

BAB III

METODE PENELITIAN

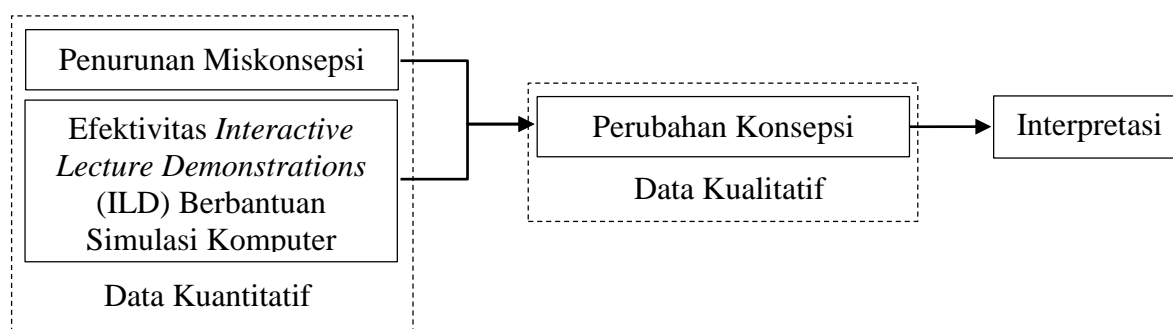
Bagian metode penelitian merupakan pembahasan mengenai metode penelitian yang digunakan dalam penelitian ini. Pembahasan di dalamnya meliputi: Desain penelitian; Populasi dan sampel; Instrumen *four-tier*; Prosedur penelitian; Analisis instrumen penelitian, dan: Analisis data.

3.1. Desain Penelitian

Metode penelitian yang digunakan adalah *mixed methods* atau metode penelitian campuran. Metode penelitian campuran adalah metode penelitian yang melibatkan pengumpulan data secara kuantitatif dan kualitatif, menyatukan dua bentuk data, dan menggunakan desain yang berbeda dengan melibatkan asumsi-asumsi dan kerangka teoritis (Creswell, 2014). Penggunaan metode penelitian campuran dilakukan untuk mendapat jawaban yang utuh dari permasalahan penelitian. Hal ini dikarenakan metode penelitian campuran akan mengurangi kelemahan dari masing-masing metode penelitian kuantitatif dan kualitatif.

Adapun desain penelitian yang digunakan adalah *explanatory sequential mixed methods*. Desain *explanatory sequential mixed methods* merupakan desain dalam metode campuran yang menarik bagi individu dengan latar belakang data kuantitatif ke pendekatan kualitatif (Creswell, 2014). Maksud keseluruhan dari desain ini adalah agar data kualitatif membantu menjelaskan secara lebih rinci hasil data kuantitatif. Desain ini memiliki dua fase dalam penerapannya, fase pertama digunakan untuk mengumpulkan data kuantitatif, sedangkan fase kedua menggunakan hasil data kuantitatif untuk merencanakan pembahasan secara kualitatif. Sehingga gagasan utamanya adalah bahwa pengumpulan data kualitatif dibangun pada hasil data kuantitatif. Almalki (2016) menyebutkan bahwa desain *explanatory sequential* biasanya digunakan dalam penelitian pendidikan yang disebut sebagai *participant selection model*.

Dalam penelitian ini, data kuantitatif digunakan untuk mengetahui penurunan kuantitas peserta didik yang mengalami miskonsepsi, dan mengetahui efektifitas penerapan model *Interactive Lecture Demonstrations* (ILD) berbantuan simulasi komputer dalam menurunkan miskonsepsi. Sedangkan data kualitatif digunakan untuk mengetahui proses perubahan konsepsi peserta didik secara umum. Desain *explanatory sequential mixed methods* secara umum dapat dilihat pada Gambar 3.1 (Almalki, 2016; Creswell, 2014).



Gambar 3. 1. *Explanatory sequential mixed methods*

Data penurunan miskonsepsi didapat dari nilai rata-rata *pre-test* dan *post-test*, untuk kelas eksperimen ataupun kelas kontrol. Kemudian hasilnya dibandingkan untuk melihat penurunan kuantitas peserta didik yang mengalami miskonsepsi. Untuk data efektifitas model *Interactive Lecture Demonstrations* (ILD) berbantuan simulasi komputer diperoleh berdasarkan nilai rata-rata *post-test* kelas eksperimen dan kelas kontrol. Sehingga data untuk penurunan kuantitas miskonsepsi dan efektifitas model *Interactive Lecture Demonstrations* (ILD) berbantuan simulasi komputer merupakan data kuantitatif. Dalam pelaksanaannya metode yang digunakan untuk data kuantitatif menggunakan jenis *quasi eksperimen* dengan desain *Nonequivalent Control Group*. Desain ini biasanya digunakan pada kelas yang sudah ada, sehingga sampel tidak dipilih secara random (Creswell, 2014). Sedangkan untuk perubahan konsepsi diperoleh berdasarkan kategori konsepsi peserta didik yang diperoleh dari nilai *pre-test* dan *post-test*. Sehingga data perubahan konsepsi merupakan data kualitatif. Setelah kedua data tersebut diperoleh, hasilnya diinterpretasikan untuk diambil kesimpulan mengenai penerapan model *Interactive Lecture Demonstrations* (ILD) berbantuan simulasi

komputer untuk mengurangi kuantitas peserta didik yang mengalami miskonsepsi dan perubahan konsepsi peserta didik pada materi gelombang cahaya.

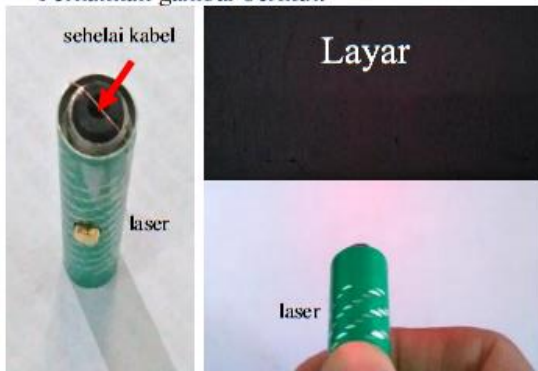
3.2. Populasi dan Sampel

Populasi dalam penelitian ini adalah kelas XI MIPA SMAN 1 Haurgeulis. Sampel diambil menggunakan teknik *cluster random sampling*, dimana randomisasi dilakukan terhadap kelompok, bukan terhadap subjek secara individual (Sugiyono, 2014; Azwar, 2010). Sampel pada penelitian ini melibatkan dua kelas XI yaitu kelas eksperimen dan kelas kontrol. Kelas eksperimen adalah kelas yang mendapatkan model ILD berbantuan simulasi komputer, sedangkan kelas kontrol adalah kelas yang mendapatkan model ILD tanpa berbantuan simulasi komputer. Sampel kelas eksperimen berjumlah 22 peserta didik (8 laki-laki dan 14 perempuan). Sedangkan sampel kelas kontrol berjumlah 24 peserta didik (10 laki-laki dan 14 perempuan). Rata-rata usia peserta didik untuk kelas eksperimen dan kelas kontrol sekitar 17 tahun.

3.3. Instrumen *Four-tier*

Instrumen *four-tier* yang digunakan dalam penelitian ini merupakan instrumen hasil pengembangan berdasarkan alternatif konsepsi peserta didik pada materi gelombang cahaya. Pengembangan instrumen tersebut menggunakan model 3D+1I (*Defining, Designing, Developing and Implementing*) (Malik, Setiawan, Suhandi, Permanasari, & Sulasman, 2018). Pada tahap *defining*, studi literatur dilakukan untuk mengetahui miskonsepsi pada materi gelombang cahaya. Tahap *designing* digunakan untuk membuat desain instrumen *four-tier (open-ended)* dan instrumen *four-tier (close-ended)*. Instrumen *four-tier (open-ended)* dirancang untuk mengumpulkan alternatif konsepsi sebelum instrumen dikembangkan (Lampiran C.1). Sedangkan instrumen *four-tier (close-ended)* dirancang untuk mengidentifikasi miskonsepsi peserta didik yang merupakan pengembangan dari instrumen *four-tier (open-ended)*. Pada tahap *developing*, pengembangan dilakukan pada instrumen *four-tier (close-ended)*. Pengembangan tersebut dapat dilihat pada Gambar 3.2 (Lampiran B.3).

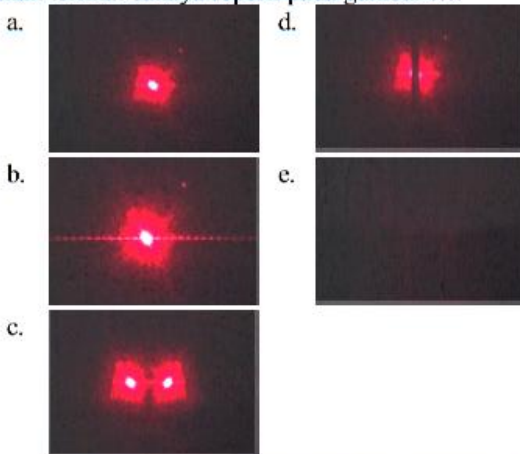
13.1 Perhatikan gambar berikut!



Gambar 1

Gambar 2

Jika sehelai kabel diletakkan di depan laser seperti pada Gambar 1, kemudian laser tersebut ditembakkan pada layar seperti pada Gambar 2, maka pada layar akan terlihat cahaya seperti pada gambar



13.2 Apakah Anda yakin dengan jawaban untuk soal 13.1?

- a. Yakin
- b. Tidak yakin

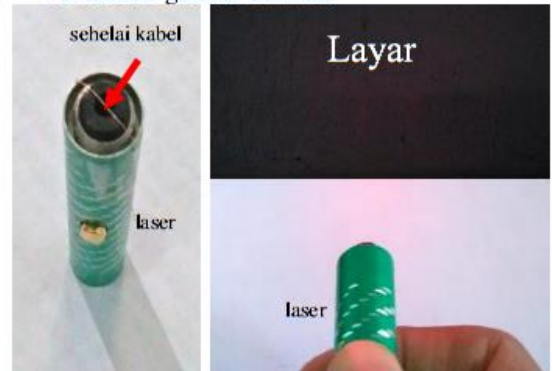
13.3 Alasan Anda untuk jawaban soal 13.1 adalah

.....

13.4 Apakah Anda yakin dengan jawaban untuk soal 13.3?

- a. Yakin
- b. Tidak yakin

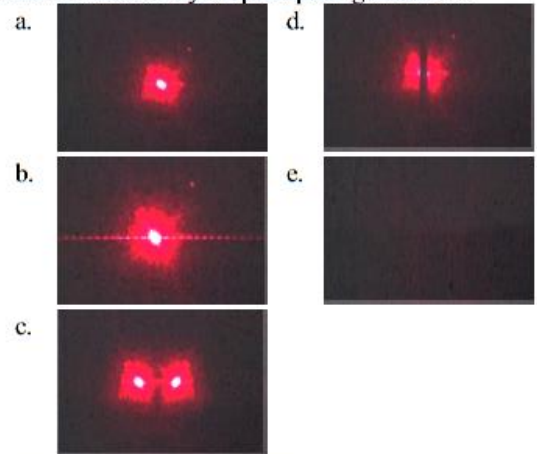
13.1 Perhatikan gambar berikut!



Gambar 1

Gambar 2

Jika sehelai kabel diletakkan di depan laser seperti pada Gambar 1, kemudian laser tersebut ditembakkan pada layar seperti pada Gambar 2, maka pada layar akan terlihat cahaya seperti pada gambar



13.2 Apakah Anda yakin dengan jawaban untuk soal 13.1?

- a. Yakin
- b. Tidak yakin

13.3 Alasan Anda untuk jawaban soal 13.1 adalah

- a. karena cahaya mengalami superposisi akibat terhalang oleh sehelai kabel
- b. karena cahaya yang dipantulkan ke layar hanya cahaya yang melalui celah kosong
- c. karena cahaya terhalang oleh sehelai kabel sehingga cahaya pada laser tidak sempurna
- d. karena sehelai kabel bukan jadi penghalang cahaya tetapi akan menyerap dan ikut membiaskan cahaya
- e. karena cahaya terhalang oleh sehelai kabel sehingga pada layar menghasilkan bayangan seperti bentuk kabel

13.4 Apakah Anda yakin dengan jawaban untuk soal 13.3?

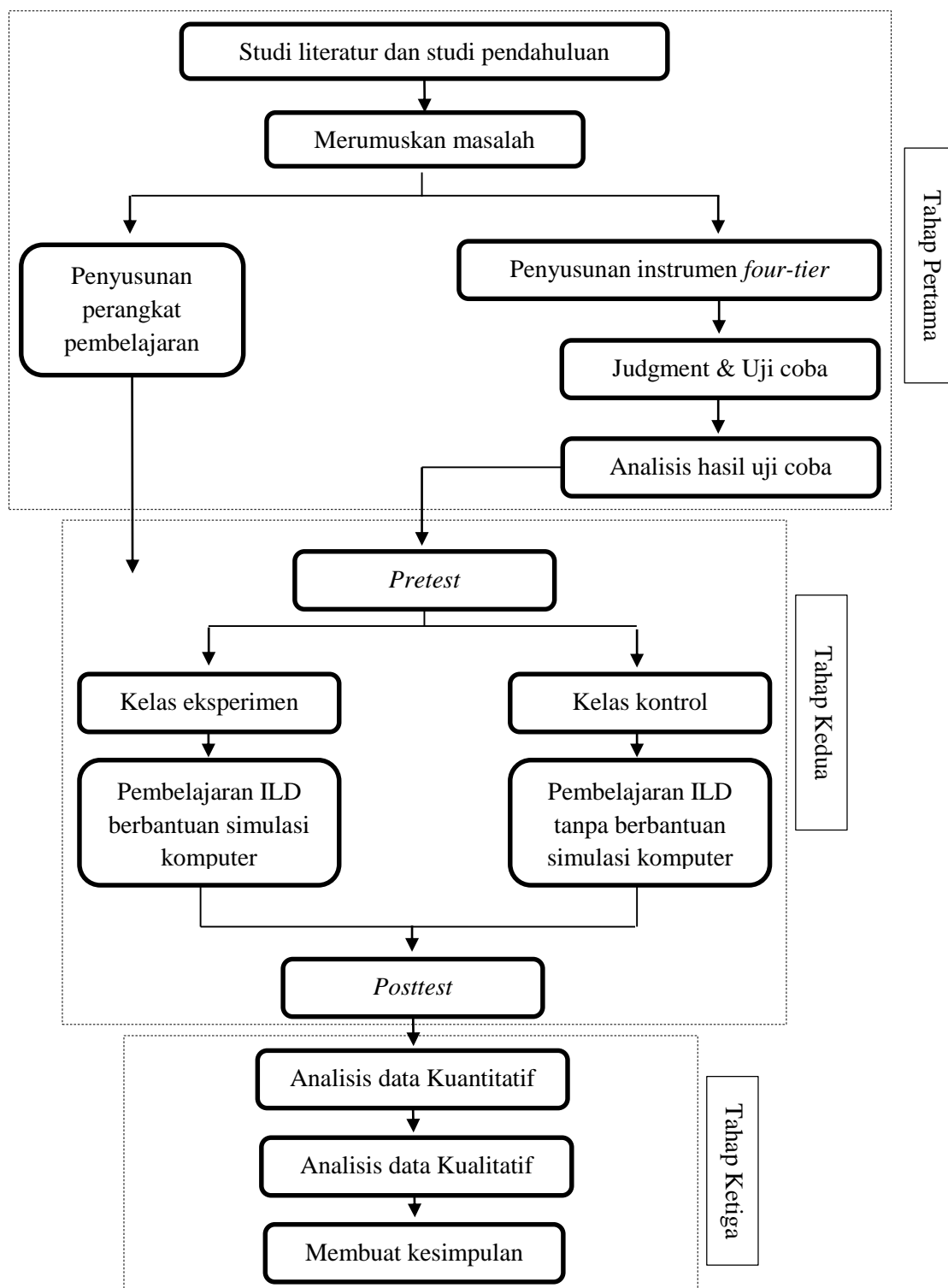
- a. Yakin
- b. Tidak yakin

Gambar 3. 2. Pengembangan Instrumen *four-tier*

Selanjutnya tahap *implementing* digunakan untuk menganalisis penggunaan instrumen *four-tier (open-ended)* dalam mengidentifikasi miskonsepsi dengan analisis Rasch. Hasil pengembangan ini kemudian digunakan untuk mengukur miskonsepsi peserta didik pada materi gelombang cahaya. Tes diagnostik ini memiliki 15 soal. Pada tingkat pertama (*first tier*) merupakan pertanyaan tentang konsep ilmiah dengan beberapa pilihan jawaban. Tingkat ke dua (*second tier*) merupakan tingkat keyakinan peserta didik dalam memilih jawaban. Tingkat ketiga (*third tier*) merupakan alasan peserta didik menjawab pertanyaan pada tingkat pertama, berupa pilihan alasan yang telah disediakan dan satu alasan terbuka. Tingkat ke empat (*four-tier*) merupakan tingkat keyakinan peserta didik dalam memilih alasan pada tingkat ketiga. Tes dilaksanakan sebanyak dua kali, yakni pada saat *pretest* dan *posttest*. Tes awal (*pretest*) dilakukan untuk melihat miskonsepsi peserta didik pada materi gelombang cahaya, dan tes akhir (*posttest*) dilakukan untuk melihat miskonsepsi peserta didik setelah diberi perlakuan. Selanjutnya data tersebut akan digunakan untuk menjawab pertanyaan penelitian yang sudah dirumuskan sebelumnya.

3.4. Prosedur Penelitian

Prosedur penelitian pada penelitian ini memiliki tiga tahapan, yakni tahap pertama, kedua, dan ketiga. Tahap pertama secara garis besar merupakan tahap persiapan sebelum penelitian. Tahapan tersebut dimulai dengan studi literature dan studi pendahuluan, dan diakhiri dengan analisis hasil uji coba serta penyusunan perangkat pembelajaran. Tahap kedua merupakan tahap pelaksanaan penelitian. Tahapan tersebut dimulai dengan melakukan *pre-test* dan diakhiri dengan memberikan *post-test*. Terakhir yakni tahap ketiga merupakan tahap setelah data penelitian didapat. Tahapan tersebut dimulai dengan menganalisis data kuantitatif hasil dari penelitian, dan diakhiri dengan membuat kesimpulan. Secara ringkas, prosedur penelitian yang akan dilakukan dapat dilihat pada Gambar 3.3.



Gambar 3. 3. Prosedur penelitian

3.5. Analisis Instrumen Penelitian

Analisis instrumen dalam penelitian ini digunakan untuk menguji instrumen yang akan digunakan. Analisis yang digunakan diantaranya: Uji validitas; Uji reliabilitas; Tingkat kesukaran, dan; Daya pembeda. Data yang digunakan dapat dilihat pada (Lampiran C.3).

3.5.1. Uji Validitas

Validitas merupakan ukuran yang menunjukkan tingkat ketepatan suatu tes dalam mengukur apa yang mau diukur. Sebuah instrumen dikatakan valid apabila mampu mengukur apa yang diinginkan dan dapat mengungkap data dari variabel yang diteliti secara tepat. Uji validitas instrumen *four-tier* dalam penelitian ini menggunakan *software* MINISTEP 4.3.1 dengan *output Table 10 Item (column): fit order* yang dianalisis dengan analisis Rasch. Pengukuran uji validitas instrumen ini dilihat berdasarkan nilai *logarithm odd unit (logit)* pada *outfit mean square (MNSQ)*, *outfit Z-standard (ZSTD)*, dan *point-measure correlation (PTMEASURE-AL COOR)*. Nilai *logit* merupakan nilai yang dihasilkan melalui perhitungan fungsi logaritma pada *software* MINISTEP. Dengan menggunakan fungsi *logit* ini, maka akan didapatkan mistar pengukuran dengan interval yang sama (Sumintono & Widhiarso, 2015). Sumintono & Widhiarso (2015) juga menyebutkan bahwa *item fit* dapat menjelaskan apakah butir soal berfungsi normal melakukan pengukuran atau tidak. Sedangkan beberapa penelitian menyebutkan bahwa *point-measure correlation* hanya digunakan untuk mengetahui daya pembeda dari suatu instrumen (Sabudin, Mansor, Meerah, & Muhammad, 2018; Smiley, 2015). Maka dalam penelitian ini, untuk mengukur validitas suatu instrumen hanya menggunakan skor *outfit mean square (MNSQ)* dan *outfit Z-standard (ZSTD)*. Hasil uji validitas instrumen *four-tier* dapat dilihat *output item (column): fit order* seperti pada Gambar 3.4.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	TOTAL MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
15	10	26	.54	.37	3.33	3.07	1.97	1.79	A .43	.28	84.6	65.7	N15
11	28	26	-.68	.20	1.75	2.04	1.93	1.97	B .48	.44	19.2	49.0	N11
12	6	26	1.22	.46	1.90	1.82	1.24	.61	C .43	.22	88.5	77.1	N12
1	46	26	-1.24	.17	1.38	1.52	1.57	1.68	D .38	.53	30.8	30.7	N1
10	21	26	-.36	.23	1.27	.77	1.45	1.05	E .24	.39	46.2	53.0	N10
14	40	26	-1.07	.17	1.30	1.21	1.31	.95	F .40	.51	30.8	35.6	N14
7	15	25	-.04	.28	.65	-.69	1.24	.64	G .33	.35	68.0	57.1	N7
13	5	25	1.39	.50	.88	-.15	.72	-.41	H .36	.21	80.0	80.3	N13
3	15	25	-.01	.28	.85	-.16	.85	-.18	g .33	.34	52.0	56.9	N3
5	12	25	.24	.33	.65	-.60	.85	-.16	f .22	.31	56.0	61.1	N5
9	5	26	1.45	.50	.84	-.26	.77	-.28	e .36	.21	80.8	81.0	N9
4	46	26	-1.24	.17	.69	-1.42	.67	-1.11	d .64	.53	34.6	30.7	N4
8	11	26	.41	.35	.58	-.78	.68	-.63	c .43	.29	73.1	63.0	N8
2	18	26	-.19	.25	.43	-1.54	.51	-1.17	b .37	.36	61.5	56.5	N2
6	22	26	-.42	.22	.26	-2.74	.33	-1.98	a .63	.40	80.8	52.5	N6
MEAN	20.0	25.7	.00	.30	1.12	.1	1.07	.2			59.1	56.7	
P.SD	13.6	.4	.86	.11	.75	1.5	.49	1.1			21.9	15.5	

Gambar 3. 4. *OUTFIT* (MNSQ) dan (ZSTD) setiap butir soal

Nilai *OUTFIT* (MNSQ) dan (ZSTD) selanjutnya dimasukkan pada kriteria untuk memeriksa kesesuaian butir soal. Kriteria tersebut dapat dilihat pada Tabel 3.1 (Sumintono & Widhiarso, 2015).

Tabel 3. 1
Kriteria nilai MNSQ dan ZSTD

Outfit	Nilai yang diterima
MNSQ	$0,5 < MNSQ < 1,5$
ZSTD	$-2,0 < ZSTD < +2,0$

Berdasarkan kriteria pada Tabel 3.1, maka validitas untuk setiap item instrumen *four-tier* dapat dilihat pada Tabel 3.2.

Tabel 3. 2
Validitas item instrumen four-tier

No Soal	Nilai MNSQ	Nilai ZSTD	Keterangan
1.	1.57	1.68	Digunakan
2.	0.51	-1.17	Digunakan
3.	0.85	-0.18	Digunakan
4.	0.64	-1.11	Digunakan
5.	0.85	-0.16	Digunakan
6.	0.33	-1.98	Digunakan
7.	1.24	0.64	Digunakan
8.	0.68	-0.63	Digunakan
9.	0.77	-0.28	Digunakan
10.	1.45	1.05	Digunakan
11.	1.93	1.97	Digunakan
12.	1.24	0.61	Digunakan
13.	0.72	-0.41	Digunakan
14.	1.31	0.95	Digunakan
15.	1.97	1.79	Digunakan

Tabel 3.2 menunjukkan semua soal memiliki validitas yang dapat diterima, sehingga semua soal pada instrumen *four-tier* dapat digunakan. Uji validitas isi juga dilakukan pada instrumen *four-tier* oleh 3 orang *expert* (3 dosen). Butir soal dikatakan valid apabila Indeks Validitas Isi $IVI \geq 0,70$ (Delgado-Rico, Carrtero-Dios, & Ruch, 2012). Dalam menentukan IVI, setiap *expert* yang memberi kriteria “sesuai” diberi skor 1 dan “tidak sesuai” diberi skor 0 (Lampiran C.2). Hasilnya akan dihitung menggunakan Persamaan 3.1.

$$IVI = \frac{\text{Jumlah expert yang memberi kriteria sesuai}}{\text{Jumlah seluruh expert}} \quad (3.1)$$

Adapun hasil uji validitas setiap butir soal adalah dapat dilihat pada Tabel 3.3.

Tabel 3.3
Hasil uji valiitas isi

No. Soal	Expert			IVI	Keputusan
	I	II	III		
1	1	1	1	1,00	Digunakan
2	1	1	1	1,00	Digunakan
3	1	1	1	1,00	Digunakan
4	1	1	1	1,00	Digunakan
5	1	1	1	1,00	Digunakan
6	1	1	1	1,00	Digunakan
7	1	1	1	1,00	Digunakan
8	1	1	1	1,00	Digunakan
9	1	1	1	1,00	Digunakan
10	1	1	1	1,00	Digunakan
11	1	1	1	1,00	Digunakan
12	1	1	1	1,00	Digunakan
13	1	1	1	1,00	Digunakan
14	1	1	1	1,00	Digunakan
15	1	1	1	1,00	Digunakan

Tabel 3.1 menunjukan bahwa 15 soal intrumen *four-tier* dapat digunakan. Hal ini dikarenakan 15 soal dikatakan valid menurut *expert*.

3.5.2. Uji Reliabilitas

Reliabilitas instrumen dalam penelitian ini mempunyai makna penting karena menunjukkan ketetapan dan kemantapan dari instrumen *four-tier*. Reliabilitas

Adam Hadiana Aminudin, 2019

PENERAPAN MODEL INTERACTIVE LECTURE DEMONSTRATIONS (ILD) BERBANTUAN SIMULASI KOMPUTER UNTUK MENURUNKAN KUANTITAS MISKONSEPSI PESERTA DIDIK DAN PENGUBAHAN KONSEPSI PADA MATERI GELOMBANG CAHAYA

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

mencerminkan ketetapan instrumen penelitian yang digunakan dalam mengukur dan menggali informasi yang diperlukan. Uji reliabilitas pada penelitian ini menggunakan *software* MINISTEP 4.3.1 dengan *output* Table 3.1 *Summary Statistics* yang dianalisis dengan analisis Rasch. Analisis Rasch dapat menampilkan beberapa nilai reliabilitas, diantaranya adalah *person reliability*, *item reliability*, dan *Cronbach alpha*. *Person reliability* menunjukkan konsistensi jawaban peserta didik, sementara *item reliability* menunjukkan kualitas item tes, sementara *Cronbach alpha* menunjukkan nilai interaksi antara person dan item dari instrumen *four-tier* secara keseluruhan. Hasil uji reliabilitas butir soal dapat dilihat pada *Output summary statistics* Gambar 3.5.

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	11.5	14.8	-1.34	.37	1.10	.22	1.07	.20
SEM	1.2	.1	.15	.03	.10	.16	.11	.19
P.SD	5.8	.5	.76	.16	.52	.80	.54	.93
S.SD	5.9	.5	.78	.16	.53	.81	.55	.95
MAX.	27.0	15.0	-.10	1.05	2.57	1.77	2.45	2.35
MIN.	1.0	13.0	-4.12	.24	.34	-1.28	.42	-1.24
REAL RMSE	.45	TRUE SD	.62	SEPARATION	1.38	Person RELIABILITY	.66	
MODEL RMSE	.40	TRUE SD	.65	SEPARATION	1.62	Person RELIABILITY	.72	
S.E. OF Person MEAN = .15								
Person RAW SCORE-TO-MEASURE CORRELATION = .89								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .66 SEM = 3.41								
SUMMARY OF 15 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	20.0	25.7	.00	.30	1.12	.14	1.07	.19
SEM	3.6	.1	.23	.03	.20	.41	.13	.30
P.SD	13.6	.4	.86	.11	.75	1.52	.49	1.14
S.SD	14.1	.5	.89	.12	.77	1.57	.50	1.18
MAX.	46.0	26.0	1.45	.50	3.33	3.07	1.97	1.97
MIN.	5.0	25.0	-1.24	.17	.26	-2.74	.33	-1.98
REAL RMSE	.37	TRUE SD	.77	SEPARATION	2.07	Item RELIABILITY	.81	
MODEL RMSE	.32	TRUE SD	.80	SEPARATION	2.50	Item RELIABILITY	.86	
S.E. OF Item MEAN = .23								

Gambar 3. 5. Nilai *person reliability*, *item reliability*, dan *Cronbach alpha*

Interpretasi dari nilai *person reliability* dan *item reliability* pada Gambar 3.5 dapat dilihat pada Tabel 3.4 (Sumintono & Widhiarso, 2015; Mohamad, Sulaiman, Sern, & Salleh, 2015).

Tabel 3. 4

Interpretasi nilai <i>person reliability</i> dan <i>item reliability</i>	
Nilai <i>person reliability</i> dan <i>item reliability</i>	Interpretasi
$0,94 \leq \text{Nilai}$	Istimewa
$0,90 \leq \text{Nilai} < 0,94$	Bagus Sekali

Nilai <i>person reliability</i> dan <i>item reliability</i>	Interpretasi
$0,80 \leq \text{Nilai} < 0,90$	Bagus
$0,67 \leq \text{Nilai} < 0,80$	Cukup
Nilai $< 0,67$	Lemah

Berdasarkan analisis Rasch maka nilai *person reliability* 0,66 termasuk dalam kategori lemah dan nilai *item reliability* 0,81 termasuk dalam kategori bagus. Sedangkan interpretasi untuk nilai *Cronbach alpha* dapat dilihat pada Tabel 3.5 (Sumintono & Widhiarso, 2015).

Tabel 3. 5
Interpretasi nilai Cronbach alpha

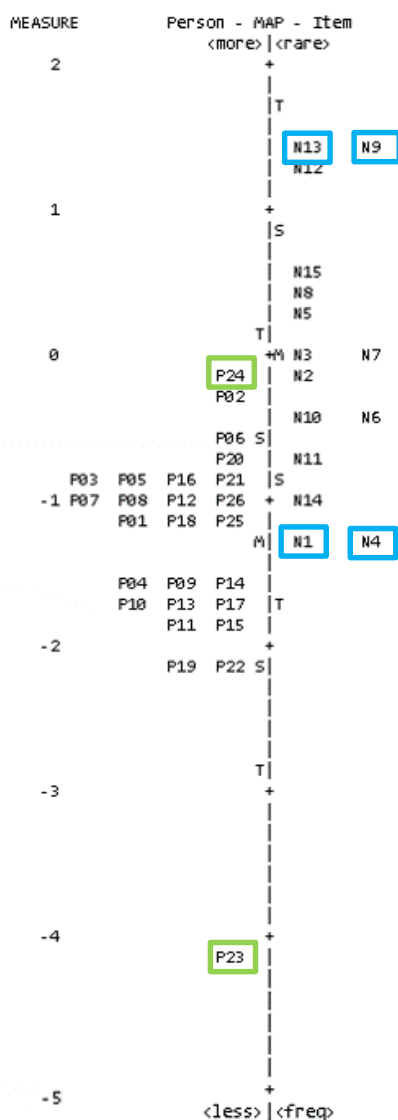
Nilai Cronbach alpha	Interpretasi
$0,8 \leq \alpha$	Bagus Sekali
$0,7 \leq \alpha < 0,8$	Bagus
$0,6 \leq \alpha < 0,7$	Cukup
$0,5 \leq \alpha < 0,6$	Jelek
$\alpha < 0,5$	Buruk

Nilai *Cronbach alpha* dari instrumen *four-tier* sebesar 0,66 sehingga masuk kategori cukup. Nilai reliabilitas orang dan nilai reliabilitas barang dari 0 hingga 1 dapat diinterpretasikan sangat mirip dengan alpha Cronbach, yang berarti bahwa nilai yang mendekati 1 menunjukkan ukuran yang lebih konsisten (Boone & Noltemeyer, 2017). Selain itu, Mohamad, Sulaiman, Sern, & Salleh (2015) menyebutkan bahwa dalam ilmu sosial, nilai α yang dapat diterima adalah 0,60. Sehingga instrumen *four-tier* dapat diterima dalam hal indeks reliabilitas.

3.5.3. Tingkat Kesukaran

Tingkat kesukaran adalah bilangan yang menunjukkan sukar dan mudahnya sesuatu instrumen soal. Tingkat kesukaran instrumen *four-ter* dalam penelitian ini menggunakan *software* MINISTEP 4.3.1 dengan *output* Table 1 *Variable (Wright) maps* dan *output* Table 13 *Item Measure* yang dianalisis dengan analisis Rasch. Penggunaan *Variable (Wright) maps* bertujuan untuk mendapat gambaran mengenai kekuatan butir soal. Boone & Noltemeyer (2017) mengungkapkan peta wright mendukung para peneliti untuk mengukur kekuatan dan kelemahan instrumen, mendokumentasikan penilaian butir tes, membandingkan teori dengan

data eksperimen, dan memberikan panduan kepada para peneliti. Sedangkan *output Table 13 Item Measure* digunakan untuk mengklasifikasikan tingkat kesulitan tiap butir soal berdasarkan nilai logit dan nilai Standar Deviasi (SD) yang didapat (Sumintono & Widhiarso, 2015). Adapun hasil dapat dilihat pada Gambar 3.6 mengenai *Variable (Wright) maps* instrumen *four-tier*.



Gambar 3. 6. Variable (Wright) maps instrumen *four-tier*

Kode *N* yang diikuti nomor pada bagian kanan Gambar 3.7 menunjukkan nomor butir soal, sedangkan kode *P* yang diikuti nomor pada bagian kiri menunjukkan nomor urut peserta didik. Gambar 3.7 menjelaskan bahwa soal yang sangat sulit ditunjukkan oleh soal dengan kode N13, dan N9 karena berada pada

posisi paling atas serta tidak ada peserta didik yang berpotensi menjawabnya. Sedangkan soal paling mudah adalah soal dengan kode N1 dan N4 karena berada pada posisi paling bawah serta dapat dijawab oleh 15 peserta didik dengan kode P24, P02, P06, P20, P05, P08, P16, P21, P03, P07, P12, P26, P01, P18, dan P25. Secara lebih rinci, tingkat kesukaran instrumen *four-tier* pada Gambar 3.7 dapat ditentukan berdasarkan nilai *logit output Table 13 Item Measure. Output Table 13 Item Measure* secara umum dapat dilihat pada Gambar 3.7.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	Item
9	5	26	1.45	.50	.84	-.26	.77	-.28	.36	.21	80.8	81.0	N9
13	5	25	1.39	.50	.88	-.15	.72	-.41	.36	.21	80.0	80.3	N13
12	6	26	1.22	.46	1.90	1.82	1.24	.61	.43	.22	88.5	77.1	N12
15	10	26	.54	.37	3.33	3.07	1.97	1.79	.43	.28	84.6	65.7	N15
8	11	26	.41	.35	.58	-.78	.68	-.63	.43	.29	73.1	63.0	N8
5	12	25	.24	.33	.65	-.60	.85	-.16	.22	.31	56.0	61.1	N5
3	15	25	-.01	.28	.85	-.16	.85	-.18	.33	.34	52.0	56.9	N3
7	15	25	-.04	.28	.65	-.69	1.24	.64	.33	.35	68.0	57.1	N7
2	18	26	-.19	.25	.43	-1.54	.51	-1.17	.37	.36	61.5	56.5	N2
10	21	26	-.36	.23	1.27	.77	1.45	1.05	.24	.39	46.2	53.0	N10
6	22	26	-.42	.22	.26	-2.74	.33	-1.98	.63	.40	80.8	52.5	N6
11	28	26	-.68	.20	1.75	2.04	1.93	1.97	.48	.44	19.2	49.0	N11
14	40	26	-1.07	.17	1.30	1.21	1.31	.95	.40	.51	30.8	35.6	N14
1	46	26	-1.24	.17	1.38	1.52	1.57	1.68	.38	.53	30.8	30.7	N1
4	46	26	-1.24	.17	.69	-1.42	.67	-1.11	.64	.53	34.6	30.7	N4
MEAN	20.0	25.7	.00	.30	1.12	.1	1.07	.2			59.1	56.7	
P.SD	13.6	.4	.86	.11	.75	1.5	.49	1.1			21.9	15.5	

Gambar 3. 7. Nilai MEASURE setiap butir soal dan Standar Deviasi (SD)

Gambar 3.7 menunjukkan bahwa item N9 memiliki nilai *measure* paling tinggi dengan *logit* sebesar 1,45, dan item N1 dan N4 memiliki nilai *measure* paling rendah dengan *logit* sebesar -1,24. Adapun nilai Standar Deviasi (SD) yang didapat sebesar 0,86, sehingga klasifikasi kategori tingkat kesulitan soal dapat dibuat berdasarkan nilai tersebut. Nilai 0,00 *logit* + 1SD adalah satu kelompok sukar, lebih besar dari +1SD adalah soal yang sangat sukar; 0,00 *logit* -1SD adalah soal yang mudah; dan lebih kecil dari -1SD soal yang sangat mudah (Sumintono & Widhiarso, 2015). Artinya terdapat empat kategori tingkat kesukaran pada instrumen *four-tier* yang dapat dilihat pada Tabel 3.6.

Tabel 3. 6
Interpretasi hasil tingkat kesukaran

Tingkat Kesukaran (TK)	Interpretasi
$0,86 < TK$	Sangat Sukar
$0,00 < TK \leq 0,86$	Sukar
$-0,86 < TK \leq 0,00$	Mudah
$TK \leq -0,86$	Sangat Mudah

Berdasarkan Gambar 3.7 dan Gambar 3.8, maka tingkat kesulitan setiap item pada instrumen *four-tier* dapat dilihat pada Tabel 3.7.

Tabel 3. 7
Hasil tingkat kesukaran tiap butir

No. Soal	Tingkat Kesukaran (TK)	Interpretasi
1.	-1,24	Sangat Mudah
2.	-0,19	Mudah
3.	-0,01	Mudah
4.	-1,24	Sangat Mudah
5.	0,24	Sukar
6.	-0,42	Mudah
7.	-0,04	Mudah
8.	0,41	Sukar
9.	1,45	Sangat Sukar
10.	-0,36	Mudah
11.	-0,68	Mudah
12.	1,22	Sangat Sukar
13.	1,39	Sangat Sukar
14.	-1,07	Sangat Mudah
15.	0,54	Sukar

Tabel 3.6 menunjukkan bahwa terdapat soal dengan tingkat kesukaran beragam, mulai dari sangat mudah, mudah, sulit, dan sangat sulit. Sehingga sebaran tingkat kesukaran instrumen *four-tier* dapat dikatakan merata dari tingkat sangat mudah hingga sangat sulit.

3.5.4. Daya Pembeda

Analisis daya pembeda digunakan untuk mengetahui kemampuan soal dalam membedakan peserta didik dengan kemampuan yang tinggi dan kemampuan yang rendah. Daya pembeda pada instrumen *four-tier* menggunakan *software* MINISTEP 4.3.1 dengan *output Table 10 Item Fit Order* yang dianalisis dengan analisis Rasch. Penggunaan *Item Fit Order* karena didalamnya menampilkan informasi mengenai *Point-measure correlation*. Seperti penjelasan sebelumnya pada bagian validitas instrumen, bahwa *point-measure correlation* digunakan untuk mengetahui daya pembeda dari suatu instrumen (Sabudin, Mansor, Meerah, & Muhammad, 2018; Smiley, 2015). Adapun nilai *Point-measure correlation (PTMEASURE-AL COOR)* untuk instrumen *four-tier* dapat dilihat pada Gambar 3.8.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL		INFIT		OUTFIT		PTMEASURE-AL		EXACT MATCH		Item
				S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%		
15	10	26	.54	.37	3.33	3.07	1.97	1.79	A	.43	.28	84.6	65.7	N15
11	28	26	-.68	.20	1.75	2.04	1.93	1.97	B	.48	.44	19.2	49.0	N11
12	6	26	1.22	.46	1.90	1.82	1.24	.61	C	.43	.22	88.5	77.1	N12
1	46	26	-1.24	.17	1.38	1.52	1.57	1.68	D	.38	.53	30.8	30.7	N1
10	21	26	-.36	.23	1.27	.77	1.45	1.05	E	.24	.39	46.2	53.0	N10
14	40	26	-1.07	.17	1.30	1.21	1.31	.95	F	.40	.51	30.8	35.6	N14
7	15	25	-.04	.28	.65	-.69	1.24	.64	G	.33	.35	68.0	57.1	N7
13	5	25	1.39	.50	.88	-.15	.72	-.41	H	.36	.21	80.0	80.3	N13
3	15	25	-.01	.28	.85	-.16	.85	-.18	g	.33	.34	52.0	56.9	N3
5	12	25	.24	.33	.65	-.60	.85	-.16	f	.22	.31	56.0	61.1	N5
9	5	26	1.45	.50	.84	-.26	.77	-.28	e	.36	.21	80.8	81.0	N9
4	46	26	-1.24	.17	.69	-1.42	.67	-1.11	d	.64	.53	34.6	30.7	N4
8	11	26	.41	.35	.58	-.78	.68	-.63	c	.43	.29	73.1	63.0	N8
2	18	26	-.19	.25	.43	-1.54	.51	-1.17	b	.37	.36	61.5	56.5	N2
6	22	26	-.42	.22	.26	-2.74	.33	-1.98	a	.63	.40	80.8	52.5	N6
MEAN	20.0	25.7	.00	.30	1.12	.1	1.07	.2				59.1	56.7	
P. SD	13.6	.4	.86	.11	.75	1.5	.49	1.1				21.9	15.5	

Gambar 3. 8. Nilai PTMEASURE-AL COOR setiap butir soal

Smiley (2015) juga memberikan interpretasi untuk setiap nilai yang diberikan pada PTMEASURE-AL COOR seperti yang terdapat pada Tabel 3.8.

Tabel 3. 8

Interpretasi nilai PTMEASURE-AL COOR

<i>PTMEASURE-AL COOR (ID)</i>	Interpretasi
0,40 < ID	Sangat baik
0,30 < ID ≤ 0,40	Baik
0,20 < ID ≤ 0,30	Kurang baik
ID ≤ 0,19	Jelek

Berdasarkan nilai PTMEASURE-AL COOR pada Gambar 3.9, maka daya pembeda untuk tiap soal dapat dilihat oleh Tabel 3.9.

Tabel 3. 9

Hasil PTMEASURE-AL COOR untuk setiap butir soal

No. Soal	PTMEASURE-AL COOR	Interpretasi
1.	0,38	Baik
2.	0,37	Baik
3.	0,33	Baik
4.	0,64	Sangat Baik
5.	0,22	Kurang Baik
6.	0,63	Sangat Baik
7.	0,33	Baik
8.	0,43	Sangat Baik
9.	0,36	Baik
10.	0,24	Kurang Baik
11.	0,48	Sangat Baik

No. Soal	PTMEASURE-AL COOR	Interpretasi
12.	0,43	Sangat Baik
13.	0,36	Baik
14.	0,40	Baik
15.	0,43	Sangat Baik

Tabel 3.9 menunjukkan bahwa tidak terdapat daya pembeda yang bernilai negatif sehingga seluruh soal dapat digunakan.

3.6. Analisis Data

Bagian ini membahas tentang cara mengolah data hasil penelitian untuk selanjutnya diambil kesimpulan sesuai rumusan pertanyaan. Adapun analisis data pada penelitian ini diantaranya: Kategori konsepsi; Penurunan kuantitas miskonsepsi peserta didik; Pengubahan konsepsi, dan; Efektivitas penerapan ILD berbantuan simulasi komputer dalam menurunkan kuantitas miskonsepsi peserta didik.

3.6.1. Kategori Konsepsi

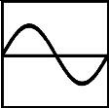
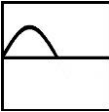
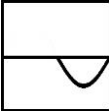
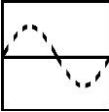

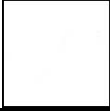
Jawaban peserta didik pada soal *four-tier* akan dikategorikan berdasarkan enam kategori konsepsi pada Tabel 2.4. Sedangkan penilaian untuk kategori tersebut diadaptasi dari penelitian Kaltakci-Gurel, Eryilmaz, & McDermott (2017) tentang penilaian miskonsepsi, dan hasilnya ditunjukkan oleh Tabel 3.10.

Tabel 3. 10
Skor miskonsepsi

Jawaban	Keterangan
<i>Tier 1</i>	Peserta didik menjawab salah pada <i>tier 1</i> mendapat skor 1
<i>Tier 3</i>	Peserta didik menjawab salah pada <i>tier 3</i> mendapat skor 1
<i>Tier 1 dan 3</i>	Peserta didik menjawab salah pada <i>tier 1</i> dan 3 mendapat skor 1
<i>All Tier 1 – 4</i>	Peserta didik menjawab salah pada <i>tier 1</i> dan 3, serta menjawab yakin pada tier 2 dan 4 mendapat skor 1

Berdasarkan Tabel 3.10 maka hasil adaptasi untuk nilai setiap kategori konsepsi ditunjukkan oleh Tabel 3.11.

Tabel 3. 11
Skor kriteria konsepsi

Kategori	Simbol	Skor
<i>Sound Understanding (SU)</i>		0
<i>Partial Positive (PP)</i>		0
<i>Partial Negative (PN)</i>		1
<i>Misconception (MC)</i>		4
<i>No Understanding (NU)</i>		3
<i>No Coding (NC)</i>		(kosong)

Penggunaan skor kriteria konsepsi pada Tabel 3.11 hanya digunakan untuk analisis Rasch. Tujuannya supaya gambaran peserta didik yang mengalami miskonsepsi dapat terlihat dengan jelas. Biasanya, skor kriteria konsepsi menampilkan skor terbesar pada kategori *Sound Understanding (SU)* dan paling kecil dengan skor 0 (nol) pada kategori *Misconception (MC)*, *No Understanding (NU)*, dan *No Coding (NC)*. Jika menggunakan penskoran biasa, peserta didik yang mengalami miskonsepsi tidak dapat diidentifikasi karena skornya sama dengan kategori NU dan NC yakni sebesar nol (0). Sedangkan dengan menggunakan penskoran yang diadaptasi dari Kaltakci-Gurel, Eryilmaz, & McDermott (2017) pada Tabel 3.9, peserta didik yang mengalami miskonsepsi dapat diidentifikasi karena memiliki skor paling besar sebesar 4. Adapun skor (kosong) digunakan pada kategori NC karena analisis Rasch membaca skor (kosong) bukan sebagai 0 (nol). Sumintono & Widhiarso (2015) menyebutkan bahwa Rasch dapat melakukan

prediksi terhadap data yang hilang (*missing data*), yang didasarkan pada pola respon yang sistematis. Sehingga skor kosong akan dibaca sebagai data yang hilang dan dihitung berdasarkan pola jawaban peserta didik pada instrumen *four-tier*. Pada proses perhitungan biasa, skor (kosong) akan terbaca sebagai 0 (nol).

3.6.2. Penurunan Kuantitas Miskonsepsi Peserta Didik

Gambaran mengenai penurunan miskonsepsi secara umum dilihat menggunakan *software* MINISTEP 4.3.1 dengan *output Table 1 Variable (Wright) maps*. Kemudian hasilnya akan disajikan dalam bentuk skalogram untuk melihat potensi peserta didik yang mengalami miskonsepsi. Perhitungan lebih lanjut dilakukan menggunakan data kategori *Misconception* (MC) yang dimiliki peserta didik. Data tersebut kemudian diolah menggunakan persamaan yang diadaptasi dari Hake (1998) tentang *gain* yang dinormalisasi. Tujuannya adalah untuk melihat penurunan miskonsepsi pada saat *pre-test* dan *post-test*. Adapun persamaan yang digunakan dapat dilihat pada Persamaan 3.2 tentang Penurunan Kuantitas Miskonsepsi (PKM).

$$\langle PKM \rangle = \frac{\% \langle S_i \rangle - \% \langle S_f \rangle}{\% \langle S_i \rangle - 0} \quad (3.2)$$

Keterangan:

$\langle PKM \rangle$ = rata-rata penurunan kuantitas miskonsepsi

$\langle \% S_i \rangle$ = rata-rata persentase miskonsepsi saat *pre-test*

$\langle \% S_f \rangle$ = rata-rata persentase miskonsepsi saat *post-test*

Kemudian nilai PKM yang didapat diinterpretasikan berdasarkan Tabel 3.12.

Tabel 3. 12

Interpretasi nilai PKM

Nilai PKM	Kategori
$0,7 < PKM$	Tinggi
$0,3 < PKM \leq 0,7$	Sedang
$PKM \leq 0,3$	Rendah

3.6.3. Pengubahan Konsepsi

Pengubahan konsepsi peserta didik menjadi analisis untuk mengetahui apakah peserta didik mengalami perubahan konsepsi atau tidak. Perubahan

konsep dilakukan dengan melihat konsep peserta didik untuk setiap konsep pada saat *pre-test* dan *post-test*. Persamaan yang digunakan untuk menentukan persentase Perubahan Konsep Setiap Kategori (PKSK) diadaptasi dari Samsudin, Suhandi, Rusdiana, Kaniawati, & Coştu (2016) pada Persamaan 3.3.

$$PKSK(\%) = \%K_{pre} - \%K_{post} \quad (3.3)$$

Keterangan:

$PKSK(\%)$ = persentase perubahan konsep untuk setiap kategori

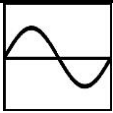
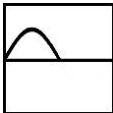
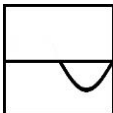
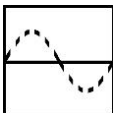
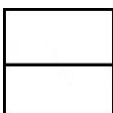
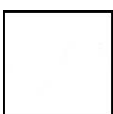
$\%K_{pre}$ = persentase konsep pada saat *pre-test*







$\%K_{post}$ = persentase konsep pada saat *post-test*

Berdasarkan Persamaan 3.3, maka dibuatlah interpretasi perubahan konsep untuk setiap kategori konsep. Interpretasi tersebut disajikan dalam Tabel 3.13.

Tabel 3. 13

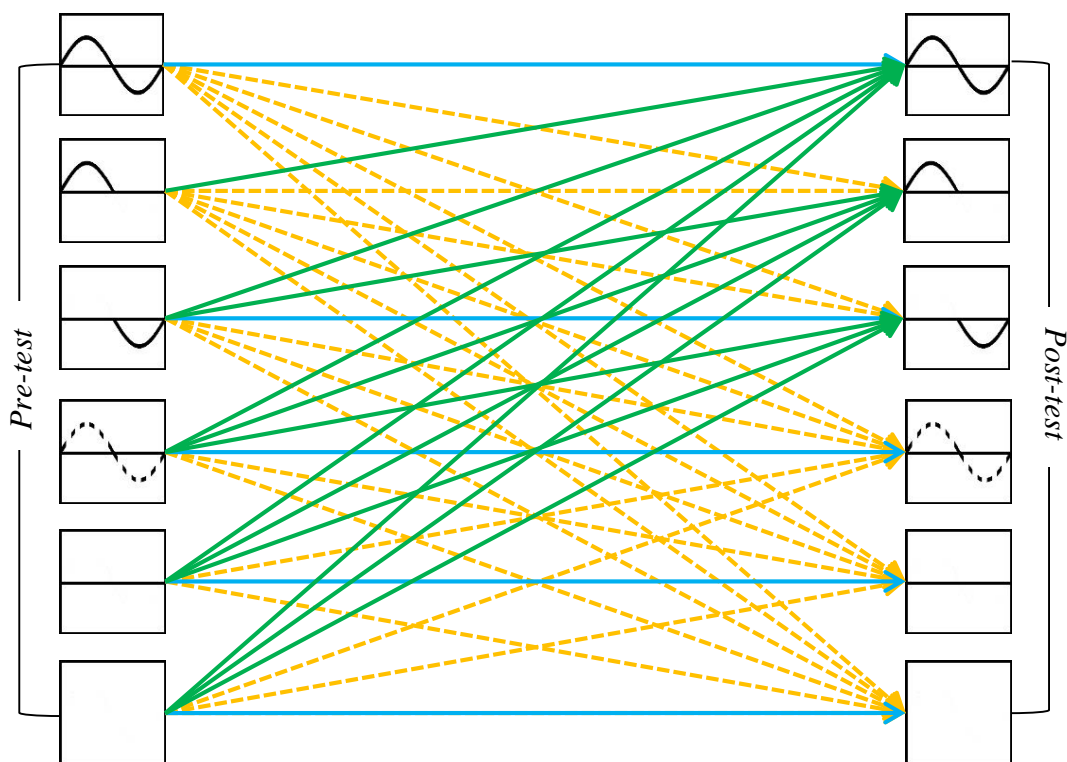
Interpretasi perubahan untuk setiap kategori konsep

No	Kategori Konsep	Perubahan	Interpretasi
1.		+ 0 -	<i>Bad Change (BC)</i> <i>No Change (NC)</i> <i>Good Change (GC)</i>
2.		+ 0 -	<i>Good Change (GC)</i> <i>No Change (NC)</i> <i>Bad Change (BC)</i>
3.		+ 0 -	<i>Good Change (GC)</i> <i>No Change (NC)</i> <i>Bad Change (BC)</i>
4.		+ 0 -	<i>Good Change (GC)</i> <i>No Change (NC)</i> <i>Bad Change (BC)</i>
5.		+ 0 -	<i>Good Change (GC)</i> <i>No Change (NC)</i> <i>Bad Change (BC)</i>
6.		+ 0 -	<i>Good Change (GC)</i> <i>No Change (NC)</i> <i>Bad Change (BC)</i>

Keterangan:  = *Sound Understanding*;  = *Partial Positive*;  = *Partial Negative*;  = *Misconception*;  = *No Understanding*;  = *No Coding*

Pada Tabel 3.13 terdapat perbedaan interpretasi untuk kategori *Sound Understanding*. Perbedaan tersebut karena perubahan (+) memiliki interpretasi *Bad Change (BC)*, sedangkan perubahan (-) memiliki interpretasi *Good Change (GC)*. Perubahan (+) memiliki interpretasi *Bad Change (BC)* karena konsepsi peserta didik lebih banyak pada saat *pre-test* dan berubah menjadi sedikit ketika *post-test*. Seharusnya perubahan pada kategori ini (-), karena konsepsi peserta didik pada kategori ini menjadi lebih besar pada saat *post-test* daripada saat *pre-test*. Sedangkan kategori konsepsi lainnya memiliki interpretasi yang sama, dimana perubahan (+) lebih baik daripada perubahan (-). Kemudian untuk mengetahui perbandingan antara perubahan pada kelas eksperimen dan kelas kontrol, maka penyajian persentase untuk setiap kategori akan disajikan.




Setelah dilakukan persentase terhadap perubahan konsepsi peserta didik, analisis lebih lanjut dilakukan untuk mengetahui ke arah mana perubahan konsepsi tersebut. Kemungkinan perubahan konsepsi yang dialami peserta didik dapat dilihat berdasarkan simbol konsepsi pada Gambar 3.9.



Gambar 3. 9. Kemungkinan perubahan konsepsi saat *pre-test* dan *post-test*

Berdasarkan arah kemungkinan perubahan konsepsi pada Gambar 3.9, maka dibuatlah kategori konsepsi berdasarkan kategori konsepsi Samsudin, Suhandi, Rusdiana, Kaniawati, & Coştu (2016). Kategori perubahan konsepsi tersebut disajikan dalam Tabel 3.14.

Tabel 3. 14.
Kategori perubahan konsepsi

Arah Perubahan Konsepsi	Kategori perubahan konsepsi
	<i>Acceptable Change (AC)</i>
	<i>No Change (NCh)</i>
	<i>Not Acceptable (NA)</i>

Kategori perubahan konsepsi pada Tabel 3.14 terdiri dari tiga jenis perubahan konsepsi, diantaranya *Acceptable Change (AC)*, *Not Acceptable (NA)*, dan *No Change (NCh)*. Kategori AC menunjukkan bahwa pemahaman peserta didik berubah dari *pre-test* ke *post-test* dengan beberapa peningkatan. Untuk kategori NA, pemahaman peserta didik berubah dari *pre-test* ke *post-test* dengan cara yang tidak menguntungkan. Sedangkan untuk kategori NCh, pemahaman peserta didik tidak berubah dari *pre-test* ke *post-test*. Sedangkan hasil perubahan konsepsi berdasarkan kategori tersebut disajikan dengan persentase Perubahan Konsepsi (PK) menggunakan Persamaan 3.4.

$$PK(\%) = \frac{\text{jumlah siswa pada suatu tipe}}{\text{jumlah seluruh siswa}} \quad (3.4)$$

3.6.4. Efektivitas Penerapan ILD Berbantuan Simulasi Komputer dalam Menurunkan Kuantitas Miskonsepsi Peserta Didik

Efektivitas penerapan ILD berbantuan simulasi komputer dapat diketahui dengan perhitungan *effect size*. Perhitungan *effect size* digunakan untuk mengetahui signifikansi suatu perlakuan dengan menganalisis perbedaan ukuran antara dua grup (Arnoldo & Víctor, 2015). Sebelum menghitung *effect size*, perlu dikelompokkan konsepsi peserta didik berdasarkan kriteria pada Tabel 2.3. Kemudian kategori konsepsi tersebut dinilai berdasarkan Tabel 3.12. Setelah diketahui skor untuk masing-masing kriteria konsepsi, maka dapat diketahui skor peserta didik ketika *pre-test* dan *post-test* baik kelas eksperimen maupun kelas

kontrol. Persamaan yang digunakan untuk menghitung *effect size* pada penelitian ini menggunakan perumusan *Glass's delta* (Δ) pada Persamaan 3.5. (Ialongo, 2016)

$$\Delta = \frac{|\bar{X}_{exp} - \bar{X}_{con}|}{SD_{con}} \quad (3.5)$$

Keterangan:

\bar{X}_{exp} = nilai rata-rata miskonsepsi kelas eksperimen

\bar{X}_{con} = nilai rata-rata miskonsepsi kelas kontrol

SD_{con} = standar deviasi kelas kontrol

Hasil perhitungan *Glass's delta* selanjutnya diinterpretasikan berdasarkan Tabel 3.15.

Tabel 3. 15
Interpretasi hasil Glass's delta

<i>Glass's delta</i>	Interpretasi
$0,70 \leq \Delta$	Besar
$0,50 \leq \Delta < 0,70$	Sedang
$0,20 \leq \Delta < 0,50$	Kecil
$0,00 \leq \Delta < 0,20$	Kurang