

BAB III

METODE PENELITIAN

3.1 Metode dan Desain Penelitian

Metode penelitian yang diterapkan pada penelitian ini adalah penelitian kuantitatif. Penelitian ini dilakukan untuk menguji teori-teori tertentu, dengan cara mengumpulkan data untuk membantah teori yang sudah ada (Creswell, 2019). Jenis penelitian yang diterapkan yaitu jenis quasi eksperimen yang berarti penelitian ini mencakup kelompok kontrol, namun tidak sepenuhnya dapat mengontrol variabel-variabel luar yang memengaruhi jalannya eksperimen (Sugiyono, 2017). Pemilihan metode quasi eksperimen dianggap sesuai untuk penelitian ini karena dalam proses pelaksanaannya, terdapat beberapa faktor (variabel eksternal) yang sulit dikendalikan dan berpotensi memengaruhi hasil penelitian. Faktor tersebut yaitu sulitnya mendapatkan izin untuk melakukan pengacakan sampel penelitian guna mendapatkan sampel yang sesuai dengan kriteria penelitian.

Desain penelitian yang digunakan dalam penelitian ini adalah *nonequivalent control group design*. Pada desain penelitian ini baik kelompok eksperimen maupun kelompok kontrol tidak dipilih secara acak (Sugiyono, 2017). Artinya, kelas yang dipilih berisi siswa yang tersedia di lapangan tanpa melakukan proses pengacakan terlebih dahulu. Penelitian ini menggunakan kelas eksperimen berisi sampel yang diberikan perlakuan (*treatment*) berupa model pembelajaran *ICARE* (*Introduction, Connection, Application, Reflection, Extension*) berbasis Simulasi PhET, sementara kelas kontrol hanya menggunakan model pembelajaran *ICARE* (*Introduction, Connection, Application, Reflection, Extension*). Sebelum diberikan perlakuan (*treatment*) kedua kelas tersebut diberikan tes awal (*pretest*) dan tes akhir (*posttest*) yang sama untuk mengetahui pengetahuan awal dan akhir siswa dalam penelitian ini. Skema desain penelitian *nonequivalent control group design* ditunjukkan pada Tabel 3. 1 berikut.

Tabel 3. 1 Desain Penelitian *Nonequivalent Control Group*

Kelas	<i>Pretest</i>	<i>Treatment</i>	<i>Posttest</i>
Eksperimen	O	X	O'
Kontrol	O	X'	O'

Keterangan:

- O : Pemberian soal *pretest* yang dilakukan sebelum diberikan pembelajaran dengan model *ICARE* berbasis PhET
- X : *Treatment* pembelajaran dengan model *ICARE* (*Introduction, Connection, Application, Reflection, Extension*) berbasis PhET
- X' : *Treatment* pembelajaran dengan model *ICARE* (*Introduction, Connection, Application, Reflection, Extension*)
- O' : Pemberian soal *post-test* yang dilakukan setelah diberikan pembelajaran dengan model *ICARE* berbasis PhET

3.2 Populasi, Sampel, dan Partisipan Penelitian

Populasi dalam penelitian ini yaitu seluruh siswa kelas XI di salah satu SMA Negeri di Kota Bandung. Populasi diartikan sebagai daerah generalisasi yang terdiri dari objek atau subjek yang memiliki kualitas dan karakteristik yang sesuai dengan penelitian untuk menjadi fokus studi dan kemudian dapat ditarik kesimpulannya (Suriani dkk., 2023). Sedangkan sampel adalah bagian dari kuantitas dan karakteristik yang dimiliki oleh populasi (Suriani dkk., 2023). Sampel dalam penelitian ini adalah dua kelas XI dari seluruh populasi. Selanjutnya, penelitian ini menggunakan teknik *purposive sampling*. Menurut Sugiyono (2017), teknik *purposive sampling* adalah metode pengambilan sampel dengan mempertimbangkan faktor-faktor tertentu. Dengan teknik sampling yang digunakan, pengambilan sampel ditentukan berdasarkan kelas yang sudah disiapkan oleh pihak sekolah.

Penelitian ini melibatkan 32 siswa dalam kelas eksperimen, 36 siswa dalam kelas kontrol, dan dua observer atau pengamat untuk mengamati keterlaksanaan model pembelajaran *ICARE* berbasis Simulasi PhET. Adapun kelas yang diambil sebagai partisipan yaitu kelas yang belum mendapatkan pembelajaran mengenai materi Fluida Dinamis.

Mega Amalia Putri, 2024

PENGUNAAN MODEL PEMBELAJARAN *ICARE* BERBASIS SIMULASI PHET UNTUK MENINGKATKAN KEMAMPUAN KOGNITIF SISWA PADA MATERI FLUIDA DINAMIS

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

3.3 Prosedur Penelitian

Penelitian ini menggunakan tiga tahap prosedur, yaitu tahap persiapan, tahap pelaksanaan, dan tahap akhir atau tahap penyelesaian. Langkah-langkah pada tiap tahap akan diuraikan sebagai berikut.

3.3.1 Tahap Persiapan

- 1) Melakukan studi literatur dari penelitian terdahulu yaitu dari buku, jurnal, skripsi, tesis, dan sumber lainnya;
- 2) Merumuskan masalah yang akan diteliti dengan menentukan pertanyaan penelitian;
- 3) Mencari solusi alternatif untuk masalah tersebut;
- 4) Merumuskan Hipotesis;
- 5) Menentukan variabel penelitian;
- 6) Menetapkan metode dan desain penelitian;
- 7) Menetapkan subjek penelitian dengan menentukan tempat penelitian, populasi, dan sampel yang akan menjadi fokus penelitian;
- 8) Menyusun instrumen penelitian yang terdiri dari Modul Ajar, Lembar Kerja Peserta Didik (LKPD), instrumen tes kemampuan kognitif, angket respon siswa, dan lembar observasi keterlaksanaan pembelajaran;
- 9) Melakukan validasi dan *judgement* instrumen oleh ahli;
- 10) Melakukan pengolahan, revisi, dan seleksi soal dari data hasil uji coba instrumen;
- 11) Menganalisis instrumen penelitian.

3.3.2 Tahap Pelaksanaan

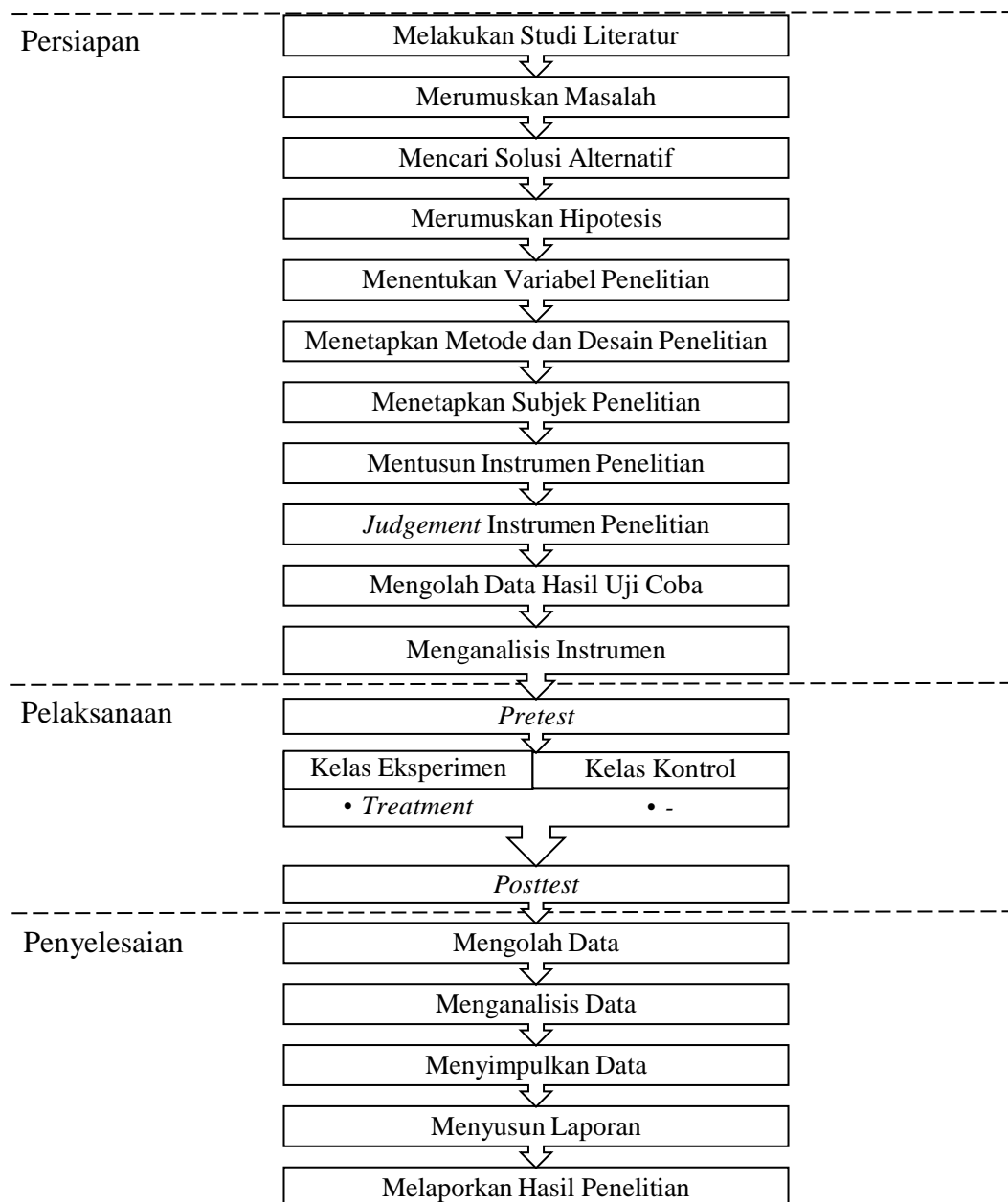
- 1) Mengadakan *pretest* kepada siswa untuk mengetahui gambaran awal terhadap kemampuan kognitif siswa sebelum diberikan *treatment*;
- 2) Melaksanakan pembelajaran dengan menggunakan model pembelajaran *ICARE* berbasis Simulasi PhET untuk kelas eksperimen dan model pembelajaran *ICARE* saja untuk kelas kontrol;
- 3) Melakukan observasi terkait keterlaksanaan kegiatan pembelajaran di kelas yang dilakukan oleh observer;

- 4) Mengadakan *posttest* kepada kedua kelas penelitian untuk mengetahui peningkatan kemampuan kognitifnya;

3.3.3 Tahap Penyelesaian

- 1) Mengolah data yang telah diperoleh dari *pretest*, *posttest*, angket respon, dan lembar observasi keterlaksanaan pembelajaran;
- 2) Menganalisis data kuantitatif hasil penelitian;
- 3) Menyimpulkan data dari hasil penelitian;
- 4) Menyusun laporan penelitian yaitu berupa skripsi;
- 5) Melaporkan hasil penelitian yang telah dilakukan.

Skema penelitian yang digunakan sebagai panduan dalam pelaksanaan penelitian ditunjukkan pada Gambar 3. 1 berikut.



Gambar 3. 1 Alur Prosedur Penelitian

3.4 Instrumen Penelitian

Instrumen penelitian adalah suatu alat yang digunakan untuk mengukur variabel penelitian yang akan diamati (Sugiyono, 2017). Instrumen penelitian yang digunakan untuk pengumpulan data diantaranya adalah lembar observasi keterlaksanaan pembelajaran, tes kemampuan kognitif, dan lembar angket respon siswa. Penjelasan instrumen penelitian yang digunakan pada penelitian ini adalah sebagai berikut.

3.4.1 Tes Kemampuan Kognitif

Kemampuan kognitif yang dimiliki siswa pada penelitian ini diukur dengan menggunakan instrumen tes berbentuk pilihan ganda yang diberikan sebelum dan setelah proses pembelajaran. Soal yang diberikan sebelum pembelajaran (*pretest*) sama dengan soal yang diberikan setelah pembelajaran (*posttest*). Hasil *pretest* dan *posttest* yang dibandingkan digunakan untuk mengukur peningkatan kemampuan kognitif siswa dalam menguasai materi pembelajaran setelah diberikan *treatment* (perlakuan) menggunakan model pembelajaran *ICARE* berbasis Simulasi PhET.

Tes kemampuan kognitif terdiri dari 23 butir soal yang mencakup materi Fluida Dinamis. Peneliti menggunakan kategori dimensi proses kognitif taksonomi bloom revisi oleh Anderson dan Krathwohl untuk membuat soal tes pilihan ganda. Aspek kemampuan kognitif pada penelitian ini dibatasi dengan menggunakan tiga aspek yaitu memahami (C2), menerapkan (C3), dan menganalisis (C4). Sebelum diterapkan pada penelitian, instrumen tes yang digunakan dalam penelitian ini diuji kelayakannya melalui penilaian ahli dan uji coba. *Judgement expert* atau penilaian oleh ahli pada instrumen tes penelitian ini dilakukan oleh 2 dosen dan 1 guru untuk menguji kelogisan konten dan kesesuaian dengan indikator yang diinginkan. Jumlah soal yang diajukan untuk *judgement expert* yaitu 27 butir soal. Setelah *judgement* oleh ahli, instrumen tes diuji cobakan pada kelompok siswa yang telah mempelajari materi Fluida Dinamis. Hasil uji coba kemudian diuji untuk validitas, reliabilitas, dan tingkat kesukaran butir soal, serta daya pembeda untuk memastikan bahwa soal tersebut layak digunakan untuk penelitian. Adapun metode yang digunakan untuk menganalisis hasil uji coba instrumen tes adalah sebagai berikut:

3.4.1.1 Validitas Butir Soal

Proses validasi dilakukan dengan memberikan umpan balik guna meningkatkan kualitas instrumen tes yang hendak digunakan. Apabila suatu instrumen dapat mengukur apa yang seharusnya diukur, maka instrumen tersebut dapat dikatakan valid (Sugiyono, 2013). Pengujian validitas instrumen pada penelitian ini dilakukan berdasarkan validitas isi (*content*) dan validitas konstruk (*construct*).

Mega Amalia Putri, 2024

PENGUNAAN MODEL PEMBELAJARAN ICARE BERBASIS SIMULASI PHET UNTUK MENINGKATKAN KEMAMPUAN KOGNITIF SISWA PADA MATERI FLUIDA DINAMIS

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

1. Uji Validitas Isi

Uji validitas isi adalah penilaian kualitas instrumen penelitian yang dilakukan oleh ahli (*judgement expert*). Validator untuk instrumen tes kemampuan kognitif pada penelitian ini terdiri dari dua dosen ahli Fisika Universitas Pendidikan Indonesia (UPI) dan satu guru mata pelajaran Fisika. *Judgement* oleh ahli dilakukan untuk menilai validitas tes terhadap kesesuaian antara bentuk soal dengan aspek kemampuan kognitif dengan menggunakan lembar validasi instrumen. Setiap validator memberi tanda *checklist* (\checkmark) pada enam aspek penilaian di setiap butir soal berupa “Sesuai” dan “Tidak sesuai”. Penelitian ini terdiri dari 27 butir soal yang divalidasi oleh ahli. Hasil penilaian dari setiap validator diperbaiki berdasarkan arahan dan catatan yang diberikan validator pada lembar validasi instrumen tes.

2. Uji Validitas Konstruk

Uji validitas konstruk dilakukan dengan uji coba instrumen tes kepada 58 siswa kelas XI MIPA di SMA Negeri di luar sekolah penelitian di Kota Bandung yang telah mempelajari materi Fluida Dinamis. Uji coba soal ini dilakukan melalui *google form*. Pengolahan hasil uji validitas konstruk dilakukan menggunakan analisis Rasch dengan *software Winstep* versi 4.5.0 berdasarkan data hasil uji coba yang diperoleh. Uji validitas yang dianalisis dengan menggunakan pemodelan Rasch disebut dengan *Unidimensionality* (Sumintono & Widhiarso, 2015). *Unidimensionality* atau unidimensionalitas suatu instrumen dapat dilihat dari *raw variance explained by measures* menggunakan *software Winstep* versi 4.5.0 yang diperoleh dari menu *output* tabel 23.0: *Item dimensionality*. Hasil uji validitas konstruk diinterpretasikan berdasarkan Tabel 3.2 berikut.

Tabel 3. 2 Interpretasi Unidimensionalitas Instrumen

<i>Raw variance explained by measures</i>	Interpretasi
$20\% \leq \textit{value} < 40\%$	Terpenuhi
$40\% \leq \textit{value} < 60\%$	Sesuai
$\textit{value} \geq 60\%$	Istimewa

(Sumintono & Widhiarso, 2015)

Hasil dari unidimensionalitas instrumen tes kemampuan kognitif yang diperoleh dari *output* tabel *Item dimensionality* pada *software Winstep* versi 4.5.0 ditunjukkan pada Gambar 3.2 berikut.

TABLE 23.0 C:\Users\admin\OneDrive - upi.edu\SKR ZOU054WS.TXT. Jan 23 2024 16:46
 INPUT: 58 Person 27 Item REPORTED: 58 Person 27 Item 2 CATS WINSTEPS 4.5.0

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

	Eigenvalue	Observed	Expected
Total raw variance in observations =	41.8650	100.0%	100.0%
Raw variance explained by measures =	14.8650	35.5%	34.6%
Raw variance explained by persons =	4.6432	11.1%	10.8%
Raw Variance explained by items =	10.2218	24.4%	23.8%
Raw unexplained variance (total) =	27.0000	64.5%	100.0%
Unexplnd variance in 1st contrast =	2.8148	13.7%	21.2%
Unexplnd variance in 2nd contrast =	2.7058	6.5%	10.0%
Unexplnd variance in 3rd contrast =	2.4648	5.9%	9.1%
Unexplnd variance in 4th contrast =	2.0763	5.0%	7.7%
Unexplnd variance in 5th contrast =	1.8470	4.4%	6.8%

Gambar 3. 2 *Output* Tabel *Item Dimensionality*

Gambar 3.2 menunjukkan nilai *raw variance explained by measures* yang diperoleh dari uji coba instrumen tes kemampuan kognitif dengan 27 butir soal adalah sebesar 35,5%. Berdasarkan Tabel 3.2, nilai tersebut termasuk dalam kategori interpretasi “Terpenuhi”. Sehingga dapat dikatakan bahwa instrumen yang digunakan pada penelitian ini dapat mengukur satu variabel tanpa dipengaruhi oleh variabel-variabel lain. Selain nilai *raw variance explained by measures*, unidimensionalitas instrumen tes juga dapat dilihat dari nilai *unexplained variance in 1st contrast*. Sebuah instrumen dapat dikatakan unidimensionalitas yang baik jika *unexplained variance in 1st contrast* bernilai kurang dari 15% (Fisher, 2007). Hasil analisis instrumen tes kemampuan kognitif pada penelitian ini menunjukkan nilai *unexplained variance in 1st contrast* sebesar 13,7% sehingga dapat dikatakan bahwa instrumen yang digunakan pada penelitian ini memiliki kuantitas unidimensionalitas instrumen tes yang baik.

Selanjutnya dilakukan uji validitas untuk setiap butir soal yang dianalisis menggunakan permodelan Rasch dengan menggunakan *software Winstep* versi 4.5.0. Uji validitas untuk setiap butir soal ini dilakukan untuk mengetahui kualitas dari setiap butir soal. Uji validitas butir soal diperoleh dari *output* tabel 10.1: *Item fit order*. Kelayakan untuk setiap butir soal dapat dilihat kualitasnya pada *item fit order* dari nilai *outfit means square* (MNSQ), *outfit Z-standard* (ZSTD), dan *point*

measure correlation (Pt Measure Corr.). Nilai dari masing-masing kriteria tersebut ditunjukkan pada Tabel 3.3 berikut.

Tabel 3. 3 Kriteria *outfit* MNSQ, ZSTD, dan *Pt Measure Corr.*

Kriteria Nilai <i>Fit-Statistic</i>	Nilai
<i>Outfit Means Square</i> (MNSQ)	$0,50 < \text{MNSQ} < 1,50$
<i>Outfit Z-Standard</i> (ZSTD)	$-2,00 < \text{ZSTD} < 2,00$
<i>Point Measure Correlation (Pt Measure Corr.)</i>	$0,40 < \text{Pt Measure Corr.} < 0,85$

(Sumintono & Widhiarso, 2015)

Setelah diperoleh hasil dari nilai masing-masing kriteria tersebut, selanjutnya dapat diinterpretasikan berdasarkan kriteria nilai *fit-statistic* pada Tabel 3.4 berikut.

Tabel 3. 4 Interpretasi Kualitas Butir Soal

Kriteria	Interpretasi
Ketiga kriteria nilai terpenuhi	Sangat sesuai
Dua dari tiga kriteria nilai terpenuhi	Sesuai
Satu dari tiga kriteria nilai terpenuhi	Kurang sesuai
Semua kriteria tidak nilai terpenuhi	Tidak sesuai

(Sumintono & Widhiarso, 2014)

Hasil uji validitas setiap butir soal kemampuan kognitif dengan menggunakan *software Winstep* versi 4.5.0 pada penelitian ini ditunjukkan pada Gambar 3.3.

TABLE 10.1 C:\Users\admin\OneDrive - upi.edu\SKR ZOU054WS.TXT, Jan 23 2024 16:46
 INPUT: 58 Person 27 Item REPORTED: 58 Person 27 Item 2 CATS WINSTEPS 4.5.0

 Person: REAL SEP.: 1.80 REL.: .76 ... Item: REAL SEP.: 3.76 REL.: .93

 Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASURE CORR.	AL EXP.	EXACT OBS%	MATCH EXP%	Item
13	47	58	-1.77	.35	1.36	1.69	2.63	2.77	A-.13	.29	81.0	81.1	S13
6	24	58	.33	.29	1.53	4.32	2.49	5.72	B-.14	.40	53.4	67.6	S6
18	28	58	.00	.29	1.79	6.49	1.89	3.76	C-.29	.39	27.6	66.8	S18
21	30	58	-.17	.29	1.79	6.44	1.83	3.36	D-.28	.39	22.4	67.2	S21
11	5	58	2.78	.52	1.13	.45	1.72	1.15	E-.22	.38	93.1	92.4	S11
27	45	58	-1.54	.34	1.38	2.05	1.66	1.46	F-.07	.31	77.6	77.9	S27
12	8	58	2.13	.42	.99	.07	1.33	.82	G-.35	.40	87.9	88.0	S12
4	18	58	.88	.31	1.02	.17	1.14	.69	H-.37	.41	75.9	73.5	S4
2	57	58	-4.50	1.02	1.02	.34	.66	.11	I-.11	.09	98.3	98.3	S2
7	54	58	-3.02	.53	1.02	.20	.86	.03	J-.17	.18	93.1	93.1	S7
16	13	58	1.41	.35	.91	-.44	.99	.07	K-.46	.41	84.5	80.7	S16
8	28	58	.00	.29	.90	-.99	.95	-.19	L-.47	.39	75.9	66.8	S8
10	31	58	-.25	.29	.81	-1.97	.94	-.24	M-.52	.39	84.5	67.4	S10
14	27	58	.08	.29	.91	-.91	.93	-.29	N-.47	.40	67.2	66.8	S14
24	23	58	.42	.29	.91	-.80	.82	-.98	m-.50	.40	70.7	68.3	S24
9	17	58	.97	.32	.87	-.88	.72	-1.24	l-.54	.41	77.6	74.8	S9
23	18	58	.88	.31	.83	-1.22	.72	-1.33	k-.56	.41	75.9	73.5	S23
17	11	58	1.67	.37	.82	-.76	.72	-.80	j-.55	.40	86.2	83.7	S17
5	47	58	-1.77	.35	.79	-1.10	.56	-1.07	i-.50	.29	81.0	81.1	S5
1	28	58	.00	.29	.78	-2.33	.72	-1.57	h-.59	.39	82.8	66.8	S1
3	36	58	-.67	.29	.78	-2.10	.71	-1.17	g-.56	.37	79.3	69.5	S3
22	26	58	.16	.29	.75	-2.62	.68	-1.85	f-.62	.40	82.8	66.8	S22
20	31	58	-.25	.29	.74	-2.75	.69	-1.60	e-.61	.39	84.5	67.4	S20
15	31	58	-.25	.29	.73	-2.88	.68	-1.64	d-.62	.39	81.0	67.4	S15
19	30	58	-.17	.29	.69	-3.51	.62	-2.08	c-.66	.39	81.0	67.2	S19
26	33	58	-.42	.29	.67	-3.62	.63	-1.80	b-.66	.38	91.4	68.2	S26
25	4	58	3.08	.58	.64	-.76	.29	-1.11	a-.63	.37	94.8	94.0	S25
MEAN	27.8	58.0	.00	.36	.98	-.3	1.00	.0			77.5	75.4	
P.SD	13.7	.0	1.59	.15	.31	2.6	.58	1.9			17.3	10.0	

Gambar 3. 3 Hasil Uji Validitas Setiap Butir Soal

Berdasarkan Gambar 3.3 diperoleh informasi mengenai nilai nilai *outfit means square* (MNSQ), *outfit z-standard* (ZSTD), dan *point measure correlation* (*Pt Measure Corr.*) pada butir soal hasil uji coba pada penelitian ini. Interpretasi data dari hasil uji validitas setiap butir soal ditunjukkan pada Tabel 3.5 berikut.

Tabel 3. 5 Hasil Uji Validitas Item Instrumen Tes Kemampuan Kognitif

Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Jumlah Kriteria Terpenuhi	Interpretasi	Simpulan
	MNSQ	ZSTD				
S1	0,72	-1,57	0,59	3 Kriteria	Sangat sesuai	Digunakan
S2	0,66	0,11	0,11	2 Kriteria	Sesuai	Digunakan
S3	0,71	-1,17	0,56	3 Kriteria	Sangat sesuai	Digunakan
S4	1,14	0,69	0,37	2 Kriteria	Sesuai	Digunakan
S5	0,56	-1,07	0,50	3 Kriteria	Sangat sesuai	Digunakan

Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Jumlah Kriteria Terpenuhi	Interpretasi	Simpulan
	MNSQ	ZSTD				
S6	2,49	5,72	-0,14	0 Kriteria	Tidak sesuai	Tidak digunakan
S7	0,86	0,03	0,17	2 Kriteria	Sesuai	Digunakan
S8	0,95	-0,19	0,47	3 Kriteria	Sangat sesuai	Digunakan
S9	0,72	-1,24	0,54	3 Kriteria	Sangat sesuai	Digunakan
S10	0,94	-0,24	0,52	3 Kriteria	Sangat sesuai	Digunakan
S11	1,72	1,15	0,22	1 Kriteria	Kurang sesuai	Digunakan
S12	1,33	0,82	0,35	2 Kriteria	Sesuai	Digunakan
S13	2,63	2,77	-0,13	0 Kriteria	Tidak sesuai	Tidak digunakan
S14	0,93	-0,29	0,47	3 Kriteria	Sangat sesuai	Digunakan
S15	0,68	-1,64	0,62	3 Kriteria	Sangat sesuai	Digunakan
S16	0,99	0,07	0,46	3 Kriteria	Sangat sesuai	Digunakan
S17	0,72	-0,80	0,55	3 Kriteria	Sangat sesuai	Digunakan
S18	1,89	3,76	-0,29	0 Kriteria	Tidak sesuai	Tidak digunakan
S19	0,62	-2,08	0,66	2 Kriteria	Sesuai	Digunakan
S20	0,69	-1,60	0,61	3 Kriteria	Sangat sesuai	Digunakan
S21	1,83	3,36	-0,28	0 Kriteria	Tidak sesuai	Tidak digunakan
S22	0,68	-1,85	0,62	3 Kriteria	Sangat sesuai	Digunakan
S23	0,72	-1,33	0,56	3 Kriteria	Sangat sesuai	Digunakan
S24	0,82	-0,98	0,50	3 Kriteria	Sangat sesuai	Digunakan
S25	0,29	-1,11	0,63	2 Kriteria	Sesuai	Digunakan
S26	0,63	-1,80	0,66	3 Kriteria	Sangat sesuai	Digunakan

Butir Soal	Nilai <i>Outfit</i>		<i>Pt Measure Corr.</i>	Jumlah Kriteria Terpenuhi	Interpretasi	Simpulan
	MNSQ	ZSTD				
S27	1,66	1,46	-0,07	2 Kriteria	Sesuai	Digunakan

Hasil uji validitas setiap butir soal yang ditunjukkan pada Tabel 3.5 diperoleh 15 butir soal yang “Sangat sesuai”, 7 butir soal yang “Sesuai”, 1 butir soal yang “Kurang sesuai”, dan 4 butir soal yang “Tidak sesuai”. Dari hasil tersebut, peneliti memutuskan untuk mereduksi sebanyak 4 butir soal yang tidak sesuai karena dianggap tidak fit dibanding soal lainnya atau tidak sesuai dengan analisis Rasch. Sedangkan untuk satu butir soal yang kurang sesuai diputuskan untuk direvisi. Dengan demikian, jumlah soal yang yang memenuhi kriteria *item-fit* dan layak digunakan dalam penelitian ini berjumlah 23 butir soal. Rincian butir soal yang digunakan dalam penelitian ditunjukkan pada Tabel 3.6 berikut.

Tabel 3. 6 Rincian Instrumen Tes Kemampuan Kognitif

No	Tingkat Kognitif	Nomor Soal	Jumlah Soal
1	Memahami (C2)	1, 2, 3, 4, 5, 12, 17, 23, 25.	9
2	Mengaplikasikan (C3)	7, 8, 9, 10, 11, 19, 20, 24, 26, 27.	10
3	Menganalisis (C4)	14, 15, 16, 22.	4

Pada Tabel 3.6, jumlah butir soal pada level C2 lebih sedikit dibandingkan dengan C3 karena hasil analisis Rasch pada uji validitas menunjukkan adanya soal yang tidak valid dari data uji coba. Distribusi butir soal pada masing-masing level kognitif dalam tes tidak dibuat merata karena didasarkan pada karakteristik setiap sub-konsep dalam materi fluida dinamis. Materi dibagi menjadi lima sub-konsep: Fluida Ideal, Debit, Asas Kontinuitas, Hukum Bernoulli, dan Penerapan Hukum Bernoulli Dalam Kehidupan dan Teknologi. Level C2, yang berfokus pada pemahaman dasar, terdapat pada semua sub-konsep karena dasar pemahaman diperlukan di setiap area. Sebaliknya, level C3 yang melibatkan perhitungan hanya diterapkan pada sub-konsep Asas Kontinuitas dan Hukum Bernoulli, di

mana proses matematis dan perhitungan diperlukan. Sementara itu, level C4, yang mengukur kemampuan analisis, terbatas pada penerapan Hukum Bernoulli dalam kehidupan sehari-hari dan teknologi yang memerlukan analisis mendalam. Soal pada level kognitif yang lebih tinggi (C3 dan C4) sering kali memerlukan lebih banyak waktu dan upaya untuk menyelesaikannya. Oleh karena itu, jumlah butir soal dikurangi untuk menjaga kualitas dan relevansi soal. Dengan cara ini, distribusi soal sesuai dengan kompleksitas dan kebutuhan evaluasi dari masing-masing sub-konsep.

3.4.1.2 Reliabilitas Butir Soal

Reliabilitas adalah suatu indeks yang menunjukkan sejauh mana suatu alat ukur dapat dipercaya atau diandalkan (Tarigan, 2022). Uji reliabilitas adalah konsistensi dari serangkaian alat ukur (Arikunto, 2017). Artinya, uji reliabilitas butir soal digunakan untuk menguji keajegan butir soal apabila diberikan berulang kali pada objek yang sama. Jika suatu instrumen tes diberikan kepada kelompok orang yang sama pada waktu yang berbeda tetapi menunjukkan hasil yang sama dengan hasil sebelumnya, maka dapat dikatakan bahwa instrumen tersebut memiliki reliabilitas yang konsisten (Sumintono & Widhiarso, 2015). Uji reliabilitas pada penelitian ini dianalisis dengan pemodelan Rasch menggunakan *software Winstep* versi 4.5.0 pada menu *output* tabel 3.1: *Summary statistic*. Data yang didapatkan pada menu *Summary statistic* diantaranya yaitu *person reliability* (p), *item reliability* (r), dan *cronbach alpha* (KR-20). *Person reliability* menunjukkan reliabilitas siswa, *item reliability* menunjukkan reliabilitas instrumen, dan *cronbach alpha* (KR-20) menunjukkan interaksi yang terjadi antara *person reliability* (p) dan *item reliability* (r). Berikut adalah interpretasi informasi *person reliability* (p), *item reliability* (r), dan *cronbach alpha* (KR-20) yang ditunjukkan pada Tabel 3.7 berikut.

Tabel 3. 7 Interpretasi *Person Realibility*, *Item Realibility*, dan *Cronbach Alpha*

Statistik	Nilai Indeks	Interpretasi
<i>Item dan person reability</i>	$r > 0,94$	Istimewa
	$0,90 < r \leq 0,94$	Sangat baik
	$0,80 < r \leq 0,90$	Baik

	$0,67 < r \leq 0,80$	Cukup
	$r \leq 0,67$	Rendah
<i>Cronbach alpha</i> (KR-20)	$KR - 20 \geq 0,80$	Sangat tinggi
	$0,70 \leq KR - 20 < 0,80$	Tinggi
	$0,60 \leq KR - 20 < 0,70$	Baik
	$0,50 \leq KR - 20 < 0,60$	Sedang
	$KR - 20 < 0,50$	Rendah

(Sumintono & Widhiarso, 2015)

Hasil uji reliabilitas instrumen tes kemampuan kognitif yang ditunjukkan pada hasil *output* tabel 3.1: *Summary statistic* ditunjukkan pada Gambar 3.4 berikut.

TABLE 3.1 C:\Users\admin\OneDrive - upi.edu\SKRI ZOU054W5.TXTp Jan 23 2024 16:46 INPUT: 58 Person 27 Item REPORTED: 58 Person 27 Item 2 CATS WINSTEPS 4.5.0									
SUMMARY OF 58 MEASURED Person									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	
MEAN	12.9	27.0	-.04	.49	.99	-.06	1.06	.12	
SEM	.6	.0	.14	.01	.03	.13	.08	.11	
P.SD	4.4	.0	1.06	.09	.21	.97	.64	.84	
S.SD	4.4	.0	1.07	.09	.21	.98	.64	.85	
MAX.	26.0	27.0	4.21	1.07	1.68	3.12	3.31	2.96	
MIN.	6.0	27.0	-1.68	.45	.65	-2.24	.53	-.90	
REAL RMSE	.52	TRUE SD	.93	SEPARATION	1.80	Person RELIABILITY	.76		
MODEL RMSE	.49	TRUE SD	.94	SEPARATION	1.91	Person RELIABILITY	.78		
S.E. OF Person MEAN = .14									
Person RAW SCORE-TO-MEASURE CORRELATION = .99									
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .76 SEM = 2.17									
SUMMARY OF 27 MEASURED Item									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	
MEAN	27.8	58.0	.00	.36	.98	-.27	1.06	.04	
SEM	2.7	.0	.31	.03	.06	.50	.11	.37	
P.SD	13.7	.0	1.59	.15	.31	2.56	.58	1.91	
S.SD	13.9	.0	1.62	.15	.32	2.61	.59	1.94	
MAX.	57.0	58.0	3.08	1.02	1.79	6.49	2.63	5.72	
MIN.	4.0	58.0	-4.50	.29	.64	-3.62	.29	-2.08	
REAL RMSE	.41	TRUE SD	1.54	SEPARATION	3.76	Item RELIABILITY	.93		
MODEL RMSE	.39	TRUE SD	1.54	SEPARATION	3.91	Item RELIABILITY	.94		
S.E. OF Item MEAN = .31									

Gambar 3.4 Hasil Analisis *Rasch Summary Statistic*

Berdasarkan Gambar 3.4 diperoleh bahwa *person reliability* bernilai 0,76 dengan interpretasi “Cukup”. Hasil interpretasi ini menunjukkan bahwa kemungkinan kurangnya jumlah sampel dan butir soal pada instrumen tes

penelitian ini. Hal ini dikarenakan untuk mendapatkan koefisien *person reliability* yang tinggi, diperlukan sampel dengan rentang kemampuan yang besar dan jumlah butir soal yang banyak (Sumintono & Widhiarso, 2015). Nilai *item reliability* diperoleh sebesar 0,93 dengan interpretasi “Sangat Baik”. Hasil interpretasi pada *item reliability* ini menunjukkan bahwa instrumen tes kemampuan kognitif yang digunakan pada penelitian ini dinyatakan sangat baik dalam hal konsistensi bobot soal dan pilihan jawabannya. Kemudian, nilai *cronbach alpha* (KR-20) diperoleh sebesar 0,76 dengan interpretasi “Tinggi”. Hasil interpretasi ini menunjukkan bahwa instrumen tes kemampuan kognitif pada penelitian ini memiliki kualitas yang baik. Sehingga instrumen tes ini dapat mengidentifikasi hubungan antara siswa (*person reliability*) dengan butir soal (*item reliability*). Dengan demikian, dapat disimpulkan bahwa instrumen tes kemampuan kognitif yang digunakan pada penelitian ini adalah valid dan reliabel.

3.4.1.3 Tingkat Kesukaran Butir Soal

Tingkat kesulitan butir soal menunjukkan kategori dari suatu butir soal, yaitu katogori sukar, sedang, atau mudah. Arikunto (2017) menyatakan bahwa tingkat kesukaran butir soal menunjukkan banyaknya responden yang dapat menjawab benar pada suatu butir soal. Analisis tingkat kesukaran butir soal pada penelitian ini dilakukan dengan menggunakan pemodelan Rasch menggunakan *software Winstep*. Tingkat kesukaran butir soal dapat ditinjau dari nilai *measure* (ME) dan standar deviasi (SD) dengan cara membandingkan nilai *logit* ME pada masing-masing item dan nilai SD (Sumintono & Widhiarso, 2015). Interpretasi dari tingkat kesukaran butir soal dapat dilihat berdasarkan kriteria pada Tabel 3.8 berikut.

Tabel 3. 8 Interpretasi Tingkat Kesukaran

Kriteria Tingkat Kesukaran	Interpretasi
$ME < -1SD$	Mudah
$-1SD \leq ME \leq +1SD$	Sedang
$ME > +1SD$	Sulit

(Sumintono & Widhiarso, 2015)

Hasil analisis tingkat kesukaran dari hasil uji coba instrumen tes dengan menggunakan *software Winstep* versi 4.5.0 pada *output* tabel *item measure* diperoleh nilai stansar deviasi (SD) sebesar 1,59. Interpretasi kriteria tingkat kesukaran butir soal ditunjukkan pada Tabel 3.9 berikut.

Tabel 3. 9 Hasil Interpretasi Tingkat Kesukaran Butir Soal

Nomor Butir Soal	Measure (ME)	Standar Deviasi (SD)	Kriteria	Interpretasi
S1	0,00	1,59	$-1,59 \leq 0,00 \leq 1,59$	Sedang
S2	-4,50	1,59	$-4,50 < -1,59$	Mudah
S3	-0,67	1,59	$-1,59 \leq -0,67 \leq 1,59$	Sedang
S4	0,88	1,59	$-1,59 \leq 0,88 \leq 1,59$	Sedang
S5	-1,77	1,59	$-1,77 < -1,59$	Mudah
S6	0,33	1,59	$-1,59 \leq 0,33 \leq 1,59$	Sedang
S7	-3,02	1,59	$-3,02 < -1,59$	Mudah
S8	0,00	1,59	$-1,59 \leq 0,00 \leq 1,59$	Sedang
S9	0,97	1,59	$-1,59 \leq 0,97 \leq 1,59$	Sedang
S10	-0,25	1,59	$-1,59 \leq -0,25 \leq 1,59$	Sedang
S11	2,78	1,59	$2,78 > 1,59$	Sulit
S12	2,13	1,59	$2,13 > 1,59$	Sulit
S13	-1,77	1,59	$-1,77 < -1,59$	Mudah
S14	0,08	1,59	$-1,59 \leq 0,08 \leq 1,59$	Sedang
S15	-0,25	1,59	$-1,59 \leq -0,25 \leq 1,59$	Sedang
S16	1,41	1,59	$-1,59 \leq 1,41 \leq 1,59$	Sedang
S17	1,67	1,59	$1,67 > 1,59$	Sulit
S18	0,00	1,59	$-1,59 \leq 0,00 \leq 1,59$	Sedang
S19	-0,17	1,59	$-1,59 \leq -0,17 \leq 1,59$	Sedang
S20	-0,25	1,59	$-1,59 \leq -0,25 \leq 1,59$	Sedang
S21	-0,17	1,59	$-1,59 \leq -0,17 \leq 1,59$	Sedang
S22	0,16	1,59	$-1,59 \leq 0,16 \leq 1,59$	Sedang
S23	0,88	1,59	$-1,59 \leq 0,88 \leq 1,59$	Sedang
S24	0,42	1,59	$-1,59 \leq 0,42 \leq 1,59$	Sedang
S25	3,08	1,59	$3,08 > 1,59$	Sulit
S26	-0,42	1,59	$-1,59 \leq -0,42 \leq 1,59$	Sedang
S27	-1,54	1,59	$-1,59 \leq -1,54 \leq 1,59$	Sedang

Berdasarkan Tabel 3.9 diperoleh bahwa butir soal nomor 2, 5, 7, dan 13 termasuk ke dalam kategori “Mudah” pada tingkat kesukaran butir soal. Butir soal

nomor 1, 3, 4, 6, 8, 9, 10, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 26, dan 27 termasuk ke dalam kategori “Sedang”. Sedangkan butir soal nomor 11, 12, 17, dan 25 termasuk ke dalam kategori “Sulit” pada tingkat kesukaran butir soal.

3.4.1.4 Daya Pembeda

Daya pembeda digunakan untuk menilai sejauh mana suatu butir soal dapat membedakan individu dengan tingkat kemampuan tinggi dan rendah (Sumintono & Widhiarso, 2015). Dalam penelitian ini, daya pembeda dianalisis menggunakan pemodelan *Rasch*, dengan melibatkan nilai *Point Measure Correlation (Pt Mean Corr)* yang mengukur korelasi antara tingkat kesulitan setiap butir soal dan tingkat kesulitan keseluruhan tes. Nilai satu menunjukkan adanya korelasi sempurna antara respons butir soal dengan kemampuan peserta tes, nilai nol menunjukkan tidak adanya hubungan antara respons butir soal tertentu dengan keseluruhan butir soal, dan nilai negatif menunjukkan bahwa butir soal tersebut memiliki kelemahan mendasar yang perlu diperiksa kembali atau dihapus dari tes (Smiley, 2015). Interpretasi daya pembeda untuk butir soal dapat dilihat pada Tabel 3.10 berikut.

Tabel 3. 10 Interpretasi Daya Pembeda Butir Soal

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

(Smiley, 2015)

Hasil analisis daya pembeda dari instrumen tes kemampuan kognitif yang dilakukan menggunakan *software Winstep* versi 4.5.0 disajikan pada Tabel 3.11 berikut.

Tabel 3. 11 Hasil Interpretasi Daya Pembeda

Nomor Butir Soal	<i>Pt Measure Corr.</i>	Interpretasi	Keterangan
S1	0,59	Sangat baik	Digunakan

Nomor Butir Soal	<i>Pt Measure Corr.</i>	Interpretasi	Keterangan
S2	0,11	Buruk	Diperbaiki-digunakan
S3	0,56	Sangat baik	Digunakan
S4	0,37	Baik	Digunakan
S5	0,50	Sangat baik	Digunakan
S6	-0,14	Buruk	Tidak digunakan
S7	0,17	Buruk	Diperbaiki-digunakan
S8	0,47	Sangat baik	Digunakan
S9	0,54	Sangat baik	Digunakan
S10	0,52	Sangat baik	Digunakan
S11	0,22	Kurang Baik	Diperbaiki-digunakan
S12	0,35	Baik	Digunakan
S13	-0,13	Buruk	Tidak digunakan
S14	0,47	Sangat baik	Digunakan
S15	0,62	Sangat baik	Digunakan
S16	0,46	Sangat baik	Digunakan
S17	0,55	Sangat baik	Digunakan
S18	-0,29	Buruk	Tidak digunakan
S19	0,66	Sangat baik	Digunakan
S20	0,61	Sangat baik	Digunakan
S21	-0,28	Buruk	Tidak digunakan
S22	0,62	Sangat baik	Digunakan
S23	0,56	Sangat baik	Digunakan
S24	0,50	Sangat baik	Digunakan
S25	0,63	Sangat baik	Digunakan
S26	0,66	Sangat baik	Digunakan
S27	-0,07	Buruk	Diperbaiki-digunakan

Berdasarkan Tabel 3.11, terdapat 17 soal yang dikategorikan dengan interpretasi “Sangat baik”, 2 soal dengan interpretasi “Baik”, 1 soal dengan interpretasi “Kurang baik” dan 7 soal dengan interpretasi “Buruk”. Interpretasi “Buruk” menunjukkan bahwa soal tersebut kurang efektif dalam membedakan peserta didik dengan kemampuan tinggi dan rendah, meskipun masih dapat digunakan setelah perbaikan dan terdapat 4 butir soal yang tidak digunakan karena memiliki daya pembeda yang buruk dan tidak fit. Dengan demikian, kesimpulannya adalah bahwa instrumen tes kemampuan kognitif untuk daya pembeda dapat digunakan sebagai alat penelitian. Setelah diuji dari segi validitas,

reliabilitas, tingkat kesukaran, dan daya pembeda, instrumen tes kemampuan kognitif dinyatakan layak untuk digunakan dalam penelitian.

3.4.2 Lembar Angket Respon Siswa

Lembar angket respon siswa digunakan untuk mengetahui respon siswa terhadap penggunaan model pembelajaran *ICARE* (*Introduction, Connection, Application, Reflection, Extension*) berbasis Simulasi PhET. Lembar angket ini digunakan kepada siswa kelompok eksperimen dengan tujuan untuk mengetahui penggunaan model pembelajaran *ICARE* (*Introduction, Connection, Application, Reflection, Extension*) berbasis Simulasi PhET untuk peningkatan kemampuan kognitif siswa pada materi Fluida Dinamis. Angket respon siswa berisi 15 pernyataan. Pernyataan ini terdiri dari 10 pernyataan positif dan 5 pernyataan negatif. Angket respon ini diterapkan kepada siswa diakhir pembelajaran setelah *posttest* melalui *google form*. Angket respon ini memiliki tiga aspek penilaian yaitu sikap siswa terhadap model pembelajaran *ICARE* berbasis Simulasi PhET, kebermanfaatan model pembelajaran *ICARE* berbasis Simulasi PhET, dan keterkaitan antara model pembelajaran *ICARE* berbasis Simulasi PhET terhadap kemampuan kognitif siswa. Adapun kis-kisi angket respon siswa dapat dilihat pada Tabel 3.12 berikut.

Tabel 3. 12 Kisi-Kisi Angket Respon Siswa

Aspek yang Diamati	Indikator	Nomor Pernyataan	Sifat Pernyataan
Sikap Siswa terhadap Model Pembelajaran <i>ICARE</i> berbasis Simulasi PhET	Model pembelajaran <i>ICARE</i> berbasis Simulasi PhET meningkatkan ketertarikan dan minat siswa dalam pembelajaran.	1	Negatif
		13	Positif
		15	Negatif
Kebermanfaatan Model Pembelajaran <i>ICARE</i> berbasis Simulasi PhET	Model pembelajaran <i>ICARE</i> berbasis Simulasi PhET dapat mempermudah siswa dalam memahami materi fisika dan menyelesaikan persoalan fisika.	4	Positif
		5	Negatif
		10	Negatif
		8	Negatif
		9	Positif

Mega Amalia Putri, 2024

PENGUNAAN MODEL PEMBELAJARAN *ICARE* BERBASIS SIMULASI PHET UNTUK MENINGKATKAN KEMAMPUAN KOGNITIF SISWA PADA MATERI FLUIDA DINAMIS

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

Aspek yang Diamati	Indikator	Nomor Pernyataan	Sifat Pernyataan
		12	Positif
Keterkaitan antara Model Pembelajaran ICARE berbasis Simulasi PhET terhadap Kemampuan Kognitif Siswa	Model pembelajaran ICARE berbasis Simulasi PhET dapat meningkatkan kemampuan kognitif siswa.	2	Positif
		3	Positif
		6	Positif
		7	Positif
		11	Positif
		14	Positif

3.5 Analisis Data Penelitian

Setelah didapatkan data pretest dan posttest, tahap selanjutnya adalah melakukan analisis data penelitian. Berikut adalah teknik analisis yang dapat digunakan dalam penelitian ini:

3.5.1 Peningkatan Kemampuan Kognitif

Peningkatan kemampuan kognitif siswa diperoleh dari hasil penelitian berupa skor *pretest* dan *posttest*. Penilaian berdasarkan data hasil *pretest* dan *posttest* kemampuan kognitif dianalisis dengan menggunakan metode *Right Only*. Metode *Right Only* adalah metode yang digunakan dengan cara memberikan skor satu pada jawaban yang benar dan memberikan skor nol pada jawaban yang salah (Faizzah, 2022). Skor hasil *pretest* dan *posttest* dihitung dengan menggunakan Persamaan (3.2) berikut.

$$Skor = \frac{\sum \text{Jawaban benar}}{\sum \text{Jumlah soal}} \times 100 \quad (3.1)$$

Setelah diperoleh skor *pretest* dan *posttest*, analisis yang digunakan untuk mengetahui peningkatan kemampuan kognitif siswa adalah analisis rerata N-Gain. Menurut Hake (1999) Uji N-Gain digunakan untuk mengetahui sejauh mana peningkatan kemampuan kognitif siswa setelah diberi perlakuan (*treatment*). Berikut adalah Persamaan (3.3) yang digunakan untuk mengetahui skor gain ternormalisasi.

$$N - Gain = \frac{\text{skor posttest} - \text{skor pretest}}{\text{skor ideal} - \text{skor pretest}} \quad (3.2)$$

Dimana kategori perolehan menurut Hake (1999) dapat dilihat pada Tabel 3.13 berikut.

Tabel 3. 13 Klasifikasi N-Gain

Nilai N-Gain	Kategori
$(\langle g \rangle) \geq 0,70$	Tinggi
$0,70 > (\langle g \rangle) \geq 0,30$	Sedang
$(\langle g \rangle) < 0,30$	Rendah

(Hake, 1999)

3.5.2 Efektivitas Penggunaan Model Pembelajaran ICARE berbasis Simulasi PhET

Efektivitas penggunaan model pembelajaran ICARE berbasis Simulasi PhET dalam meningkatkan kemampuan kognitif siswa, dianalisis dengan menggunakan uji normalitas, uji homogenitas, uji hipotesis, dan *effect size*. Proses pengolahan dan analisis data dilakukan menggunakan *software* IBM SPSS *Statistics* 26.

3.4.1.1 Uji Normalitas

Uji normalitas digunakan untuk menentukan apakah data yang diperoleh memiliki distribusi yang normal atau tidak (Kadir, 2019). Dalam penelitian ini, uji normalitas dilakukan menggunakan uji Kolmogorov-Smirnov-*Normality test*. Adapun kriteria interpretasi uji Kolmogorov-Smirnov ditunjukkan pada Tabel 3.14 berikut.

Tabel 3. 14 Interpretasi Hasil Uji Normalitas Data

Probabilitas	Kesimpulan
Sig. > 0,05	Terdistribusi normal
Sig. < 0,05	Tidak terdistribusi normal

(Kadir, 2019)

3.4.1.2 Uji Homogenitas

Uji homogenitas digunakan untuk menentukan apakah data yang diambil berasal dari populasi yang homogen atau tidak (Kadir, 2019). Dalam penelitian

ini, uji homogenitas dilakukan menggunakan uji *homogeneity of variances*. Kriteria interpretasi uji homogenitas ditunjukkan pada Tabel 3.15 berikut.

Tabel 3. 15 Interpretasi Hasil Uji Homogenitas Data

Probabilitas	Kesimpulan
Sig. > 0,05	Homogen
Sig. < 0,05	Tidak Homogen

(Kadir, 2019)

3.4.1.3 Uji Hipotesis

Uji Hipotesis yang digunakan dalam penelitian ini adalah uji Mann-Whitney (U). Uji Mann-Whitney (U) adalah salah satu uji non-parametrik yang kuat dan digunakan sebagai alternatif untuk uji-t (Kadir, 2019). Uji ini digunakan karena data tidak terdistribusi secara normal atau tidak homogen. Analisis uji Mann-Whitney dilakukan menggunakan *software IBM SPSS Statistics 26*. Kriteria interpretasi uji hipotesis ditunjukkan pada Tabel 3.16 berikut.

Tabel 3. 16 Interpretasi Hasil Uji Hipotesis Data

Probabilitas	Kesimpulan
Sig. > 0,05	H ₀ diterima H ₁ ditolak
Sig. < 0,05	H ₀ ditolak H ₁ diterima

3.4.1.4 Effect Size

Menurut Cohen (1988) *effect size* digunakan untuk untuk mengetahui besarnya pengaruh setelah diberikan perlakuan. Dalam hal ini, *effect size* digunakan untuk mengetahui seberapa besar efek atau perbedaan antara kelas eksperimen dan kelas kontrol dilakukan. Oleh karena itu, penelitian ini dilakukan dengan menghitung Cohen's d menggunakan rumus *effect size* dari Cohen's d untuk mengukur sejauh mana efektivitas pada model pembelajaran ICARE berbasis Simulasi PhET untuk meningkatkan kemampuan kognitif siswa pada materi Fluida Dinamis. Persamaan 3.4 berikut adalah persamaan yang digunakan untuk menghitung *effect size* menurut Thalheimer, W. dan Cook, S (2002).

$$d = \frac{\bar{x}_t - \bar{x}_c}{S_{pooled}} \quad (3. 3)$$

$$S_{pooled} = \sqrt{\frac{(n_t-1)S_t^2 + (n_c-1)S_c^2}{n_t+n_c}} \quad (3.4)$$

Keterangan:

- d = Cohen's d effect size
 \bar{x}_t = Nilai rata-rata N-Gain kelas eksperimen
 \bar{x}_c = Nilai rata-rata N-Gain kelas kontrol
 S_{pooled} = Standar deviasi gabungan
 n_t = Jumlah siswa kelas eksperimen
 n_c = Jumlah siswa kelas kontrol
 S_t = Standar deviasi kelas eksperimen
 S_c = Standar deviasi kelas kontrol

Jika besarnya *effect size* sudah diketahui, maka selanjutnya dilakukan suatu interpretasi berdasarkan nilai *effect size*. Adapun kriteria interpretasi nilai *effect size* menurut Becker (2000) ditunjukkan pada Tabel 3.17 berikut.

Tabel 3. 17 Interpretasi Nilai Effect Size

Effect Size	Interpretasi
$0,20 \leq d < 0,50$	Rendah
$0,50 \leq d < 0,80$	Sedang
$0,80 \leq d \leq 2,00$	Tinggi

(Becker, 2000)

3.5.3 Angket Respon Siswa

Lembar angket respon siswa dilaksanakan diakhir pembelajaran setelah *posttest* melalui *google form*. Angket respon siswa ini bertujuan untuk mengumpulkan respon atau tanggapan siswa mengenai model pembelajaran yang digunakan dan untuk memperkuat hasil data kemampuan kognitif siswa. Pada penelitian ini, angket respon siswa terdiri dari pernyataan positif dan pernyataan negatif. Selain itu, teknik analisis data yang digunakan berdasarkan skala likert. Berikut adalah perolehan skor jawaban siswa pada pernyataan positif dan negatif yang ditunjukkan pada Tabel 3.18 dan Tabel 3.19 berikut.

Tabel 3. 18 Skor Angket Respon Pernyataan Positif

Jawaban	Skor
Sangat tidak setuju	1
Tidak setuju	2
Setuju	3
Sangat setuju	4

(Gola dkk., 2022)

Tabel 3. 19 Skor Angket Respon Pernyataan Negatif

Jawaban	Skor
Sangat tidak setuju	4
Tidak setuju	3
Setuju	2
Sangat setuju	1

(Gola dkk., 2022)

Persentase respon siswa tiap butir pernyataan dihitung menggunakan Persamaan 3.6 berikut.

$$R_s(\%) = \frac{\text{Jumlah respon siswa}}{\text{Banyak aspek}} \times 100\% \quad (3.5)$$

Persentase skor respon siswa yang diperoleh, kemudian dikonversikan ke pernyataan penilaian. Skala penilaian respon siswa dapat dilihat pada Tabel 3.20 berikut.

Tabel 3. 20 Kriteria Respon Siswa

Interval Skor (%)	Kriteria
$75 \leq P \leq 100$	Sangat baik
$50 \leq P \leq 75$	Baik
$25 \leq P \leq 50$	Cukup Baik
$0 \leq P \leq 25$	Kurang baik

(Gola dkk., 2022)