

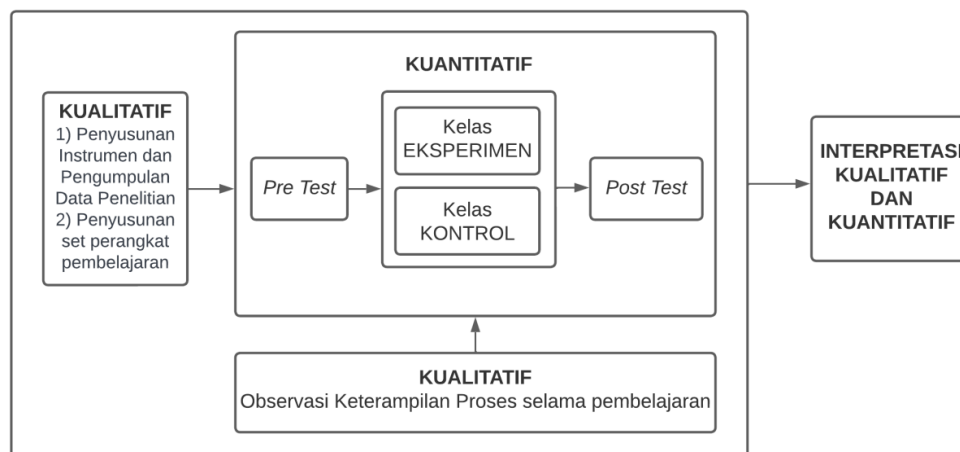
BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Desain yang diaplikasikan untuk penelitian yaitu desain penelitian metode campuran (*mix methods research design*) yang menggabungkan pendekatan kualitatif dan kuantitatif, untuk menjawab pertanyaan penelitian. Desain ini digunakan karena kemampuannya untuk lebih memberikan pemahaman jawaban atas pertanyaan penelitian, memperoleh perspektif alternatif, dan memperoleh informasi yang lebih spesifik dan rinci lebih dari yang bisa diperoleh melalui hasil analisis statistik (Cresswell, 2015). Desain penelitian campuran yang digunakan yaitu desain *embedded*. Desain *embedded* merupakan salah satu metode campuran dimana satu metode penelitian akan ditanamkan pada metode penelitian lainnya (Cresswell dan Clark, 2018). Dalam desain *embedded*, satu metode penelitian dijadikan sebagai penelitian utama (*primary role*), sedangkan metode lainnya dijadikan sebagai model pendukung (*secondary role*) yang bertujuan agar kekurangan data yang didapatkan dari metode penelitian utama dapat digali lebih dalam pada metode penelitian pendukung sehingga didapatkan hasil penelitian yang tepat mampu menjawab pertanyaan penelitian, apalagi jika peneliti tidak memiliki cukup waktu dalam pengumpulan data (Iskandar, dkk., 2021).

Pada penelitian ini digunakan desain *embedded* seperti terlihat dalam Gambar 3.1.

Gambar 3.1 *Embedded Design*

Fokus dari struktur desain ini ialah dalam menghimpun data berdimensi kuantitatif, sementara data dimensi kualitatif bertindak sebagai komplementer bagi data kuantitatif. Dalam rangka penelitian kuantitatif ini, strategi metode eksperimen semu (Quasi Experiment) digunakan, dan dengan merancang model non-equivalent control group design (Cresswell, 2015), dimana penelitian melibatkan tiga kelompok, yakni satu kelompok eksperimen dan dua kelompok kontrol. Desain penelitian kuantitatif *Non-equivalent control group design* ditunjukkan pada Tabel 3.1.

Tabel 3.1
Non-equivalent control group design

Kelas	Tes awal	Treatment/Intervensi	Tes akhir
Eksperimen	O ₁	X ₁	O ₂
Kontrol I	O ₃	X ₂	O ₄
Kontrol II	O ₅	X ₃	O ₆

Keterangan :

- O₁ = Tes awal kelas eksperimen
- O₂ = Tes akhir kelas eksperimen
- O₃ = Tes awal kelas kontrol I
- O₄ = Tes akhir kelas kontrol I
- O₅ = Tes awal kelas kontrol II
- O₆ = Tes akhir kelas kontrol II
- X₁ = Model *Scientific Group Inquiry Learning* (SGIL) berorientasi

Profil Pelajar Pancasila

X₂ = Model *Scientific Group Inquiry Learning* (SGIL)

X₃ = Model *Direct Instruction*

Pengumpulan data kualitatif ditanamkan selama intervensi. Pada tahap awal penulis akan memberikan tes awal keterampilan proses untuk mengidentifikasi nilai keterampilan proses peserta didik sebelum intervensi. Kemudian diberikan intervensi berupa model SGIL berorientasi Profil Pelajar Pancasila pada kelas eksperimen, serta model SGIL dan *Direct Instruction* pada kelas kontrol. Selama intervensi, observer akan melakukan observasi keterampilan proses kepada peserta didik berpanduan rubrik. Kemudian dilanjutkan dengan pemberian tes akhir keterampilan proses untuk mengidentifikasi ketercapaian keterampilan proses peserta didik, kemudian keseluruhan data baik kuantitatif maupun kualitatif akan digunakan untuk menginterpretasi ketercapaian pembelajaran keterampilan proses.

3.2 Populasi dan Sampel

Populasi dan sampel penelitian ini adalah seluruh peserta didik kelas X di salah satu sekolah negeri Kabupaten Indragiri Hilir tahun pelajaran 2022/2023, dimana terdapat empat rombongan belajar (rombel) untuk kelas X tersebut. Sampel pada penelitian ini dipilih menggunakan sampling random sederhana (*simple random sampling*). Sampling random sederhana ini dilaksanakan secara acak tanpa mengindahkan strata/tingkatan dalam populasi itu, serta sifat homogen dari populasi tersebut (Sundayana, 2020).

3.3 Instrumen dan Pengumpulan Data Penelitian

a. Tes Keterampilan Proses Sains

Instrumen yang digunakan dalam penelitian ini adalah Tes Keterampilan Proses, berbentuk essay (*essay item*), dengan rincian Aspek dan Jumlah soal Instrumen Tes KPS pada Tabel 3.2.

Tabel 3.2
Aspek dan Jumlah Soal Instrumen Tes KPS

Aspek Keterampilan Proses	Nomor Soal	Jumlah Item Soal
1. Mengamati	1	1
2. Mempertanyakan dan memprediksi	2,3,4	3
3. Merencanakan dan melakukan penyelidikan	5,6,7	3
4. Memproses dan menganalisis data dan informasi	8,9	2
5. Mencipta	10	1
6. Mengevaluasi dan merefleksi	11,12,13	3
7. Mengomunikasikan hasil	14,15	2
Total item soal		15

Tes Keterampilan Proses Sains yang dikembangkan oleh Nisfullail dkk. (2023) tersebut dipergunakan dalam penelitian Pembelajaran *Scientific Group Inquiry Learning* (SGIL) Berorientasi Profil Pelajar Pancasila untuk Meningkatkan Keterampilan Proses pada Materi Energi terbarukan yang telah dilaksanakan.

b. Lembar Observasi Keterampilan Proses dan Lembar Keterlaksanaan

Sebagai alat pengumpul data kualitatif dalam penelitian digunakan instrumen non tes berupa Lembar Observasi Keterampilan Proses dengan rubrik berpanduan pada buku Panduan Pembelajaran dan Asesmen Kurikulum Merdeka 2022. Interpretasi secara kualitatif dilakukan dengan panduan rubrik.

Lembar observasi tersebut akan digunakan selama proses pembelajaran berlangsung, dimana observasi dilakukan per individu dalam kelompok. Guna mengatasi kelemahan hasil observasi oleh pengamat, maka observasi akan dilakukan oleh empat orang pengamat. Selain itu, selama pembelajaran juga menggunakan lembar keterlaksanaan pembelajaran pada guru dan peserta didik. Lembar keterlaksanaan ini berupa daftar cek keterlaksanaan pembelajaran.

c. Lembar Validasi

Instrumen tes dan rubrik keterampilan proses yang digunakan dalam penelitian, sebelum diberikan kepada peserta didik dalam pembelajaran kelas eksperimen maupun kontrol, maka akan divalidasi terlebih dahulu. Validasi

akan dilakukan oleh tiga orang dosen Program Magister Pendidikan Fisika dan satu orang widyaiswara (fasilitator sekolah penggerak, magister pendidikan fisika), serta lima orang guru fisika Sekolah Menengah Atas, sehingga untuk keperluan validasi tersebut, maka telah disusun lembar validasi ahli yang digunakan untuk memvalidasi Instrumen Tes dan Rubrik Keterampilan Proses.

d. Analisis Instrumen Tes

Sebelum penerapannya, instrumen telah melalui tahap pengujian validitas dan reliabilitas. Validitas konten dan reliabilitas Instrumen Tes Keterampilan Proses telah diuji oleh penulis, memastikan keabsahan dan keandalan instrumen tes tersebut. Dengan demikian, instrumen ini telah terbukti valid dan reliabel, dan siap digunakan dalam kerangka penelitian selanjutnya.

d.1 Validitas Konten

Proses validasi konten Tes Keterampilan Proses dan Rubrik Penilaian Keterampilan Proses melibatkan penilaian oleh sejumlah ahli (*expert judgement*), yaitu tiga dosen Program Magister Pendidikan Fisika dan seorang widyaiswara (fasilitator sekolah penggerak, dengan latar belakang magister pendidikan fisika), serta lima guru fisika Sekolah Menengah Atas. Tujuan penilaian ini adalah untuk menilai sejauh mana kesesuaian antara soal-soal dalam tes dengan indikator-indikator keterampilan proses yang diukur, dan untuk mendapatkan masukan berupa kritik serta saran terhadap instrumen tes dan lembar observasi keterampilan proses yang telah dikembangkan. Keputusan mengenai kecocokan instrumen tes dan non tes akan dibuat oleh para ahli tersebut, yang kemudian memutuskan apakah instrumen tersebut dapat langsung digunakan tanpa perlu perubahan, membutuhkan perbaikan tertentu, atau mungkin memerlukan revisi menyeluruh (Widoyoko, 2020).

Hasil dari proses penilaian oleh ahli kemudian dianalisis menggunakan rumus Content Validity Ratio (CVR) yang dikembangkan oleh Lawshe. CVR merupakan pendekatan validasi konten yang digunakan untuk mengevaluasi kesesuaian item soal tes dengan komponen yang diukur, berdasarkan penilaian

dari para ahli/*expert judgement* (Wilson, dkk., 2012). Rumusan CVR tersebut sebagai berikut :

$$CVR = \frac{n_e^{-\frac{N}{2}}}{\frac{N}{2}} \dots\dots\dots(1)$$

Keterangan :

CVR = *Content Validity Ratio* (CVR)

n_e = Jumlah ahli yang menyatakan “layak digunakan tanpa revisi atau layak digunakan dengan revisi sesuai saran”

N = Total ahli yang menilai validitas konten

Instrumen dapat dinyatakan valid apabila nilai CVR hitung yang diperoleh lebih besar dari nilai CVR kritis (Liliawati, dkk., 2020). Panduan nilai CVR kritis berdasarkan tabel Schipper nilai CVR minimal untuk variasi jumlah ahli seperti pada Tabel 3.3.

Tabel 3.3
Nilai CVR minimal untuk variasi jumlah ahli

Jumlah Ahli	Nilai CVR Minimal	Jumlah Ahli	Nilai CVR Minimal
5	0,736	13	0,456
6	0,672	14	0,440
7	0,622	15	0,425
8	0,582	16	0,368
9	0,548	17	0,329
10	0,520	18	0,300
11	0,496	19	0,287
12	0,475	20	0,260

(Wilson, dkk., 2012)

Hasilnya kemudian diinterpretasi untuk mengetahui kategori validitasnya dengan kategori skor CVR pada Tabel 3.4.

Tabel 3.4
Kategori Skor CVR (Liliawati, dkk., 2020)

Rentang Skor CVR	Kategori
0,68 – 1,00	Sangat sesuai
0,34 – 0,67	Sesuai
0,00 – 0,33	Tidak sesuai

Berikut ini disajikan Rekapitulasi Hasil Validasi Konten oleh Validator Ahli pada Tabel 3.5.

Tabel 3.5
Rekapitulasi Hasil Validasi Konten oleh Validator Ahli

Kode Butir Soal	Kode Aspek Penilaian	N	n_e	CVR	Kategori	Kode Butir Soal	Kode Aspek Penilaian	N	n_e	CVR	Kategori
E1	A	7	7	1	Sangat Sesuai	E9	A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai
	D	7	7	1	Sangat Sesuai		D	7	7	1	Sangat Sesuai
E2	A	7	7	1	Sangat Sesuai	E10	A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai
	D	7	7	1	Sangat Sesuai		D	7	7	1	Sangat Sesuai
E3	A	7	7	1	Sangat Sesuai	E11	A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai
	D	7	7	1	Sangat Sesuai		D	7	7	1	Sangat Sesuai
E4	A	7	7	1	Sangat Sesuai	E12	A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai
	D	7	7	1	Sangat Sesuai		D	7	7	1	Sangat Sesuai
E5	A	7	7	1	Sangat Sesuai	E13	A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai
	D	7	7	1	Sangat Sesuai		D	7	7	1	Sangat Sesuai
E6	A	7	7	1	Sangat Sesuai	E14	A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai

Kode Butir Soal	Kode Aspek Penilaian	N	n_e	CVR	Kategori	Kode Butir Soal	Kode Aspek Penilaian	N	n_e	CVR	Kategori
E7	D	7	7	1	Sangat Sesuai	E15	D	7	7	1	Sangat Sesuai
	A	7	7	1	Sangat Sesuai		A	7	7	1	Sangat Sesuai
	B	7	7	1	Sangat Sesuai		B	7	7	1	Sangat Sesuai
	C	7	7	1	Sangat Sesuai		C	7	7	1	Sangat Sesuai
E8	D	7	7	1	Sangat Sesuai		D	7	7	1	Sangat Sesuai
	A	7	7	1	Sangat Sesuai						
	B	7	7	1	Sangat Sesuai						
	C	7	7	1	Sangat Sesuai						
	D	7	7	1	Sangat Sesuai						

Keterangan :

CVR = *Content Validity Ratio* (CVR)

n_e = Kumulasi ahli yang *menjudge* “layak digunakan tanpa revisi atau layak digunakan dengan revisi sesuai saran”

N = Kumulasi ahli yang menilai validitas konten

A = Aspek materi

B = Aspek konstruksi

C = Aspek bahasa

D = Aspek rubrik penilaian

Tabel 3.5 menunjukkan bahwa rata-rata jumlah ahli yang menyatakan bahwa instrumen tes keterampilan proses sains layak digunakan tanpa revisi atau layak digunakan dengan revisi sesuai saran sebanyak tujuh ahli dari total tujuh ahli yang menilai validitas konten. Selanjutnya dihitung skor CVR, dan rata-rata per item soal diperoleh skor CVR sebesar 1. Hasil ini kemudian diinterpretasikan untuk mengetahui validitas konten berdasarkan Tabel 3.4 kategori skor CVR (Liliawati, dkk., 2020) sehingga disimpulkan bahwa validitas konten instrumen tes keterampilan proses berikut rubrik penilaiannya termasuk dalam kategori *sangat sesuai*. Namun demikian, masih terdapat beberapa saran perbaikan rubrik agar dapat lebih spesifik mengakomodasi

semua jawaban peserta didik, tidak hanya bersifat umum, khususnya pada rubrik soal nomor 5, 10, 12, 14, dan 15.

d.2 Uji Coba Instrumen

Setelah melalui proses validasi oleh para ahli, peneliti melakukan penyesuaian pada instrumen Tes Keterampilan Proses sesuai dengan masukan dan saran yang diberikan. Selanjutnya, peneliti mengimplementasikan uji coba instrumen tersebut kepada siswa kelas XI yang telah mengikuti pembelajaran tentang Energi Terbarukan. Data yang dihasilkan dari uji coba ini akan dianalisis menggunakan pendekatan pemodelan Rasch melalui perangkat lunak Winstep.

d.2.1 Validitas Item

Dalam pandangan Sumintono dan Widhiarso (2015), validitas suatu item dapat dievaluasi melalui analisis kesesuaian item (item fit), yang mengacu pada sejauh mana butir pertanyaan sesuai dengan model yang digunakan. Suatu item dianggap memiliki kesesuaian item (item fit) jika terpenuhi parameter kesesuaian butir (item fit) pada Tabel 3.6.

Tabel 3.6
Parameter Kesesuaian Butir (Item Fit)

Parameter	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)	$0,4 < \text{Pt Measure Corr} < 0,85$

Suatu butir soal memiliki relevansi sangat baik jika skor MNSQ, ZSTD dan Measured Correlation yang diperoleh berada dalam rentang skor yang dapat diterima. Item masih dapat diterima (tidak dimodifikasi) jika memenuhi setidaknya salah satu dari parameter ini. Namun, elemen item dianggap tidak sesuai jika tidak memenuhi ketiga parameter tersebut, sehingga kategorinya harus diubah atau diganti (Sumintono dan Widhiarso, 2015). Selain dapat mempertimbangkan item yang tidak sesuai, model Rasch juga dapat mempertimbangkan responden yang tidak sungguh-sungguh dalam melakukan

tes. Setelah soal diuji coba kepada satu kelas berjumlah 20 orang peserta didik, yang meliputi tiga orang laki-laki dan tujuh belas orang perempuan, maka jawaban siswa dianalisis menggunakan Rasch model, dan hasil Tampilan hasil Item Statistics : Misfit Order ditunjukkan pada Gambar 3.2.

TABLE 10.1 C:\Users\User\Desktop\HASIL UJI VALID ZOU015WS.TXTA May 9 2023 17:30
INPUT: 20 Person 15 Item REPORTED: 20 Person 15 Item 4 CATS WINSTEPS 5.4.3.0

Person: REAL SEP.: 1.94 REL.: .79 ... Item: REAL SEP.: 4.51 REL.: .95

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
11	53	20	-2.29	.40	2.39	3.58	2.32	3.40	A .55	.56	40.0	58.9	E11
5	35	20	.91	.43	1.96	2.92	2.02	3.07	B .47	.55	25.0	61.6	E5
8	66	20	-4.39	.41	1.43	1.34	1.30	.80	C .54	.44	50.0	64.1	E8
15	31	20	1.66	.44	1.34	1.35	1.34	1.38	D .65	.55	65.0	62.6	E15
1	57	20	-2.92	.40	.95	-.07	1.20	.69	E .16	.54	75.0	61.2	E1
9	34	20	1.09	.43	1.07	.34	1.08	.41	F .54	.55	40.0	61.2	E9
3	49	20	-1.63	.41	.85	-.41	.82	-.52	G .55	.58	70.0	62.8	E3
7	24	20	3.50	.64	.84	-.19	.74	-.26	H .79	.57	90.0	86.0	E7
2	51	20	-1.96	.40	.75	-.82	.80	-.64	I .43	.57	60.0	60.6	E2
10	24	20	3.50	.64	.72	-.47	.44	-.97	J .88	.57	90.0	86.0	E10
14	41	20	-.19	.43	.69	-.88	.63	-1.09	K -.03	.57	85.0	70.7	E14
4	44	20	-.75	.43	.45	-1.89	.40	-2.08	L .72	.58	90.0	69.9	E4
6	44	20	-.75	.43	.35	-2.39	.35	-2.33	M .77	.58	90.0	69.9	E6
12	22	20	4.63	.89	.34	-1.05	.12	-.98	N .87	.55	95.0	92.6	E12
13	42	20	-.38	.43	.10	-4.41	.10	-4.42	O .87	.58	100.0	71.1	E13
MEAN	41.1	20.0	.00	.48	.95	-.20	.91	-.24			71.0	69.3	
P.SD	12.5	.0	2.47	.13	.60	1.94	.63	1.94			22.7	10.3	

Gambar 3.2 Tampilan hasil Item Statistics : Misfit Order

Hasil Item Statistics : Misfit Order di atas dapat diringkas menjadi ringkasan hasil *item statistics* : misfit pada Tabel 3.7.

Tabel 3.7

Ringkasan Hasil Item Statistics : Misfit Order

Kode Butir Soal	Nilai <i>Outfit Mean Square</i> (MNSQ)	Nilai <i>Outfit Z-Standard</i> (ZSTD)	Nilai <i>Point Measure Correlation</i> (Pt Measure Corr)	Kesimpulan Hasil Validitas Item Soal
E1	1,20	0,69	0,16	Valid
E2	0,80	-0,64	0,43	Valid
E3	0,82	-0,52	0,55	Valid
E4	0,40	-2,08	0,72	Valid
E5	2,02	3,07	0,47	Valid
E6	0,35	-2,33	0,77	Valid
E7	0,74	-0,26	0,79	Valid
E8	1,30	0,80	0,54	Valid
E9	1,08	0,41	0,54	Valid

Kode Butir Soal	Nilai <i>Outfit Mean Square</i> (MNSQ)	Nilai <i>Outfit Z-Standard</i> (ZSTD)	Nilai <i>Point Measure Correlation</i> (Pt Measure Corr)	Kesimpulan Hasil Validitas Item Soal
E10	0,44	-0,97	0,88	Valid
E11	2,32	3,40	0,55	Valid
E12	0,12	-0,98	0,87	Valid
E13	0,10	-4,42	0,87	Valid
E14	0,63	-1,09	-0,03	Valid
E15	1,34	1,38	0,65	Valid

Soal kode E1 nilai *Outfit Mean Square* sebesar 1,20 dan nilai *Outfit Z-Standard* sebesar 0,69, dikategorikan nilainya dapat diterima, namun nilai *Point Measure Correlation* sebesar 0,16 tidak dapat diterima, sehingga disimpulkan butir soal 1 (E1) memiliki kesesuaian butir soal yang baik dan soal masih bisa diterima (tidak diubah) karena memenuhi dua dari parameter (*Outfit Mean Square* dan nilai *Outfit Z-Standard*). Soal kode E5 nilai *Outfit Mean Square* sebesar 2,02 dan nilai *Outfit Z-Standard* sebesar 3,07, dikategorikan nilainya tidak dapat diterima, namun nilai *Point Measure Correlation* sebesar 0,16 dapat diterima, sehingga disimpulkan butir soal 5 (E5) memiliki kesesuaian butir soal yang baik dan soal masih bisa diterima (tidak diubah) karena memenuhi satu dari parameter (*Point Measure Correlation*). Soal kode E6 nilai *Outfit Mean Square* sebesar 0,35 dan nilai *Point Measure Correlation* sebesar 0,77, dikategorikan nilainya dapat diterima, namun nilai *Outfit Z-Standard* sebesar -2,33 tidak dapat diterima, sehingga disimpulkan butir soal 6 (E6) memiliki kesesuaian butir soal yang baik dan soal masih bisa diterima (tidak diubah) karena memenuhi dua dari parameter (*Outfit Mean Square* dan nilai *Point Measure Correlation*).

Soal kode E11 nilai *Outfit Mean Square* sebesar 2,32 dan nilai *Outfit Z-Standard* sebesar 3,40, dikategorikan nilainya tidak dapat diterima, namun nilai *Point Measure Correlation* sebesar 0,55 dapat diterima, sehingga disimpulkan butir soal 11 (E11) memiliki kesesuaian butir soal yang baik dan soal masih bisa diterima (tidak diubah) karena memenuhi satu dari parameter (*Point Measure Correlation*). Soal kode E13 nilai *Outfit Mean Square* sebesar 0,10

dan nilai *Point Measure Correlation* sebesar 0,87, dikategorikan nilainya dapat diterima, namun nilai *Outfit Z-Standard* sebesar -4,42 tidak dapat diterima, sehingga disimpulkan butir soal 6 (E6) memiliki kesesuaian butir soal yang baik dan soal masih bisa diterima (tidak diubah) karena memenuhi dua dari parameter (*Outfit Mean Square* dan nilai *Point Measure Correlation*). Soal kode E14 nilai *Outfit Mean Square* sebesar 0,63 dan nilai *Outfit Z-Standard* sebesar -1,09, dikategorikan nilainya dapat diterima, namun nilai *Point Measure Correlation* sebesar -0,03 tidak dapat diterima, sehingga disimpulkan butir soal 14 (E14) memiliki kesesuaian butir soal yang baik dan soal masih bisa diterima (tidak diubah) karena memenuhi dua dari parameter (*Outfit Mean Square* dan nilai *Outfit Z-Standard*).

Berdasarkan Tabel 3.7, dapat disimpulkan bahwa semua butir soal instrumen tes keterampilan proses yang diberikan bisa dipahami dengan baik oleh para peserta didik, dan tidak ada soal yang miskonsepsi.

d.2.2 Reliabilitas

Reliabilitas mengindikasikan sejauh mana suatu pengukuran dapat diulang dengan hasil yang konsisten dalam percobaan berulang (Sumintono dan Widhiarso, 2015). Arikunto (2020) mengungkapkan bahwa reliabilitas berkaitan dengan tingkat kepercayaan, sehingga sebuah tes dianggap memiliki tingkat kepercayaan yang tinggi ketika hasil yang diperoleh tetap atau meskipun terdapat variasi, perubahan tersebut dianggap tidak signifikan.

Dalam rangka menganalisis reliabilitas dalam penelitian ini, pendekatan pemodelan Rasch dengan perangkat lunak Winstep akan digunakan. Analisis ini akan melibatkan evaluasi terhadap nilai Alpha Cronbach yang mencerminkan kualitas interaksi antara responden dan item dalam instrumen. Informasi mengenai reliabilitas dan nilai Alpha Cronbach akan tercatat dalam *Output Tables: Summary Statistics*. Penafsiran atas nilai reliabilitas dan nilai Alpha Cronbach, sesuai dengan pandangan Sumintono & Widhiarso (2015), interpretasi nilai *person reliability* dan *item reliability* dalam analisis model Rasch pada Tabel 3.8 dan kategori nilai alpha cronbach pada Tabel 3.9.

Tabel 3.8
Interpretasi Nilai *Person Reliability* dan *Item Reliability* dalam Analisis Model Rasch

Koefisien Reliabilitas	Kategori
Nilai > 0,94	Istimewa
$0,91 \leq \text{Nilai} \leq 0,94$	Sangat Baik
$0,81 \leq \text{Nilai} \leq 0,90$	Baik
$0,67 \leq \text{Nilai} \leq 0,80$	Cukup
Nilai < 0,67	Lemah

Tabel 3.9
Kategori Nilai Alpha Cronbach

Koefisien Reliabilitas Item	Kategori
$\alpha \geq 0,8$	Sangat baik
$0,7 \leq \alpha < 0,8$	Baik
$0,6 \leq \alpha < 0,7$	Cukup
$0,5 \leq \alpha < 0,6$	Kurang
$\alpha < 0,5$	Sangat kurang

Melalui aplikasi winstep didapatkan keluaran berupa tampilan hasil *Output Tables: Summary Statistics* pada Gambar 3.3.

TABLE 3.1 C:\Users\User\Desktop\HASIL UJI VALIDI ZOU015WS.TXTB May 9 2023 17:30
INPUT: 20 Person 15 Item REPORTED: 20 Person 15 Item 4 CATS WINSTEPS 5.4.3.0

SUMMARY OF 20 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	30.9	15.0	-1.63	.52	1.01	.00	.91	-.10
SEM	1.1	.0	.28	.00	.11	.25	.11	.18
P.SD	4.9	.0	1.23	.02	.48	1.09	.46	.81
S.SD	5.1	.0	1.26	.02	.49	1.12	.47	.83
MAX.	47.0	15.0	2.31	.55	2.62	3.04	2.20	1.73
MIN.	26.0	15.0	-2.94	.48	.39	-1.95	.33	-1.39

REAL RMSE	.56	TRUE SD	1.09	SEPARATION	1.94	Person RELIABILITY	.79
MODEL RMSE	.52	TRUE SD	1.11	SEPARATION	2.14	Person RELIABILITY	.82
S.E. OF Person MEAN = .28							

Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .83 SEM = 2.00
STANDARDIZED (50 ITEM) RELIABILITY = .94

SUMMARY OF 15 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	41.1	20.0	.00	.48	.95	-.20	.91	-.24
SEM	3.3	.0	.66	.04	.16	.52	.17	.52
P.SD	12.5	.0	2.47	.13	.60	1.94	.63	1.94
S.SD	12.9	.0	2.56	.14	.62	2.00	.65	2.01
MAX.	66.0	20.0	4.63	.89	2.39	3.58	2.32	3.40
MIN.	22.0	20.0	-4.39	.40	.10	-4.41	.10	-4.42

REAL RMSE	.54	TRUE SD	2.41	SEPARATION	4.51	Item RELIABILITY	.95
MODEL RMSE	.50	TRUE SD	2.42	SEPARATION	4.84	Item RELIABILITY	.96
S.E. OF Item MEAN = .66							

Item RAW SCORE-TO-MEASURE CORRELATION = -.99
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

Gambar 3.3 Tampilan hasil *Output Tables: Summary Statistics*

Berdasarkan hasil analisis *Summary Statistics*, ditemukan bahwa nilai reliabilitas individu (*person reliability*) sebesar 0,79, menunjukkan bahwa konsistensi jawaban peserta didik dapat dianggap cukup. Sementara itu, nilai reliabilitas item (*item reliability*) mencapai 0,95, mengindikasikan mutu butir-butir soal instrumen memiliki kategori sangat baik. Oleh karena itu, dapat diambil kesimpulan bahwa jawaban peserta didik memiliki tingkat konsistensi yang memadai, dan butir-soal dalam aspek reliabilitas instrumen memiliki kualitas yang sangat baik. Selanjutnya, nilai alpha Cronbach sebesar 0,83, yang terkategori sebagai sangat baik, memberikan indikasi bahwa reliabilitas instrumen tes keterampilan proses telah mencapai tingkat yang sangat memuaskan.

d.2.3 Tingkat Kesukaran

Tahapan berikutnya melibatkan penentuan Tingkat Kesukaran. Sebuah pertanyaan yang baik adalah pertanyaan yang memiliki tingkat kesukaran yang seimbang, tidak terlalu sederhana atau terlalu kompleks (Arikunto, 2020, hal.

232). Dalam konteks penelitian ini, Tingkat Kesukaran butir soal diidentifikasi menggunakan pemodelan Rasch melalui software Winstep. Evaluasi Tingkat Kesukaran dari setiap butir soal dapat dilakukan melalui Output Tables: Item Measure. Dalam kolom Entry Number, butir-butir soal diurutkan berdasarkan skor logit, dari yang tertinggi hingga yang terendah.

Kemudian, berdasarkan pandangan Sumintono dan Widhiarso (2015), skor logit dalam kolom JMLE Measure merefleksikan derajat kesulitan suatu soal. Interpretasi umum adalah semakin tinggi skor logit, semakin sulit soal tersebut, dan sebaliknya. Dalam kolom item measure, informasi tentang standar deviasi juga disajikan. Berdasarkan nilai standar deviasi dan rata-rata skor logit (kolom JMLE measure), tingkat kesulitan butir-butir soal dapat dikategorikan. Nilai logit +1 SD mengindikasikan kategori kesulitan "sangat sulit", lebih dari +1 SD berarti "sangat sulit", nilai logit -1 SD mengartikan "mudah", dan kurang dari -1 SD diartikan "sangat mudah" (Sumintono dan Widhiarso, 2015). Temuan ini menunjukkan empat kategori tingkat kesulitan: sangat sulit, sulit, mudah, dan sangat mudah, pada instrumen Tes Keterampilan Proses yang digunakan.

Hasil dari analisis dengan menggunakan aplikasi Winstep menghasilkan informasi berupa tampilan tabel *Item Statistics: Measure Order* pada Gambar 3.4.

TABLE 13.1 C:\Users\User\Desktop\HASIL UJI VALID ZOU015WS.TXTA May 9 2023 17:30
 INPUT: 20 Person 15 Item REPORTED: 20 Person 15 Item 4 CATS WINSTEPS 5.4.3.0

Person: REAL SEP.: 1.94 REL.: .79 ... Item: REAL SEP.: 4.51 REL.: .95

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S. E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
12	22	20	4.63	.89	.34	-1.05	.12	-.98	.87	.55	95.0	92.6	E12
7	24	20	3.50	.64	.84	-.19	.74	-.26	.79	.57	90.0	86.0	E7
10	24	20	3.50	.64	.72	-.47	.44	-.97	.88	.57	90.0	86.0	E10
15	31	20	1.66	.44	1.34	1.35	1.34	1.38	.65	.55	65.0	62.6	E15
9	34	20	1.09	.43	1.07	.34	1.08	.41	.54	.55	40.0	61.2	E9
5	35	20	.91	.43	1.96	2.92	2.02	3.07	.47	.55	25.0	61.6	E5
14	41	20	-.19	.43	.69	-.88	.63	-1.09	-.03	.57	85.0	70.7	E14
13	42	20	-.38	.43	.10	-4.41	.10	-4.42	.87	.58	100.0	71.1	E13
4	44	20	-.75	.43	.45	-1.89	.40	-2.08	.72	.58	90.0	69.9	E4
6	44	20	-.75	.43	.35	-2.39	.35	-2.33	.77	.58	90.0	69.9	E6
3	49	20	-1.63	.41	.85	-.41	.82	-.52	.55	.58	70.0	62.8	E3
2	51	20	-1.96	.40	.75	-.82	.80	-.64	.43	.57	60.0	60.6	E2
11	53	20	-2.29	.40	2.39	3.58	2.32	3.40	.55	.56	40.0	58.9	E11
1	57	20	-2.92	.40	.95	-.07	1.20	.69	.16	.54	75.0	61.2	E1
8	66	20	-4.39	.41	1.43	1.34	1.30	.80	.54	.44	50.0	64.1	E8
MEAN	41.1	20.0	.00	.48	.95	-.20	.91	-.24			71.0	69.3	
P. SD	12.5	.0	2.47	.13	.60	1.94	.63	1.94			22.7	10.3	

Gambar 3.4 Tampilan hasil Table Item Statistics : Measure Order

Berdasarkan Gambar 3.4, dapat dilihat bahwa nilai rata-rata item selalu 0,00 logit, sehingga tingkat kesulitan butir-butir soal dapat dikelompokkan menjadi pengelompokan soal berdasarkan tingkat kesukaran soal pada Tabel 3.10.

Tabel 3.10
Pengelompokan Soal Berdasarkan Tingkat Kesukaran Soal

Kode Butir Soal	JMLE Measure	Kategori Tingkat Kesukaran Soal
E1	-2,92	sangat mudah
E2	-1,96	sangat mudah
E3	-1,63	sangat mudah
E4	-0,75	mudah
E5	0,91	sukar
E6	-0,75	mudah
E7	3,50	sangat sukar
E8	-4,39	sangat mudah
E9	1,09	sangat sukar
E10	3,50	sangat sukar
E11	-2,29	sangat mudah
E12	4,63	sangat sukar
E13	-0,38	mudah
E14	-0,19	mudah
E15	1,66	sangat sukar

Berdasarkan Tabel 3.10 dapat disimpulkan bahwa pada instrumen tes keterampilan proses yang telah dirancang dan diuji coba, terdiri atas lima soal dengan kategori sangat mudah, empat soal dengan kategori mudah, satu soal dengan kategori sukar, dan lima soal dengan kategori sangat sukar.

d.2.4 Daya Pembeda

Usai menguji tingkat kesulitan, langkah selanjutnya adalah mengidentifikasi daya pembeda. Daya pembeda suatu pertanyaan mengindikasikan kemampuan pertanyaan tersebut dalam memisahkan antara peserta didik yang memiliki tingkat kemampuan tinggi dan peserta didik dengan kemampuan rendah (Arikunto, 2020, hlm. 235). Dalam penelitian ini, daya pembeda diestimasi melalui model Rasch dengan menggunakan perangkat lunak Winstep.

Hasil uji daya pembeda butir soal dapat dilihat melalui Output Tables: Item

Fit. Menurut penjelasan Purwanto dan rekan-rekannya (dalam Athiyyah, 2022), nilai MNSQ dan ZSTD digunakan sebagai pertimbangan dalam menilai kevalidan butir soal, sementara nilai Point Measure Correlation dianalisis untuk menilai daya pembedanya. Pedoman interpretasi nilai Point Measure Correlation diuraikan oleh Smiley (dalam Athiyyah, 2022) berupa interpretasi nilai Point Measure Correlation dalam Tabel 3.11.

Tabel 3.11
Interpretasi Nilai *Point Measure Correlation*

Daya Pembeda	Interpretasi
PT-Measure Corr. > 0,40	Sangat Baik
0,30 < PT-Measure Corr. ≤ 0,40	Baik
0,20 < PT-Measure Corr. ≤ 0,30	Cukup
PT-Measure Corr. ≤ 0,20	Tidak Cukup

Berdasarkan Gambar 3.4 Tampilan hasil Table Item Statistics : Measure Order, diperoleh daya pembeda dengan hasil interpretasi nilai daya pembeda dalam Tabel 3.12.

Tabel 3.12
Hasil Interpretasi Nilai Daya Pembeda

Kode Butir Soal	Daya Pembeda / Nilai <i>Point Measure Correlation</i> (Pt Measure Corr)	Interpretasi Nilai
E1	0,16	tidak cukup
E2	0,43	sangat baik
E3	0,55	sangat baik
E4	0,72	sangat baik
E5	0,47	sangat baik
E6	0,77	sangat baik
E7	0,79	sangat baik
E8	0,54	sangat baik
E9	0,54	sangat baik
E10	0,88	sangat baik
E11	0,55	sangat baik
E12	0,87	sangat baik
E13	0,87	sangat baik
E14	-0,03	tidak cukup
E15	0,65	sangat baik

Sehingga dapat disimpulkan bahwa soal dengan kode E2, E3, E4, E5, E6, E7, E8, E9, E10, E11, E12, E13, dan E15 mampu membedakan peserta didik

kelompok atas dan bawah, sedangkan soal dengan kode E1 dan E14 tidak mampu membedakan peserta didik kelompok atas dan bawah, sehingga harus direvisi sebelum diujikan ke peserta didik.

Berdasarkan pembahasan hasil uji validitas konten dan item, reliabilitas, tingkat kesukaran, dan daya pembeda soal instrumen tes keterampilan proses, dapat disimpulkan keputusan penggunaan instrumen tes keterampilan proses dalam Tabel 3.13.

Tabel 3.13
Keputusan Penggunaan Instrumen Tes Keterampilan Proses

Nomor Butir Soal	Kode Butir Soal	Validitas Konten	Validitas Item	Reliabilitas	Tingkat Kesukaran	Daya Pembeda	Keputusan
1	E1	Valid	Valid	Reliabel	Sangat Mudah	Tidak Cukup	Dapat Digunakan dengan Revisi
2	E2	Valid	Valid	Reliabel	Sangat Mudah	Sangat Baik	Dapat Digunakan
3	E3	Valid	Valid	Reliabel	Sangat Mudah	Sangat Baik	Dapat Digunakan
4	E4	Valid	Valid	Reliabel	Mudah	Sangat Baik	Dapat Digunakan
5	E5	Valid	Valid	Reliabel	Sukar	Sangat Baik	Dapat Digunakan
6	E6	Valid	Valid	Reliabel	Mudah	Sangat Baik	Dapat Digunakan
7	E7	Valid	Valid	Reliabel	Sangat Sukar	Sangat Baik	Dapat Digunakan
8	E8	Valid	Valid	Reliabel	Sangat Mudah	Sangat Baik	Dapat Digunakan
9	E9	Valid	Valid	Reliabel	Sangat Sukar	Sangat Baik	Dapat Digunakan
10	E10	Valid	Valid	Reliabel	Sangat Sukar	Sangat Baik	Dapat Digunakan
11	E11	Valid	Valid	Reliabel	Sangat Mudah	Sangat Baik	Dapat Digunakan
12	E12	Valid	Valid	Reliabel	Sangat Sukar	Sangat Baik	Dapat Digunakan
13	E13	Valid	Valid	Reliabel	Mudah	Sangat Baik	Dapat Digunakan
14	E14	Valid	Valid	Reliabel	Mudah	Tidak Cukup	Dapat Digunakan dengan Revisi
15	E15	Valid	Valid	Reliabel	Sangat Sukar	Sangat Baik	Dapat Digunakan

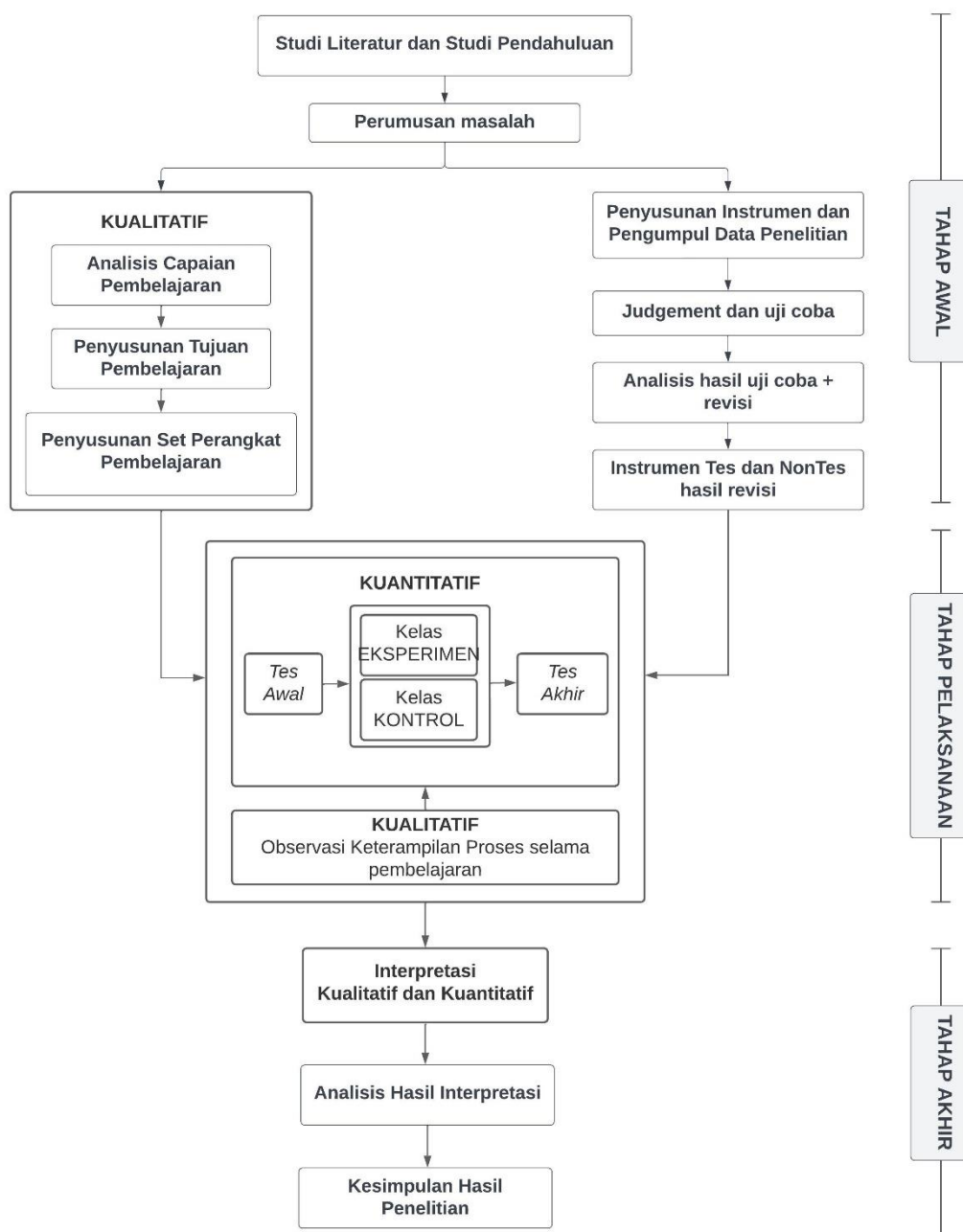
Pada Tabel 3.13 dapat dilihat bahwa seluruh soal valid dan reliabel, namun soal berkode E1 (soal bernomor 1) memiliki daya pembeda pada kategori tidak cukup (jelek) dengan nilai Pt Measure Corr 0,16 sehingga butir soal E1 perlu

direvisi sebelum digunakan. Demikian juga dengan soal berkode E14 (soal bernomor 14) memiliki daya pembeda pada kategori tidak cukup (jelek) dengan nilai P_t Measure Corr $-0,03$ sehingga butir soal E14 tidak dapat digunakan atau dapat digunakan namun harus direvisi. Kesimpulannya seluruh soal tes keterampilan proses yang berjumlah lima belas butir soal dapat digunakan untuk mengukur keterampilan proses pada materi energi terbarukan.

3.4 Prosedur Penelitian

Penelitian diawali dengan studi literatur dan studi pendahuluan, hasilnya kemudian digunakan untuk merumuskan masalah penelitian. Selanjutnya dibuat set perangkat pembelajaran beserta instrumen Tes Keterampilan Proses Sains, instrumen non tes berupa Lembar Observasi Keterampilan Proses Peserta Didik, Lembar Observasi Keterlaksanaan *Scientific Group Inquiry Learning* (SGIL) Berorientasi Profil Pelajar Pancasila Materi Energi Terbarukan Pada Peserta Didik, dan Lembar Observasi Keterlaksanaan Keterlaksanaan *Scientific Group Inquiry Learning* (SGIL) Berorientasi Profil Pelajar Pancasila Materi Energi Terbarukan Pada Guru serta dilanjutkan dengan menguji validitas dan reliabilitas instrumen, jika ada kritik dan saran dari para ahli kemudian direvisi dan mendapatkan instrumen hasil revisi yang valid dan reliabel. Instrumen tersebut digunakan dalam penelitian pada kelas eksperimen dan kontrol sehingga didapatkan data kuantitatif dari hasil tes awal dan tes akhir, beserta data kualitatif dari hasil observasi selama pembelajaran. Seluruh data yang diperoleh tersebut kemudian diinterpretasikan melalui pengolahan data, analisis, lalu akan diperoleh kesimpulan hasil penelitian.

Prosedur yang dilakukan dalam penelitian ini dapat dilihat dalam alur penelitian pada Gambar 3.5.



Gambar 3.5 Alur Penelitian

3.5 Analisis Data

Pengolahan dan analisis data hasil observasi keterlaksanaan pembelajaran *Scientific Group Inquiry Learning* (SGIL) Berorientasi Profil Pelajar Pancasila materi Energi Terbarukan pada guru dan peserta didik diolah melalui perhitungan persentase (%) kegiatan-kegiatan pada setiap fase SGIL

berorientasi Profil Pelajar Pancasila pada pembelajaran pertama dan pembelajaran kedua. Perhitungan dilakukan dengan menggunakan persamaan 3. berikut :

$$PK(\%) = \frac{JKT}{JSK} \times 100\% \dots\dots\dots(2)$$

Keterangan :

PK (%) = Persentase keterlaksanaan kegiatan

JKT = Jumlah kegiatan yang terlaksana

JSK = Jumlah seluruh kegiatan dalam pembelajaran

selanjutnya dikategorikan sesuai kategori keterlaksanaan pembelajaran pada guru dan peserta didik pada Tabel 3.14.

Tabel 3.14
Kategori Keterlaksanaan Pembelajaran pada Guru dan Peserta Didik
(Riduwan, 2012)

Keterlaksanaan Pembelajaran	Kategori
0	Seluruh kegiatan tidak terlaksana
0 s.d 24	Sebagian kecil kegiatan terlaksana
25 s.d 49	Hampir setengah kegiatan terlaksana
50	Setengah dari kegiatan terlaksana
50 s.d 75	Sebagian besar kegiatan terlaksana
76 s.d 99	Hampir semua kegiatan terlaksana
100	Semua kegiatan terlaksana

Keterangan : PK = Persentase keterlaksanaan kegiatan

Profil keterampilan proses (secara kualitatif) dianalisis setelah dilakukan implementasi model SGIL berorientasi Profil Pelajar Pancasila melalui hasil penilaian non tes (dari lembar observasi). Skor non tes tersebut langsung diinterpretasikan sesuai rubrik sehingga dapat dijelaskan profil keterampilan proses peserta didik. Skor hasil observasi kemudian dikonversi dalam persentase, lalu dihitung nilai rata-ratanya kemudian interpretasikan sesuai kriteria penilaian profil keterampilan proses pada Tabel 3.15.

Tabel 3.15
Kriteria Penilaian Profil Keterampilan Proses (Pusat Asesmen, 2022)

Kriteria Penilaian Profil Keterampilan Proses	Kategori
81 - 100	Sangat baik
71 - 80	Baik
61 - 70	Cukup
0 - 60	Perlu Bimbingan

Langkah selanjutnya yaitu menganalisis peningkatan keterampilan proses (secara kuantitatif) setelah dilakukan implementasi model SGIL berorientasi Profil Pelajar berdasarkan skor Gain yang dinormalisasi, diperoleh dari skor tes awal dan skor tes akhir. Gain yang dinormalisasi ialah perbandingan antara gain yang dicapai oleh peserta didik dengan gain maksimum yang dapat dicapai (Hake, 1998). Perhitungan nilai rata-rata gain ternormalisasi ($N-gain$) atau $\langle g \rangle$ menurut Hake (1998), yaitu sebagai berikut:

$$\langle g \rangle = \frac{\% \langle S_{post} \rangle - \% \langle S_{pre} \rangle}{100 - \% \langle S_{pre} \rangle} \dots \dots \dots (3)$$

Keterangan :

$\langle g \rangle$ = skor rerata gain ternormalisasi

$\langle S_{post} \rangle$ = skor rerata tes akhir

$\langle S_{pre} \rangle$ = skor rerata tes awal

Interpretasi hasil perhitungan $N-gain$ dapat dilihat dalam kategori peningkatan keterampilan proses menurut Hake (1998) yang dimodifikasi oleh Sundayana (2020) pada Tabel 3.16.

Tabel 3.16
Kategori Peningkatan Keterampilan Proses (Sundayana, 2020)

$\% \langle g \rangle$	Kategori
$0,7 \leq N - gain \leq 1,0$	Tinggi (<i>High-g</i>)
$0,3 \leq N - gain < 0,7$	Sedang (<i>Medium-g</i>)
$0,0 < N - gain < 0,3$	Rendah (<i>Low-g</i>)
$N - gain = 0,0$	Tetap
$-1,0 \leq N - gain < 0,0$	Terjadi penurunan

Setelah mengetahui ada atau tidaknya peningkatan keterampilan proses peserta didik melalui pembelajaran SGIL berorientasi Profil Pelajar Pancasila, maka selanjutnya akan dilakukan uji efektivitas model pembelajaran SGIL

berorientasi Profil Pelajar Pancasila melalui uji hipotesis skor N-gain untuk mengetahui apakah terdapat perbedaan yang signifikan antara skor kelas eksperimen dan kelas kontrol. Pengujian hipotesis dapat dilakukan dengan dua cara, yaitu uji statistik parametrik dan uji statistik non parametrik. Untuk menentukan uji statistik yang sesuai, terlebih dahulu dilakukan uji normalitas dan uji homogenitas, kemudian kemungkinan dilanjutkan dengan uji hipotesis/uji beda dua rata-rata.

Uji normalitas dilakukan untuk mengetahui apakah data tes awal dan tes akhir serta nilai rata-rata gain ternormalisasi pada kelas eksperimen dan kontrol berdistribusi normal. Pada penelitian ini uji normalitas dilakukan dengan uji statistik Kolmogorov-Smirnov, data dapat dikatakan berdistribusi normal jika kriteria tingkat signifikansi pengujian lebih besar dari 0,05. Setelah melakukan uji normalitas, dilakukan uji homogenitas dengan tujuan untuk mengetahui data tes awal, tes akhir dan nilai rata-rata N-gain pada kelas eksperimen dan kontrol dengan varians sama atau tidak. Uji homogenitas dilakukan dengan menggunakan Levene test, data dikatakan homogen jika kriteria uji signifikansi lebih besar dari 0,05.

Jika data memenuhi persyaratan normal dan homogen, metode parametrik untuk perbedaan dua nilai rata-rata dapat dilakukan. Namun, jika data tidak memenuhi kedua syarat tersebut, maka uji beda dua rata-rata akan dilakukan dengan menggunakan metode non parametrik. Uji statistik non-parametrik dilakukan dengan menggunakan uji U Mann Whitney untuk sampel independen. Hipotesis statistik diuji sebagai berikut:

H_0 : Pembelajaran dengan menggunakan model *Scientific Group Inquiry Learning* (SGIL) berorientasi Profil Pelajar Pancasila tidak efektif meningkatkan keterampilan proses.

H_1 : Pembelajaran dengan menggunakan model *Scientific Group Inquiry Learning* (SGIL) berorientasi Profil Pelajar Pancasila efektif meningkatkan keterampilan proses.

Kriteria yang digunakan dalam uji hipotesis statistik ini adalah bahwa apabila nilai signifikansi (sig.) lebih kecil dari α , dengan α bernilai 0,05, maka

hipotesis alternatif (H_1) dapat diterima. Jika dalam uji perbedaan dua rata-rata ditemukan bahwa tidak terdapat perbedaan yang signifikan dalam keterampilan proses siswa melalui penerapan model SGIL berorientasi Profil Pelajar Pancasila, langkah selanjutnya adalah mengukur dampak efek (*effect size*). Untuk menghitung nilai efek treatment dalam pembelajaran, dapat menggunakan rumus Cohen's d seperti berikut:

$$d = \frac{\bar{x}_t - \bar{x}_c}{S_{pooled}} \dots\dots\dots(4)$$

Keterangan :

- d = Cohen's d effect size
- \bar{x}_t = Nilai rerata N-gain kelas eksperimen
- \bar{x}_c = Nilai rerata N-gain kelas kontrol
- S_{pooled} = Standar deviasi gabungan

Standar deviasi gabungan (S_{pooled}) jika ukuran sampelnya sama dihitung menggunakan rumus berikut (Maher, Markey, & Ebert-May, 2013):

$$S_{pooled} = \sqrt{\frac{(n_t-1)s_t^2 + (n_c-1)s_c^2}{n_t+n_c}} \dots\dots\dots(5)$$

Standar deviasi gabungan (S_{pooled}) jika ukuran sampelnya berbeda dihitung menggunakan rumus berikut (Maher, Markey, & Ebert-May, 2013):

$$S_{pooled} = \sqrt{\frac{(n_t-1)s_t^2 + (n_c-1)s_c^2}{n_t+n_c-2}} \dots\dots\dots(6)$$

Keterangan :

- n_t = Jumlah peserta didik kelas eksperimen
- n_c = Jumlah peserta didik kelas kontrol
- s_t = Standar deviasi N-gain kelas eksperimen
- s_c = Standar deviasi N-gain kelas kontrol

Berikut disajikan kategori interpretasi nilai *Cohen's d* pada Tabel 3.17.

Tabel 3.17
Interpretasi Nilai *Cohen's d*

<i>Effect Size</i>	Interpretasi
$1.00 \leq d$	Sangat Tinggi
$0.8 \leq d < 1.00$	Tinggi
$0.5 \leq d < 0.8$	Sedang
$0.2 \leq d < 0.5$	Rendah
$d < 0.20$	Sangat rendah

Semua perhitungan statistik baik parametrik maupun non parametrik menggunakan program IBM SPSS *Statistics* versi 27 pada taraf signifikansi 5% atau tingkat kepercayaan 95%.