

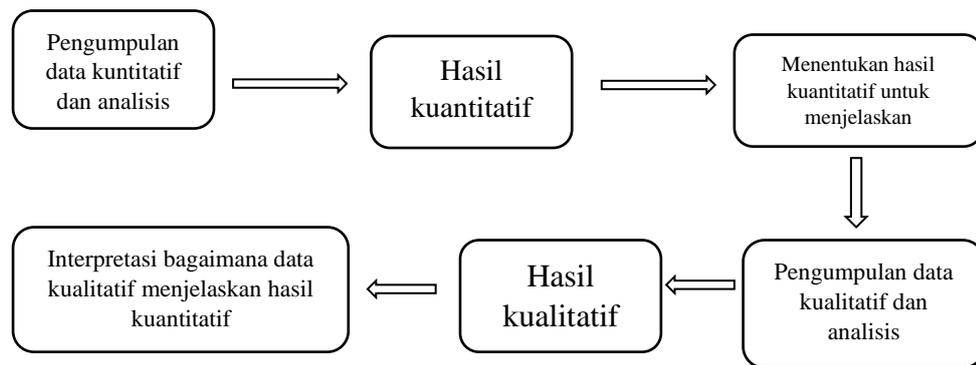
BAB III

METODE PENELITIAN

Bab III merupakan pembahasan mengenai metode penelitian yang digunakan dalam penelitian ini. Pembahasan di dalamnya meliputi: Desain penelitian; Populasi dan sampel; Instrumen penelitian; Prosedur dan alur 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 supaya didapatkan jawaban yang utuh dari permasalahan penelitian. Hal ini dikarenakan metode penelitian campuran akan mengurangi kelemahan dari masing-masing metode penelitian kuantitatif dan kualitatif. Penelitian ini dirancang sebagai intervensi dalam mengajar momentum dan impuls. Namun, selain untuk mengukur pengaruh dari intervensi terhadap peningkatan kemampuan memahami peserta didik dan penurunan miskonsepsi yang ditunjuk hanya dengan cara pengukuran kuantitatif, juga akan dianalisis secara kualitatif untuk memperoleh informasi lebih mendalam terkait faktor-faktor yang mempengaruhi hasil penelitian. Oleh karena itu, model yang dipilih dalam metode campuran ini adalah *sequential explanatory* dengan desain seperti pada gambar berikut:



Gambar 3.1. Desain *sequential explanatory*

Pada Gambar 3.1, dapat dilihat bahwa pengumpulan data kuantitatif dan kualitatif yang didapatkan digunakan untuk menjelaskan dan menginterpretasikan hasil data yang diperoleh. Data kuantitatif diperoleh secara langsung pada saat sebelum dan sesudah penerapan strategi *Predict Discuss Explain Observe Discuss Explain* (PDEODE) berbantuan simulasi komputer yaitu dengan menghitung nilai rata-rata *pre-test* dan *post-test*, sedangkan data kualitatif yang diperoleh didapatkan secara tidak langsung.

Penurunan kuantitas miskonsepsi serta peningkatan kemampuan memahami diperoleh dari nilai rata-rata peserta didik pada saat *pre-test* dan *post-test*. Efektivitas strategi *Predict Discuss Explain Observe Discuss Explain* (PDEODE) berbantuan simulasi komputer diperoleh dengan menghitung *effect size* yang kemudian diinterpretasikan.

3.2 Populasi dan Sampel

Populasi dalam penelitian ini adalah kelas X MIPA-1 SMAN Haurgeulis yang berlokasi di Indramayu. Sampel diambil menggunakan teknik *cluster random sampling*, dimana randomisasi dilakukan terhadap kelompok, bukan terhadap subjek secara individual (Azwar, 2010). Sampel pada penelitian ini berjumlah 22 peserta didik (12 perempuan dan 10 laki-laki) dengan rata-rata usia peserta didik sekitar 16 tahun.

3.3 Instrumen Penelitian

Instrumen yang digunakan dalam penelitian ini adalah sebagai berikut:

3.3.1 Lembar Observasi

Lembar Observasi digunakan untuk melihat keterlaksanaan kegiatan pembelajaran pada materi momentum dan impuls. Lembar observasi model pembelajaran dengan strategi *Predict Discuss Explain Observe Discuss Explain* (PDEODE) berbantuan simulasi komputer di kelas berisi tahap-tahap pembelajaran yang telah disesuaikan. Selanjutnya diisi oleh observer yang sebelumnya sudah mengetahui tentang tatacara pengisian lembar observasi.

Kemudian observer memberikan tanda checklist (√) pada kolom yang telah disediakan, dan memberi komentar terhadap keterlaksanaan kegiatan pembelajaran. Sehingga diharapkan lembar observasi ini dapat memberikan gambaran mengenai keterlaksanaan kegiatan guru dan peserta didik saat proses pembelajaran berlangsung.

3.3.2 Instrumen *four-tier test*

Instrumen ini digunakan untuk mengukur miskonsepsi dan kemampuan memahami peserta didik pada materi momentum dan impuls. Tes ini dikembangkan berdasarkan miskonsepsi peserta didik pada materi tersebut dan berbentuk pilihan ganda empat tingkat. Tingkat pertama (*first-tier*) merupakan pertanyaan tentang konsep ilmiah dengan beberapa pilihan jawaban. Pada tingkat pertama ini dapat digunakan untuk mengukur kemampuan memahami peserta didik. 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 diagnostik ini memiliki 15 soal. Tes dilaksanakan sebanyak dua kali, yakni pada saat *pre-test* dan *post-test*. Tes awal (*pre-test*) dilakukan untuk melihat miskonsepsi peserta didik pada materi momentum dan impuls, dan tes akhir (*post-test*) dilakukan untuk melihat miskonsepsi peserta didik setelah diberi perlakuan menggunakan strategi *Predict*

Discuss Explain Observe Discuss Explain (PDEODE) berbantuan simulasi komputer. Selanjutnya data tersebut akan digunakan untuk menjawab pertanyaan penelitian yang sudah dirumuskan sebelumnya.

Instrumen *four-tier* yang digunakan dikembangkan 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 momentum dan impuls. Tahap *designing* digunakan untuk membuat desain instrumen *four-tier (open-ended)* dan instrumen *four-tier (closed-ended)*. Instrumen *four-tier (open-ended)* dirancang untuk mengumpulkan alternatif konsepsi sebelum instrumen dikembangkan. Sedangkan instrumen *four-tier (closed-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 (closed-ended)*. Pengembangan tersebut dapat dilihat pada Gambar 3.2.

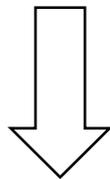
5.1 Perhatikan Tabel 1 berikut ini!

Tabel 1 Kendaraan Transportasi.

No.	Nama Benda
1	Mobil Sedan
2	Sepeda
3	Motor
4	Kereta
5	Mobil Bus

Pada Tabel 1, ketika keempat benda bergerak dengan kecepatan yang sama, benda yang mempunyai momentum terbesar adalah

- Mobil
 - Sepeda
 - Motor
 - Kereta
 - Mobil Bus
- 5.2 Apakah Anda yakin dengan jawaban untuk soal 5.1?
a. Yakin b. Tidak Yakin
- 5.3 Penjelasan terbaik dari jawaban Anda untuk pertanyaan di atas adalah
.....
.....
.....
- 5.4 Apakah Anda yakin dengan jawaban untuk soal 5.3?
a. Yakin b. Tidak Yakin



(Four-tier open-ended)

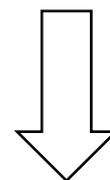
5.1 Perhatikan Tabel 1 berikut ini!

Tabel 1 Kendaraan Transportasi.

No.	Nama Benda
1	Mobil Sedan
2	Sepeda
3	Motor
4	Kereta
5	Mobil Bus

Pada Tabel 1, ketika keempat benda bergerak dengan kecepatan yang sama, benda yang mempunyai momentum terbesar adalah

- Mobil
 - Sepeda
 - Motor
 - Kereta
 - Mobil Bus
- 5.2 Apakah Anda yakin dengan jawaban untuk soal 5.1?
a. Yakin b. Tidak Yakin
- 5.3 Penjelasan terbaik dari jawaban Anda untuk pertanyaan di atas adalah
a. Benda yang mempunyai kecepatan paling konstan maka momentumnya paling besar
b. Karena harus menyeimbangi massa yang paling besar
c. Massa benda berbanding terbalik dengan momentumnya
d. Kelima benda tersebut mempunyai momentum yang besarnya sama karena kecepatannya sama
e. Semakin besar massa suatu benda maka semakin besar pula momentumnya
- 5.4 Apakah Anda yakin dengan jawaban untuk soal 5.3?
a. Yakin b. Tidak Yakin



(Four-tier closed-ended)

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 momentum dan impuls.

3.4 Prosedur dan Alur Penelitian

Penelitian ini melalui tiga tahapan, yaitu:

3.4.1 Tahap Persiapan

Kegiatan yang dilakukan dalam pada tahap persiapan meliputi:

- a. Studi literatur, dilakukan untuk memperoleh teori dan informasi mengenai strategi pembelajaran PDEODE dan simulasi komputer yang *ter-up to date* sehingga dapat dijadikan landasan kuat yang akan diterapkan dalam penelitian.
- b. Menentukan lokasi dan sampel penelitian yang mendukung ketercapaian tujuan penelitian dan meminta izin kepada sekolah yang bersangkutan.
- c. Menentukan materi pembelajaran untuk materi yang digunakan dalam penelitian
- d. Menyusun RPP dan skenario pembelajaran berdasarkan strategi pembelajaran PDEODE sehingga alur dari pembelajarannya tersusun dengan rapih.
- e. Membuat surat izin penelitian.
- f. Membuat instrumen penelitian (tes diagnostik *for-tier test*, Lembar Kerja Peserta Didik (LKPD), dan lembar observasi keterlaksanaan pembelajaran)
- g. Men-*judgement* instrumen tes kepada 3 dosen yang ahli dibidangnya.
- h. Merevisi instrumen
- i. Uji coba instrument kemudian menganalisisinya.
- j. Melatih observer untuk mengetahui cara pengisian lembar observasi tentang keterlaksanaan pembelajaran.

3.4.2 Tahap Pelaksanaan

Kegiatan yang dilakukan dalam tahap pelaksanaan meliputi:

- a. Melakukan tes sebelum perlakuan (*pre-test*).
- b. Memberikan perlakuan pada peserta didik berupa pembelajaran dengan strategi PDEODE berbantuan simulasi komputer pada materi momentum dan impuls.

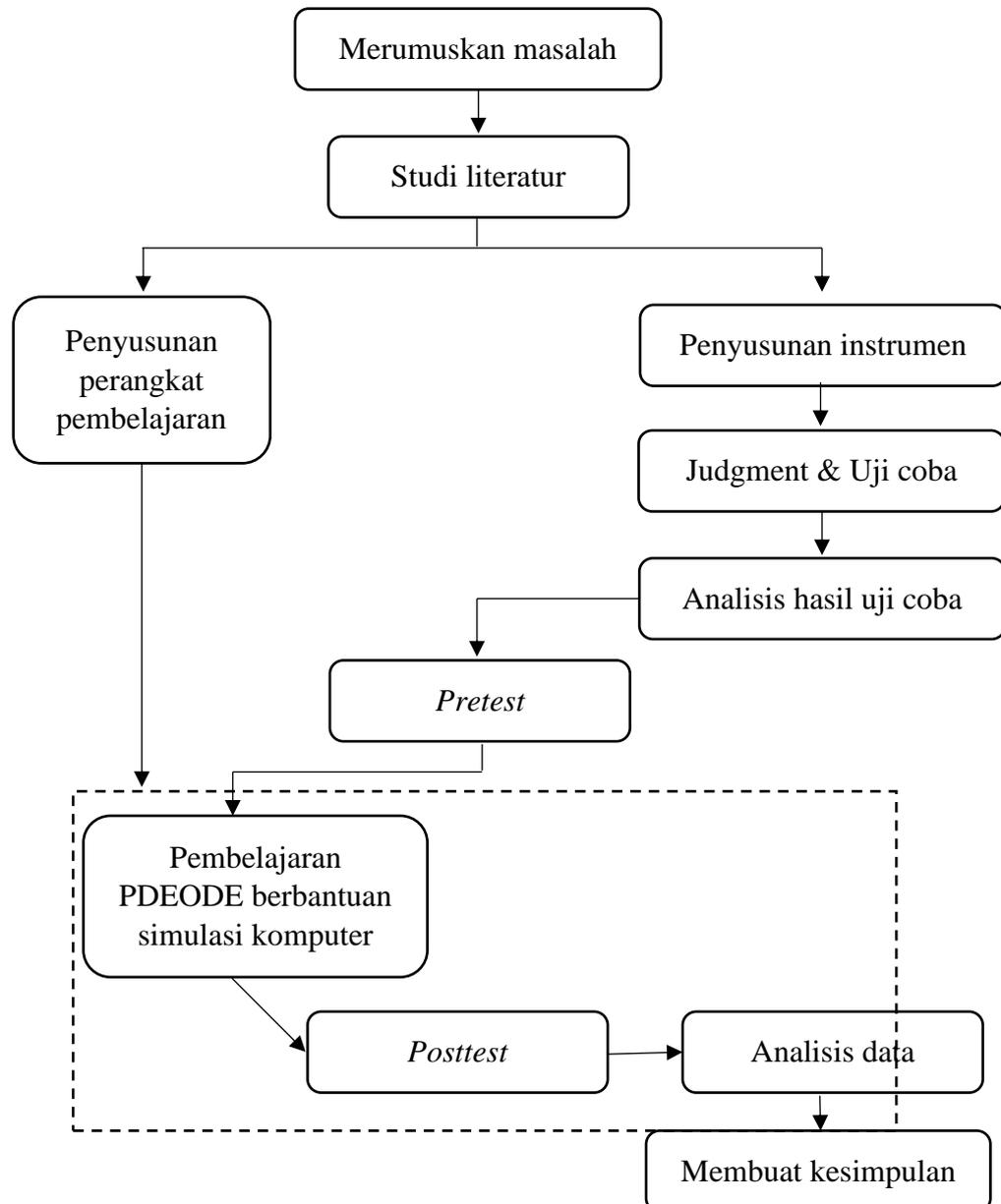
- c. Selama proses pembelajaran berlangsung, observer melakukan observasi tentang keterlaksanaan pembelajaran PDEODE atau POE.
- d. Melakukan tes setelah perlakuan (*post-test*).

3.4.3 Tahap Analisis dan Pembahasan

Kegiatan yang dilakukan dalam tahap ini meliputi:

- a. Mengolah data hasil *pre-test* dan *post-test*.
- b. Menganalisis data hasil penelitian.
- c. Membuat kesimpulan.

Adapun langkah-langkah penelitian tersebut ditunjukkan pada gambar berikut



Gambar 3. 3. Alur Proses Penelitian

3.5 Analisis Instrumen Penelitian

Sebelum digunakan, tes *four-tier* dan tes kemampuan memahami harus diuji dahulu validitas, reliabilitas, tingkat kesukaran, dan daya pembedanya supaya baik, sah, dan dapat dipercaya. Hal-hal spesifik yang harus diuji adalah dari segi validitas, reliabilitas, tingkat kesukaran dan daya pembeda.

3.5.1. Uji Validitas

Validitas adalah suatu ukuran yang menunjukkan tingkat-tingkat kevalidan atau kesahihan suatu instrumen. Suatu instrumen yang valid mempunyai validitas yang tinggi, sedangkan instrumen yang kurang valid memiliki validitas yang rendah. Sebuah instrumen dikatakan valid apabila mampu mengukur apa yang diinginkan dan dapat mengungkap data dari variabel yang diteliti secara tepat. Uji validitas yang digunakan pada instrumen *four-tier* ini merupakan validitas isi yang meliputi kesesuaian butir soal dengan konten serta kesesuaian butir soal dengan kunci jawaban.

Uji validitas instrumen *four-tier* dapat diketahui dengan menggunakan menu *output 10. Item (column): fit order* pada *software* MINISTEP. Pengukuran uji validitas instrumen ini dilihat dari hasil skor pada menu *Outfit mean square (MNSQ)*, *Outfit Z-standard (ZSTD)*, dan *Point Measure Correlation (Pt Mean Corr)*. Menurut Boone dkk (2014), kriteria tersebut digunakan untuk mengetahui tingkat kesesuaian butir (*item fit*). Sedangkan penelitian yang dilakukan oleh Sabudin dkk (2018) menyatakan bahwa *Point Measure Correlation (Pt Mean Corr)* hanya digunakan untuk mengetahui daya pembeda dari suatu instrumen. Maka untuk mengukur kevalidan atau kesahihan suatu instrumen hanya menggunakan skor pada menu *Outfit mean square (MNSQ)* dan *Outfit Z-standard (ZSTD)* pada *software* MINISTEP untuk dianalisis dengan Rasch. Hasil uji validitas instrumen *four-tier* dapat dilihat seperti pada Gambar 3.4 berikut.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
					MNSQ	ZSTD	MNSQ	ZSTD					
10	12	28	1.00	.30	1.55	1.1	1.43	1.0	A .34	.27	67.9	62.5	Q10
8	34	28	-.02	.17	1.43	1.6	1.32	1.0	B .26	.47	28.6	38.7	Q8
6	61	28	-.68	.15	1.39	1.7	1.31	1.2	C .37	.54	17.9	22.9	Q6
9	44	28	-.29	.16	1.09	.5	1.32	1.1	D .38	.51	25.0	31.7	Q9
11	44	28	-.29	.16	1.27	1.2	1.16	.6	E .55	.51	17.9	31.7	Q11
13	10	28	1.20	.34	.92	.0	1.26	.7	F-.32	.25	53.6	66.7	Q13
2	45	28	-.31	.16	1.14	.7	1.03	.2	G .68	.52	10.7	28.5	Q2
3	52	28	-.48	.15	1.01	.1	.97	.0	g .50	.54	32.1	28.9	Q3
4	30	28	.10	.18	.98	.0	.85	-.3	f .79	.44	17.9	39.8	Q4
7	24	28	.31	.20	.85	-.3	.77	-.5	e .35	.40	46.4	46.0	Q7
1	64	28	-.75	.15	.85	-.7	.74	-1.0	d .61	.54	32.1	26.4	Q1
5	60	28	-.66	.15	.61	-2.0	.66	-1.4	c .56	.54	39.3	24.6	Q5
14	19	28	.53	.22	.29	-2.4	.46	-1.4	b .42	.35	64.3	54.1	Q14
12	23	28	.35	.20	.36	-2.4	.42	-1.7	a .58	.39	60.7	46.5	Q12
MEAN	37.3	28.0	.00	.19	.98	-.1	.98	.0			36.7	39.2	
S.D.	17.6	.0	.59	.06	.36	1.3	.32	1.0			18.3	13.6	

Gambar 3. 4. Menu *Output item (column): fit order*

Pada Gambar 3.4 terlihat bahwa soal nomor 10 memiliki skor 1,55 pada kolom *INFIT MNSQ* dan skor 1,43 pada *OUTFIT MNSQ* dan seterusnya untuk nomor soal yang berbeda. Untuk mengetahui kevalidan dari tiap soal dapat dicari dengan interpretasi menurut Boone (2014) sebagai berikut.

Tabel 3. 1.

Interpretasi nilai Output MNSQ dan ZSTD

Output item	Skor	Keterangan
Outfit MNSQ	$0,5 < \text{MNSQ} < 1,5$	Diterima
Outfit ZSTD	$-2,0 < \text{ZSTD} < +2,0$	Diterima

Berdasarkan Gambar 3.4 dan Tabel 3.1, instrumen soal *four-tier* dapat dikategorikan dan diinterpretasikan pada Tabel 3.2 berikut.

Tabel 3. 2.

Interpretasi kevalidan item instrumen four-tier

No Soal	Skor MNSQ	Skor ZSTD	Keterangan
1	0,74	-1,0	Digunakan
2	1,03	0,2	Digunakan
3	0,97	0,0	Digunakan
4	0,85	-0,3	Digunakan
5	0,66	-1,4	Digunakan
6	1,31	1,2	Digunakan
7	0,77	-0,5	Digunakan
8	1,32	1,0	Digunakan
9	1,32	1,1	Digunakan
10	1,43	1,0	Digunakan
11	1,16	0,6	Digunakan
12	0,42	-1,7	Digunakan
13	1,26	0,7	Digunakan
14	0,46	-1,4	Digunakan

Maka dari hasil interpretasi uji kevalidan pada Tabel 3.2 dapat dikatakan bahwa semua nomor pada soal instrumen sudah layak digunakan.

3.5.2. Uji Reliabilitas

Reliabilitas instrumen mempunyai pengertian bahwa suatu instrumen cukup dapat dipercaya atau digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik dan tetap. Sebuah instrumen yang *reliable* berapa kali digunakan dalam penelitian apapun hasilnya akan relatif sama sehingga hasil yang didapatkan dari instrumen tersebut dapat dipercaya. Uji reliabilitas pada penelitian ini menggunakan *software* MINISTEP 4.3.1. Perangkat lunak MINISTEP ini merupakan sebuah aplikasi pendukung untuk mengetahui kriteria dan kualitas suatu instrumen soal berdasarkan analisis Rasch. Salah satu menu *output* pada *software* ini yaitu *output 3.1 Summary Statistics* yang fungsinya yaitu untuk menampilkan beberapa nilai reliabilitas, diantaranya adalah *person reliability*, *item reliability*, dan *Cronbach alpha*. *Person reliability* menunjukkan konsistensi jawaban peserta didik, artinya kita dapat mengetahui peserta didik yang konsisten/ajeg dengan jawaban yang mereka berikan saat menjawab soal tersebut. *Item reliability* menunjukkan kualitas item tes, artinya kita dapat mengetahui kualitas sebuah instrumen yang digunakan, dalam penelitian ini instrumen *four-tier*. Sementara *Cronbach alpha* menunjukkan nilai interaksi antara konsistensi jawaban peserta didik dan item soal dari instrumen *four-tier* secara keseluruhan. Hasil uji reliabilitas butir soal dapat dilihat pada Gambar 3.5 sebagai berikut:

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	18.6	14.0	-.51	.26	1.00	.0	.98	.0
S.D.	8.1	.0	.49	.05	.34	.8	.36	.7
MAX.	34.0	14.0	.32	.38	1.88	2.1	1.95	1.8
MIN.	7.0	14.0	-1.37	.22	.32	-1.3	.42	-1.0
REAL RMSE	.28	TRUE SD	.40	SEPARATION	1.41	Person RELIABILITY		.67
MODEL RMSE	.27	TRUE SD	.42	SEPARATION	1.57	Person RELIABILITY		.71
S.E. OF Person MEAN = .09								
Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .72								
SUMMARY OF 14 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	37.3	28.0	.00	.19	.98	-.1	.98	.0
S.D.	17.6	.0	.59	.06	.36	1.3	.32	1.0
MAX.	64.0	28.0	1.20	.34	1.55	1.7	1.43	1.2
MIN.	10.0	28.0	-.75	.15	.29	-2.4	.42	-1.7
REAL RMSE	.21	TRUE SD	.55	SEPARATION	2.56	Item RELIABILITY		.87
MODEL RMSE	.20	TRUE SD	.56	SEPARATION	2.77	Item RELIABILITY		.89
S.E. OF Item MEAN = .16								
UMEAN=.0000 USCALE=1.0000								
Item RAW SCORE-TO-MEASURE CORRELATION = -.98								
392 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 815.93 with 348 d.f. p=.0000								
Global Root-Mean-Square Residual (excluding extreme scores): 1.0737								

Gambar 3. 5. Menu output Summary Statistics

Pada Tabel 3.5 terdapat beberapa nilai reliabilitas penting pada kotak yang berwarna yaitu *Person Reliability*, *Cronbach Alpha (KR-20)* dan *Item Reliability*. Interpretasi untuk nilai *person reliability* dan *item reliability* pada Gambar 3.4 dapat dilihat pada Tabel 3.3 (Sumintono & Widhiarso, 2015; Mohamad, Sulaiman, Sern, & Salleh, 2015). Sedangkan interpretasi untuk nilai *Cronbach alpha* dapat dilihat pada Tabel 3.4 (Sumintono & Widhiarso, 2015).

Tabel 3. 3.

Interpretasi nilai person reliability dan item reliability

Nilai <i>person reliability</i> dan <i>item reliability</i>	Interpretasi
$0,94 \leq$ Nilai	Istimewa
$0,91 \leq$ Nilai $< 0,94$	Bagus Sekali
$0,81 \leq$ Nilai $< 0,90$	Bagus
$0,67 \leq$ Nilai $< 0,80$	Cukup
Nilai $< 0,67$	Lemah

Tabel 3. 4.

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

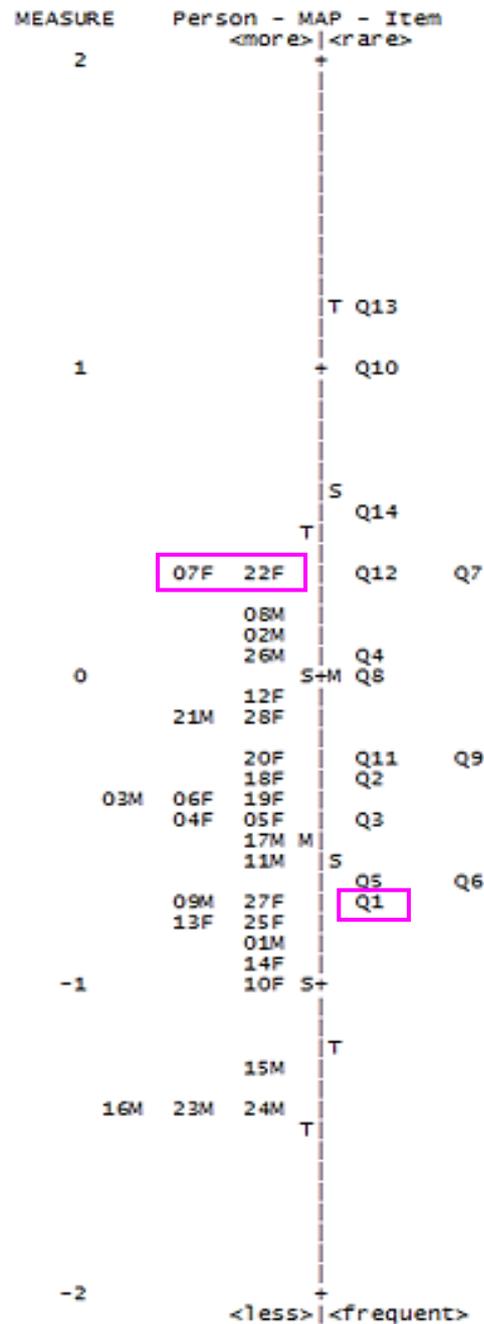
Berdasarkan Gambar 3.5 dan Tabel 3.3 dengan menggunakan analisis Rasch, nilai *person reliability* yang muncul adalah sebesar 0,67, artinya konsistensi jawaban peserta didik berada dalam kategori cukup. Kemudian untuk nilai *item reliability* yang muncul adalah sebesar 0,87, artinya kualitas item tes pada instrumen *four-tier* yang digunakan berada dalam kategori bagus. Nilai *person reliability* dan *item reliability* yang dipakai adalah *REAL RMSE* karena nilai ini merupakan kondisi terburuk reliabilitas batas bawah berdasarkan instrumen yang dipakai (Sumintono & Widhiarso, 2015).

Selanjutnya, nilai *Cronbach Alpha* (KR-20) dari instrumen *four-tier* diketahui sebesar 0,72. Berdasarkan Tabel 3.4 mengenai interpretasi nilai *Cronbach Alpha*, maka nilai KR-20 ini berada dalam kategori bagus. Nilai *person reliability* dan nilai *item reliability* dari nol hingga satu dapat diinterpretasikan sangat mirip dengan alpha Cronbach, yang berarti bahwa nilai yang mendekati angka 1 menunjukkan ukuran yang lebih konsisten (Boone & Noltemeyer, 2017). Hasil perhitungan reliabilitas menunjukkan bahwa secara keseluruhan, instrumen *four-tier* yang diuji sudah *reliable* dan dapat dipercaya sehingga sudah layak digunakan untuk mengukur penurunan miskonsepsi dan peningkatan kemampuan memahami konsep peserta didik pada materi momentum dan impuls.

3.5.3. Tingkat Kesukaran

Tingkat kesukaran bertujuan untuk membedakan soal yang terlalu mudah, mudah, sulit, dan sangat sulit. Hal ini diperlukan untuk menghasilkan distribusi soal

yang baik. Tingkat kesukaran instrumen *four-tier* dapat diketahui dengan menggunakan *software* MINISTEP 4.3.1 pada menu *output Table 1 Variable (Wright) maps* dan *output Table 13 Item Measure* yang dianalisis dengan Rasch. *Output Variable (Wright) maps* pada *software* ini bertujuan untuk mendapat gambaran mengenai kekuatan setiap butir soal. Sesuai dengan penjelasan Boone & Noltemeyer (2017) bahwa *wright maps* 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 *Item Measure* digunakan untuk mengklasifikasikan tingkat kesulitan tiap butir soal berdasarkan nilai logit dan nilai Standar Deviasi (SD) yang didapat (Sumintono & Widhiarso, 2015). Hasil menu *output Table 1 Variable (Wright) maps* instrumen *four-tier* setelah diolah dalam MINISTEP dapat dilihat pada Gambar 3.6 sebagai berikut.



Gambar 3. 6. Menu *Output Variable (Wright) maps*

Gambar 3.6 secara garis besar hanya menggambarkan kekuatan instrumen soal *four-tier* yang sudah dijawab oleh peserta didik. Pada bagian kanan di dalam gambar, kode Q (*Question*) yang diikuti angka menunjukkan nomor butir soal. Selanjutnya pada bagian kiri, kode M (*Male*) atau F (*Female*) secara berurutan menunjukkan peserta didik laki-laki dan perempuan. Di gambar terlihat bahwa soal

yang berada di atas 22F (peserta didik perempuan nomor urut 22) dan 07F merupakan soal yang berpotensi tidak dapat dijawab peserta didik, soal tersebut antara lain adalah Q10, Q13, dan Q14. Kemudian untuk soal selain itu yakni Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, Q11, dan Q12 berpotensi masih bisa dijawab oleh peserta didik. Menu *output Variable (Wright) maps* disini untuk menunjukkan secara umum bagaimana potensi instrumen soal tersebut dijawab oleh peserta didik, untuk lebih jelasnya dapat dilihat melalui menu *output Table 13 Item Measure* pada Gambar 3.7.

Hasil analisis tingkat kesukaran instrumen soal *four-tier* setelah diolah dalam MINISTEP dapat dilihat pada menu *output Table 13 Item Measure* seperti sebagai berikut:

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL		INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Item
				S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%		
13	10	28	1.20	.34	.92	.0	1.26	.7	-.32	.25	53.6	66.7	Q13	
10	12	28	1.00	.30	1.55	1.1	1.43	1.0	.34	.27	67.9	62.5	Q10	
14	19	28	.53	.22	.29	-2.4	.46	-1.4	.42	.35	64.3	54.1	Q14	
12	23	28	.35	.20	.36	-2.4	.42	-1.7	.58	.39	60.7	46.5	Q12	
7	24	28	.31	.20	.85	-.3	.77	-.5	.35	.40	46.4	46.0	Q7	
4	30	28	.10	.18	.98	.0	.85	-.3	.79	.44	17.9	39.8	Q4	
8	34	28	-.02	.17	1.43	1.6	1.32	1.0	.26	.47	28.6	38.7	Q8	
9	44	28	-.29	.16	1.09	.5	1.32	1.1	.38	.51	25.0	31.7	Q9	
11	44	28	-.29	.16	1.27	1.2	1.16	.6	.55	.51	17.9	31.7	Q11	
2	45	28	-.31	.16	1.14	.7	1.03	.2	.68	.52	10.7	28.5	Q2	
3	52	28	-.48	.15	1.01	.1	.97	.0	.50	.54	32.1	28.9	Q3	
5	60	28	-.66	.15	.61	-2.0	.66	-1.4	.56	.54	39.3	24.6	Q5	
6	61	28	-.68	.15	1.39	1.7	1.31	1.2	.37	.54	17.9	22.9	Q6	
1	64	28	-.75	.15	.85	-.7	.74	-1.0	.61	.54	32.1	26.4	Q1	
MEAN	37.3	28.0	.00	.19	.98	-.1	.98	.0			36.7	39.2		
S.D.	17.6	.0	.59	.06	.36	1.3	.32	1.0			18.3	13.6		

Gambar 3. 7. Menu *Output Item Measure*

Gambar 3.7 diatas merupakan hasil output dari 14 instrumen soal yang diolah menggunakan *software* MINISTEP. Pada gambar tersebut menunjukkan bahwa soal nomor 13 pada kolom *ENTRY NUMBER* memiliki nilai *measure* paling tinggi dengan *logit (logarithm odd unit)* sebesar 1,20, sedangkan soal nomor 1 memiliki nilai *measure* paling rendah dengan *logit* sebesar -0,75. Hasil berikutnya pada output ini mengenai nilai Standar Deviasi (SD) yaitu menunjukkan angka 0.59, artinya dari standar deviasi tersebut dapat dibuat klasifikasi kategori tingkat kesulitan instrumen soalnya. Menurut Sumintono & Widhiarso (2015) menyatakan

bahwa nilai $0,00 \text{ logit} + 1\text{SD}$ adalah satu kelompok sukar, lebih besar dari $+1\text{SD}$ adalah soal yang sangat sukar; $0,00 \text{ logit} - 1\text{SD}$ adalah soal yang mudah; dan lebih kecil dari -1SD soal yang sangat mudah. Klasifikasi kategori tingkat kesukaran pada instrumen *four-tier* ini dapat diinterpretasikan seperti pada Tabel 3.5.

Tabel 3. 5.
Interpretasi hasil tingkat kesukaran dengan analisis Rasch

Tingkat Kesukaran (TK)	Interpretasi
$0,59 < \text{TK}$	Sangat Sukar
$0,00 < \text{TK} \leq 0,59$	Sukar
$-0,59 \leq \text{TK} \leq 0,00$	Mudah
$\text{TK} < -0,59$	Sangat Mudah

Tingkat kesulitan setiap item pada instrumen *four-tier* berdasarkan Gambar 3.5 dan Gambar 3.6 dapat diinterpretasikan seperti pada Tabel 3.6 dengan menggunakan analisis Rasch sebagai berikut.

Tabel 3. 6.
Hasil tingkat kesukaran tiap butir tes dengan analisis Rasch

No. Soal	Tingkat Kesukaran (TK)	Interpretasi
1.	-0,75	Sangat Mudah
2.	-0,31	Mudah
3.	-0,48	Mudah
4.	0,10	Sukar
5.	-0,66	Sangat Mudah
6.	-0,68	Sangat Mudah
7.	0,31	Sukar
8.	-0,02	Mudah
9.	-0,29	Mudah
10.	1,00	Sangat Sukar
11.	-0,29	Mudah
12.	0,35	Sukar
13.	1,20	Sangat Sukar
14.	0,53	Sukar

Hasil interpretasi tiap butir soal pada Tabel 3.6 menunjukkan bahwa terdapat soal dengan tingkat kesukaran berbeda-beda, mulai dari sangat mudah; soal nomor 1, 5, dan 6, mudah; soal nomor 2, 3, 8, 9 dan 11, sukar; soal nomor 4, 7, dan 12, dan 14, sangat sukar; soal nomor 10 dan 13. Namun hasil ini berbeda dengan hasil tingkat kesukaran yang dihitung secara manual. Tingkat kesulitan butir dengan analisis Rasch pada dasarnya sama dengan taraf kesukaran teori tes klasik, yaitu perbandingan antara jumlah jawaban benar dengan jumlah soal yang diujikan. Perbedaannya, pada analisis Rasch nilai peluang diskalakan dengan memasukkan fungsi logaritma, sehingga dapat menghasilkan suatu pengukuran dengan interval yang sama karena terjadi transformasi data *odd ratio* dengan logaritma (Sumintono & Widhiarso, 2015).

3.5.4. Daya Pembeda

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* dapat diketahui menggunakan *software* MINISTEP 4.3.1 dengan menu *output Table 10 Item Fit Order* yang dianalisis dengan Rasch. Smiley (2015) menyatakan bahwa korelasi *point-measure* mengacu pada hubungan antara tingkat kesulitan item tiap individu dan tingkat kesulitan soal secara keseluruhan, dimana nilai satu menunjukkan kemampuan peserta didik yang tinggi menjawab soal dengan benar dan sebaliknya yang menunjukkan korelasi sempurna menurut Rasch.

Penggunaan *Point Measure Correlation* seperti yang diungkapkan sebelumnya dapat memberikan informasi mengenai daya pembeda dari suatu instrumen soal yang digunakan untuk membedakan kemampuan peserta didik. Adapun nilai *Point Measure Correlation (Pt Mean Corr)* untuk instrumen *four-tier* dapat dilihat pada Gambar 3.8 sebagai berikut.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
10	12	28	1.00	.30	1.55	1.1	1.43	1.0	A .34	.27	67.9	62.5	Q10
8	34	28	-.02	.17	1.43	1.6	1.32	1.0	B .26	.47	28.6	38.7	Q8
6	61	28	-.68	.15	1.39	1.7	1.31	1.2	C .37	.54	17.9	22.9	Q6
9	44	28	-.29	.16	1.09	.5	1.32	1.1	D .38	.51	25.0	31.7	Q9
11	44	28	-.29	.16	1.27	1.2	1.16	.6	E .55	.51	17.9	31.7	Q11
13	10	28	1.20	.34	.92	.0	1.26	.7	F .32	.25	53.6	66.7	Q13
2	45	28	-.31	.16	1.14	.7	1.03	.2	G .68	.52	10.7	28.5	Q2
3	52	28	-.48	.15	1.01	.1	.97	.0	g .50	.54	32.1	28.9	Q3
4	30	28	.10	.18	.98	.0	.85	-.3	f .79	.44	17.9	39.8	Q4
7	24	28	.31	.20	.85	-.3	.77	-.5	e .35	.40	46.4	46.0	Q7
1	64	28	-.75	.15	.85	-.7	.74	-1.0	d .61	.54	32.1	26.4	Q1
5	60	28	-.66	.15	.61	-2.0	.66	-1.4	c .56	.54	39.3	24.6	Q5
14	19	28	.53	.22	.29	-2.4	.46	-1.4	b .42	.35	64.3	54.1	Q14
12	23	28	.35	.20	.36	-2.4	.42	-1.7	a .58	.39	60.7	46.5	Q12
MEAN	37.3	28.0	.00	.19	.98	-.1	.98	.0			36.7	39.2	
S.D.	17.6	.0	.59	.06	.36	1.3	.32	1.0			18.3	13.6	

Gambar 3. 8. Menu Output item (column): fit order

Untuk dapat mengetahui kategori daya pembeda tiap butir soal instrumen *four-tier* pada kolom *PT-MEASURE CORR.*, Smiley (2015) memberikan interpretasi untuk setiap nilai yang diberikan seperti yang terdapat pada Tabel 3.7.

Tabel 3. 7.

Interpretasi nilai Point Measure Correlation

<i>Pt Mean Corr</i>	Interpretasi
$0,40 < ID$	Sangat baik
$0,30 \leq ID \leq 0,40$	Baik
$0,20 \leq ID < 0,30$	Kurang baik
$ID < 0,20$	Jelek

Berdasarkan Gambar 3.8 dan Tabel 3.7, daya pembeda untuk tiap butir soal dengan mengacu pada nilai *Pt Mean Corr* dapat diinterpretasikan seperti Tabel 3.8. berikut.

Tabel 3. 8.

Interpretasi daya pembeda untuk setiap butir soal

No. Soal	Pt Mean Corr	Interpretasi
1.	0,61	Sangat Baik
2.	0,68	Sangat Baik
3.	0,50	Sangat Baik
4.	0,79	Sangat Baik
5.	0,56	Sangat Baik
6.	0,37	Baik
7.	0,35	Baik
8.	0,26	Kurang Baik
9.	0,38	Baik
10.	0,34	Baik
11.	0,55	Sangat Baik
12.	0,58	Sangat Baik
13.	0,32	Baik
14.	0,42	Sangat Baik

Berdasarkan rekapitulasi Tabel 3.8 dapat dianalisis bahwa dari 14 soal yang diujicobakan, ada satu soal yang memiliki daya pembeda dengan kriteria kurang baik yaitu soal nomor 8. Soal tersebut dapat digunakan, namun diperbaiki terlebih dahulu baik dari segi isi maupun tata bahasanya.

3.6 Teknik Analisis Data

Ada beberapa jenis data yang dikumpulkan dalam penelitian ini yaitu hasil *pre-test*, *post-test*, observasi aksivitas guru berkaitan dengan keterlaksanaan pembelajaran strategi PDEODE dengan berbantuan simulasi komputer dan observasi aktivitas peserta didik. Langkah-langkah dalam analisis data penelitian ini dilakukan dengan langkah sebagai berikut.

3.6.1. Perhitungan Konsepsi Peserta Didik

Terdapat enam kriteria konsepsi peserta didik berdasarkan jawaban peserta didik pada soal *four-tier* yaitu *Misconception* (MC), *Sound Understanding* (SU), *Partial Positive* (PP), *Partial Negative* (PN), *No Understanding* (NU), dan *No Coding* (NC). Penilaian untuk kategori tersebut diadaptasi dari penelitian Kaltakci-Gurel, Eryilmaz, & McDermott (2017) tentang penilaian miskonsepsi, dan hasilnya ditunjukkan oleh Tabel 3.8.

Tabel 3. 9.

Skor Miskonsepsi

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

Berdasarkan Tabel 3.9 maka hasil adaptasi untuk nilai setiap kategori konsepsi ditunjukkan oleh Tabel 3.10.

Tabel 3. 10.

Skor Kriteria Miskonsepsi

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

Skor kriteria konsepsi di atas digunakan untuk analisis Rasch yang nantinya dimasukan kedalam *software* MINISTEP. Tujuannya untuk melihat gambaran potensi 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 Miskonsepsi Peserta Didik

Penurunan kuantitas peserta didik yang mengalami miskonsepsi hanya dilakukan pada kategori konsepsi peserta didik berupa kategori *Misconception* (MC). Kemudian hasilnya diolah menggunakan Persamaan 3.1 tentang Penurunan Kuantitas Miskonsepsi (PKM) hasil adaptasi dari persamaan gain yang dinormalisasi (Kurniawan, Suhandi, & Hasanah, 2016).

$$PKM = \frac{\% \text{ pretest} - \% \text{ posttest}}{\% \text{ pretest} - \% \text{ ideal}} \times 100\% \quad (3.1.)$$

Kemudian nilai PKM dapat diinterpretasikan pada Tabel 3.11.

Tabel 3. 11.

Interpretasi nilai PKM

Nilai PKM	Kategori
$70 < PKM \leq 100$	Tinggi
$30 < PKM \leq 70$	Sedang
$0 < PKM \leq 30$	Rendah

3.6.3. Perubahan Konsepsi

Perubahan konsepsi peserta didik menjadi analisis untuk mengetahui apakah peserta didik mengalami perubahan konsepsi atau tidak. Perubahan konsepsi dilakukan dengan melihat konsepsi peserta didik untuk setiap konsep pada saat *pre-test* dan *post-test*. Persamaan yang digunakan untuk menentukan persentase Perubahan Konsepsi Setiap Kategori (PKSK) diadaptasi dari Samsudin, Suhandi, Rusdiana, Kaniawati, & Coştu (2016) dan ditunjukkan oleh Persamaan 3.2.

$$PK(\%) = \%K_{pre} - \%K_{post} \quad (3.2)$$

Keterangan:

$PK(\%)$ = persentase perubahan konsepsi

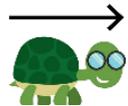
$\%K_{pre}$ = persentase konsepsi pada saat *pre-test*

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

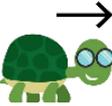
Berdasarkan Persamaan 3.2, maka dibuatlah interpretasi perubahan konsepsi untuk setiap kategori konsepsi. Interpretasi tersebut disajikan dalam Tabel 3.12.

Tabel 3.12

Interpretasi perubahan untuk setiap kategori konsepsi

No	Kategori Konsepsi	Perubahan	Interpretasi
1.		+	<i>Bad Change (BC)</i>
		0	<i>No Change (NC)</i>
		-	<i>Good Change (GC)</i>

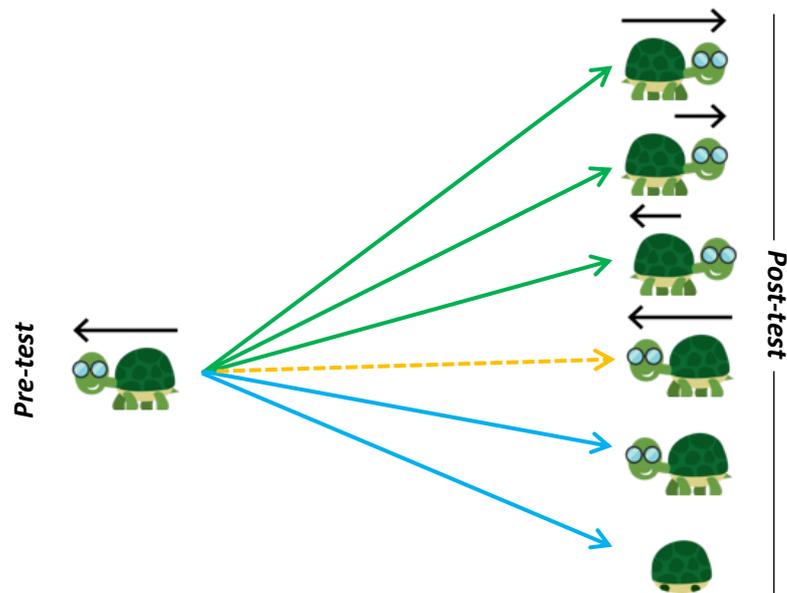
(lanjutan Tabel 3.12)

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

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

Pada Tabel 3.12 terdapat perbedaan interpretasi untuk kategori . 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 (-).

Setelah dilakukan presentase terhadap perubahan konsepsi peserta didik, analisis lebih lanjut dilakukan untuk mengetahui kea rah mana perubahan konsepsi tersebut. Kemungkinan perubahan konsepsi yang dialami peserta didik dapat dilihat berdasarkan simbol konsepsi pada Gambar 3.9 sebagai berikut.



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 3.13.

Tabel 3.13

Kategori perubahan konsepsi

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

Kategori perubahan konsepsi pada Tabel 3.13 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 3.3.

$$PKP(\%) = \frac{\text{jumlah siswa pada suatu tipe}}{\text{jumlah seluruh siswa}} \quad (3.3)$$

3.6.4. Menghitung Gain Skor yang Dinormalisasi

Peningkatan kemampuan memahami konsep peserta didik disajikan menggunakan 2 tahap yaitu menggunakan analisis Rasch dan dengan menghitung rata-rata gain yang dinormalisasi menurut Hake (1999). Untuk mengetahui peningkatan hasil belajar peserta didik pada ranah kognitif, digunakan data hasil *pre-test* dan *post-test* peserta didik. Dari kedua data tes tersebut kemudian dilihat peningkatan atau gain nya yaitu dengan mengurangi data hasil *post-test* dengan data *pre-test*. Gain yang dinormalisasi merupakan perbandingan antara skor gain aktual yaitu skor gain yang diperoleh peserta didik dengan skor gain maksimum yaitu skor gain tertinggi yang mungkin diperoleh peserta didik (Hake, 1999). Kemudian hasilnya diolah menggunakan persamaan 3.2 sebagai berikut.

$$\langle g \rangle = \frac{\% \text{ posttest} - \% \text{ pretest}}{100 - \% \text{ pretest}} \quad (3.4.)$$

Gain yang dinormalisasi tiap peserta didik akan digunakan untuk menghitung rata-rata gain yang dinormalisasi. Data gain yang dinormalisasi tersebut akan menunjukkan seberapa besar usaha yang dibutuhkan untuk meningkatkan kognitif peserta didik. Data gain yang dinormalisasi kemudian dianalisis sesuai dengan kriteria yang diberikan. Menurut Hake (1999), interpretasi rata-rata gain yang

dinormalisasi terhadap peningkatan hasil belajar pada aspek kognitif pada suatu pembelajaran dibagi ke dalam tiga kriteria sebagai mana tercantum pada Tabel 3.14.

Tabel 3. 14.

Kriteria Peningkatan Gain yang Dinormalisasi

Rata-rata gain yang dinormalisasi	Kriteria
$0,00 < \langle g \rangle \leq 0,30$	Rendah
$0,30 < \langle g \rangle \leq 0,70$	Cukup
$0,70 < \langle g \rangle \leq 1,00$	Tinggi

3.6.5. Menghitung Efektifitas Peningkatan Kemampuan Konsep

Efektivitas penerapan strategi PDEODE berbantuan simulasi komputer dapat diketahui dengan perhitungan *effect size*. Efektifitas pembelajaran ini dapat dihitung setelah mengetahui skor untuk masing-masing kriteria konsepsi. Semua desain grup berkorelasi memiliki sifat statistik yang sama, namun untuk kesederhanaan, metode dijelaskan dalam tindakan berulang *pre-test* dan *post-test* desain (Gibbons, Hedeker & Davis, 1993). Dalam menghitung *effect size* (d_{RM}) dapat digunakan perumusan Morris & DeShon (2002) seperti yang ditunjukkan oleh persamaan 3.3 berikut.

$$d_{RM} = \frac{|M_{post} - M_{pre}|}{SD_{pre}} \quad (3.3.)$$

Keterangan:

M_{post} = perbedaan rata-rata antara skor *posttest*

M_{pre} = perbedaan rata-rata antara skor *pretest*

SD_{pre} = standar deviasi skor *pretest*

Hasil perhitungan *effect size* selanjutnya diinterpretasikan berdasarkan Tabel 3.15 seperti berikut.

*Tabel 3. 15.**Interpretasi Hasil effect size*

<i>Effect size</i>	Interpretasi
$0,00 \leq d < 0,20$	Kecil
$0,20 \leq d < 0,50$	Sedang
$0,50 \leq d < 0,80$	Besar
$d \geq 0,80$	Sangat Besar