

BAB III

METODE PENELITIAN

3.1. Metode Penelitian

Penelitian ini menggunakan metode penelitian kuantitatif, yang dapat didefinisikan sebagai metode untuk menguji tujuan penelitian dengan memeriksa hubungan antara variabel yang terlibat dalam penelitian (Creswell & Creswell, 2023). Variabel dalam metode ini dapat diukur melalui berbagai bentuk alat ukur seperti instrumen penelitian, sehingga di dapatkan data dalam bentuk angka yang dapat dianalisis menggunakan prosedur statistik (Creswell & Creswell, 2023). Variabel yang terlibat dalam penelitian ini meliputi variabel independen berupa penggunaan model *argument driven inquiry* dalam proses pembelajaran dan variabel dependen berupa kemampuan kognitif serta keterampilan berpikir kritis peserta didik.

3.2. Desain Penelitian

Desain penelitian yang digunakan dalam penelitian ini adalah *quasi-experiment design* dengan bentuk *pre-test post-test nonequivalent control group design*. Desain penelitian *quasi-experiment design between* yang dilakukan dengan memanipulasi satu atau lebih variabel untuk mengevaluasi bagaimana manipulasi ini mempengaruhi hasil yang diamati (Creswell & Creswell, 2023). Kelas eksperimen dan kontrol yang terlibat dalam penelitian ini tidak ditentukan secara acak (Sugiyono, 2018, hlm. 179), melainkan ditentukan berdasarkan keadaan kolektif subjek penelitian yang berkumpul secara alami (Campbell dkk., 1963, hlm. 47) seperti di sekolah, di mana dua kelas digunakan sebagai kelas eksperimen dan kontrol. Melalui *pre-test post-test nonequivalent control group design*, kedua kelas mengikuti *pre-test* sebelum kelas eksperimen diberi perlakuan dan *post-test* setelah diberi perlakuan pada kelas eksperimen. Kelas eksperimen pada penelitian ini diberi perlakuan dengan menggunakan model *argument driven inquiry*, sedangkan kelas kontrol tanpa menggunakan model *argument driven inquiry*.

Tabel 3. 1 *pre-test post-test nonequivalent control group design*

Kelas	Pre-Test	Perlakuan	Post-Test
Eksperimen	O	X	O'
Kontrol	O	—	O'

Keterangan:

O = *pre-test* pada kelas eksperimen dan kelas kontrol

O' = *post-test* pada kelas eksperimen dan kelas kontrol

X = Pembelajaran fisika menggunakan model *argument driven inquiry*

— = Pembelajaran fisika tanpa menggunakan model *argument driven inquiry*

3.3. Populasi dan Sampel

Populasi dalam penelitian ini merupakan peserta didik SMA Swasta di Kabupaten Bandung. Sampel merupakan bagian yang dipilih dari populasi untuk diamati dalam penelitian. Sampel pada penelitian ini terdiri atas 56 peserta didik dengan 35 peserta didik kelas kontrol dan 37 peserta didik kelas eksperimen. Pengambilan sampel dilakukan menggunakan teknik *convenient sampling*, di mana sampel penelitian dipilih berdasarkan ketersediaan sampel di sekolah.

3.4. Prosedur Penelitian

Prosedur penelitian yang dilakukan pada penelitian ini dibagi kedalam tiga tahapan penelitian yang meliputi tahap persiapan, tahap pelaksanaan, dan tahap penyelesaian, yang dijabarkan sebagai berikut.

3.4.1. Tahap Persiapan

1. Studi literatur dari berbagai sumber, seperti jurnal-jurnal, buku, skripsi, tesis, serta sumber lainnya, dan melakukan studi pendahuluan.
2. Merumuskan masalah yang akan diteliti dengan menentukan pertanyaan penelitian.
3. Menentukan metode dan desain penelitian.
4. Menentukan subjek penelitian dengan memilih tempat, populasi, dan sampel yang akan dijadikan fokus utama penelitian.

Cellina Zsa Zsenca Vitania Amanda, 2025

EFEKTIVITAS MODEL ARGUMENT DRIVEN INQUIRY DALAM MENINGKATKAN KEMAMPUAN KOGNITIF DAN KETERAMPILAN BERPIKIR KRITIS PESERTA DIDIK PADA MATERI KALOR
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

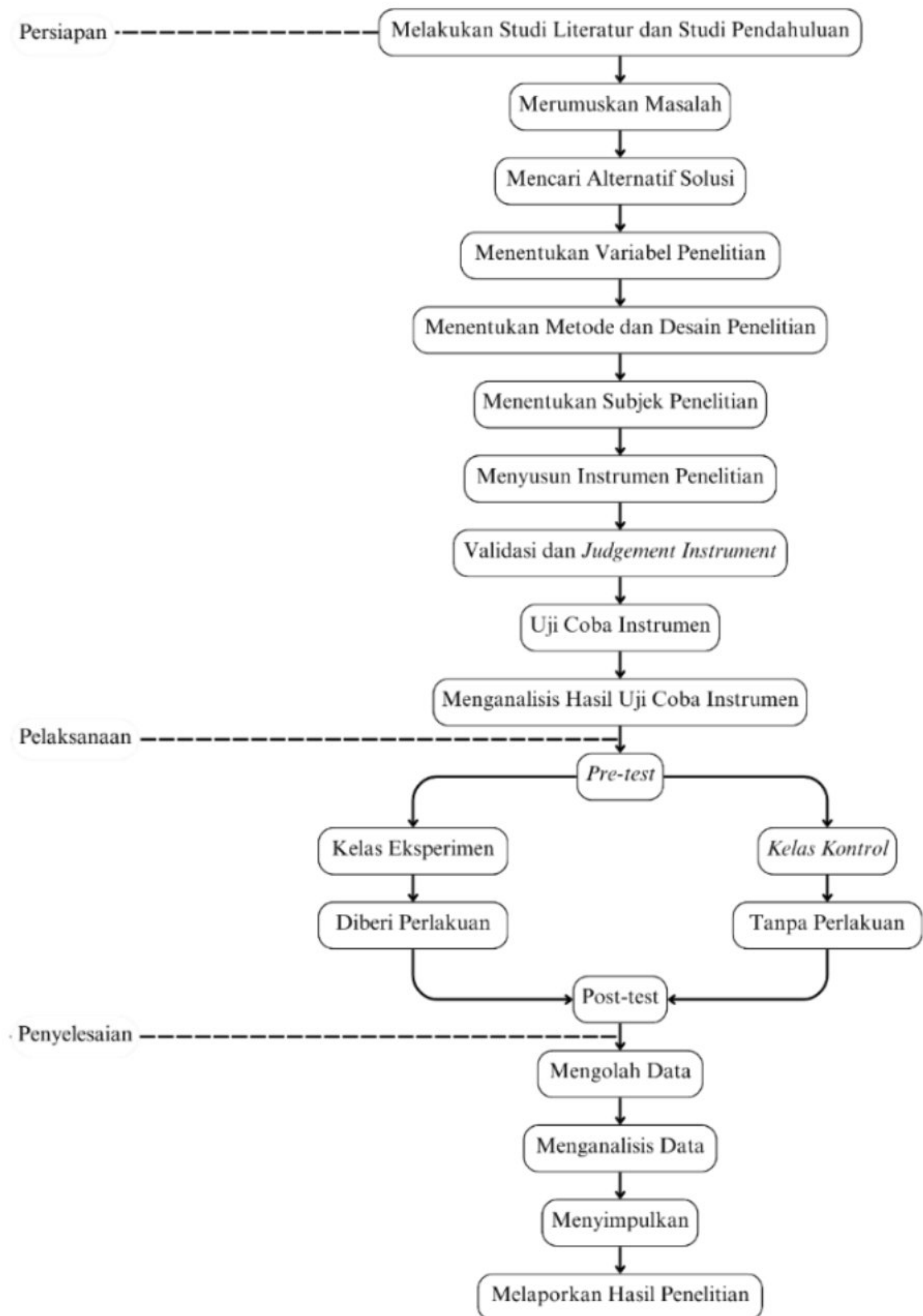
5. Menyusun perangkat pembelajaran dan instrumen penelitian seperti modul ajar, lembar kerja peserta didik, instrumen tes, serta lembar observasi keterlaksanaan pembelajaran.
6. Melakukan validasi dan judgment instrumen tes oleh ahli.
7. Melakukan uji coba instrumen tes untuk mengetahui validitas, reliabilitas, dan tingkat kesukaran instrumen tes penelitian.
8. Melakukan analisis instrumen penelitian, dilakukan untuk menentukan butir soal yang layak untuk dijadikan instrumen penelitian.

3.4.2. Tahap Pelaksanaan

4. Melakukan pre-test berupa tes keterampilan berpikir kritis dan kemampuan kognitif kepada peserta didik di dalam kelas eksperimen dan kontrol.
5. Melaksanakan pembelajaran dengan memberikan perlakuan pada kelas eksperimen menggunakan model *argument driven inquiry* sedangkan pada kelas kontrol tanpa menggunakan model *argument driven inquiry*.
6. Melakukan observasi keterlaksanaan kegiatan pembelajaran di kelas oleh observer.
7. Melakukan post-test berupa tes keterampilan berpikir kritis dan kemampuan kognitif kepada kedua kelas penelitian.

3.4.3. Tahap Penyelesaian

3. Mengelola data hasil penelitian dari skor *pre-test* dan *post-test* peserta didik.
4. Menganalisis data hasil penelitian.
5. Menyimpulkan data hasil penelitian.
6. Menyusun temuan hasil penelitian.



Gambar 3. 1 Skema Penelitian

Cellina Zsa Zsenca Vitania Amanda, 2025

EFEKTIVITAS MODEL ARGUMENT DRIVEN INQUIRY DALAM MENINGKATKAN KEMAMPUAN KOGNITIF DAN KETERAMPILAN BERPIKIR KRITIS PESERTA DIDIK PADA MATERI KALOR
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

3.5. Instrumen Penelitian

3.5.1. Lembar Observasi Keterlaksanaan Pembelajaran

Lembar observasi keterlaksanaan pembelajaran digunakan untuk mengukur keterlaksanaan model *argument driven inquiry* dalam dua pertemuan, seperti yang ditunjukkan pada Tabel 3.2. Lembar observasi yang digunakan dalam penelitian ini berupa kotak *checklist* (✓) dengan pilihan “Ya” ketika aktivitas guru dan peserta didik sesuai dan “Tidak” ketika aktivitas guru dan peserta didik tidak sesuai dengan tahapan pembelajaran yang direncanakan.

Tabel 3. 2 Rincian Pembahasan pada setiap Pertemuan

Pertemuan Ke-	Pembahasan Konsep
1	Kalor dan Perubahan Wujud Zat
2	Pemuaian dan Perpindahan Kalor

3.5.2. Tes Keterampilan Berpikir Kritis

Instrumen tes keterampilan berpikir kritis digunakan untuk mengukur keterampilan berpikir kritis peserta didik sebelum diberi perlakuan (*pre-test*) dan sesudah diberi perlakuan (*post-test*) pada kelas eksperimen dan kontrol. Instrumen tes keterampilan berpikir kritis terdiri dari 8 butir soal dalam bentuk pilihan ganda pada materi kalor. Indikator keterampilan berpikir kritis menurut Ennis (2011) yang digunakan pada instrumen tes penelitian ini terdiri atas *basic clarification*, *bases for a decision*, *inference*, *advanced clarification*, dan *supposition and integration*. Sebaran butir soal pada setiap indikator berpikir kritis ditunjukkan pada Tabel 3.3.

Tabel 3. 3 Sebaran Butir Soal Keterampilan Berpikir Kritis

Indikator Keterampilan Berpikir Kritis	Butir Soal
<i>Basic Clarification</i>	S1, S2, dan S3
<i>Bases for a Decision</i>	S4 dan S5
<i>Inference</i>	S6
<i>Advanced Clarification</i>	S7
<i>Supposition and Integration</i>	S8

3.5.3. Tes Kemampuan Kognitif

Instrumen tes kemampuan kognitif digunakan untuk mengukur kemampuan kognitif peserta didik sebelum diberi perlakuan (*pre-test*) dan sesudah diberi perlakuan (*post-test*) pada kelas eksperimen dan kontrol. Instrumen tes kemampuan kognitif terdiri dari 14 butir soal dalam bentuk pilihan ganda pada materi kalor. Dimensi proses kognitif taksonomi Bloom revisi (Anderson dkk., 2001) yang digunakan pada instrumen tes penelitian ini terdiri atas dimensi mengingat (C1), memahami (C2), menerapkan (C3), dan menganalisis (C4). Sebaran butir soal pada setiap dimensi proses kemampuan kognitif ditunjukkan pada Tabel 3.4.

Tabel 3. 4 Sebaran Butir Soal Kemampuan Kognitif

Dimensi Proses Kognitif	Butir Soal
Mengingat (C1)	S1 dan S6
Memahami (C2)	S2, S7, S8, S12, S13, dan S16
Menerapkan (C3)	S3, S9, S14, dan S18
Menganalisis (C4)	S4 dan S10

3.6. Analisis Instrumen Penelitian

3.6.1. Validitas Isi

Uji validitas isi merupakan bentuk validitas yang dilakukan untuk mengevaluasi apakah item dalam instrumen penelitian mewakili keterukuran perilaku atau kualitas yang hendak diukur (Taherdoost, 2018). Evaluasi pada proses validasi ini dilakukan dengan melibatkan para ahli untuk mengevaluasi setiap bentuk ekspresi dalam instrumen penelitian (Taherdoost, 2018). Para ahli yang dilibatkan dalam proses validasi ini meliputi lima dosen ahli program studi Pendidikan Fisika Universitas Pendidikan Indonesia. Uji validasi isi dilakukan terhadap setiap butir soal yang mengacu pada tiga pernyataan aspek materi, tiga pernyataan aspek konstruksi, dan tiga pernyataan aspek bahasa.

Setiap ahli diberikan instrumen penilaian dan lembar validasi yang dapat diisi dengan memberikan penilaian berupa *rating scale* mulai dari 1 yang menunjukkan bahwa soal tidak sesuai, hingga 3 yang menandakan bahwa soal

tersebut sesuai. Hasil uji validitas isi dilakukan menggunakan analisis indeks Aiken (koefisien validitas Aiken) yang diperoleh melalui Persamaan 3.1.

$$V = \frac{S}{[n(c-1)]} = \frac{\sum(r - I_0)}{[n(c-1)]} \quad 3.1$$

Keterangan:

V = koefisien validitas aiken

r = angka yang diberikan ahli/penilai

I_0 = angka penilaian validasi terendah

c = angka penilaian validasi tertinggi

n = jumlah ahli/penilai

3.6.1.1. Validitas Isi Instrumen tes Kemampuan Kognitif

Pada validasi isi penelitian ini, butir soal dinyatakan valid ketika koefisien Aiken atau $V \geq 0,90$ untuk butir soal yang dinilai oleh 5 orang ahli, dan $V \geq 1,00$ untuk butir soal yang dinilai oleh 4 orang ahli. Perbedaan koefisien Aiken tabel yang digunakan pada penelitian ini disebabkan karena butir soal S6, S15, dan S16 tidak diberikan penilaian oleh seorang ahli karena narasi pada indikator soal yang menyerupai tujuan pembelajaran.

Tabel 3. 5 Hasil Analisis Validitas Isi Instrumen Tes Kemampuan Kognitif

Butir Soal	Aspek Penilaian			Rata-Rata	Keterangan
	Isi	Konstruksi	Bahasa		
S1	0,90	1,00	0,97	0,96	Valid/Digunakan
S2	0,93	0,97	0,97	0,96	Valid/Digunakan
S3	0,93	0,83	0,97	0,91	Valid/Digunakan
S4	0,90	0,93	0,97	0,93	Valid/Digunakan
S5	1,00	1,00	1,00	1,00	Valid/Digunakan
S6	1,00	1,00	1,00	1,00	Valid/Digunakan
S7	0,97	1,00	0,97	0,98	Valid/Digunakan
S8	0,97	1,00	0,97	0,98	Valid/Digunakan
S9	0,87	1,00	1,00	0,96	Valid/Digunakan
S10	1,00	1,00	1,00	1,00	Valid/Digunakan
S11	0,97	1,00	0,97	0,98	Valid/Digunakan
S12	0,97	0,97	0,97	0,97	Valid/Digunakan

Cellina Zsa Zsenca Vitania Amanda, 2025

EFEKTIVITAS MODEL ARGUMENT DRIVEN INQUIRY DALAM MENINGKATKAN KEMAMPUAN KOGNITIF DAN KETERAMPILAN BERPIKIR KRITIS PESERTA DIDIK PADA MATERI KALOR
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

S13	0,97	1,00	0,93	0,97	Valid/Digunakan
S14	0,93	0,97	1,00	0,97	Valid/Digunakan
S15	0,88	1,00	1,00	0,96	Tidak Valid/ Tidak Digunakan
S16	1,00	1,00	1,00	1,00	Valid/Digunakan
S17	0,93	0,97	1,00	0,97	Valid/Digunakan
S18	1,00	0,97	0,93	0,97	Valid/Digunakan

Berdasarkan hasil analisis data validitas isi pada Tabel 3.5 menunjukkan bahwa terdapat 17 soal dengan kategori valid dan 1 soal dengan kategori tidak valid. Satu butir soal dengan kategori tidak valid yaitu S15 tidak digunakan dalam uji coba instrumen, sehingga uji coba instrumen tes kemampuan kognitif yang digunakan berjumlah 17 soal.

3.6.1.2. Validitas Isi Instrumen Tes Keterampilan Berpikir Kritis

Pada validasi isi penelitian ini, butir soal dinyatakan valid ketika koefisien Aiken atau $V \geq 0,90$ untuk butir soal yang dinilai oleh 5 orang ahli.

Tabel 3. 6 Hasil Analisis Validitas Isi Instrumen Tes Keterampilan Berpikir Kritis

Butir Soal	Aspek Penilaian			Rata- Rata	Keterangan
	Isi	Konstruksi	Bahasa		
S1	1,00	0,93	0,90	0,94	Valid/Digunakan
S2	0,87	1,00	0,93	0,93	Valid/Digunakan
S3	1,00	1,00	0,93	0,98	Valid/Digunakan
S4	0,90	0,90	0,97	0,92	Valid/Digunakan
S5	0,97	0,93	0,97	0,96	Valid/Digunakan
S6	1,00	1,00	0,90	0,97	Valid/Digunakan
S7	1,00	1,00	0,90	0,97	Valid/Digunakan
S8	0,87	1,00	1,00	0,96	Valid/Digunakan

Berdasarkan hasil analisis data validitas isi pada Tabel 3.6 menunjukkan terdapat 8 soal dengan kategori valid, sehingga uji coba instrumen tes keterampilan berpikir kritis yang digunakan berjumlah 8 soal.

3.6.2. Validitas Konstruk

Validitas konstruk merupakan jenis validasi yang dilakukan untuk mengetahui apakah instrumen secara akurat mengukur konstruksi hipotesis atau konsep secara tepat (Creswell & Creswell, 2023). Uji validitas konstruk dilakukan pada instrumen tes yang terdiri dari 17 soal dengan melakukan uji coba kepada 68 peserta didik kelas XI yang telah mempelajari kalor. Uji validitas konstruk dalam penelitian ini dilakukan melalui analisis *Rasch Model* menggunakan MINISTEP 5.10.0.0 untuk mengetahui *outfit* MNSQ-ZSTD butir soal, korelasi butir soal, dan *unidimensionality* instrumen tes (Planinic dkk., 2019). Uji *unidimensionality* dilakukan untuk mengetahui bahwa setiap butir soal pada instrumen tes yang digunakan hanya mengukur satu dimensi dari fenomena yang sedang diamati. Uji *unidimensionality* didapatkan berdasarkan keluaran pada tabel 23. *Item: dimensionality*. Hasil uji *unidimensionality* kemudian dikategorikan berdasarkan kriteria yang terpenuhi pada Tabel 3.7.

Tabel 3. 7 Interpretasi *Unidimensionality*

<i>Essential Unidimensionality</i> (%)	Kategori
20,1 – 39,9	Cukup
40 – 59,9	Baik
60 – 100	Sangat Baik

(Sumintono & Widhiarso dalam Susanti dkk., 2023)

Outfit MNSQ-ZSTD dan korelasi item didapatkan berdasarkan hasil keluaran tabel 10. *Item (column): fit order*. Menurut Sumintono & Widhiarso (dalam Susanti dkk., 2023) suatu item dikatakan valid ketika memenuhi kriteria sebagai berikut:

1. Nilai *Outfit Mean Square* (MNSQ) yang diterima yaitu $0,5 \leq MNSQ < 1,50$;
2. Nilai *Outfit Z-Standard* (ZSTD) yang diterima yaitu $-2,00 \leq ZSTD < 2,00$;
3. Nilai *Point Measure Correlation* (PTMEASURE CORR.) yang diterima yaitu $0,4 < PTMEASURE CORR. < 0,85$.

Menurut Linares (2002), nilai *outfit mean square* pada rentang $0,5 \leq MNSQ < 1,50$, menunjukkan bahwa butir soal termasuk kedalam alat ukur yang memberikan hasil akurat dan relevan sesuai dengan tujuan pengukuran. Sedangkan nilai *outfit Z-standart* pada rentang $-2,00 \leq ZSTD < 2,00$, menunjukkan bahwa data hasil pengukuran butir soal memiliki tingkat prediktibilitas yang bervariasi, mulai dari sangat tidak terduga hingga sangat mudah ditebak.

3.6.2.1. Validitas Konstruk Instrumen Tes Kemampuan Kognitif

Hasil *unidimensionality* instrumen tes kemampuan kognitif pada penelitian ini ditunjukkan pada Gambar 3.2.

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

	Eigenvalue	Observed	Expected
Total raw variance in observations =	24.9080	100.0%	100.0%
Raw variance explained by measures =	7.9080	31.7%	31.1%
Raw variance explained by persons =	3.6941	14.8%	14.5%
Raw Variance explained by items =	4.2139	16.9%	16.6%
Raw unexplained variance (total) =	17.0000	68.3%	68.9%
Unexplnd variance in 1st contrast =	2.8873	11.6%	17.0%
Unexplnd variance in 2nd contrast =	2.4775	9.9%	14.6%
Unexplnd variance in 3rd contrast =	2.3034	9.2%	13.5%
Unexplnd variance in 4th contrast =	1.6871	6.8%	9.9%
Unexplnd variance in 5th contrast =	1.2851	5.2%	7.6%

Essential Unidimensionality (Rasch/Common variance) = 45.8%

Gambar 3. 2 Keluaran Tabel *Item: Dimensionality* Instrumen Tes Kemampuan Kognitif

Berdasarkan Gambar 3.2, nilai *Essential Unidimensionality* yang diperoleh dari uji coba instrumen tes penelitian ini adalah sebesar 45,8% yang termasuk pada kategori “Baik”. Berdasarkan hasil tersebut, dapat diartikan bahwa instrumen tes pada penelitian ini dapat dengan baik mengukur satu dimensi dari fenomena yang sedang diamati, yaitu kemampuan kognitif peserta didik.

Hasil *outfit* MNSQ-ZSTD dan korelasi butir sial instrumen tes kemampuan kognitif pada penelitian ini ditunjukkan pada Gambar 3.3.

Item STATISTICS: MISFIT ORDER

ENTRY	TOTAL	TOTAL	JMLE	MODEL	INFIT		OUTFIT		PTMEASUR-AL		EXACT MATCH		
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	Item
17	24	68	1.67	.28	1.58	4.65	2.57	5.13	A-.19	.39	50.0	70.2	17
11	24	68	1.67	.28	1.39	3.27	2.26	4.36	B-.02	.39	55.9	70.2	11
5	58	68	-1.23	.38	.96	-.08	2.11	2.14	C .29	.40	91.2	87.1	5
4	60	68	-1.54	.41	1.30	1.06	.95	.07	D .23	.38	82.4	89.2	4
16	52	68	-.51	.32	.96	-.17	1.23	.85	E .42	.43	85.3	80.3	16
6	48	68	-.14	.30	1.11	.76	1.01	.10	F .37	.43	67.6	75.7	6
8	56	68	-.96	.35	.99	.03	.83	-.37	G .44	.41	82.4	84.8	8
10	34	68	.94	.27	.95	-.51	.87	-.76	H .47	.42	67.6	67.6	10
13	30	68	1.23	.27	.88	-1.33	.79	-1.26	I .52	.41	64.7	67.3	13
14	50	68	-.32	.31	.87	-.79	.72	-1.15	h .55	.43	82.4	78.0	14
15	28	68	1.37	.27	.87	-1.40	.78	-1.22	g .52	.40	67.6	67.8	15
7	56	68	-.96	.35	.86	-.58	.76	-.59	f .52	.41	88.2	84.8	7
3	50	68	-.32	.31	.85	-.93	.68	-1.36	e .57	.43	76.5	78.0	3
1	58	68	-1.23	.38	.83	-.65	.67	-.74	d .53	.40	91.2	87.1	1
2	52	68	-.51	.32	.82	-1.00	.65	-1.36	c .59	.43	79.4	80.3	2
9	28	68	1.37	.27	.78	-2.50	.68	-1.86	b .60	.40	73.5	67.8	9
12	52	68	-.51	.32	.77	-1.32	.70	-1.09	a .60	.43	85.3	80.3	12
MEAN	44.7	68.0	.00	.32	.99	-.09	1.07	.05			76.0	77.4	
P.SD	12.9	.0	1.08	.04	.22	1.70	.60	1.96			11.7	7.5	

Gambar 3. 3 Keluaran Tabel *Item (column): Fit Order* Instrumen Tes Kemampuan Kognitif

Berdasarkan Gambar 3.3, analisis lebih lanjut dilakukan untuk mengetahui validitas setiap butir soal. Hasil analisis pada setiap butir soal disajikan pada Tabel 3.8.

Tabel 3. 8 Hasil Analisis Validitas Konstruk Instrumen Tes Kemampuan Kognitif

No	Butir Soal	Outfit		Pt. Measure Corr.	Keterangan	
		MNSQ	ZSTD			
1.	S1	0,67	-0,74	0,53	Ketiga kriteria terpenuhi	Digunakan
2.	S2	0,65	-1,36	0,59	Ketiga kriteria terpenuhi	Digunakan
3.	S3	0,68	-1,36	0,57	Ketiga kriteria terpenuhi	Digunakan

4.	S4	0,95	0,07	0,23	Dua dari tiga kriteria terpenuhi	Digunakan
5.	S5	2,11	2,14	0,29	Ketiga kriteria tidak terpenuhi	Tidak Digunakan
6.	S6	1,01	0,1	0,37	Dua dari tiga kriteria terpenuhi	Digunakan
7.	S7	0,76	-0,59	0,52	Ketiga kriteria terpenuhi	Digunakan
8.	S8	0,83	-0,37	0,44	Ketiga kriteria terpenuhi	Digunakan
9.	S9	0,68	-1,86	0,6	Ketiga kriteria terpenuhi	Digunakan
10.	S10	0,87	-0,76	0,47	Ketiga kriteria terpenuhi	Digunakan
11.	S11	2,26	4,36	-0,02	Ketiga kriteria tidak terpenuhi	Tidak Digunakan
12.	S12	0,7	-1,09	0,6	Ketiga kriteria terpenuhi	Digunakan
13.	S13	0,79	-1,26	0,52	Ketiga kriteria terpenuhi	Digunakan
14.	S14	0,72	-1,15	0,55	Ketiga kriteria terpenuhi	Digunakan
15.	S16	0,78	-1,22	0,52	Ketiga kriteria terpenuhi	Digunakan
16.	S17	1,23	0,85	0,42	Ketiga kriteria terpenuhi	Digunakan
17.	S18	2,57	5,13	-0,19	Ketiga kriteria tidak terpenuhi	Tidak Digunakan

Berdasarkan Tabel 3.8, menunjukkan bahwa terdapat 12 soal memenuhi ketiga kriteria, 2 soal memenuhi dua kriteria, dan 3 soal tidak memenuhi ketiga kriteria kriteria item dikatakan valid menurut Sumintono & Widhiarso (dalam Susanti dkk., 2023). Tiga butir soal yang tidak memenuhi ketiga kriteria yaitu nilai $0,5 \leq MNSQ < 1,50$; $-2,00 \leq ZSTD < 2,00$; dan $0,4 < PTMEASURE CORR. < 0,85$ pada butir soal S5, S11, dan S17 tidak digunakan dalam instrumen penelitian, sehingga instrumen tes kemampuan kognitif yang digunakan pada penelitian ini berjumlah 14 soal.

3.6.2.2. Validitas Konstruk Instrumen Tes Keterampilan Berpikir Kritis

Hasil *unidimensionality* instrumen tes keterampilan berpikir kritis pada penelitian ini ditunjukkan pada Gambar 3.4

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

	Eigenvalue	Observed	Expected
Total raw variance in observations =	11.8347	100.0%	100.0%
Raw variance explained by measures =	3.8347	32.4%	32.7%
Raw variance explained by persons =	2.7455	23.2%	23.4%
Raw Variance explained by items =	1.0892	9.2%	9.3%
Raw unexplained variance (total) =	8.0000	67.6%	67.3%
Unexplnd variance in 1st contrast =	1.5811	13.4%	19.8%
Unexplnd variance in 2nd contrast =	1.4991	12.7%	18.7%
Unexplnd variance in 3rd contrast =	1.3415	11.3%	16.8%
Unexplnd variance in 4th contrast =	1.1895	10.1%	14.9%
Unexplnd variance in 5th contrast =	1.1123	9.4%	13.9%

Essential Unidimensionality (Rasch/Common variance) = 46.4%

Gambar 3. 4 Keluaran Tabel *Item: Dimensionality* Instrumen Tes Keterampilan Berpikir Kritis

Berdasarkan hasil keluaran tabel 23. *Item: dimensionality*, nilai *Essential Unidimensionality* yang diperoleh dari uji coba instrumen tes keterampilan berpikir kritis adalah sebesar 46,4% yang termasuk pada kategori “Baik”. Berdasarkan hasil tersebut, dapat diartikan bahwa instrumen tes keterampilan berpikir kritis pada penelitian ini dapat dengan baik mengukur satu dimensi dari fenomena yang sedang diamati, yaitu keterampilan berpikir kritis peserta didik.

Hasil *outfit* MNSQ-ZSTD dan korelasi butir soal instrumen tes keterampilan berpikir kritis ditunjukkan pada Gambar 3.5.

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-AL CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
1	39	68	.21	.29	1.22	1.49	1.42	2.08	A .36	.53	70.6	74.7	1
2	40	68	.12	.29	1.28	1.87	1.38	1.84	B .34	.53	69.1	74.6	2
3	42	68	-.05	.29	1.35	2.30	1.25	1.19	C .31	.52	66.2	75.0	3
4	40	68	.12	.29	1.16	1.10	1.26	1.34	D .41	.53	69.1	74.6	4
5	62	68	-2.33	.45	.95	-.08	.49	-.66	d .36	.29	91.2	91.1	5
8	26	68	1.28	.29	.78	-1.76	.65	-1.85	c .66	.51	79.4	73.4	8
6	34	68	.62	.29	.69	-2.45	.61	-2.62	b .74	.53	83.8	74.4	6
7	41	68	.04	.29	.59	-3.41	.49	-3.17	a .79	.52	88.2	74.7	7
MEAN	40.5	68.0	.00	.31	1.00	-.12	.94	-.23			77.2	76.6	
P.SD	9.5	.0	.97	.05	.27	2.03	.39	1.98			9.1	5.5	

Gambar 3. 5 Keluaran Tabel *Item (column): Fit Order* Instrumen Tes Keterampilan Berpikir Kritis

Berdasarkan hasil keluaran tabel 10. *Item (column): fit order*, analisis lebih lanjut dilakukan untuk mengetahui validitas setiap butir soal. Hasil analisis setiap butir soal disajikan pada Tabel 3.9.

Tabel 3. 9 Hasil Analisis Validitas Konstruk Instrumen Tes Keterampilan Berpikir Kritis

Butir Soal	Outfit		Pt. Measure Corr.	Keterangan	
	MNSQ	ZSTD			
S1	1,42	2,08	0,36	Satu dari tiga kriteria terpenuhi	Digunakan
S2	1,38	1,84	0,34	Dua dari tiga kriteria terpenuhi	Digunakan
S3	1,25	1,19	0,31	Dua dari tiga kriteria terpenuhi	Digunakan
S4	1,26	1,34	0,41	Ketiga kriteria terpenuhi	Digunakan
S5	0,49	-0,66	0,36	Satu dari tiga kriteria terpenuhi	Digunakan
S6	0,61	-2,62	0,74	Dua dari tiga kriteria terpenuhi	Digunakan
S7	0,49	-3,17	0,79	Satu dari tiga kriteria terpenuhi	Digunakan
S8	0,65	-1,85	0,66	Ketiga kriteria terpenuhi	Digunakan

Berdasarkan Tabel 3.9, menunjukkan bahwa terdapat 2 soal memenuhi ketiga kriteria, 3 soal memenuhi dua kriteria, dan 3 soal memenuhi satu kriteria item dikatakan valid menurut Sumintono & Widhiarso (dalam Susanti dkk., 2023), sehingga instrumen tes keterampilan berpikir kritis yang digunakan pada penelitian ini berjumlah 8 soal dengan perbaikan pada soal S1, S5, dan S7, sebelum digunakan.

3.6.3. Uji Reliabilitas

Reliabilitas dalam penelitian kuantitatif dapat diartikan sebagai konsistensi dari waktu ke waktu atas instrumen dan hasil pengukuran (Cohen dkk., 2007). Suatu instrumen dikatakan reliabel, ketika instrumen dapat menghasilkan hasil yang presisi dan akurat ketika dilakukan pengukuran pada subjek dan konteks serupa (Cohen dkk., 2007). Uji reliabilitas dilakukan pada instrumen tes yang terdiri dari 17 soal dengan melakukan uji coba kepada 68 peserta didik kelas XI yang telah mempelajari kalor. Uji reliabilitas penelitian ini dilakukan melalui analisis *Rasch Model* menggunakan MINISTEP 5.10.0.0 *person reliability* (p), *item reliability* (r), dan *crobbach alpha* ($KR - 20$). Uji reliabilitas didapatkan berdasarkan keluaran pada tabel 3.1 *Summary statistic*. Hasil uji reliabilitas kemudian dikategorikan berdasarkan kriteria yang terpenuhi pada Tabel 3.10.

Tabel 3. 10 Interpretasi Uji Reliabilitas

Uji Reliabilitas	Hasil Uji Reliabilitas	Kategori
Pearson & Item reliability	$< 0,67$	Rendah
	$0,67 - 0,80$	Cukup
	$0,81 - 0,90$	Baik
	$0,91 - 0,94$	Sangat Baik
	$> 0,94$	Luas Biasa
Cronbach alpha ($KR - 20$)	$< 0,50$	Rendah
	$0,50 - 0,60$	Cukup
	$0,61 - 0,70$	Baik
	$0,71 - 0,80$	Sangat Baik
	$> 0,80$	Luas Biasa

(Sumintono & Widhiarso dalam Susanti dkk., 2023)

3.6.3.1. Uji Reliabilitas Instrumen Tes Kemampuan Kognitif

Hasil *person reliability* (p), *item reliability* (r), dan *crobbach alpha* ($KR - 20$) instrumen tes kemampuan kognitif pada penelitian ini ditunjukkan pada Gambar 3.6.

SUMMARY OF 68 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	11.2	17.0	.90	.63	.98	.00	1.07	.12
SEM	.4	.0	.13	.01	.03	.11	.09	.13
P.SD	3.0	.0	1.09	.10	.24	.86	.76	1.04
S.SD	3.1	.0	1.10	.10	.24	.87	.77	1.05
MAX.	16.0	17.0	3.28	1.06	1.45	1.68	4.37	2.76
MIN.	3.0	17.0	-1.88	.55	.57	-1.50	.35	-1.46
REAL RMSE	.66	TRUE SD	.87	SEPARATION	1.31	Person	RELIABILITY	.63
MODEL RMSE	.64	TRUE SD	.89	SEPARATION	1.39	Person	RELIABILITY	.66
S.E. OF Person MEAN = .13								
Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .69 SEM = 1.69								
STANDARDIZED (50 ITEM) RELIABILITY = .85								
SUMMARY OF 17 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	44.7	68.0	.00	.32	.99	-.09	1.07	.05
SEM	3.2	.0	.27	.01	.06	.43	.15	.49
P.SD	12.9	.0	1.08	.04	.22	1.70	.60	1.96
S.SD	13.3	.0	1.12	.04	.23	1.75	.61	2.02
MAX.	60.0	68.0	1.67	.41	1.58	4.65	2.57	5.13
MIN.	24.0	68.0	-1.54	.27	.77	-2.50	.65	-1.86
REAL RMSE	.33	TRUE SD	1.03	SEPARATION	3.12	Item	RELIABILITY	.91
MODEL RMSE	.32	TRUE SD	1.04	SEPARATION	3.25	Item	RELIABILITY	.91
S.E. OF Item MEAN = .27								
Item RAW SCORE-TO-MEASURE CORRELATION = -1.00								
Global statistics: please see Table 44.								
UMEAN=.0000 USCALE=1.0000								

Gambar 3. 6 Keluaran Tabel *Summary Statistic* Instrumen Tes Kemampuan Kognitif

Berdasarkan Gambar 3.6, nilai *person reliability* (p) yang diperoleh dari uji coba instrumen tes kemampuan kognitif adalah sebesar 0,63 yang termasuk pada kategori “Baik”; *item reliability* (r) sebesar 0,91 dengan kategori “Baik”; dan *crobbach alpha* ($KR - 20$) sebesar 0,69 dengan kategori “Cukup”. Berdasarkan hasil tersebut, dapat diartikan bahwa instrumen tes kemampuan kognitif pada penelitian ini memiliki konsistensi yang baik dalam menghasilkan

hasil yang akurat dan presisi ketika dilakukan pengukuran pada subjek dan konteks serupa.

3.6.3.2. Uji Reliabilitas Instrumen Tes Keterampilan Berpikir Kritis

Hasil *person reliability* (p), *item reliability* (r), dan *crobbach alpha* ($KR - 20$) instrumen tes keterampilan berpikir kritis pada penelitian ini ditunjukkan pada Gambar 3.7.

SUMMARY OF 68 MEASURED Person									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	4.8	8.0	.60	.90	1.01	.10	.94	.13	
SEM	.2	.0	.16	.02	.03	.08	.04	.07	
P.SD	1.9	.0	1.27	.13	.26	.67	.31	.57	
S.SD	1.9	.0	1.28	.13	.26	.67	.31	.58	
MAX.	7.0	8.0	2.20	1.09	1.67	1.78	1.89	1.78	
MIN.	2.0	8.0	-1.24	.76	.70	-1.03	.42	-.63	
REAL RMSE	.96	TRUE SD	.84	SEPARATION	.87	Person	RELIABILITY	.43	
MODEL RMSE	.91	TRUE SD	.89	SEPARATION	.98	Person	RELIABILITY	.49	
S.E. OF Person MEAN = .16									
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00									
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .57 SEM = 1.23									
STANDARDIZED (50 ITEM) RELIABILITY = .86									
SUMMARY OF 8 MEASURED Item									
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		
					MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	40.5	68.0	.00	.31	1.00	-.12	.94	-.23	
SEM	3.6	.0	.37	.02	.10	.77	.15	.75	
P.SD	9.5	.0	.97	.05	.27	2.03	.39	1.98	
S.SD	10.1	.0	1.04	.06	.29	2.17	.42	2.11	
MAX.	62.0	68.0	1.28	.45	1.35	2.30	1.42	2.08	
MIN.	26.0	68.0	-2.33	.29	.59	-3.41	.49	-3.17	
REAL RMSE	.33	TRUE SD	.91	SEPARATION	2.76	Item	RELIABILITY	.88	
MODEL RMSE	.31	TRUE SD	.92	SEPARATION	2.92	Item	RELIABILITY	.90	
S.E. OF Item MEAN = .37									
Item RAW SCORE-TO-MEASURE CORRELATION = -.99									
Global statistics: please see Table 44.									
UMEAN=.0000 USCALE=1.0000									

Gambar 3. 7 Keluaran Tabel *Summary Statistic* Instrumen Tes Keterampilan Berpikir Kritis

Berdasarkan Gambar 3.7, nilai *person reliability* (p) yang diperoleh melalui uji coba instrumen tes keterampilan berpikir kritis adalah sebesar 0,43 yang termasuk pada kategori “Rendah”; *item reliability* (r) sebesar 0,88 dengan kategori “Sangat Baik”; dan *crobbach alpha* ($KR - 20$) sebesar 0,57 dengan

kategori “Baik”. Berdasarkan hasil tersebut, dapat diartikan bahwa instrumen tes kemampuan kognitif pada penelitian ini memiliki konsistensi yang baik dalam menghasilkan hasil yang akurat dan presisi ketika dilakukan pengukuran pada subjek dan konteks serupa.

3.6.4. Tingkat Kesukaran

Tingkat kesukaran merupakan suatu nilai yang menunjukkan kesukaran butir soal, di mana soal yang baik bukanlah soal yang mudah atau sukar tetapi berada diantaranya (Arikunto, 2015, hlm. 222). Pada penelitian ini tingkat kesukaran butir soal dianalisis melalui analisis *Rasch Model* menggunakan MINISTEP 5.10.0.0, di mana tingkat kesukaran item dapat diamati berdasarkan hasil JMLE MEASURE *item* yang kemudian dikategorikan kedalam kategori pada tabel 3.11.

Tabel 3. 11 Interpretasi Tingkat Kesukaran Butir Soal

<i>Measure Logit Butir Soal</i>	Kategori
$Measure\ Logit > SD$	Sangat Sulit
$0,00 < Measure\ Logit \leq SD$	Sulit
$-SD \leq Measure\ Logit \leq 0,00$	Mudah
$Measure\ Logit < -SD$	Sangat Mudah

3.6.4.1. Tingkat Kesukaran Instrumen Tes Kemampuan Kognitif

Nilai *JMLE Measure* butir soal instrumen tes kemampuan kognitif pada penelitian ini ditunjukkan pada Gambar 3.3. Berdasarkan Gambar 3.3, analisis lebih lanjut pada *JMLE Measure* setiap butir soal dilakukan untuk mengetahui tingkat kesukaran butir soal. Hasil analisis setiap butir soal disajikan pada Tabel 3.12.

Tabel 3. 12 Hasil Analisis Tingkat Kesukaran Instrumen Tes Kemampuan Kognitif

Butir Soal	<i>JMLE Measure</i>	<i>Standard Deviation</i>	Keterangan
S1	-1,23	1,08	Sangat Mudah
S2	-0,51	1,08	Mudah
S3	-0,32	1,08	Mudah
S4	-1,54	1,08	Sangat Mudah
S6	0,14	1,08	Sulit

Butir Soal	<i>JMLE Measure</i>	<i>Standard Deviation</i>	Keterangan
S7	-0,96	1,08	Mudah
S8	-0,96	1,08	Mudah
S9	1,37	1,08	Sangat Sulit
S10	0,94	1,08	Sulit
S12	-0,51	1,08	Mudah
S13	1,23	1,08	Sangat Sulit
S14	-0,32	1,08	Mudah
S16	1,37	1,08	Sangat Sulit
S18	-0,51	1,08	Mudah

Berdasarkan Tabel 3.12, menunjukkan bahwa terdapat 2 soal dengan kategori mudah, 9 soal dengan kategori sedang, dan 3 soal dengan kategori sulit.

3.6.4.2. Tingkat Kesukaran Instrumen Tes Keterampilan Berpikir Kritis

Nilai *JMLE Measure* butir soal instrumen tes kemampuan kognitif pada penelitian ini ditunjukkan pada Gambar 3.5. Berdasarkan Gambar 3.5, analisis lebih lanjut pada *JMLE Measure* setiap butir soal dilakukan untuk mengetahui tingkat kesukaran butir soal. Hasil analisis setiap butir soal disajikan pada Tabel 3.13.

Tabel 3. 13 Hasil Analisis Tingkat Kesukaran Instrumen Tes Keterampilan Berpikir Kritis

Butir Soal	<i>JMLE Measure</i>	<i>Standard Deviation</i>	Keterangan
S1	0,21	0,97	Sulit
S2	0,12	0,97	Sulit
S3	-0,05	0,97	Mudah
S4	0,12	0,97	Sulit
S5	-2,33	0,97	Sangat Mudah
S6	0,62	0,97	Sulit
S7	0,04	0,97	Sulit
S8	1,28	0,97	Sangat Sulit

Berdasarkan Tabel 3.13, menunjukkan bahwa terdapat 1 soal dengan kategori mudah, 6 soal dengan kategori sedang, dan 1 soal dengan kategori sulit.

3.7. Teknik Analisis Data

Pada penelitian ini, pengumpulan data dilakukan dengan memberikan *pre-test* dan *post-test* berbentuk pilihan ganda untuk mengukur kemampuan kognitif dan keterampilan berpikir kritis peserta didik serta melalui lembar keterlaksanaan model argument driven inquiry selama proses pembelajaran oleh observer. Data penelitian yang diperoleh kemudian dianalisis dengan uji statistik untuk data yang bersifat kuantitatif yang dijabarkan sebagai berikut.

3.7.1. Keterlaksanaan Model *Argument Driven Inquiry*

Lembar observasi digunakan sebagai alat ukur untuk mengetahui keterlaksanaan model *argument driven inquiry* pada proses pembelajaran. Hasil penilaian keterlaksanaan pembelajaran oleh observer kemudian diakumulasikan dalam persentasi seperti pada persamaan berikut.

$$x = \frac{\text{Jumlah kegiatan yang terlaksana}}{\text{Jumlah total kegiatan}} \times 100\% \quad 3.2$$

Interpretasi keterlaksanaan model *argument driven inquiry* dijelaskan pada tabel 3.14.

Tabel 3. 14 Interpretasi keterlaksanaan model argument driven inquiry

Keterlaksanaan(%)	Kriteria
$90 < x \leq 100$	Sangat baik
$75 < x \leq 91$	Baik
$50 < x \leq 75$	Sedang
$25 < x \leq 50$	Kurang
$0 < x \leq 25$	Sangat Kurang

(Ramadhana & Hadi, 2021)

3.7.2. Peningkatan Keterampilan Berpikir Kritis dan Kemampuan Kognitif

Peningkatan keterampilan berpikir kritis dan kemampuan kognitif peserta didik diperoleh berdasarkan skor *pretest* dan *posttest* yang kemudian dianalisis

menggunakan *Normalized Gain* atau *N-Gain score* melalui persamaan berikut (Hake, 1998).

$$N - Gain = \frac{< Skor Posttest > - < Skor Pretest >}{Skor Maksimum - < Skor Pretest >} \quad 3.3$$

Skor *N-Gain* yang diperoleh kemudian diinterpretasikan berdasarkan kategori pada tabel 3.15.

Tabel 3. 15 Interpretasi Skor N-Gain

Nilai <i>N-Gain</i>	Kategori
$\langle g \rangle \geq 0,7$	Tinggi
$0,7 > \langle g \rangle \geq 0,3$	Sedang
$\langle g \rangle < 0,3$	Rendah

(Hake, 1998)

3.7.3. Efektivitas Penggunaan Model *argument driven inquiry*

Untuk menganalisis efektivitas penggunaan *model argument driven inquiry* dalam meningkatkan kemampuan kognitif dan keterampilan berpikir kritis peserta didik, dilakukan uji normalitas dan homogenitas data pretest dan posttest menggunakan aplikasi SPSS.

3.7.3.1. Uji Normalitas

Uji normalitas digunakan untuk mengetahui distribusi data sebagai syarat penggunaan uji statistik parametrik atau non-parametrik pada analisis data penelitian. Ketika data terdistribusi normal, maka digunakan metode statistik parametrik, sedangkan ketika data tidak terdistribusi normal, maka digunakan metode statistik non-parametrik. Uji Shapiro wilk merupakan metode statistik yang digunakan dalam penelitian ini. Dalam pengambilan keputusan uji normalitas didasarkan pada harga Sig. dengan kriteria sebagai berikut:

- Apabila *Sig.* < 0,05 maka dapat dikatakan bahwa data tidak terdistribusi normal;
- Apabila *Sig.* > 0,05 maka dapat dikatakan bahwa data terdistribusi normal.

3.7.3.2. Uji Homogenitas

Uji homogenitas dilakukan untuk mengetahui asal sampel dari populasi variansi yang homogen atau tidak. Uji homogenitas dapat dilakukan melalui uji Bartlet serta uji Levene. Uji Bartlet merupakan prosedur yang digunakan untuk menilai kesetaraan varians populasi yang berbeda pada data yang terdistribusi normal. Sedangkan uji Levene merupakan alternatif yang dapat digunakan ketika data tidak terdistribusi normal.

- Apabila, nilai probabilitas signifikan secara statistik atau $Sig. < 0,05$ maka sampel tidak homogen.
- Apabila nilai probabilitas signifikan secara statistik atau $Sig. > 0,05$ maka sampel homogen

3.7.3.3. Uji Hipotesis

Uji hipotesis dilakukan untuk mengetahui parameter yang akan diuji kebenarannya berdasarkan data yang diperoleh pada penelitian. Melalui uji hipotesis diperkenalkan ukuran perbedaan dan cara menghitung perbedaan yang dimulai dengan membandingkan hipotesis nol, yang menyatakan bahwa ‘tidak ada perbedaan yang signifikan secara statistik antara kedua kelompok’, atau ‘tidak ada pengaruh signifikan yang terjadi pada kelompok terhadap perlakuan yang diberikan’, dan, jika ini tidak didukung, maka hipotesis alternatif didukung, yaitu ditemukannya ‘perbedaan yang signifikan secara statistik antara dua (atau lebih) kelompok’ atau ‘terdapat pengaruh signifikan secara statistik yang terjadi pada kelompok yang diberikan perlakuan’ (Cohen dkk., 2007, hlm. 542).

Tabel 3. 16 Uji Statistik Parametrik dan Non-Parametrik

Statistik Parametrik	Statistik Non-Parametrik
Uji t sampel independen	Uji Mann Whitney U
H_0 = tidak ada perbedaan signifikan secara statistik kemampuan kognitif/KBK peserta didik antara kelas eksperimen dan kelas kontrol	

H_A = terdapat perbedaan signifikan secara statistik kemampuan kognitif/KBK peserta didik antara kelas eksperimen dan kelas kontrol		
$Sig. (2 - tailed) > 0,05 \rightarrow$	maka	$Asymp. sig. (2 - tailed) > 0,05 \rightarrow$
H_0 diterima dan H_A ditolak		maka H_0 diterima dan H_A ditolak
$Sig. (2 - tailed) < 0,05 \rightarrow$	maka	$Asymp. sig. (2 - tailed) < 0,05 \rightarrow$
H_0 ditolak dan H_A diterima		maka H_0 ditolak dan H_A diterima

3.7.3.4. *Effect size*

Dalam penelitian eksperimen, *effect size* digunakan untuk membandingkan kekuatan relatif variabel independen yang berbeda pada variabel dependen yang sama, dalam penelitian ini penggunaan model argument driven inquiry berperan sebagai variabel independen dan kemampuan kognitif serta keterampilan berpikir kritis peserta didik berperan sebagai variabel dependen. Sehingga hasil yang ditunjukkan pada perhitungan *effect size* pada penelitian ini dapat digunakan untuk membantu memutuskan apakah perbedaan yang dibuat variabel independen pada variabel dependen cukup kuat untuk merekomendasikan implementasinya dalam praktik. Coe (Cohen dkk., 2007) mengembangkan penilaian *effect size* dengan mengurangi rata-rata kelompok kontrol dari rata-rata kelompok eksperimen dan membaginya dengan standar deviasi gabungan, seperti pada persamaan berikut.

$$d = \frac{\bar{X}_e - \bar{X}_c}{SD_{pooled}} \quad 3.4$$

$$SD_{pooled} = \sqrt{\frac{(N_E - 1)SD_E^2 + (N_C - 1)SD_C^2}{N_E + N_C - 2}} \quad 3.5$$

Keterangan:

d = *effect size* untuk perbedaan antara rata-rata

$\bar{X}_1$ = rata-rata kelompok kontrol

$\bar{X}_2$ = rata-rata kelompok eksperimen

SD_{pooled} = standar deviasi kelompok control

N_E = jumlah peserta didik kelompok eksperimen

N_C = jumlah peserta didik kelompok kontrol

SD_E = standar deviasi kelompok eksperimen

SD_C = standar deviasi kelompok kontrol

Harga *effect size* yang diperoleh kemudian diinterpretasikan berdasarkan kategori pada Tabel 3.17.

Tabel 3. 17 Interpretasi *effect size*

<i>Effect Size</i>	Kategori
$0 < d \leq 0,20$	Efek lemah
$0,20 < d \leq 0,50$	Efek sederhana
$0,50 < d \leq 1,00$	Efek sedang
$1,00 < d$	Efek kuat

Coe (Cohen dkk., 2007)

3.7.4. Hubungan Keterampilan Berpikir Kritis dan Kemampuan Kognitif

Hubungan keterampilan berpikir kritis dan kemampuan kognitif pada penelitian ini dianalisis melalui uji korelas, uji regresi, dan uji determinasi menggunakan aplikasi SPSS.

3.7.4.1. Uji Korelasi

Uji korelasi dilakukan untuk menentukan bagaimana arah hubungan antara dua variabel (Suresh & Raju, 2022). Arah hubungan antara dua variabel ditunjukkan melalui harga koefisien korelasi, di mana ketika koefisien korelasi bernilai positif maka peningkatan nilai suatu variabel diikuti dengan peningkatan nilai variabel lainnya atau ketika penurunan nilai suatu variabel diikuti dengan penurunan nilai variabel lainnya. Jika koefisien korelasi bernilai negatif maka peningkatan nilai suatu variabel diikuti dengan penurunan nilai variabel lainnya, begitu pula sebaliknya. Pada penelitian ini uji korelasi digunakan untuk menguji hubungan antara keterampilan berpikir kritis dan kemampuan kognitif peserta didik. Analisis dilakukan melalui korelasi momen produk Pearson ketika data terdistribusi normal atau korelasi urutan peringkat

Spearman ketika data tidak terdistribusi normal. Analisis korelasi momen produk Pearson dihitung melalui Persamaan 3.6.

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N\{\sum X^2 (\sum X)^2\}\{\sum Y^2 (\sum Y)^2\}}} \quad 3.6$$

Keterangan:

r_{xy} = Koefisien korelasi antara variabel X dan Y

N = Jumlah data

X = Nilai kemampuan kognitif

Y = Nilai keterampilan berpikir kritis

Harga koefisien korelasi yang diperoleh kemudian diinterpretasikan berdasarkan kategori pada Tabel 3.18.

Tabel 3. 18 Interpretasi koefisien korelasi

Koefisien Korelasi	Kategori
0,00 – 0,10	Dapat diabaikan
0,10 – 0,39	Korelasi lemah
0,40 – 0,69	Korelasi sedang
0,70 – 0,89	Korelasi kuat
0,90 – 1,00	Korelasi sangat kuat

(Schober & Schwarte, 2018)