dilakukan literatur *review* terlebih dahulu untuk mengetahui gambaran konsepsi alternatif peserta didik terkait konsep alat optik serta menyusun pembelajaran terkait RPP POEAW, LKPD, dan simulasi virtual materi alat optik. Metode yang digunakan dalam penelitian kuantitatif adalah metode eksperimen, dengan desain penelitian yang digunakan adalah *Quasi-Experimental Design.* Desain ini melibatkan dua kelompok yang dipilih untuk dijadikan sebagai kelompok eksperimen dan kelompok kontrol. Kedua kelompok diberikan *pre-test* dan *post-test*. Kelompok eksperimen diberi *treatment* berupa pembelajaran POEAW dengan berbantuan simulasi virtual, sedangkan pada kelompok kontrol diberikan pembelajaran POEAW tanpa berbatuan simulasi virtual.

Setelah intervensi, penelitian data kualitatif dijadikan kajian dan analisis secara mendalam menggunakan kodefikasi konsepsi peserta didik. Seluruh data kualitatif diperoleh dari analisis banyaknya peserta didik yang mengalami kondisi miskonsepsi tersebut melalui analisis data level konsepsi. Selanjutnya, data kualitatif ini diolah sesuai kodefikasi kategori dan tipe perubahan miskonsepsi. Serta literasi peningkatkan fisika peserta didik di kategorikan pada tingkat rendah, sedang, dan tinggi.

3.2 Partisipan

Partisipan penelitian ini terdiri dari peserta didik kelas X IPA dari salah satu SMA di Kota Bandung. Para peserta didik memiliki usia antar 15 sampai 17 tahun dan berasal dari latar belakang keluarga yang hampir sama. Sebagian besar peserta didik berasal dari suku Sunda, Jawa, dan beberapa suku lainnya. Sampel penelitian ini sebanyak 54 peserta didik yang dibagi menjadi dua kelompok. Kelompok eksperimen sebanyak 27 peserta didik dan kelompok kontrol sebanyak 27 peserta didik.

3.3 Populasi dan Sampel Penelitian

Populasi penelitian ini yaitu seluruh peserta didik kelas XI IPA semester genap 2021-2022 SMA di Kota Bandung. Sebagain besar populasi dalam penelitian ini berasal dari suku Sunda dan beberapa suku lainnya seperti Jawa, Batak, Minang, dan lain-lain. Setelah populasi dipilih, maka dilakukan pemilihan sampel penelitian.

Sampel yang digunakan adalah dua kelas XI yang dibagi menjadi dua kelompok, yaitu kelompok eksperimen dan kelompok kontrol. Pada kelompok eksperimen pembelajaran dilakukan dengan menggunakan POEAW berbantuan simulasi virtual, sedangkan pada kelompok kontrol pembelajaran dilakukan dengan POEAW tanpa berbantuan simulasi virtual. Sampel pada penelitian ini dipilih dengan menggunakan teknik *random sampling.*

3.4 Instrumen Penelitian dan Media Pembelajaran

Guna memperoleh data yang mendukung penelitian ini, sehingga dilakukan penyusunan dan penggunaan beberapa instrumen yang meliputi: (1) tes literasi fisika dan (2) *four-tiers test*. Media yang digunakan dalam pembelajaran yaitu simulasi virtual materi alat optik. Masing-masing instrumen dan media pembelajaran diurai dan dijelaskan sebagai berikut:

3.4.1. Tes Literasi Fisika

Tes literasi fisika pada penelitian ini disajikan berbentuk pilihan ganda untuk *pre-test* dan *post-test*. Tes literasi fisika disusun sesuai dengan aspek pengetahuan dan apsek kompetensi sesuai dengan PISA. Tes literasi fisika pada *pre-test* berguna untuk mengetahui literasi fisika awal peserta didik sebelum melakukan pembelajaran dengan menggunakan POEAW berbantuan simulasi virtual atau pembelajaran POEAW tanpa berbantuan simulasi virtual. Sedangkan tes literasi fisika pada *post-test* dilakukan untuk mengetahui pencapaian literasi fisika peserta didik setelah diberikan pembelajaran POEAW berbantuan simulasi virtual atau pembelajaran POEAW tanpa berbantuan simulasi virtual.

Tes literasi fisika yang dikembangkan sesuai dengan aspek literasi fisika. Aspek kompetensi literasi fisika terdiri dari: (1) penjelasan fenomena atau masalah fisika secara ilmiah; (2) evaluasi dan rancangan penyelidikan ilmiah fisika; dan (3) interpretasi data dan bukti secara ilmiah sesuai dengan konsep, prinsip, dan teori fisika. Aspek pengetahuan terdiri dari pengetahuan konten, prosedural, dan epistemik. Tes literasi fisika yang disusun fokus pada materi alat optik sejumlah 15 butir soal. Sub materi tes literasi fisika terdistribusi pada: (1) mata, kamera, dan kacamata; (2) lup; (3) mikroskop; dan (4) teropong. Sebaran soal tes literasi fisika seperti yang ditunjukkan pada Tabel 3.1 dan soal tes literasi fisika dilampirkan pada Lampiran A.1.

Tabel 3.1. Sebaran soal tes literasi fisika

No	Sub Materi	Aspek Kompetensi dan Pengetahuan (Literasi Fisika)	No Soal
1.	Mata, kacamata, dan kamera	Penjelasan fenomena atau masalah alat optik secara ilmiah (pengetahuan konten).	1
			2
		Evaluasi dan rancangan penyelidikan ilmiah alat optik (pengetahuan prosedural).	3
			4
		Interpretasi data dan bukti secara ilmiah. (pengetahuan epistemik).	5
			6
2.	Lup	Penjelasan fenomena atau masalah alat optik secara ilmiah (pengetahuan konten).	7
		Evaluasi dan rancangan penyelidikan ilmiah alat optik (pengetahuan prosedural).	8
		Interpretasi data dan bukti secara ilmiah (pengetahuan epistemik).	9
3.	Mikroskop	Penjelasan fenomena atau masalah alat optik secara ilmiah (pengetahuan konten).	10
		Evaluasi dan rancangan penyelidikan ilmiah alat optik (pengetahuan prosedural).	11
		Interpretasi data dan bukti secara ilmiah (pengetahuan epistemik).	12
4.	Teropong	Penjelasan fenomena atau masalah alat optik secara ilmiah (pengetahuan konten).	13
		Evaluasi dan rancangan penyelidikan ilmiah alat optik (pengetahuan prosedural).	14
		Interpretasi data dan bukti secara ilmiah (pengetahuan epistemik).	15

Butir soal literasi fisika dikembangkan dalam bentuk pilihan ganda. Soal tes literasi yang telah disusun kemudian dilakukan validasi ahli kepada beberapa ahli. Kemudian, butir soal literasi fisika divalidasi empiris dengan melakukan uji coba soal tes literasi fisika kepada peserta didik yang telah mempelajari materi alat opik. Masing-masing hasil dan analisis dari validasi ahli dan validasi empiris diuraikan sebagai berikut.

3.4.1.1 Validasi Ahli (*Expert Judgment*)

Validasi dilakukan oleh lima orang ahli, terdiri dari dosen pendidikan fisika, guru mata pelajaran fisika, dan juga peneliti dibidang serupa. Uji validasi dilakukan terhadap setiap butir soal yang mengacu pada rubrik lembar validasi (Lampiran B.1). Lembar validasi berupa daftar cek yang terdiri dari beberapa pernyataan pada setiap butir soal dengan kriteria “valid tanpa revisi”, “valid dengan revisi”, dan “tidak valid”. Setiap butir soal dinilai berdasarkan sembilan indikator penilaian seperti yang ditampilkan pada Tabel 3.2 dan dilampirkan pada Lampiran B.1.

Tabel 3.2. Indikator Penilaian Butir Literasi Fisika

Nomor Indikator penilaian	Deskripsi Indikator
1	Butir soal yang dibuat sesuai dengan indikator soal dan aspek literasi fisika.
2	Kesesuaian konsep dalam butir soal dengan konsep yang dikemukakan oleh para ahli.
3	Butir soal dibuat untuk mengetahui kemampuan literasi fisika peserta didik.
4	Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia.
5	Bahasa yang digunakan mudah dimengerti oleh peserta didik.
6	Pilihan jawaban homogen serta logis dari segi materi.
7	Hanya ada satu kunci jawaban.
8	Soal tidak memberikan petunjuk jawaban ke arah jawaban yang benar.
9	Pilihan jawaban tidak menggunakan pernyataan "semua jawaban benar" atau "jawaban salah".

Tes literasi fisika yang telah disusun kemudian divalidasi oleh ahli (*expert judgment*). Hasil penilaian ahli kemudian dianalisis dengan *Content Validity Ratio* (CVR). CVR dihitung dengan menggunakan persamaan berikut:

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \tag{3.1}$$

dimana

n_e= jumlah validator yang memberikan penilaian "valid tanpa revisi*"* atau "valid dengan revisi"

N = jumlah keseluruhan validator.

Hasil perhitungan nilai CVR diinterpretasi sesuai dengan yang disajikan pada Tabel 3.3.

Tabel 3.3.Nilai minimum CVR uji satu pihak $\boldsymbol{\alpha}$ =0,5

Jumlah Validator	CVR_{kritis}
5	0,736
6	0,672
7	0,622
8	0,582
9	0,548
10	0,520

(Wilson et al., 2012)

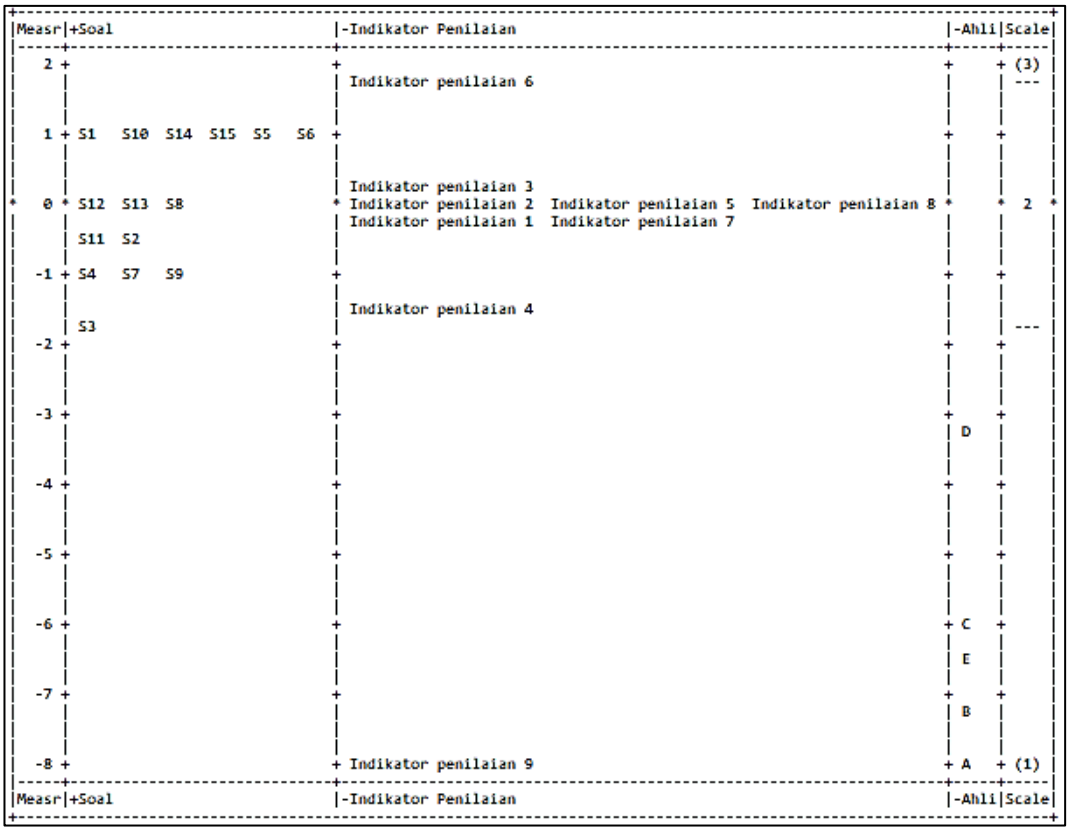
Analisis dari validasi ahli menggunakan CVR dihasilkan seperti pada Tabel 3.4 berikut.

Tabel 3.4. Hasil Validasi Ahli Menggunakan CVR

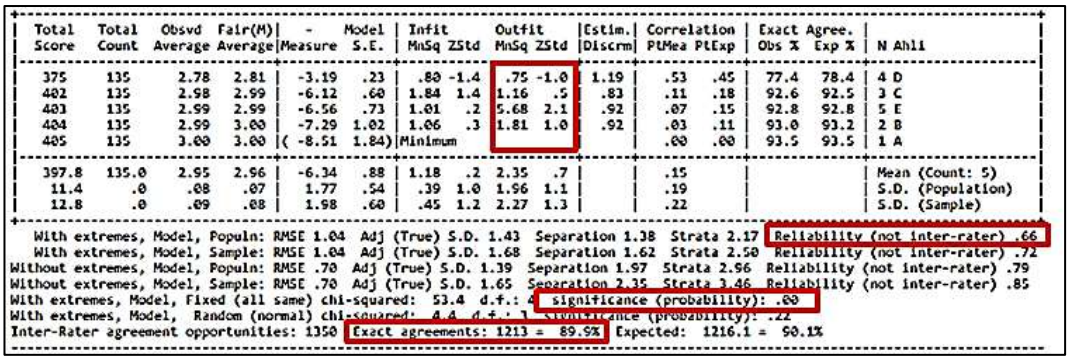
Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi	Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi
S1	1	5	5	1	Sesuai	S9	1	5	5	1	sesuai
	2	5	5	1	Sesuai		2	5	5	1	sesuai
	3	5	5	1	Sesuai		3	5	5	1	sesuai
	4	5	5	1	Sesuai		4	5	5	1	sesuai
	5	5	5	1	Sesuai		5	5	5	1	sesuai
	6	5	5	1	Sesuai		6	5	5	1	sesuai
	7	5	5	1	Sesuai		7	5	5	1	sesuai
	8	5	5	1	Sesuai		8	5	5	1	sesuai
	9	5	5	1	Sesuai		9	5	5	1	sesuai
S2	1	5	5	1	Sesuai	S10	1	5	5	1	sesuai
	2	5	5	1	Sesuai		2	5	5	1	sesuai
	3	5	5	1	Sesuai		3	5	5	1	sesuai
	4	5	5	1	Sesuai		4	5	5	1	sesuai
	5	5	5	1	Sesuai		5	5	5	1	sesuai
	6	5	5	1	Sesuai		6	5	5	1	sesuai
	7	5	5	1	Sesuai		7	5	5	1	sesuai
	8	5	5	1	Sesuai		8	5	5	1	sesuai
	9	5	5	1	Sesuai		9	5	5	1	sesuai
S3	1	5	5	1	Sesuai	S11	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	4	0,6	diperbaiki		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S4	1	5	5	1	sesuai	S12	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S5	1	5	5	1	sesuai	S13	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S6	1	5	5	1	sesuai	S14	1	5	5	1	sesuai

Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi	Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S7	1	5	5	1	sesuai	S15	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S8	1	5	5	1	sesuai						
	2	5	5	1	sesuai						
	3	5	5	1	sesuai						
	4	5	5	1	sesuai						
	5	5	5	1	sesuai						
	6	5	5	1	sesuai						
	7	5	5	1	sesuai						
	8	5	5	1	sesuai						
	9	5	5	1	sesuai						

Validasi ahli yang dianalisis dengan CVR menghasilkan bahwa butir soal nomor 1, 2 ,4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, dan 15 memiliki nilai CVR indeks yang lebih dari 0,736. Nilai CVR indeks yang lebih dari CVR kritis dapat diinterpretasikan bahwa butir soal "sesuai" sehingga dapat digunakan. Sedangkan butir soal nomor 3 perlu diperbaiki sesuai indikator penilaian nomor tujuh dan sesuai saran dari ahli. Hasil penilaian ahli juga dilakukan analisis multi-rater. Hasil penilaian ahli yang dianalisis menggunakan multi-rater ditampilkan pada Gambar 3.2 dan Gambar 3.3.



Gambar 3.2.Hasil Multi-rater Validasi Ahli Soal Tes Literasi Fisika



Gambar 3.3. Hasil Pengukuran Ahli pada Soal Tes Literasi Fisika

Gambar 3.2 di atas menunjukkan terdapat sembilan butir soal yang dinilai memenuhi syarat oleh para ahli. Butir soal tersebut terdiri dari S12, S13, S8, S1, S10, S14, S15, S5, dan S6. Sementara itu, ahli menilai enam butir soal yang memenuhi syara yaitu S3, S4, S7, S9, S11, dan S2. Kolom ketiga pada Gambar 3.2 menunjukkan informasi tingkat kesulitan indikator penilaian. Pembagian indikator penilaian ditunjukkan pada kolom ketiga. Semakin rendah nilai logit dari indikator penilaian maka indikator penilaian tersebut semakin mudah dipenuhi oleh suatu butir. Semakin tinggi nilai logit dari suatu indikator penilaian,

maka semakin sulit indikator penilaian tersebut dipenuhi oleh suatu butir soal yang dinilai oleh validator (ahli). Nilai logit indikator penilaian sama memiliki tingkat kesulitan yang sama. Menurut Gambar 3.2 indikator penilaian 9 (pilihan jawaban tidak menggunakan pernyataan "semua jawaban benar" atau "jawaban salah") merupakan indikator penilaian yang paling mudah karena menurut penilian ahli semua butir soal memenuhi indikator penilaian ini. Sedangkan indikator penilaian 6 (pilihan jawaban dan alasan homogen dan logis dari segi materi) merupakan indikator penilaian yang paling sulit untuk dipenuhi oleh butir soal menurut penilaian ahli.

Berdasarkan Gambar 3.3, kualitas penilaian ahli yang diurutkan berdasarkan tingkat keparahan butir soal. Ahli D dan Ahli C adalah yang paling konsisten kriteria kevalidan (Boone et al., 2014; Darmana et al., 2021). Ahli C dan D memiliki nilai memenuhi rentang 0,5 sampai 1,5 (Outfit MNSQ) dan -2 sampai +2 (Outfit ZSTD). Ahli E dan B adalah menunjukkan penilaian yang kurang valid. Sedangkan ahli A merupakan ahli yang terburuk karena memiliki nilai infit yang paling rendah. Reliabilitas penilaian antar ahli cukup yaitu 0,66, menunjukkan bahwa para ahli memberikan penilaian yang cukup berbeda, tetapi ada juga yang sama (Koçak, 2020). Kecenderungan penilai dapat mempengaruhi nilai reliabilitas (Bond & Fox, 2013; Darmana et al., 2021). Data yang dihasilkan konsisten ditunjukkan dengan pengukuran yang didukung oleh nilai uji Chi-square ($p<0,01$). Kesamaan penilaian oleh lima ahli (kesepakatan yang benar) mencapai 89,9%, menunjukkan bahwa terdapat sedikit perbedaan dari penilaian semua butir soal oleh kelima ahli. Beberapa studi literatur juga menunjukkan kecenderungan penilaian penilai berbeda dan perilaku penilai seperti kelonggaran dan keparahan dapat mempengaruhi keandalan penilai (Brookhart et al., 2006; Güler, 2014; Darmana et al., 2021).

3.4.1.2 Validitas Instrumen

a) Validitas butir soal

Validitas digunakan untuk mengetahui tingkat kevalidan suatu instrumen. Suatu instrumen dikatakan valid apabila instrumen yang digunakan dapat mengukur apa yang hendak diukur. Hasil dari uji coba instrumen dianalisis dari analisis butir soal dengan menggunakan pemodelan Rasch melalui *software*

MINISTEP. Analisis ini menggunakan *software* MINISTEP dengan mengolah data hasil uji coba. Analisis validitas butir soal dengan menggunakan *Rasch Model* pada *butir fit order* memiliki tiga kriteria yaitu *outfit MNSQ*, *outfit ZSTD*, dan *Pt mean corr* (Bonee et al., 2014; Bond & Fox, 2015; Sumintono & Widhiarso, 2015). Rincian masing-masing kriteria seperti pada Tabel 3.5.

Tabel 3.5. Kriteria Validitas Butir Soal

Outfit	**Nilai yang diterima**
Outfit mean square (MNSQ)	0,5 < MNSQ < 1,5
Outfit Z-standart (ZSTD)	-2,00 < ZSTD < +2.00
Point Measure Correlation (Pt Mean Corr)	0,4 < *Pt Measure Corr* <0,85

(Sumintono & Widhiarso, 2014)

Hasil uji coba soal literasi fisika yang telah di analisis menggunakan *Rasch Model* ditunjukkan pada Gambar 3.4.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
4	40	61	-1.51	.31	1.11	.89	1.66	1.58	A .41	.50	70.5	73.8	S4
11	17	61	.92	.36	1.03	.22	1.49	1.27	B .56	.59	83.6	82.7	S11
7	35	61	-1.02	.31	1.22	1.60	1.44	1.40	C .42	.55	67.2	74.2	S7
9	24	61	.09	.33	1.26	1.34	1.14	.59	D .51	.61	67.2	79.9	S9
12	21	61	.43	.34	1.18	.92	1.16	.61	E .53	.61	77.0	81.3	S12
15	18	61	.79	.35	.91	-.41	1.10	.40	F .62	.60	85.2	82.5	S15
13	23	61	.20	.33	1.07	.41	.96	-.04	G .59	.61	80.3	80.4	S13
14	15	61	1.18	.37	.98	-.01	1.04	.23	H .58	.58	83.6	83.0	S14
3	29	61	-.43	.32	.97	-.15	.98	.02	g .60	.59	83.6	76.5	S3
1	40	61	-1.51	.31	.89	-.88	.96	.01	f .54	.50	77.0	73.8	S1
6	17	61	.92	.36	.94	-.21	.85	-.31	e .62	.59	83.6	82.7	S6
8	24	61	.09	.33	.93	-.31	.82	-.65	d .65	.61	77.0	79.9	S8
5	21	61	.43	.34	.88	-.59	.84	-.46	c .66	.61	83.6	81.3	S5
10	28	61	-.33	.32	.81	-1.14	.69	-1.29	b .69	.59	85.2	77.3	S10
2	27	61	-.23	.32	.67	-2.18	.61	-1.76	a .76	.60	86.9	78.0	S2
MEAN	25.3	61.0	.00	.34	.99	-.03	1.05	.11			79.5	79.2	
P.SD	7.7	.0	.82	.02	.15	.96	.29	.91			6.4	3.2	

Gambar 3.4. *OUTFIT MNSQ*, *OUTFIT ZSTD*, dan *PT MEASURE CORR* setiap butir soal literasi fisika

Suatu butir dikatakan "sangat sesuai" dan layak digunakan jika tiga kriteria terpenuhi (Dewi et al., 2021; Nurdini et al., 2020). Apabila hanya satu atau dua kriteria yang terpenuhi, maka butir dianggap "sesuai" dan dapat dipertahankan (Sumintono & Widhiarso, 2015; Nurdini et al., 2020; Dewi et al., 2021). Namun, apabila ketiga kriteria tidak terpenuhi, maka butir tersebut "tidak sesuai" dan perlu

diperbaiki (Nurdini et al., 2020; Dewi et al., 2021). Berdasarkan kriteria dari validitas yang ditetapkan, hasil analisis butir soal literasi fisika yang ditunjukkkan pada Gambar 3.4 dapat dilihat bahwa butir soal S4 hanya tidak memenuhi kriteria *OUTFIT* MNSQ (bernilai 1,66), namun memenuhi kriteria pada *OUTFIT ZSTD* dan *PT MEASURE CORR*. Sehingga soal S4 tetap dipertahankan dan dikatakan valid atau sesuai. Sedangkan butir soal yang lain memenuhi semua kriteria dan dapat dikatakan sangat sesuai. Sehingga kesimpulan akhir yaitu tidak ada butir soal yang perlu diubah dan semua butir soal tes literasi fisika dapat digunakan.

b) Unidimensionalitas instrumen

Unidimensionalitas instrumen tes literasi fisika merupakan ukuran yang penting untuk mengevaluasi apakah instrumen tes literasi fisika yang dikembangkan mampu mengukur kemampuan literasi fisika peserta didik (Sumintono & Widhiarso, 2015). Pengujian unidimensionalitas instrumen dilakukan dengan menganalisis nilai *raw variance explained by measures* dan *unexplained variance* (Sumintono & Widhiarso, 2015; Nurdini et al., 2020). Persyaratan untuk uji unidimensionalitas yaitu memiliki nilai kurang dari 3 pada *eigen value unexplained variance*. Sedangkan *unexplained variance* pada *observed value* memiliki nilai kurang dari 15%. Unidimensionalitas berdasarkan *raw vaiance explained by measures* diinterpretasikan seperti pada Tabel 3.6.

Tabel 3.6. Interpretasi dari Unidimensionalitas Instrumen

Nilai raw variance explained by measures (V)	Interpretasi
$V < 20\%$	Tidak terpenuhi
$20\% \leq V \leq 40\%$	Terpenuhi
$40\% < V \leq 60\%$	Bagus
$V > 60\%$	Istimewa

Hasil analisis unidimentionalitas instrumen menggunakan *software* MINISTEP ditampilkan pada Gambar 3.5.

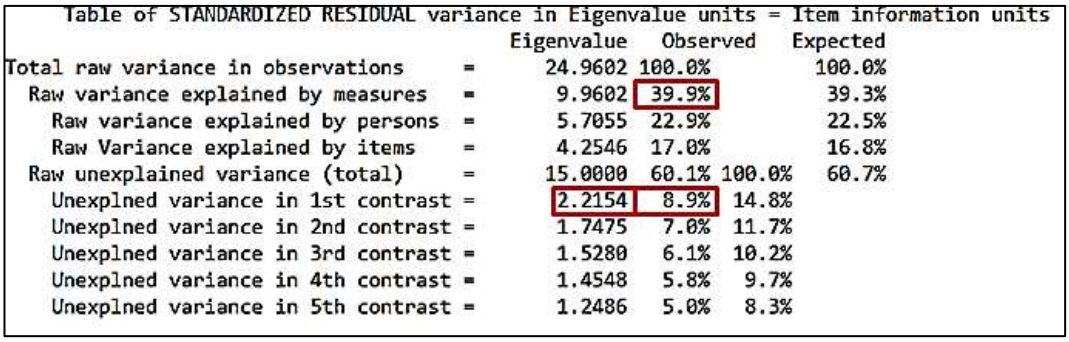


Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

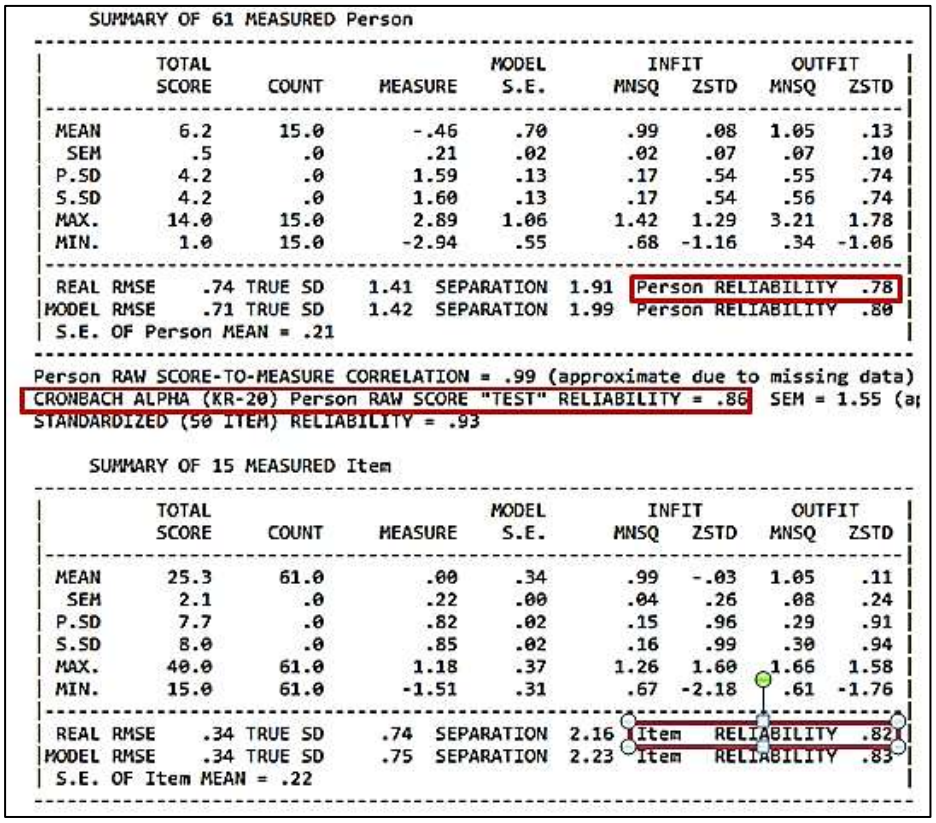
	Eigenvalue	Observed		Expected
Total raw variance in observations =	24.9602	100.0%		100.0%
Raw variance explained by measures =	9.9602	39.9%		39.3%
Raw variance explained by persons =	5.7055	22.9%		22.5%
Raw Variance explained by items =	4.2546	17.0%		16.8%
Raw unexplained variance (total) =	15.0000	60.1%	100.0%	60.7%
Unexplned variance in 1st contrast =	2.2154	8.9%	14.8%	
Unexplned variance in 2nd contrast =	1.7475	7.0%	11.7%	
Unexplned variance in 3rd contrast =	1.5280	6.1%	10.2%	
Unexplned variance in 4th contrast =	1.4548	5.8%	9.7%	
Unexplned variance in 5th contrast =	1.2486	5.0%	8.3%	

Gambar 3.5.Hasil analisis unidimentionalitas instrumen literasi fisika

Berdasarkan Gambar 3.5 menunjukkan bahwa nilai *raw variance explaianed by measures* butir tes soal literasi fisika memiliki kriteria terpenuhi sesuai pada Tabel 3.6. Sedangkan pada *unexplained variance* memiliki *eigen value* kurang dari 15% dan memiliki nilai *observed* kurang dari 3. *Unexplained variance* butir tes literasi fisika memenuhi kriteria pada uji unidimensionalitas. Berdasarkan analisis unidimensionalitas dapat dikatakan bahwa tes literasi fisika valid untuk mengukur kemampuan literasi fisika peserta didik.

3.4.1.3 Uji Reliabilitas

Uji reliabilitas digunakan untuk mengetahui tingkat kepercayaan suatu instrumen apabila dipakai berulang-ulang mampu memberikan hasilnya sama atau untuk untuk mengetahui konsistensi instrumen tes yang digunakan dalam penelitian (Sumintono, 2015). Uji reliabilitas mencerminkan ketepatan instrumen penelitian yang digunakan dalam mengukur informasi yang digunakan. Uji reliabilitas penelitian ini menggunakan *software* MINISTEP pada *output summary statistics* yang dianalisiis dengan *Rasch Model*. Analisis menggunakan *Rasch Model* menampilkan beberapa nilai yaitu *person reliability*, *butir reliability*, dan *Cronbach Alpha*. *Person reliability* menunjukkan konsistensi dari jawaban peserta didik. *Butir reliability* menunjukkan kualitas butir-butir soal dalam instrumen. Sedangkan *Cronbach Alpha* menunjukkan interaksi antara *person* dan butir soal secara keseluruhan. Hasil uji reliabilitas soal tes literasi fisika ditampilkan pada Gambar 3.6.



SUMMARY OF 61 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	6.2	15.0	-.46	.70	.99	.08	1.05	.13
SEM	.5	.0	.21	.02	.02	.07	.07	.10
P.SD	4.2	.0	1.59	.13	.17	.54	.55	.74
S.SD	4.2	.0	1.60	.13	.17	.54	.56	.74
MAX.	14.0	15.0	2.89	1.06	1.42	1.29	3.21	1.78
MIN.	1.0	15.0	-2.94	.55	.68	-1.16	.34	-1.06

REAL RMSE .74 TRUE SD 1.41 SEPARATION 1.91 Person RELIABILITY .78
MODEL RMSE .71 TRUE SD 1.42 SEPARATION 1.99 Person RELIABILITY .80
S.E. OF Person MEAN = .21

Person RAW SCORE-TO-MEASURE CORRELATION = .99 (approximate due to missing data)
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .86 SEM = 1.55 (a
STANDARDIZED (50 ITEM) RELIABILITY = .93

SUMMARY OF 15 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	25.3	61.0	.00	.34	.99	-.03	1.05	.11
SEM	2.1	.0	.22	.00	.04	.26	.08	.24
P.SD	7.7	.0	.82	.02	.15	.96	.29	.91
S.SD	8.0	.0	.85	.02	.16	.99	.30	.94
MAX.	40.0	61.0	1.18	.37	1.26	1.60	1.66	1.58
MIN.	15.0	61.0	-1.51	.31	.67	-2.18	.61	-1.76

REAL RMSE .34 TRUE SD .74 SEPARATION 2.16 Item RELIABILITY .82
MODEL RMSE .34 TRUE SD .75 SEPARATION 2.23 Item RELIABILITY .83
S.E. OF Item MEAN = .22

Gambar 3.6. *Person reliability, butir reliability,* dan *Cronbach Alpha*

Hasil analisis uji reliabilitas menggunakan *Rasch Model* yang ditunjukkan pada Gambar 3.6 dapat diinterpetasi melalui kriteria nilai *person reliability* dan *butir reliability* sesuai pada Tabel 3.7.

Tabel 3.7. Kategori Reliabilitas

Nilai *person reliability* dan *butir reliability*	**Kategori**
r < 0,67	Lemah
0,67 ≤ r < 0,80	Cukup
0,80 ≤ r < 0,90	Bagus
0,90 ≤ r < 0,94	Bagus Sekali
0,94 ≤ r	Istimewa

(Sumintono & Widhiarso, 2015)

Berdasarkan hasil analisis pada Gambar 3.6, nilai *person reliability* menunjukkan nilai 0,78 tergolong dalam kategori cukup dan nilai *butir reliability* menunjukkan nilai 0,82 tergolong dalam kategori bagus. Hasil analisis uji reliabilitas menggunakan *Rasch Model* juga dapat diinterpetasi melalui kriteria nilai *Cronbach Alpha* sesuai Tabel 3.8.

Tabel 3.8. Kriteria nilai *Alpha Cronbach*

Nilai *Alpha Cronbach*	**Kategori**
r < 0,50	Sangat buruk
0,50 ≤ r < 0,60	Buruk
0,60 ≤ r < 0,70	Cukup
0,70 ≤ r < 0,80	Bagus
0,80 ≤ r	Bagus Sekali

(Sumintono & Widhiarso, 2015)

Berdasarkan hasil pada Gambar 3.6, nilai *Cronbach Alpha* menunjukkan nilai sebesar 0,86 yang dapat dikategorikan bagus sekali. Hasil analisis uji coba soal tes literasi fisika dapat disimpulkan bahwa konsitensi jawaban peserta didik yang cukup, kualitas butir soal yang bagus, dan interaksi peserta didik dan butir secara keseluruhan pada kategori bagus sekali. Sehingga intrumen tes literasi fisika dapat digunakan dan dikatakan reliabel.

3.4.1.4 Taraf Kesukaran

Taraf kesukaran tes adalah tingkatan yang menunjukkan sukar dan mudahnya suatu instrumen soal tes. Taraf kesukaran instrumen tes literasi fisika dalam penelitian ini menggunakan *Rasch Model* pada program MINISTEP. *Output* dari *Butir Measure* digunakan untuk mengklasifikasi taraf kesukaran tiap butir soal berdasarkan nilai *measure logit* dan nilai Standar Deviasi (SD) yang diperoleh (Sumintono & Widhiarso, 2015). Semakin tinggi *logit* suatu butir soal, semakin tinggi pula tingkat kesukaran suatu butir soal. Butir soal dapat dikelompokkan menjadi sangat mudah, mudah, sulit, dan sangat sulit. Berikut interpretasi taraf kesuraan butir ditunjukkan pada Tabel 3.9.

Tabel 3.9. Interpretasi Taraf Kesukaran Butir

Interpretasi	**Nilai *Measure logit***
Sangat Mudah	$M < -1SD$
Mudah	$-1SD \leq M \leq 0$
Sulit	$0 \leq M \leq SD$
Sangat Sulit	$M > SD$

(Sumintono & Widhiarso, 2015)

Hasil *output* dari *item measure* untuk taraf kesukaran soal ditunjukkan pada Gambar 3.7 berikut.

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
14	15	61	1.18	.37	.98	-.01	1.04	.23	.58	.58	83.6	83.0	S14
6	17	61	.92	.36	.94	-.21	.85	-.31	.62	.59	83.6	82.7	S6
11	17	61	.92	.36	1.03	.22	1.49	1.27	.56	.59	83.6	82.7	S11
15	18	61	.79	.35	.91	-.41	1.10	.40	.62	.60	85.2	82.5	S15
5	21	61	.43	.34	.88	-.59	.84	-.46	.66	.61	83.6	81.3	S5
12	21	61	.43	.34	1.18	.92	1.16	.61	.53	.61	77.0	81.3	S12
13	23	61	.20	.33	1.07	.41	.96	-.04	.59	.61	80.3	80.4	S13
8	24	61	.09	.33	.93	-.31	.82	-.65	.65	.61	77.0	79.9	S8
9	24	61	.09	.33	1.26	1.34	1.14	.59	.51	.61	67.2	79.9	S9
2	27	61	-.23	.32	.67	-2.18	.61	-1.76	.76	.60	86.9	78.0	S2
10	28	61	-.33	.32	.81	-1.14	.69	-1.29	.69	.59	85.2	77.3	S10
3	29	61	-.43	.32	.97	-.15	.98	.02	.60	.59	83.6	76.5	S3
7	35	61	-1.02	.31	1.22	1.60	1.44	1.40	.42	.55	67.2	74.2	S7
1	40	61	-1.51	.31	.89	-.88	.96	.01	.54	.50	77.0	73.8	S1
4	40	61	-1.51	.31	1.11	.89	1.66	1.58	.41	.50	70.5	73.8	S4
MEAN	25.3	61.0	.00	.34	.99	-.03	1.05	.11			79.5	79.2	
P.SD	7.7	.0	.82	.02	.15	.96	.29	.91			6.4	3.2	

Gambar 3.7.*Measure logit* dan standar deviasi pada keluaran *item measure*

Besesuaian dengan Gambar 3.7, sehingga hasil *output butir measure* dapat dianalisis seperti pada Tabel 3.10 berikut.

Tabel 3.10. Hasil Analisis Taraf Kesukaran Soal Tes Literasi Fisika

Nilai *Measure logit*	Interpretasi Taraf Kesukaran	Butir Soal	Jumlah Butir Soal
$M < -0{,}82$	Sangat Mudah	S4, S1, S7	3
$-0{,}82 \leq M \leq 0$	Mudah	S3, S10, S2	3
$0 \leq M \leq 0{,}82$	Sulit	S9, S8, S13, S12, S5, S15	6
$0{,}82 > SD$	Sangat Sulit	S11, S6, S14	3

Berdasarkan data pada Tabel 3.10 menunjukkan bahwa tingkat kesulitaan soal tes literasi fisika memiliki interpretasi yang bervariasi. Butir soal dengan taraf kesukaran "sangat sulit" memiliki persentase 20%, taraf kesukaran "sulit" memiliki persentase 40%, taraf kesukaran "mudah" memiliki persentase 20%, dan taraf kesukaran " sangat mudah" memiliki presentase 20%. Soal yang paling banyak yaitu berada pada taraf kesukaran "sulit".

3.4.2. *Four-tiers test*

Instrumen *four-tier test* yang digunakan dalam penelitian ini disusun berdasarkan konsepsi alternatif peserta didik pada materi alat optik. Langkah awal yang dilakukan dalam menyusun *four-tiers test* yaitu dengan melakukan studi literatur untuk mengetahui miskonsepsi yang dialami peserta didik pada materi

alat optik. Setelah melakukan studi literatur, peneliti menyusun instrumen berupa tes *four-tiers* (*open-ended*) untuk mengumpulkan konsepsi alternatif sebelum instrumen dikembangkan menjadi tes *four-tier* (*close-ended*). Sebaran butir soal *four-tier* sesuai dengan miskonsepsi yang dialami peserta didik dari hasil studi literatur yang kemudian didistribusikan pada 15 butir soal seperti pada Tabel 3.11.

Tabel 3.11. Sebaran soal *fou-tier*

No	Sub-Materi	Miskonsepsi	No Soal
1.	Mata, kacamata, dan kamera	Titik dekat hipermetropi lebih jauh dari mata normal sehingga benda harus diletakkan lebih dekat dari 25 cm.	1
		Semakin besar diameter *aperature* suatu lensa, semakin banyak cahaya yang masuk sehingga bayangan semakin besar.	2
		Semakin bertambah usia, semakin bertambah baik daya akomodasi mata.	3
		Pupil pada mata manusia memiliki fungsi yang sama dengan diafragma, yaitu mengatur intensitas cahaya yang masuk.	4
		Titik dekat mata penderita miopi lebih dekat dari titik dekat mata normal (PP<25cm).	5
		Jarak kamera terhadap objek lebih dekat ketika kamera berposisi *landscape* daripada posisi *portrait*.	6
2.	Lup	Jika sebagian lensa lup ditutup, maka bayangan yang dibentuk adalah separuh dari benda.	7
		Lensa cembung digunakan sebagai lup karena memiliki sifat menyebarkan cahaya sehingga bayangan suatu benda diperbesar dari ukuran aslinya.	8
		Kekuatan lup tidak dipengaruhi oleh medium tempat lup digunakan, artinya kekuatan lup di udara dan di air sama.	9
3.	Mikroskop	Mikroskop berfungsi untuk melihat benda-benda kecil supaya terlihat besar dan jelas.	10
		Pengamatan menggunakan mikroskop secara akomodasi maksimum terjadi apabila bayangan yang dibentuk oleh lensa objektif berada tepat di fokus lensa okuler.	11
		Cermin cekung dan lensa cembung memiliki sifat menyebarkan cahaya.	12
4.	Teropong	Semua lensa pada teropong panggung adalah lensa cembung yang dapat mengumpulkan cahaya.	13
		Pada pengamatan tidak berakomodasi, perbesaran bayangan dari teropong bintang dipengaruhi oleh panjang tabung teropong.	14
		Semakin banyak lensa suatu teropong, semakin besar perbesaran angularnya.	15

Four-tiers test yang digunakan dalam penelitian ini terdiri dari empat tingkatan. Pada tingkat pertama (*first-tier*) berupa pertanyaan terkait konsep

ilmiah alat optik dengan beberapa pilihan jawaban. Pada tingkat kedua (*second-tier*) berupa tingkat keyakinan peserta didik memilih jawaban pada tingkat pertama. Pada tingkat ketiga (*third-tier*) berupa alasan dari peserta didik dalam menjawab pertanyaan dari tingkat pertama. Pada tingkat keempat (*fourth-tier*) berupa tingkat keyakinan dari peserta didik memilih alasan jawaban di tingkat ketiga.

Pilihan jawaban pada tingkat ketiga (*third-tier*) diperoleh dari tes *four-tiers* (*open-ended*) yang disebarkan pada peserta didik. Pada *four-tiers* (*open-ended*), soal pada tingkat ketiga berbentuk terbuka sehingga peserta didik dapat menulisaan masing-masing alasan terkait konsep yang peserta didik pahami. Jawaban peserta didik pada tingkatan ketiga kemudian dipilah dan diklasifikasikan sehingga dapat digunakan pada soal *four-tiers* (*closed-ended*). Contoh hasil pengembangan *four-tiers* (*open-ended*) menjadi *four-tiers* (*closed-ended*) seperti yang ditunjukkan pada Gambar 3.8 dan Lampiran A.2.

Soal 6.1
Ana hendak memotret suatu objek berupa bunga dengan menggunakan kamera. Ketika memotret, Ana mencoba dua posisi kamera yaitu *landscape* dan *portrait* dengan kamera yang sama. Selanjutnya Ana mendapatkan gambar potret seperti Gambar (a) dan (b).

Gambar (a). Hasil foto pada posisi *landscape*

Gambar (b). Hasil foto pada posisi *portrait*

Berdasarkan hasil potret yang dilakukan Ana, menunjukkan bahwa jarak kamera terhadap objek ketika kamre berposisi ….

A. *landscape* lebih dekat daripada posisi *portrait*
B. *portrait* lebih dekat daripada posisi *landscape*
C. *portrait* dan *landscape* adalah sama
D. *portrait* dan *landscape* adalah tidak dapat ditentukan
E. *portrait* dan *landscape* adalah sejajar

Soal 6.2
Apakah Anda yakin dengan jawaban Anda untuk soal 6.1?

A. Yakin
B. Tidak yakin

Soal 6.3
Alasan untuk jawaban soal 6.1:
…….....
…….....

Question 6.4
Apakah Anda yakin dengan jawaban Anda untuk soal 6.4?

A. Yakin
B. Tidak yakin

Kunci Jawaban: 6.1: C dan 6.3: E

Gambar 3.8 (a)

Soal 6.1
Ana hendak memotret suatu objek berupa bunga dengan menggunakan kamera. Ketika memotret, Ana mencoba dua posisi kamera yaitu *landscape* dan *portrait* dengan kamera yang sama. Selanjutnya Ana mendapatkan gambar potret seperti Gambar (a) dan (b).

Gambar (a). Hasil foto pada posisi *landscape*

Gambar (b). Hasil foto pada posisi *portrait*

Berdasarkan hasil potret yang dilakukan Ana, menunjukkan bahwa jarak kamera terhadap objek ketika kamre berposisi ….

A. *landscape* lebih dekat daripada posisi *portrait*
B. *portrait* lebih dekat daripada posisi *landscape*
C. *portrait* dan *landscape* adalah sama
D. *portrait* dan *landscape* adalah tidak dapat ditentukan
E. *portrait* dan *landscape* adalah sejajar

Soal 6.2
Apakah Anda yakin dengan jawaban Anda untuk soal 6.1?

A. Yakin
B. Tidak yakin

Soal 6.3
Alasan untuk jawaban soal 6.1:

A. karena sisi *landscape* memiliki sisi vertikal lebih pendek dari sisi horizontal
B. karena pada kamera *portrait* objek yang ditangkapnya terlihat lebih dekat dan tidak memuat latar yang terlalu banyak sehingga fokus kamera lebih kepada objek saja
C. karena pada gambar *portrait*, menghasilkan gambar yang terlihat lebih dekat
D. karena *portrait* dan *landscape* tidak dapat ditentukan
E. karena yang membedakan adalah *portrait* berorientasi vertikal sedangkan *landscape* berorientasi horizontal. Jadi, jarak kamera terhadap objek adalah sama

Question 6.4
Apakah Anda yakin dengan jawaban Anda untuk soal 6.4?

A. Yakin
B. Tidak yakin

Kunci Jawaban: 6.1: C dan 6.3: E

Gambar 3.8 (b)

Gambar 3.8. (a) *Four-tiers* (*open-ended*) dan (b) *Four-tiers* (*closed-ended*)

Sebelum digunakan sebagai instrumen dalam penelitian, tes *four-tier* yang telah dikembangkan kemudian dilakukan validasi ahli dan validasi empiris. Pada

penelitian ini, instrumen tes *four-tier* digunakan sebanyak dua kali, yakni pada saat *pre-test* dan *post-test.* Pada saat *pre-test, four-tiers test* digunakan untuk mengetahui konsepsi awal peserta didik sebelum dilakukan pembelajaran POEAW, sedangkan pada saat *post-test, four-tiers test* digunakan untuk mengetahui konsepsi peserta didik setelah dilakukan pembelajaran.

3.4.2.1 Validasi Ahli (*Expert Judgment*)

Validasi ahli pada instrumen soal *four-tier* dilakukan oleh lima orang ahli. Ahli tersebut yaitu dosen pendidikan fisika, guru fisika, dan peneliti dibidang yang sama. Validasi yang dilakukan ahli terhadap setiap butir soal yang mengacu pada lembar validasi pada Lampiran B.2. Lembar validasi pada instrumen *four-tier* berbentuk sama dengan instrumen literasi fisika, yaitu daftar cek dengan kriteria "valid tanpa revisi", "valid dengan revisi", dan juga "tidak valid". Setiap butir soal *four-tier* juga dinilai berdasarkan sembilan indikator seperti Tabel 3.12.

Tabel 3.12. Indikator Penilaian Butir *Four-tier*

Nomor Indikator penilian	Deskripsi Indikator
1	Butir soal yang dibuat sesuai dengan miskonsepsi
2	Kesesuaian konsep dalam butir soal dengan konsep yang dikemukakan oleh para ahli.
3	Butir soal dibuat untuk mengetahui pemahaman konsep peserta didik.
4	Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia.
5	Bahasa yang digunakan mudah dimengerti oleh peserta didik.
6	Pilihan jawaban dan alasan homogen serta logis dari segi materi.
7	Hanya ada satu kunci jawaban.
8	Soal tidak memberikan petunjuk jawaban ke arah jawaban yang benar.
9	Pilihan jawaban tidak menggunakan pernyataan "semua jawaban benar" atau "jawaban salah".

Hasil penilaian ahli kemudian dianalisis dengan *Content Validity Ratio* (CVR). CVR dihitung dengan menggunakan persamaan berikut:

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \tag{3.2}$$

Simbol n_e pada persamaan di atas bermakna jumlah validator yang menilai "valid tanpa revisi*"* atau "valid dengan revisi" pada butir soal *four-tier*, sedangkan N memiliki makna jumlah seluruh validator. Nilai CVR kritis uji satu pihak pada

jumlah validator lima orang yaitu 0,736 dengan signifikansi 0,05. Jika nilai CVR indeks yang diperoleh lebih dari CVR kritis, maka dapat diinterpretasikan "sesuai" (Wilson, et al., 2012). Hasil analisis dari validasi ahli menggunakan CVR dihasilkan seperti pada Tabel 3.13 berikut.

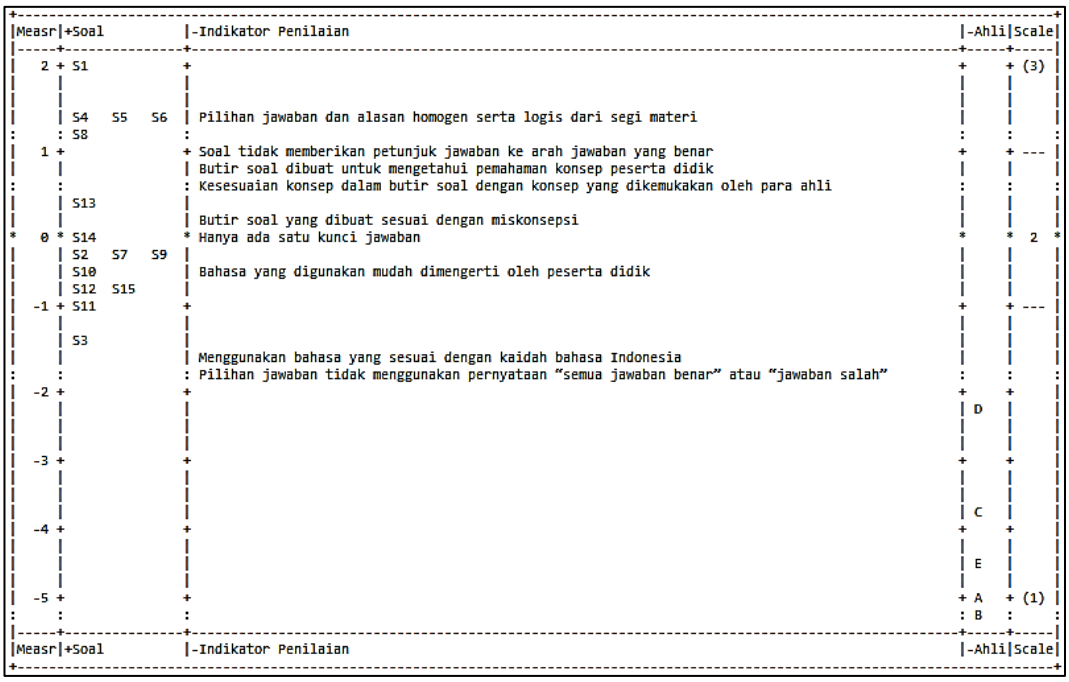
Tabel 3.13. Hasil Validasi Ahli pada Butir Four-tier Menggunakan CVR

Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi	Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi
S1	1	5	5	1	sesuai	S9	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S2	1	5	5	1	sesuai	S10	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	4	0,6	diperbaiki
	9	5	5	1	sesuai		9	5	5	1	sesuai
S3	1	5	5	1	sesuai	S11	1	5	5	1	sesuai
	2	5	4	0,6	diperbaiki		2	5	5	1	sesuai
	3	5	4	0,6	diperbaiki		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	4	0,6	diperbaiki		8	5	4	0,6	diperbaiki
	9	5	5	1	sesuai		9	5	5	1	sesuai
S4	1	5	5	1	sesuai	S12	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai

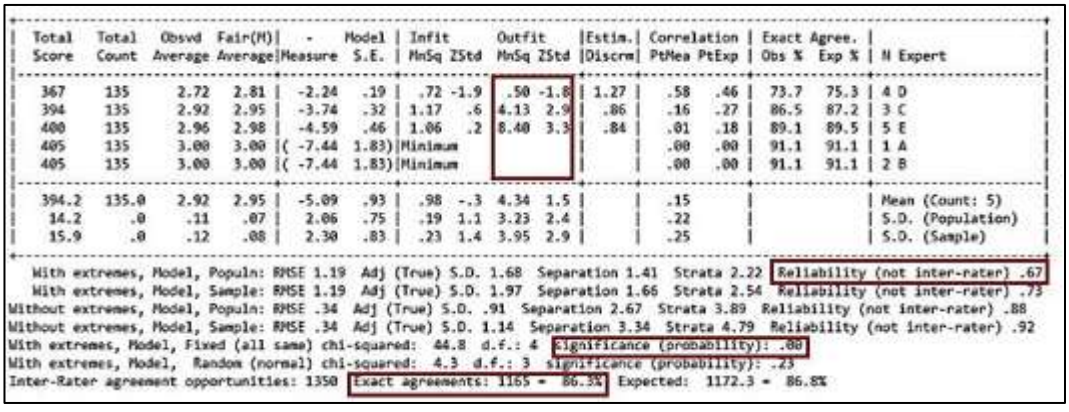
Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi	Butir	Nomor Aspek Penilaian	N	n_e	CVR indeks	Interpretasi
	9	5	5	1	sesuai		9	5	5	1	sesuai
S5	1	5	5	1	sesuai	S13	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S6	1	5	5	1	sesuai	S14	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	5	1	sesuai
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S7	1	5	5	1	sesuai	S15	1	5	5	1	sesuai
	2	5	5	1	sesuai		2	5	5	1	sesuai
	3	5	5	1	sesuai		3	5	4	0,6	diperbaiki
	4	5	5	1	sesuai		4	5	5	1	sesuai
	5	5	5	1	sesuai		5	5	5	1	sesuai
	6	5	5	1	sesuai		6	5	5	1	sesuai
	7	5	5	1	sesuai		7	5	5	1	sesuai
	8	5	5	1	sesuai		8	5	5	1	sesuai
	9	5	5	1	sesuai		9	5	5	1	sesuai
S8	1	5	5	1	sesuai						
	2	5	5	1	sesuai						
	3	5	5	1	sesuai						
	4	5	5	1	sesuai						
	5	5	5	1	sesuai						
	6	5	5	1	sesuai						
	7	5	5	1	sesuai						
	8	5	5	1	sesuai						
	9	5	5	1	sesuai						

Berdasarkan hasil pada Tabel 3.13 menunjukkan bahwa pada butir soal nomor S1, S2, S4, S5, S6, S7, S8, S9, S12, S13, dan S14 memiliki nilai CVR indeks lebih dari 0,736 pada semua indikator penilaian. Oleh karena itu, butir

nomor S1, S2, S4, S5, S6, S7, S8, S9, S12, S13, dan S14 dapat diinterpretasikan "sesuai" dan dapat digunakan tanpa adanya perbaikan (Wilson, et al., 2012). Sedangkan pada butir soal nomor S3, S10, S11, dan S15 memerlukan perbaikan sebab terdapat beberapa indikator penilaian pada butir yang memiliki nilai CVR indeks kurang dari CVR kritis. Butir S3, S10, S11, dan S15 perlu diperbaiki sesuai dengan saran dari ahli sebelum dilakukan validasi empiris. Selain dilakukan analisis melalui CVR, hasil penilaian ahli juga dianalisis mengunakan multi-rater. Hasil analisis penilaian ahli mengunakan multi-rater ditampilkan pada Gambar 3.9 dan Gambar 3.10.



Gambar 3.9. Hasil Multi-rater Validasi Ahli pada *Four-Tier Test*



Gambar 3.10. Hasil Pengukuran Ahli pada *Four-Tier Test*

Gambar 3.9 menunjukkan bahwa terdapat delapan butir soal yang dinilai kurang baik oleh para ahli, yaitu butir soal nomor S2, S7, S9, S10, S12, S15, S11, dan S3. Sementara itu, ahli menilai tujuh butir soal yang memenuhi syarat, yaitu S14, S13, S8, S4, S5, S6, dan S1. Kolom ketiga pada Gambar 3.9 di atas menunjukkan informasi tingkat kesulitan indikator penilaian. Pembagian indikator penilaian ditunjukkan pada kolom ketiga. Semakin rendah nilai logit indikator penilaian maka indikator penilaian semakin mudah memenuhi suatu butir. Semakin tinggi nilai logit suatu indikator penilaian, maka semakin sulit indikator penilaian tersebut terpenuhi pada suatu butir soal yang dievaluasi oleh validator (ahli). Indikator penilaian dengan nilai logit sama memiliki tingkat kesulitan yang sama.

Berdasarkan Gambar 3.9 indikator penilaian 4 (menggunakan bahasa yang mengikuti kaidah bahasa Indonesia) dan 9 (pilihan jawaban tidak menggunakan pernyataan; semua jawaban benar atau jawaban salah) merupakan indikator penilaian yang paling mudah karena semua soal memuaskan menurut peniliaan ahli. Sedangkan indikator penilaian 6 (pilihan jawaban dan alasan homogen dan logis dari segi materi) merupakan indikator penilaian yang paling sulit, mayoritas soal tidak memenuhinya menurut pendapat ahli. Menurut semua pendapat ahli, I1 merupakan satu-satunya soal yang memenuhi semua indikator penilaian.

Gambar 3.10 menggambarkan kualitas penilaian ahli yang diurutkan berdasarkan tingkat keparahan butir. Ahli D adalah yang paling konsisten kriteria kevalidan (Boone et al., 2014; Darmana et al., 2021). Ahli D memiliki nilai Outfit MNSQ dan Outfit ZSTD dalam rentang yang sesuai yaitu 0,5-1,5 (Outfit MNSQ) dan -2 hingga +2 (Outfit ZSTD). Ahli A dan B adalah yang terburuk karena memiliki nilai infit yang paling rendah. Reliabilitas antar ahli cukup yaitu 0,67, menunjukkan bahwa para ahli memberikan skor yang cukup berbeda, tetapi ada juga yang sama (Koçak, 2020). Kecenderungan penilai mempengaruhi nilai reliabilitas (Bond & Fox, 2013; Darmana et al., 2021). Data yang dihasilkan konsisten dengan model pengukuran yang didukung oleh nilai uji *Chi-square* ($p<0,01$). Kesamaan penilaian oleh lima ahli (kesepakatan yang benar) mencapai 86,3%, menunjukkan bahwa ada sedikit perbedaan penilaian semua item oleh kelima ahli. Beberapa studi dalam literatur juga menunjukkan bahwa

kecenderungan penilaian penilai berbeda dan bahwa perilaku penilai seperti kelonggaran dan keparahan mempengaruhi keandalan penilai (Brookhart et al., 2006; Güler, 2014; Darmana et al., 2021).

3.4.2.2 Validitas Instrumen

a) Validitas butir soal

Validitas butir soal pada instrumen *four-tier* dilakukan dengan tujuan untuk mengetahui butir *four-tier* yang telah dikembangkan dapat mengkur apa yang sedang diukur. Analisis validitas butir soal pada *four-tier* dilakukan dengan menggunakan *Rasch Model* dengan bantuan perangkat lunak MINISTEP. Analisis validitas butir soal dengan menggunakan *Rasch Model* yaitu dengan memperhatikan keluaran dari *item fit order*. Keluaran *item fit order* dianalisis sesuai kriteria *outfit MNSQ*, *outfit ZSTD*, dan *Pt mean corr* (Bonee et al., 2014; Bond & Fox, 2015; Sumintono & Widhiarso, 2015). Ketiga kriteria tersebut dirinci seperti pada Tabel 3.14.

Tabel 3.14. Kriteria Validitas Butir Soal

Outfit	**Nilai yang diterima**
Outfit mean square (MNSQ)	0,5 < MNSQ < 1,5
Outfit Z-standart (ZSTD)	-2,00 < ZSTD < +2.00
Point Measure Correlation (Pt Mean Corr)	0,4 < *Pt Measure Corr* <0,85

(Sumintono & Widhiarso, 2014)

Hasil keluaran *item fit order* dari instrumen *four-tier* yang telah di analisis dengan *Rasch Model* ditunjukkan seperti Gambar 3.11.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
1	72	75	.03	.12	1.15	.74	1.46	1.57	A .29	.47	45.3	48.0	S1
6	49	75	.47	.16	1.44	1.40	1.39	1.37	B .49	.37	62.7	58.0	S6
8	87	75	-.18	.11	1.08	.47	1.42	1.51	C .62	.52	37.3	45.9	S8
14	103	75	-.37	.10	.95	-.23	1.34	1.30	D .53	.56	41.3	41.6	S14
9	120	75	-.54	.10	1.25	1.45	1.31	1.27	E .49	.59	40.0	40.5	S9
3	209	75	-1.35	.10	1.25	1.59	1.15	.58	F .51	.59	37.3	35.6	S3
7	114	75	-.48	.10	1.23	1.35	1.14	.62	G .50	.58	42.7	40.3	S7
13	131	75	-.65	.10	1.10	.68	1.20	.89	H .56	.61	37.3	35.5	S13
11	54	75	.35	.15	1.07	.35	1.03	.20	g .48	.39	57.3	56.9	S11
12	88	75	-.19	.11	.99	-.02	.96	-.07	f .61	.52	37.3	46.0	S12
15	31	75	1.10	.22	.91	-.21	.96	-.08	e .51	.29	80.0	64.3	S15
5	31	75	1.10	.22	.85	-.42	.94	-.20	d .21	.29	64.0	64.3	S5
4	75	75	-.02	.12	.81	-.91	.89	-.32	c .49	.48	50.7	47.4	S4
10	49	75	.47	.16	.64	-1.31	.66	-1.35	b .57	.37	70.7	58.0	S10
2	58	75	.27	.14	.52	-2.21	.64	-1.42	a .40	.41	60.0	55.8	S2
MEAN	84.7	75.0	.00	.13	1.02	.18	1.10	.39			50.9	49.2	
P.SD	44.9	.0	.63	.04	.24	1.06	.25	.94			13.5	9.4	

Gambar 3.11. *OUTFIT MNSQ, OUTFIT ZSTD,* dan *PT MEASURE CORR* setiap butir soal *four-tier*

Butir soal dapat dikatakan sangat layak atau "sangat sesuai" apabila ketiga kriteria *outfit* MNSQ, *outfit* ZSTD, dan *Pt measure corr* telah terpenuhi. Sedangkan jika hanya salah satu saaja kriteria yang terpenuhi, maka butir dianggap masih "sesuai" dan dipertahankan (Sumintono & Widhiarso, 2015; Nurdini et al., 2020; Dewi et al., 2021). Namun, jika semua kriteria tidak terpenuhi, maka butir tidak layak atau "tidak sesuai". Butir-butir yang tidak sesuai tersebut perlu diperbaiki (Nurdini et al., 2020; Dewi et al., 2021).

Berdasarkan Gambar 3.11 dengan mengacu pada kriteria di Tabel 3.14, butir S1, S5, dan S2 hanya memenuhi dua kriteria yaitu pada *outfit* MNSQ dan *outfit* ZSTD, sehingga butir tersebut dikatakan "sesuai" dan dapat dipertahankan. Sedangkan butir lainnya memenuhi ketiga kriteria, sehingga butir-butir tersebut dapat dikatakan "sangat layak". Dengan demikian dapat disimpulkan bahwa semua butir *four-tier* dapat digunakan dan tidak perlu diubah.

b) Unidimensionalitas instrumen

Unidimensionalitas instrumen *four-tier* yaitu ukuran yang penting untuk mengevaluasi apakah instrumen *four-tier* mampu mendeteksi miskonsepsi peserta didik. Pengujian unidimensionalitas dilakukan dengan memperhatikan *raw variance explained by measures* dan *unexplained variance* (Sumintono & Widhiarso, 2015; Nurdini et al, 2020). *Eigen value unexplained variance* pada uji unidimensionalitas harus bernilai kurang dari 3. Sedangkan *unexplained variance*

pada *observed value* bernilai kurang dari 15%. Unidimensionalitas pada *raw vaiance explained by measures* diinterpretasikan seperti pada Tabel 3.15.

Tabel 3.15. Interpretasi *raw vaiance explained by measures*

Nilai raw variance explained by measures (V)	Interpretasi
$V < 20\%$	Tidak terpenuhi
$20\% \leq V \leq 40\%$	Terpenuhi
$40\% < V \leq 60\%$	Bagus
$V > 60\%$	Istimewa

Hasil keluaran untuk analisis unidimentionalitas instrumen *four-tier* menggunakan *software* MINISTEP ditampilkan pada Gambar 3.12.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units

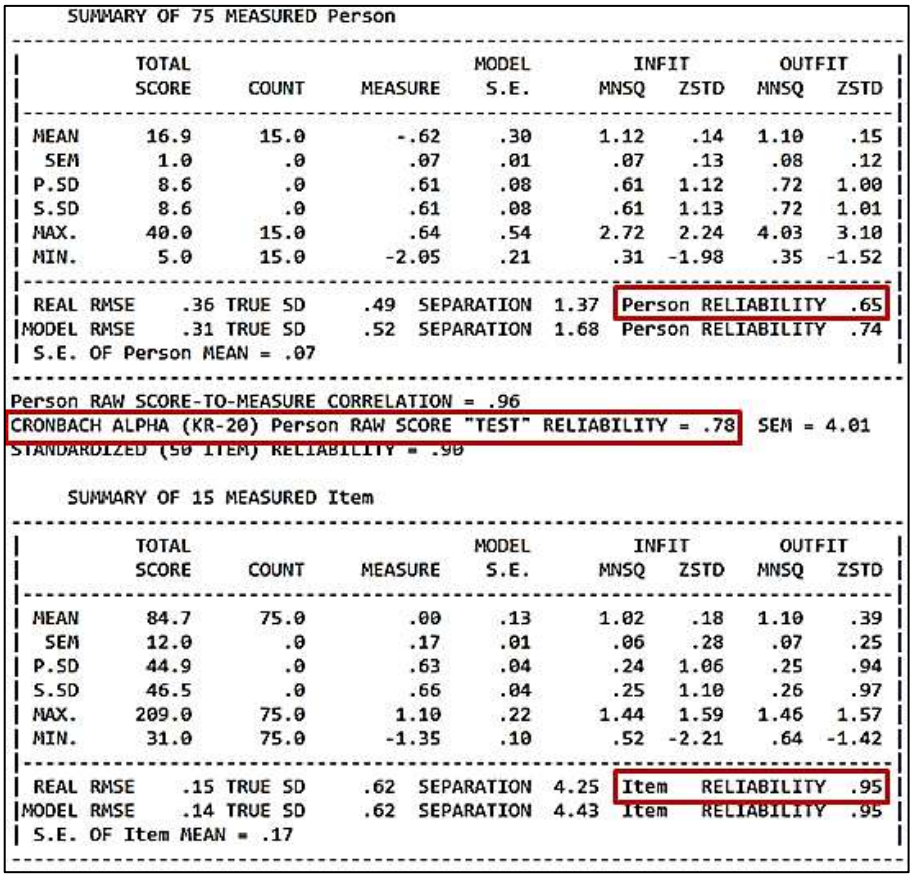
	Eigenvalue	Observed		Expected
Total raw variance in observations =	26.3313	100.0%		100.0%
Raw variance explained by measures =	11.3313	43.0%		46.1%
Raw variance explained by persons =	1.4882	5.7%		6.1%
Raw Variance explained by items =	9.8432	37.4%		40.1%
Raw unexplained variance (total) =	15.0000	57.0%	100.0%	53.9%
Unexplned variance in 1st contrast =	2.1271	8.1%	14.2%	
Unexplned variance in 2nd contrast =	1.9267	7.3%	12.8%	
Unexplned variance in 3rd contrast =	1.5892	6.0%	10.6%	
Unexplned variance in 4th contrast =	1.4282	5.4%	9.5%	
Unexplned variance in 5th contrast =	1.1879	4.5%	7.9%	

Gambar 3.12. Hasil analisis unidimentionalitas *four-tier*

Berdasarkan Gambar 3.12, nilai *raw variance explaianed by measures* pada instrumen *four-tier* memiliki nilai 43,0% sehingga dapat dinterpretasikan bagus. Sedangkan semua *eigen value unexplained variance* bernilai kurang dari 3 dan *observed* juga berniali kurang dari 15%. Dengan demikian, *unexplained variance* pada *four-tier* memenuhi kriteria pada uji dimensionalitas. Simpulan akhir dari uji unidimensionalitas adalah instrumen *four-tier* valid untuk mengetahui miskonsepsi peserta didik.

3.4.2.3 Uji Reliabilitas

Uji reliabilitas pada instrumen *four-tier* berbantuan *software* MINISTEP pada keluaran *summary statistics* yang dianalisiis dengan *Rasch Model*. Analisis reliabilitas memperhatikan beberapa nilai yaitu *item reliability, person reliability,* dan *Cronbach Alpha*. Hasil uji reliabilitas soal instrumen *four-tier* disajikan pada Gambar 3.13.



SUMMARY OF 75 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	16.9	15.0	-.62	.30	1.12	.14	1.10	.15
SEM	1.0	.0	.07	.01	.07	.13	.08	.12
P.SD	8.6	.0	.61	.08	.61	1.12	.72	1.00
S.SD	8.6	.0	.61	.08	.61	1.13	.72	1.01
MAX.	40.0	15.0	.64	.54	2.72	2.24	4.03	3.10
MIN.	5.0	15.0	-2.05	.21	.31	-1.98	.35	-1.52

REAL RMSE .36 TRUE SD .49 SEPARATION 1.37 Person RELIABILITY .65
MODEL RMSE .31 TRUE SD .52 SEPARATION 1.68 Person RELIABILITY .74
S.E. OF Person MEAN = .07

Person RAW SCORE-TO-MEASURE CORRELATION = .96
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .78 SEM = 4.01
STANDARDIZED (50 ITEM) RELIABILITY = .90

SUMMARY OF 15 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	84.7	75.0	.00	.13	1.02	.18	1.10	.39
SEM	12.0	.0	.17	.01	.06	.28	.07	.25
P.SD	44.9	.0	.63	.04	.24	1.06	.25	.94
S.SD	46.5	.0	.66	.04	.25	1.10	.26	.97
MAX.	209.0	75.0	1.10	.22	1.44	1.59	1.46	1.57
MIN.	31.0	75.0	-1.35	.10	.52	-2.21	.64	-1.42

REAL RMSE .15 TRUE SD .62 SEPARATION 4.25 Item RELIABILITY .95
MODEL RMSE .14 TRUE SD .62 SEPARATION 4.43 Item RELIABILITY .95
S.E. OF Item MEAN = .17

Gambar 3.13.*Person reliability, Item reliability,* dan *Cronbach Alpha* pada instrumen *four-tier*

Hasil *person reliability* dan *butir reliability* dapat diinterpretasikan sesuai pada Tabel 3.16.

Tabel 3.16. Kategori Reliabilitas

Nilai *person reliability* dan *item reliability*	**Kategori**
r < 0,67	Lemah
0,67 ≤ r < 0,80	Cukup
0,80 ≤ r < 0,90	Bagus
0,90 ≤ r < 0,94	Bagus Sekali
0,94 < r	Istimewa

(Sumintono & Widhiarso, 2015)

Berdasarkan Gambar 3.13, *person reliability* bernilai 0,65 dan termasuk pada kategori lemah, sedangkan *item reliability* bernilai 0,95 dan termasuk pada kategori istimewa. Nilai *Cronbach Alpha* juga dapat tafsirkan ke dalam beberapa kriteria sesuai Tabel 3.17.

Tabel 3.17. Kriteria nilai *Alpha Cronbach*

Nilai *Alpha Cronbach*	**Kategori**
r < 0,50	Sangat buruk
0,50 ≤ r < 0,60	Buruk
0,60 ≤ r < 0,70	Cukup
0,70 ≤ r < 0,80	Bagus
0,80 ≤ r	Bagus Sekali

(Sumintono & Widhiarso, 2015)

Bersesuaian dengan Gambar 3.13 dan Tabel 3.17, *Cronbach Alpha* bernilai 0,86 dan masuk kategori bagus sekali. Hasil uji reliabilitas instrumen tes *four-tier* dapat disimpulkan bahwa peserta didik memiliki konsistensi jawaban pada kategori cukup, kualitas butir soal pada *four-tier* bagus, serta interaksi antara peserta didik dan butir secara keseluruhan menunjukkan bagus sekali. Dengan demikian, instrumen tes *four-tier* dapat digunakan dan dikatakan reliabel.

3.4.2.4 Taraf Kesukaran

Taraf kesukaran instrumen *four-tier* dianalisis menggunakan *Rasch Model* dengan program MINISTEP. Taraf kesuakaran butir soal diklasifikasikan dengan memperhatikan keluaran dari *item measure* pada bagian berdasarkan nilai *measure logit* dan nilai Standar Deviasi (SD) (Sumintono & Widhiarso, 2015). Taraf kesukaran butir soal bergantung pada nilai *logit.* Semakin tinggi nilai *logit* suatu butir, maka butir tersebut memiliki taraf kesukaran yang semakin tinggi juga. Klasifikasi dari taraf kesukaran butir dapat interpretasi seperti pada Tabel 3.18.

Tabel 3.18. Interpretasi Taraf Kesukaran Butir Soal

Interpretasi	**Nilai *Measure logit***
Sangat Mudah	$M < -1SD$
Mudah	$-1SD \leq M \leq 0$
Sukar	$0 \leq M \leq SD$
Sangat Sukar	$M > SD$

(Sumintono & Widhiarso, 2015)

Hasil *item measure* untuk taraf kesukaran soal *four-tier* ditunjukkan pada Gambar 3.14 berikut.

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
5	31	75	1.10	.22	.85	-.42	.94	-.20	.21	.29	64.0	64.3	S5
15	31	75	1.10	.22	.91	-.21	.96	-.08	.51	.29	80.0	64.3	S15
6	49	75	.47	.16	1.44	1.40	1.39	1.37	.49	.37	62.7	58.0	S6
10	49	75	.47	.16	.64	-1.31	.66	-1.35	.57	.37	70.7	58.0	S10
11	54	75	.35	.15	1.07	.35	1.03	.20	.48	.39	57.3	56.9	S11
2	58	75	.27	.14	.52	-2.21	.64	-1.42	.40	.41	60.0	55.8	S2
1	72	75	.03	.12	1.15	.74	1.46	1.57	.29	.47	45.3	48.0	S1
4	75	75	-.02	.12	.81	-.91	.89	-.32	.49	.48	50.7	47.4	S4
8	87	75	-.18	.11	1.08	.47	1.42	1.51	.62	.52	37.3	45.9	S8
12	88	75	-.19	.11	.99	-.02	.96	-.07	.61	.52	37.3	46.0	S12
14	103	75	-.37	.10	.95	-.23	1.34	1.30	.53	.56	41.3	41.6	S14
7	114	75	-.48	.10	1.23	1.35	1.14	.62	.50	.58	42.7	40.3	S7
9	120	75	-.54	.10	1.25	1.45	1.31	1.27	.49	.59	40.0	40.5	S9
13	131	75	-.65	.10	1.10	.68	1.20	.89	.56	.61	37.3	35.5	S13
3	209	75	-1.35	.10	1.25	1.59	1.15	.58	.51	.59	37.3	35.6	S3
MEAN	84.7	75.0	.00	.13	1.02	.18	1.10	.39			50.9	49.2	
P.SD	44.9	.0	.63	.04	.24	1.06	.25	.94			13.5	9.4	

Gambar 3.14.Hasil *logit measure* dan standar deviasi (SD) pada *item measure* soal tes *four-tier*

Besesuaian dengan *logit measure* dan standar devisai pada Gambar 3.14, sehingga butir soal *four-tier* dapat diklasifikasikan seperti Tabel 3.19 berikut.

Tabel 3.19. Hasil Analisis Taraf Kesukaran Tes *Four-tier*

Nilai *Measure logit*	Interpretasi Taraf Kesukaran	Butir Soal	Jumlah Butir Soal
$M < -0{,}63$	Sangat Mudah	S3, S13	2
$-0{,}63 \leq M \leq 0$	Mudah	S9, S7, S14, S12, S8, S4	6
$0 \leq M \leq 0{,}63$	Sukar	S1, S2, S11, S10, S6	5
$0{,}63 > SD$	Sangat Sukar	S15, S5	2

Berdasarkan klasifikasi butir pada Tabel 3.6 menerangkan bahwa taraf kesukaran soal tes *four-tier* bervariasi. Butir soal dengan taraf kesukaran "sangat sukar" sebanyak 13,33%, taraf kesukaran "sukar" sebanyak 33,33%, taraf kesukaran "mudah" sebanyak 40%, dan taraf kesukaran "sangat mudah" sebanyak 13,33%. Butir soal yang paling banyak yaitu butir dengan taraf kesukaran "mudah" dan "sulit".

3.4.3. Simulasi Virtual Materi Alat Optik

Simulasi virtual materi alat optik yang digunakan dalam penelitian ini dikembangkan sesuai dengan miskonsepsi yang terjadi pada peserta didik. Simulasi ini dikembangkan dengan berbasis aplikasi android pada *smartphone.* Simulasi ini juga mampu terhubung dengan PhET dan aplikasi *Ray Optics.*

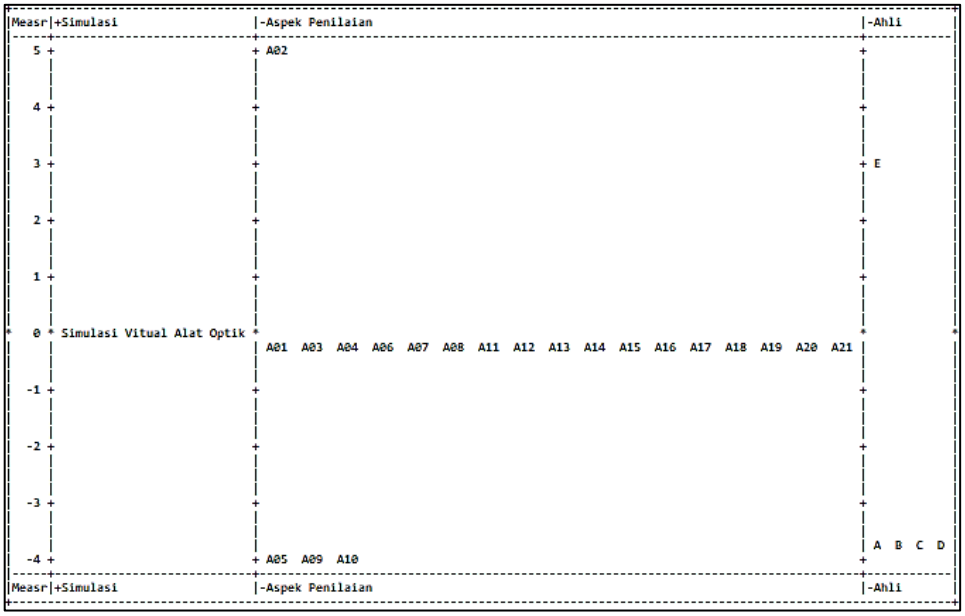
Sebelum digunakan sebagai media yang membantu pembelajaran dalam penelitian ini, simulasi virtual materi alat optik divalidasi oleh ahli terlebih dahulu. Simulasi virtual materi alat optik divalidasi oleh lima orang ahli. Ahli tersebut terdiri atas dosen pendidikan fisika, guru fisika, dan peneliti dibidang yang sama. Validasi yang dilakukan ahli terhadap simulasi virtual materi alat optik mengacu pada lembar validasi (Lampiran B.5). Lembar validasi ini berbentuk daftar cek dengan kriteria "sangat baik", "baik", "kurang", dan juga "sangat kurang". Setiap butir soal *four-tier* juga dinilai berdasarkan aspek pengoprasian, aspek tampilan, aspek materi, dan aspek bahasa. Setiap aspek memiliki beberapa aspek penilian yang dirinci seperti pada Tabel 3.20.

Tabel 3.20. Aspek Penilaian Simulasi Virtual Alat Optik

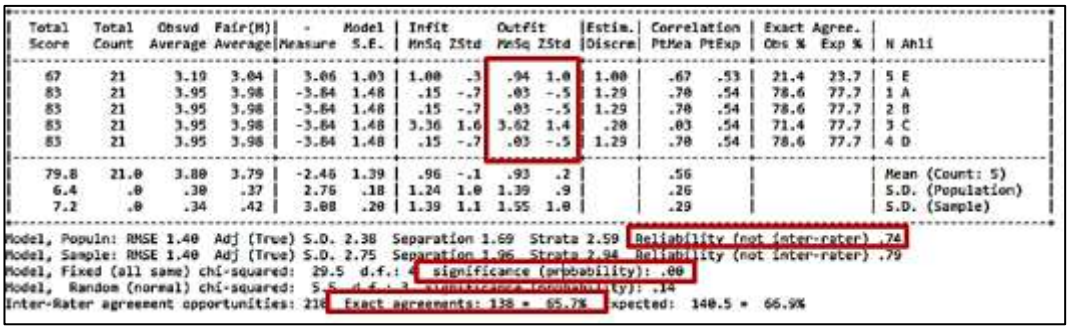
Kode Aspek Penilaian	**Aspek yang dinilai**
Aspek Pengoperasian	
A01	Kelancaran Simulasi Virtual Alat Optik saat digunakan
A02	Kejelasan petunjuk penggunaan Simulasi Virtual Alat Optik
A03	Kemudahan dalam menjalankan Simulasi Virtual Alat Optik melalui *icon* yang disediakan
A04	Kemudahan dalam penggunaan Simulasi Virtual Alat Optik yang dikembangkan (*user friendly*)
Aspek Tampilan	
A05	Kejelasan pembagian materi pada Simulasi Virtual Alat Optik
A06	Urutan dan sistematika penyajian Simulasi Virtual Alat Optik
A07	Pengaturan ruang/tata letak gambar dan wacana simulasi saling berkaiatan
A08	Pengaturan tampilan desain *layout* Simulasi Virtual Alat Optik menarik perhatian peserta didik
A09	Kesesuaian pemilhan gambar/animasi dalam Simulasi Vitual Alat Optik
A10	Kualitas gambar/animasi dalam Simulasi Virual Alat Optik
A11	Kesesuaian pemilihan warna pada Simulasi Virtual Alat Optik
A12	Kesesuaian pemilihan huruf (*font*) pada Simulasi Virtual Alat Optik
Aspek Materi	
A13	Kesesuaian Simulasi Virtual Alat Optik dengan LKPD Materi Alat Optik
A14	Ketersediaan gambar dan simulasi pada Simulasi Virtual Alat Optik mendukung pembelajaran materi alat optik
A15	Keakuratan konsep pada setiap Simulasi Virtual Alat Optik yang disusun
A16	Kemudahan peserta didik dalam menggunakan Simulasi Virtual Alat Optik
A17	Relevansi Simulasi Virtual Alat Optik dengan muatan kurikulum yang berlaku pada pembelajaran fisika

Kode Aspek Penilaian	Aspek yang dinilai
Aspek Bahasa	
A18	Ketepatan struktur kalimat dalam Simulasi Virtual Alat Optik
A19	Bahasa yang digunakan baku sesuai kaidah penulisan dalam Bahasa Indonesia
A20	Tata bahasa dalam Simulasi Virtual Alat Optik mudah dipahami dan tidak membingungkan pembaca
A21	Bahasa yang dimuat pada Simulasi Virtual Alat Optik sesuai dengan tingkat keterbacaan atau jenjang pendidikan SMA

Hasil penilaian ahli kemudian dianalisis mengGunakan Multi-rater. Hasil analisis penilaian ahli menggunakan multi-rater ditampilkan pada Gambar 3.15 dan Gambar 3.16.



Gambar 3.15. Hasil Multi-rater Validasi Ahli pada Simulasi Virtual Materi Alat Optik



Gambar 3.16. Hasil Pengukuran Ahli pada Simulasi Virtual Materi Alat Optik

Gambar 3.15 menunjukkan bahwa pada simulasi virtual menunjukkan memenuhi semua aspek penilaian yaitu A01, A03, A04, A05, A06, A07, A08, A09, A10, A11, A12, A13, A14, A15, A16, A17, A18, A19, A20, dan A21. Namun simulasi virtual belum memenuhi aspek A02 (kejelasan petunjuk penggunaan simulasi virtual alat optik). Sehingga sesuai saran perbaikan dari ahli, kejelasan petunjuk penggunaan pada simulasi diperbaiki terlebih dahulu sebelum digunakan dalam pembelajaran. Kolom ketiga pada Gambar 3.15 di atas menunjukkan informasi dari tingkat kesulitan aspek penilaian. Semakin rendah nilai logit dari suatu aspek penilian maka aspek tersebut semakin mudah dipenuhi oleh simulasi virtual yang dikembangkan. Semakin tinggi nilai logit suatu aspek penilaian, maka semakin sulit aspek penilaian tersebut terpenuhi oleh simulasi virtual sesuai penilian dari validator (ahli). Aspek penilaian yang memiliki nilai logit, maka memiliki tingkat kesulitan yang sama. Berdasarkan Gambar 3.15 aspek penilaian A05, A09, dan A010 merupakan aspek penilaian yang paling mudah karena semua soal memuaskan menurut peniliaan ahli. Sedangkan aspek A02 merupakan aspek penilaian yang paling sulit dipenuhi oleh simulasi virtual.

Gambar 3.16 menggambarkan kualitas penilaian ahl. Ahli A, B, D, dan E adalah ahli yang konsisten kriteria pada kevalidan. Ahli A, B, D, dan E memiliki nilai *Outfit* MNSQ dan *Outfit* ZSTD dalam rentang yang sesuai yaitu 0,5-1,5 (*Outfit* MNSQ) dan -2 hingga +2 (*Outfit* ZSTD) (Boone et al., 2014; Darmana et al., 2021). Ahli C menunjukkan nilai *Outfit* MNSQ yang berada diluar rentang, namun ahli C memiliki *Outfit* ZSTD memenuhi rentang sehingga ahli ini masih dapat dikatakan valid. Reliabilitas antar ahli cukup yaitu 0,74, menunjukkan bahwa para ahli memberikan skor yang cukup berbeda, tetapi ada juga yang sama (Koçak, 2020). Kecenderungan penilai mempengaruhi nilai reliabilitas (Bond & Fox, 2013; Darmana et al., 2021). Data yang dihasilkan konsisten dengan model pengukuran yang didukung oleh nilai uji Chi-square ($p<0,01$). Kesamaan penilaian oleh lima ahli (kesepakatan yang benar) mencapai 65,75%, menunjukkan bahwa ada sedikit perbedaan penilaian dari para ahli. Beberapa studi dalam literatur juga menunjukkan bahwa kecenderungan penilaian penilai berbeda dan bahwa perilaku penilai seperti kelonggaran dan keparahan

mempengaruhi keandalan penilai (Brookhart et al., 2006; Güler, 2014; Darmana et al., 2021).

3.5 Prosedur Penelitian

Prosedur penelitian ini dibagi menjadi tiga tahapan, yaitu tahap persiapan, tahap pelaksanaan, dan tahap analisis. Tahap pada prosedur penelitian masing-masing dideskripsikan sebagai berikut.

1) Tahap Awal

Pada tahap awal, peneliti menggali informasi dengan melakukan studi literatur terkait kemampuan literasi fisika, miskonsepsi, serta melakukan penelaahan kurikulum terhadap materi fisika di SMA berdasarakan kurikulum 2013 revisi. Selain studi literatur, peneliti juga melakukan studi pendahuluan dengan melakukan observasi secara langsung kepada peserta didik dengan menyebarkan soal tes literasi fisika dan *four-tier* instrumen. Hasil studi pendahuluan digunakan untuk mengidentifikasi kondisi kemampuan literasi fisika dan miskonsepsi peserta didik. Kemudian merumuskan inti masalah sesuai dengan kajian pustaka dan hasil yang diperoleh dari studi pendahuluan. Selanjutnya, peneliti merumuskan solusi yang dapat diterapkan dalam pembelajaran dari permasalahan yang dihadapi pada literasi fisika dan miskonsepsi melalui pembelajaran POEAW berbantuan simulasi virtual. Penulis juga mengembangkan RPP dengan model POEAW berbantuan simulasi virtual materi alat optik, instrumen literasi fisika materi alat optik, instrumen *four-tier* materi alat optik, LKPD POEAW materi alat optik, serta simulasi virtual materi alat optik.

Penyusunan instrumen terdiri dari perumusan indikator, sebaran kontruksi soal, penyusunana butir sesuai dengan indikator dan kontruksi soal. Setelah instrumen disusun, kemudian instrumen divalidasi ahli oleh lima ahli yang terdiri dari dosen pendidikan fiska, guru fisika, dan peneliti pada bidang yang bersesuaian. Kemudian instrumen yang telah divalidasi ahli diperbaiki sesuai dengan saran dan masukan dari ahli. Setelah dilakukan perbaikan, instrumen divalidasi empiris melalui uji coba soal kepada peserta didik. Peserta didik dalam uji coba yaitu peserta didik yang telah mempelajarai materi alat optik. Hasil uji coba dilakukan analisis instrumen yang terdiri dari validitas instrumen, reliabilitas

instrumen, dan taraf kesukaran butir. Sesuai dengan analisis instrumen, kemudian instrumen dievaluasi dan diperbaiki.

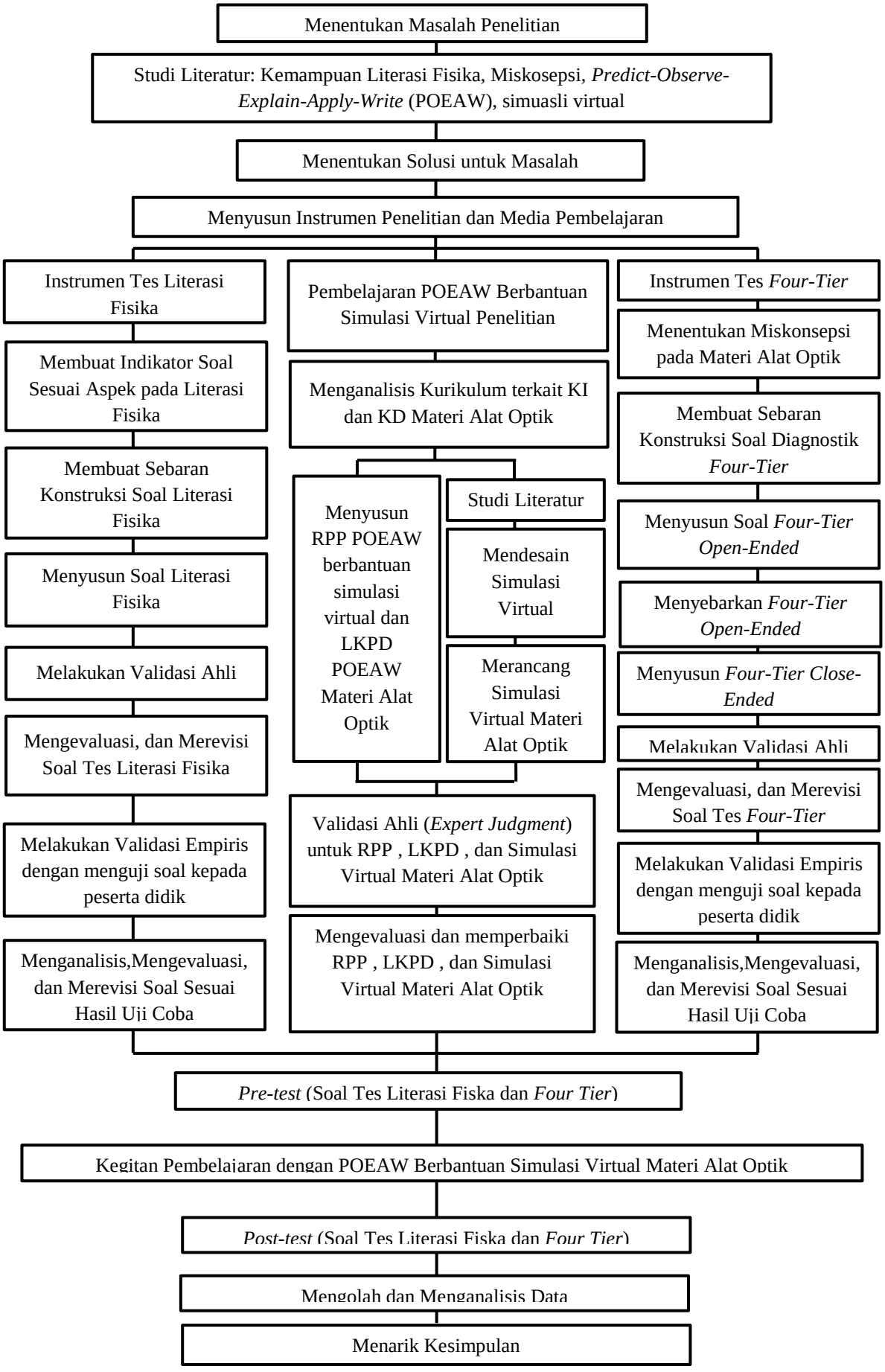
Pengembangan simulasi virtual materi alat optik disesuaikan dengan KI dan KD materi alat optik, serta miskonsepsi yang dialami oleh peserta didik berdasarkan studi literatur. Setiap miksonsepsi dikelompokkan berdasarkan sub materi. Kemudian setiap simulasi pada setiap miskonsepsi ditentukan dan didesain pada powerpoint untuk mempermudah. Hasil desain kemudian dirancang menjadi sebuah aplikasi berbasis android pada *smartphone*. Aplikasi juga dirancang dan mampu dihubungkan pada PhET dan *Ray Optics*. Hasil simulasi virtual materi alat optik kemudian divalidasi oleh lima ahli. Bersesuaian dengan hasil validasi ahli, simulasi virtual materi alat optik diperbaiki.

2) Tahap Pelaksanaan

Tahap pelaksanaan dimulai dari melakukan *pre-test* kepada peserta didik untuk mendapatkan data kemampuan awal literasi fisika dan konsepsi awal peserta didik kepada kelompok eksperimen dan kelompok kontrol. Kemudian memberikan perlakuan kepada peserta didik. Kelompok eksperimen diberi pembelajaran POEAW dengan berbantuan simulasi virtual, sedangkan kelompok kontrol diberi pembelajaran POEAW tanpa simulasi virtual. Selama perlakuan, terdapat pengamat yang mengamati pembelajaran POEAW yang dilaksanankan dan pengamat juga mengisi lembar keterlaksanaan pembelajaran yang dilakukan. Setelah dilakukan perlakuan, kelompok eksperimen dan kelompok kontrol diberikan *post-test* untuk mengetahui kemampuan literasi fisika dan konsepsi akhir peserta didik. Kelompok kontrol dan kelompok eksperimen bertujuan untuk mengukur seberapa besar efektivitas model POEAW berbantuan simulasi virtual terhadap miskonsepsi peserta didik.

3) Tahap Analisis

Pada tahap analisis, peneliti melakuan pengolahan data kemampuan literasi fisika peserta didik, miskonsepsi peserta didik, dan karakteristik perubahan miskonsepsi peserta didik. Selanjutnya, melakukan analisis hasil pengolahan data penelitian. Hasil analisis kemudian digunakan untuk membuat kesimpulan, implikasi, dan rekomendasi.



Gambar 3.17. Alur prosedur penelitian

3.6 Teknik Analisis Data

Analisis data dilakukan guna memperoleh interpretasi dari data yang telah didapatkan. Sebagai upaya memperoleh hasil analisis diperlukan beberapa teknik sebagai berikut:

3.6.1. Analisis Peningkatan Literasi Fisika

Data yang telah diperoleh dari instrumen tes literasi fisika dari kegiatan *pre-test* dan *post-test* digunakan untuk menentukan *n-gain*. Analisis *n-gain* dilakukan guna mengetahui besar peningkatan literasi fisika peserta didik. Skor *n-gain* dihitung dengan menggunakan rumus berikut:

$$\langle g \rangle = \frac{\%<S_f> - \%<S_i>}{100 - \%<S_i>} \quad (3.4)$$

(Hake,1999)

Keterangan:

$\langle g \rangle$ = skor *n-gain* peningkatan literasi fisika

$\%<S_i>$ = rata-rata persentase skor *pre-test* literasi fisika

$\%<S_f>$ = rata-rata persentase skor *post-test* literasi fisika

Skor *n-gain* yang telah diperoleh dikategorikan sesuai Tabel 3.21.

Tabel 3.21. Kategori Skor *n-gain*

Skor	Kategori
$\langle g \rangle \geq 0{,}7$	Tinggi
$0{,}3 \leq \langle g \rangle < 0{,}7$	Sedang
$\langle g \rangle < 0{,}3$	Rendah

3.6.2. Analisis Kuantitas Penurunan Miskonsepsi Peserta Didik

Data yang diperoleh dari *four-tiers test* yang dilakukan pada *pre-test* dan *post-test* terlebih dahulu yaitu dengan melakukan pengkategorian konsepsi yang dimiliki peserta didik. Setiap kategori dinilai berdasarkan nilai miskonsepsi (Aminudin et. al, 2019). Kemudian data tersebut diolah menggunakan persamaan *gain* ternormalisasi yang telah diadaptasi dari Hake (1999) dan Kurniawan, Suhandi, & Hasanah (2016) dan Muliyani (2018) untuk mengetahui penurunan miskonsepsi saat *post-test* dibandingkan saat *pre-test*. Berikut persamaan kuantitas penurunan dari miskonsepsi:

$$< KPM >= \frac{\%<S_i>-\%<S_f>}{\%<S_i>-0} \quad (3.5)$$

Keterangan:

$< KPM >$ = rata-rata kuantitas penurunan miskonsepsi

$\% < S_i >=$ rata-rata persentase miskonsepsi *pre-test*

$\% < S_f >=$ rata-rata persentase miskonsepsi *post-test*

Penurunan kuantitas miskonsepsi dikategorisasikan seperti pada Tabel 3.22.

Tabel 3.22. Kategori Penurunan Miskonsepsi

Nilai Penurunan Miskonsepsi	Kategori
$KPM \geq 0{,}7$	Tinggi
$0{,}3 \leq KPM < 0{,}7$	Sedang
$KPM < 0{,}3$	Rendah

3.6.3. Analisis Karakteristik Penurunan Miskonsepsi

Karakteristik penurunan miskonsepsi dianalisis dengan melihat perubahan konsepsi yang dialami oleh peserta didik. Perubahan konsepsi dilihat berdasarkan kodefikasi konsepsi peserta didik untuk setiap konsep alat optik pada konsepsi awal dan konsepsi akhir. Perubahan konsepsi perserta didik memiliki beberapa kemungkinan seperti pada Tabel 3.23.

Tabel 3.23. Kemungkinan perubahan konsepsi

Konsepsi Awal	Arah Perubahan Konsepsi	Konsepsi Akhir
Misconception (MC)		*Sound Understanding* (SU)
		Partial Positive (PP)
		Partial Negative (PN)

Konsepsi Awal	Arah Perubahan Konsepsi	Konsepsi Akhir
		Misconception (MC)
		No Understanding (NU)
		No Coding (NC)

Berdasarkan kemungkinan arah perubahan konsepsi sehingga perubahan konsepsi peserta didik dapat dikategorikan seperti yang disajikan pada Tabel 3.24.

Tabel 3.24. Kategori perubahan konsepsi

Arah Perubahan Konsepsi	Kategori Perubahan Konsepsi
	Acceptable Change (AC)
	Not Acceptable (NA)
	No Change (NCh)

(Samsudin, Suhandi, Rusdiana, Kaniawati, & Costu, 2016)

Kategori perubahan konsepsi terdiri dari *Acceptable Change* (AC), *Not Acceptable* (NA), dan *No Change* (NCh). Kategori *Acceptable Change* (AC) menunjukkan bahwa dari konsepsi awal ke konsepsi akhir peserta didik mengalami peningkatan pemahaman konsep menjadi lebih baik dan miskonsepsi yang dialami mengalami penurunan. Pada kategori *Not Acceptable* (NA) menunjukkan dari konsepsi awal ke konsepsi akhir peserta didik mengalami penurunan pemahaman konsep. Sedangkan pada kategori *No Change* (NCh) peserta didik tidak mengalami perubahan pemahaman dan peserta didik tetap mengalamii miskonsepsi.

3.6.4. Analisis Efektivitas Penerapan Pembelajaran POEAW Berbantuan Simulasi Virtual Terhadap Peningkatan Literasi Fisika Peserta Didik

Efektivitas penerapan pembelajaran POEAW berbantuan simulasi virtual dibanding pembelajaran tanpa simulasi virtual terhadap peningkatan literasi fisika

peserta didik dianalisis dengan menggunakan perhitungan *effect size* dan uji beda rata-rata dari kelompok eksperimen dengan kelompok kontrol. Berikut uraian masing-masing analisis yang digunakan:

a. Analisis Uji Beda

1. Uji Normalitas

Uji normalitas dilakukan sebagai uji prasyarat. Uji normalitas dilakukan guna mengetahui data kuantitas penurunan miskonsepsi terdistribusi normal atau tidak. Pada penelitian ini menggunakan uji Shapiro-Wilk sebagai uji normalitas pada SPSS. Hasil uji normalitas *n-gain* peningkatan literasi fisika peserta didik pada kelompok kontrol dan eksperimen dihitung menggunakan uji Shapiro-Wilk. Uji Shapiro-Wilk adalah salah satu metode uji normalitas yang mana dapat digunakan apabila jumlah sampel kurang dari lima puluh. Uji normalitas pada *n-gain* peningkatan literasi fisika menggunakan taraf signifikansi $\alpha = 0{,}05$. Jika nilai signifikansi (sig) > 0,05 maka data *n-gain* peningkatan literasi fisika tersebut memiliki distribusi normal dan sebaliknya jika nilai signifikansi < 0,05 maka data tersebut tidak memiliki distribusi normal. Hasil yang diperoleh uji normlitas melalui uji Shapiro-Wilk pada kelompok kontrol adalah sebesar 0,563 dan dan pada kelompok eksperimen adalah 0,067. Kedua hasil tersebut memiliki nilai signifikansi lebih dari dari 0,05 sehingga dapat disimpulkan bahwa data *n-gain* peningkatan literasi fisika pada kelompok kontrol dan kelompok eksperimen berdistribusi normal. Berdasarkan hasil uji normalitas kelompok kontrol dan kelompok eksperimen yang berdistribusi normal maka uji selanjutnya dapat dilakukan adalah uji parametrik.

2. Uji Homogenitas

Uji homogenitas dilakukan guna mengetahui data kuantitas penurunan miskonsepsi penelitian ini homogen atau tidak. Pengunjian homogenitas pada data *n-gain* peningkatan literasi fisika pada kelompok kontrol dan eksperimen menggunakan uji Levene. Pengujian homogentitas pada data *n-gain* peningkatan literasi fisika menggunakan taraf signifikansi $\alpha = 0{,}05$. Jika nilai dari signifikansi (sig) > 0,05, maka data *n-gain* peningkatan literasi fisika tersebut homogen dan sebaliknya jika nilai dari signifikansi < 0,05 maka data *n-gain* peningkatan literasi fisika tersebut tidak homogen. Nilai probabilitas hasil dari uji homogenitas

menggunakan uji Levene pada kolom signifikansi adalah 0,03 yang bernilai tidak lebih dari 0,05. Sebab nilai hasil uji homogenitas menunjukkan siginifikansi < 0,05, maka dapat disimpulkan bahwa data *n-gain* peningkatan literasi fisika peserta didik kelompok kontrol dan kelompok eksperimen adalah tidak homogen.

3. Uji-t'

Uji beda rata-rata melalui uji-t' karena data bersifat homogen dan berdistribusi normal serta bersifat independen. Uji-t' dilakukan dengan menggunakan SPSS. Uji-t' dilakukan untuk mengetahui signifikan atau tidak perbedaan rata-rata peningkatan literasi fisika kelompok eksperimen yang menggunakan pembelajaran POEAW berbantuan simulasi virtual dengan kelompok kontrol yang menggunakan pembelajaran POEAW tanpa berbantuan simulasi virtual. Hipotesis pada uji beda adalah sebagai berikut:

H_0: $\mu_1 = \mu_2$: (Tidak terdapat perbedaan signifikan antara rata-rata peningkatan literasi fisika kelompok ekperimen dan kelompok kontrol)

H_1: $\mu_1 \neq \mu_2$: (Terdapat perbedaan signifikan antara rata-rata antara peningkatan literasi fisika kelompok eksperimen dan kelompok kontrol)

Pengambilan keputusan sesuai dengan kriteria sebagai berikut:

H_0 diterima jika nilai probabilitas (Sig) > 0,05

H_0 ditolak jika nilai probabilitas (Sig) < 0,05

b. Analisis *Effect Size*

Perhitungan *effect size* berguna mengetahui signifikansi perlakuan dengan menganalisis perbedaan ukuran antara dua kelompok (Arnoldo & Victor, 2015). Untuk mendapatkan nilai *effect size* dapat dihitung nilai *cohen's d* (Samsudin et al., 2017). Persamaan untuk menghitung nilai *cohen's d* ditunjukkan sebagai berikut:

$$\Delta = \frac{\bar{x}_{experimen\ group} - \bar{x}_{control\ group}}{SD_{pooled}} \quad (3.6)$$

Keterangan:

Δ = Nilai *cohen's d*

$\bar{x}_{experimen\ group}$ = rata-rata *n-gain* peningkatan literasi fisika kelompok eksperimen

$\bar{x}_{control\ group}$ = rata-rata *n-gain* peningkatan literasi fisika kelompok kontrol

SD = Standar deviasi

Untuk memperoleh nilai SD_{pooled}, dilakukan perhitungan menggunakan nilai standar deviasi kedua kelompok. Rumus untuk menghitung SD_{pooled} ditampilkan sebagai berikut:

$$SD_{pooled} = \sqrt{\frac{(N_E-1)SD_E{}^2+(N_C-1)SD_C{}^2}{N_E+N_C-2}} \quad (3.7)$$

Nilai dari *cohen's d* kemudian diinterpretasi sesuai Tabel 3.25.

Tabel 3.25. Interpretasi nilai *cohen's d*

Nilai *cohen's d* pada *effect size*	Makna
$0{,}00 \leq \Delta \leq 0{,}20$	Tidak berarti
$0{,}20 \leq \Delta \leq 0{,}50$	Kecil
$0{,}50 \leq \Delta \leq 0{,}80$	Sedang
$\Delta \geq 0{,}80$	Besar

3.6.5. Analisis Efektivitas Penerapan Pembelajaran POEAW Berbantuan Simulasi Virtual Terhadap Penurunan Miskonsepsi Peserta Didik

Efektivitas penerapan pembelajaran POEAW berbantuan simulasi virtual dibanding pembelajaran tanpa simulasi virtual terhadap penurunan miskonsepsi dapat dianalisis dengan menggunakan perhitungan *effect size* dan uji beda rata-rata dari kelompok eksperimen dengan kelompok kontrol. Berikut uraian masing-masing analisis yang digunakan:

a. Analisis Uji Beda

1. Uji Normalitas

Uji normalitas dilakukan sebagai uji prasyarat. Uji normalitas dilakukan guna mengetahui data kuantitas penurunan miskonsepsi terdistribusi normal atau tidak. Uji normalitas Kuantitas Penurunan Miskonsepsi pada kelompok kontrol dan eksperimen dihitung menggunakan uji Shapiro-Wilk. Uji Shapiro-Wilk adalah salah satu metode uji normalitas yang mana dapat digunakan apabila jumlah sampel kurang dari lima puluh. Uji normalitas pada Kuantitas Penurunan Miskonsepsi (KPM) menggunakan taraf signifikansi α = 0,05. Jika nilai signifikansi (sig) > 0,05 maka data Kuantitas Penurunan Miskonsepsi (KPM) tersebut memiliki distribusi normal dan sebaliknya jika nilai signifikansi < 0,05

maka data tersebut tidak memiliki distribusi normal. Hasil yang diperoleh uji normlitas melalui uji Shapiro-Wilk pada kelompok kontrol adalah sebesar 0,021 dan dan pada kelompok eksperimen adalah 0,000. Kedua hasil tersebut memiliki nilai signifikansi kurang dari 0,05 sehingga dapat disimpulkan bahwa data Kuantitas Penurunan Miskonsepsi (KPM) pada kelompok kontrol dan kelompok eksperimen tidak berdistribusi normal. Berdasarkan hasil uji normalitas kelompok kontrol dan kelompok eksperimen yang menghasilkan tidak berdistribusi normal maka uji selanjutnya dapat dilakukan adalah uji nonparametrik.

2. Uji Homogenitas

Uji homogenitas dilakukan guna mengetahui data kuantitas penurunan miskonsepsi penelitian ini homogen atau tidak. Pengujian homogenitas pada data Kuntitas Penurunan Miskonsepsi (KPM) pada kelompok kontrol dan eksperimen menggunakan uji Levene. Pengujian homogentitas pada data KPM menggunakan taraf signifikansi α = 0,05. Jika nilai dari signifikansi (sig) > 0,05, maka data KPM tersebut homogen dan sebaliknya jika nilai dari signifikansi < 0,05 maka data KPM tersebut tidak homogen. Nilai probabilitas hasil dari uji homogenitas menggunakan uji Levene pada kolom signifikansi adalah 0,737 yang bernilai lebih dari 0,05. Sebab nilai hasil uji homogenitas menunjukkan siginifikansi > 0,05, maka dapat disimpulkan bahwa data kuantitas penurunan miskonsepsi kelompok kontrol dan kelompok eksperimen adalah homogen.

3. Uji Mann-Whiney U

Uji beda rata-rata melalui uji Mann-Whiney U karena data bersifat homogen dan berdistribusi normal serta bersifat independen. Uji Mann-Whiney U dilakukan dengan menggunakan SPSS. Uji Mann-Whiney U dilakukan untuk mengetahui signifikan atau tidak perbedaan rata-rata kelompok eksperimen yang menggunakan pembelajaran POEAW berbantuan simulasi virtual dengan kelompok kontrol yang mengguankan pembelajaran POEAW tanpa berbantuan simulasi virtual. Hipotesis pada uji beda sebagai berikut:

H_0: $\mu_1 = \mu_2$: (Tidak terdapat perbedaan signifikan antara kuantitas penurunan miskonsepsi kelompok ekperimen dan kelompok kontrol)

H_1: $\mu_1 \neq \mu_2$: (Terdapat perbedaan signifikan antara kuantitas penurunan miskonsepsi kelompok eksperimen dan kelompok kontrol)

Pengambilan keputusan sesuai dengan kriteria sebagai berikut:

H_0 diterima jika nilai probabilitas (Sig) > 0,05

H_0 ditolak jika nilai probabilitas (Sig) < 0,05

b. Analisis *Effect Size*

Perhitungan *effect size* berguna mengetahui signifikansi perlakuan dengan menganalisis perbedaan ukuran antara dua kelompok (Arnoldo & Victor, 2015). Untuk mendapatkan nilai *effect size* dapat dihitung nilai *cohen's d* (Samsudin et al., 2017). Persamaan untuk menghitung nilai *cohen's d* ditunjukkan sebagai berikut:

$$\Delta = \frac{\bar{x}_{experimen\ group} - \bar{x}_{control\ group}}{SD_{pooled}} \quad (3.8)$$

Keterangan:

Δ = Nilai *cohen's d*

$\bar{x}_{experimen\ group}$ = rata-rata kuantitas penurunan miskonsepsi kelompok eksperimen

$\bar{x}_{control\ group}$ = rata-rata kuantitas penurunan miskonsepsi kelompok kontrol

SD = Standar deviasi

Untuk memperoleh nilai SD_{pooled}, dilakukan perhitungan menggunakan nilai standar deviasi kedua kelompok. Rumus untuk menghitung SD_{pooled} ditampilkan sebagai berikut:

$$SD_{pooled} = \sqrt{\frac{(N_E - 1){SD_E}^2 + (N_C - 1){SD_C}^2}{N_E + N_C - 2}} \quad (3.9)$$

Nilai dari *cohen's d* kemudian diinterpretasi sesuai Tabel 3.26.

Tabel 3.26. Interpretasi nilai *cohen's d*

Nilai *cohen's d* pada *effect size*	Makna
$0{,}00 \leq \Delta \leq 0{,}20$	Tidak berarti
$0{,}20 \leq \Delta \leq 0{,}50$	Kecil
$0{,}50 \leq \Delta \leq 0{,}80$	Sedang
$\Delta \geq 0{,}80$	Besar

Berdasarkan instrumen-instrumen dan teknik analisis yang digunakan dalam penelitian ini memiliki matriks yang sesuai pada Tabel 3.27.

Tabel 3.27. Matriks Instrumen Penelitian dan Teknik Analisis Data

No	Sumber Data	Instrumen	Teknik Pengumpulan Data	Teknik Analisis Data
1	Hasil tes literasi fisika	Tes literasi fisika	*Pre-test* dan *post-test*	• *n-gain* • *Rasch model* • Uji Beda • *effect size*
2	Hasil tes *four-tiers* untuk miskonsepsi	*Four-tier test*	*Pre-test* dan *post-test*	• *n-gain* • *Rasch model* • Uji Beda • *effect size*

BAB IV
TEMUAN DAN PEMBAHASAN

Temuan dan pembahasan mengemukakan hasil penelitian terkait penerapan POEAW untuk meningkatkan literasi fisika dan menurunkan miskonsepsi peserta didik pada materi alat optik. Bagian temuan dan pembahasan menjawab beberapa pertanyaan penelitian pada bagian rumusan masalah penelitian. Berikut masing-masing temuan dan pembahasan.

4.1 Peningkatan Literasi Fisika Peserta Didik Setelah Pembelajaran POEAW Berbantuan Simulasi Virtual pada Materi Alat Optik

Sebelum penerapan pembelajaran POEAW, peserta didik diberikan tes awal (*pre-test*) untuk mengetahui kemampuan awal literasi fisika peserta didik pada materi alat optik. Beberapa sub materi pada alat optik terdiri dari: (1) mata, kamera, dan kacamata; (2) lup; (3) mikroskop; dan (4) teropong. Pembelajaran POEAW dilakukan pada kelompok kontrol sebanyak empat kali pertemuan sesuai dengan sub materi yang dibahas pada setiap pertemuan. Sedangkan, pada kelompok eksperimen dilakukan pembelajaran POEAW berbantuan simulasi virtual juga dilakukan sebanyak empat kali pertemuan. Setelah peserta didik menerima pembelajaran, peserta didik pada kelompok kontrol dan kelompok eksperimen diberikan tes akhir (*post-test*) untuk mengetahui kemampuan akhir literasi fisika peserta didik. Hasil dari *pre-test* dan *post-test* peserta didik digunakan untuk menganalisis peningkatan kemampuan literasi fisika peserta didik sesuai pada rumusan masalah.

Rumusan masalah yang pertama pada penelitian ini adalah “Bagaimana peningkatan literasi fisika peserta didik setelah melakukan pembelajaran POEAW berbantuan simulasi virtual pada materi alat optik?”. Uji *N-gain* dilakukan dengan tujuan mengetahui ukuran peningkatan dari kemampuan literasi fisika peserta didik pada sebelum dan sesudah dilakukan pembelajaran. Hasil dari rekapitulasi perhitungan menggunakan *N-gain* kelompok kontrol dan kelompok eksperimen tertera pada Tabel 4.1.