapatkan Gambaran kualitas butir soal.

Aiken (1985) merumuskan formula Aiken's V untuk menghitung *content-validity coefficient* yang didasarkan pada hasil penilaian dari panel ahli sebanyak n orang terhadap suatu item dari segi sejauh mana item tersebut mewakili konstruk yang diukur. Rumus validitas yang dikemukakan oleh Aiken yaitu sebagai berikut:

$$V = \sum S / [n (c - 1)] \quad (3.1)$$

(Aiken, 1985)

Keterangan:

$$S = r - l_o$$

r = angka yang diberikan penilai

l_o = angka penilaian validitas terendah

n = jumlah penilai

c = angka penilaian validitas tertinggi

Nilai dari koefisien V adalah berada antara 0 sampai 1 (Aiken, 1985). Analisis Aiken V menghitung kesepakatan proporsional berdasarkan distribusi skor. Aiken V memiliki

tabel kritis yang berisi nilai-nilai batas yang digunakan untuk menentukan apakah indeks V yang diperoleh dari perhitungan sudah cukup tinggi untuk menunjukkan bahwa item tersebut valid. Tabel tersebut adalah sebagai berikut:

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

Gambar 3.4 Batas Ambang Aiken V

Pada tabel, kolom kiri menunjukkan jumlah rater (n), sedangkan baris atas menunjukkan jumlah kategori skala penilaian (c). Nilai koefisien Aiken V dinyatakan valid jika memenuhi syarat nilai $V >$ ambang batas (Aiken, 1985). Misalnya skala Likert 1–5 berarti jumlah kategori adalah 5. Untuk jumlah penilai 5 maka dipilih baris dengan angka 5 pada kolom *Raters* (n). Lalu sesuaikan dengan jumlah kategori

penilaian yang dipakai. Contoh jika jumlah kategori 5 ($c=5$) dan jumlah penilai 5 maka nilai Aiken's V minimum yang dianggap signifikan = 0,80 dengan $p = 0,040$.

b. Validitas Konstruk (*Construct Validity*)

Item butir soal yang telah dinyatakan valid dari segi validitas isi kemudian di uji cobakan kepada peserta didik. Hasil uji coba kemudian dianalisis menggunakan pendekatan Rasch Model untuk mengetahui validitas instrumen dari segi konstruksinya. Uji validitas konstruk pada penelitian ini dianalisis melalui *item dimensionality*. Suatu instrumen dinyatakan memiliki konstruk yang memadai apabila butir-butir soal di dalamnya hanya mengukur satu dimensi utama (Sumintono & Widhiarso, 2014). Analisis validitas konstruk dilakukan dengan meninjau nilai *raw variance explained by measures* serta *unexplained variance in the first contrast*. Kriteria hasil uji dimensionalitas ditetapkan berdasarkan kriteria menurut Sumintono & Widhiarso (2014) sebagaimana tertera pada tabel berikut ini.

Tabel 3. 3 Kriteria Analisis Unidimensionalitas

Nilai	Interpretasi (%)
<i>Raw variance explained by measures</i>	$> 20\% - 40\%$ (diterima)
	$41\% - 59\%$ (lebih baik)
	$60\% - 100\%$ (istimewa)
<i>Unexplained variance 1st contrast</i>	$eigenvalue \leq 3$
	$observed \leq 15\%$

2. Uji Reliabilitas

Reliabilitas dapat diartikan sebagai taraf kepercayaan sebuah instrumen. Menurut Sumintono & Widhiarso (2014) reliabilitas menjelaskan seberapa jauh pengukuran yang dilakukan berkali-kali akan menghasilkan informasi yang sama. Uji reliabilitas dengan

pendekatan Rasch Model dianalisis melalui tiga indikator utama, yakni *person reliability*, *item reliability*, dan *Cronbach's Alpha*. *Person reliability* merefleksikan tingkat konsistensi atau kestabilan respons yang diberikan oleh peserta didik, sementara *item reliability* menggambarkan mutu serta konsistensi antar butir dalam instrumen. Adapun *Cronbach's Alpha* mencerminkan derajat keterkaitan antara peserta didik dengan keseluruhan butir soal yang diujikan. Adapun kriteria reliabilitas menurut Sumintono & Widhiarso (2014) adalah sebagaimana **Tabel 3.4** dan **Tabel 3.5** berikut ini.

Tabel 3.4 Interpretasi Nilai *Item Reliability* dan *Person Reliability*

Nilai	Interpretasi
$> 0,94$	Istimewa
$0,91 - 0,94$	Bagus Sekali
$0,81 - 0,90$	Bagus
$0,67 - 0,80$	Cukup
$< 0,67$	Lemah

Tabel 3.5 Interpretasi Nilai Cronbach Alpha

Nilai	Interpretasi
$0,80 < \alpha$	Bagus Sekali
$0,70 \leq \alpha \leq 0,80$	Bagus
$0,60 \leq \alpha < 0,70$	Cukup
$0,50 \leq \alpha < 0,60$	Jelek
$\alpha < 0,50$	Jelek Sekali

3. Tingkat Kesukaran

Tingkat kesukaran atau *index difficulty* merupakan bilangan yang menunjukkan sukar atau mudahnya suatu soal. Menurut Arikunto, (2015) tingkat kesukaran menunjukkan sukar atau mudahnya suatu soal dalam mengukur kemampuan peserta didik dan ditunjukan dalam indeks kesukaran. Soal yang baik adalah soal yang tidak terlalu mudah ataupun terlalu sukar. Besarnya nilai tingkat kesukaran berkisar antara 0 sampai dengan 1.

Tingkat kesukaran dalam *Rasch Model* diukur melalui parameter *item difficulty* yang dinyatakan dalam satuan logit. Distribusi tingkat kesukaran yang seimbang, yaitu adanya butir mudah, sedang, dan sulit, diperlukan agar instrumen dapat membedakan kemampuan peserta didik pada seluruh spektrum keterampilan yang diukur. Dalam *Rasch Model*, tingkat kesukaran butir (*item difficulty*) dihitung dalam skala logit berdasarkan persamaan probabilitas jawaban benar. Pengelompokan tingkat kesukaran butir soal dalam penelitian ini mengacu pada kriteria yang ditetapkan oleh Sumintono & Widhiarso (2014) yaitu sebagaimana tertera pada tabel berikut ini.

Tabel 3.6 Interpretasi Tingkat Kesukaran

Syarat Nilai	Tingkat Kesukaran
$1SD < Measure$	Sangat Sukar
$0 < Measure \leq 1SD$	Sukar
$-1SD \leq Measure \leq 0$	Mudah
$Measure < -1SD$	Sangat Mudah

4. Distraktor

Distraktor atau pengecoh adalah pilihan jawaban salah pada soal pilihan ganda yang dirancang untuk mengalihkan perhatian peserta dengan kemampuan rendah. Dalam *Rasch Model*, analisis keberfungsian distraktor dilakukan dengan melihat distribusi pemilihan opsi oleh peserta dengan berbagai tingkat kemampuan. Distraktor yang efektif seharusnya lebih banyak dipilih oleh peserta didik berkemampuan rendah, sementara peserta didik berkemampuan tinggi cenderung memilih jawaban benar. Distraktor yang jarang dipilih atau justru dipilih oleh peserta didik berkemampuan tinggi dinilai tidak berfungsi dengan baik dan memerlukan revisi (Hambleton R.K., Swaminathan H., 1991).

Analisis distraktor dilakukan dengan menghitung proporsi peserta pada setiap kategori kemampuan yang memilih opsi tertentu. Analisis distraktor ditinjau berdasarkan persentase peserta didik yang

menjawab opsi. Mengacu pada Arikunto (2016) distraktor pada soal pilihan ganda dikatakan berfungsi efektif apabila dipilih oleh minimal 5% peserta tes.

5. *Different Item Functioning (DIF)*

Analisis DIF bertujuan untuk mengidentifikasi potensi bias suatu butir terhadap kelompok responden dengan karakteristik berbeda namun kemampuan setara. Deteksi bias diperlukan untuk mengetahui apakah item-item yang diberikan mempunyai bias dalam kategori responden tertentu atau tidak (Sumintono & Widhiarso, 2014). Dalam penelitian ini, uji DIF dilakukan berdasarkan variabel gender dan demografi asal sekolah (kota atau desa). Suatu butir dikategorikan bias apabila perbedaan tingkat kesulitannya signifikan pada kelompok yang dibandingkan ($p\text{-value} < 0,05$). Temuan ini penting untuk memastikan instrumen bersifat adil dan tidak menguntungkan atau merugikan kelompok tertentu. Interpretasi hasil DIF menurut Sumintono & Widhiarso (2014) adalah sebagai berikut:

Tabel 3. 7 Interpretasi DIF

Nilai	Kriteria
$p\text{ value} > 0,05$	Tidak ada Bias
$p\text{ value} \leq 0,05$	Ada Bias

6. Variabel (Wright) Map

Variabel Wright Map adalah representasi visual yang menempatkan distribusi kemampuan peserta didik dan tingkat kesulitan butir soal pada satu skala logit yang sama. Pada peta ini, kemampuan peserta didik biasanya ditampilkan di sisi kiri, sedangkan tingkat kesulitan item ditampilkan di sisi kanan. Wright Map memudahkan peneliti untuk mengevaluasi kesesuaian antara kemampuan responden dan tingkat kesulitan soal. Distribusi yang baik akan menunjukkan adanya ketercocokan antara rentang kemampuan peserta didik dan variasi tingkat kesulitan butir, sehingga instrumen dapat mengukur kemampuan seluruh peserta secara optimal (Wright., 1987).