

BAB III

METODELOGI PENELITIAN

A. Metode Penelitian

Dalam penelitian ini, metode yang digunakan adalah kuasi eksperimen. Penelitian kuasi eksperimen bertujuan untuk “memperoleh informasi yang merupakan perkiraan bagi informasi yang dapat diperoleh dengan eksperimen sebenarnya dalam keadaan tidak memungkinkan untuk mengontrol atau memanipulasi semua variabel yang relevan kecuali dari beberapa variabel-variabel tersebut” (Panggabean, 1996: 27). Jadi digunakannya kuasi eksperimen ini karena tidak semua variabel lain, selain variabel yang diteliti, yang dapat mempengaruhi hasil penelitian dapat dikontrol. Selain itu juga karena pada kenyataannya di lokasi yang dijadikan sebagai tempat penelitian sulit untuk mendapatkan kelompok kontrol yang benar-benar mempunyai karakteristik yang homogen dengan kelompok eksperimen.

B. Desain Penelitian

Desain penelitian yang digunakan dalam penelitian ini adalah *one group time series design*. Secara bagan dapat digambarkan sebagai berikut:

Tabel 3. 1
Desain Penelitian

<i>Pretest</i>	Perlakuan	<i>Posttest</i>
$T_1T_2T_3$	X	$T_4T_5T_6$

dengan:

T_1, T_2, T_3 = *pretest* seri I, *pretest* seri II dan *pretest* seri III.

X = perlakuan

T_4, T_5, T_6 = *posttest* seri I, *posttest* seri II dan *posttest* seri III.

Desain ini digunakan dengan pertimbangan bahwa dalam *one group time series design*, setiap seri terdiri atas *pretest*, penerapan perlakuan, dan *posttest*. Dengan dilakukannya *pretest* dan *posttest* dalam satu seri, maka faktor-faktor eksternal yang dapat mempengaruhi sampel penelitian selain penerapan model pembelajaran dapat diminimalisir. Dalam desain ini, kelompok yang digunakan untuk penelitian tidak dipilih secara random, tetapi ditentukan berdasarkan pertimbangan-pertimbangan tertentu.

C. Populasi dan Sampel Penelitian

Menurut Arikunto (2006: 130), “populasi adalah keseluruhan subjek penelitian”, sedangkan sampel adalah sebagian dari keseluruhan subjek yang diteliti yang dianggap merepresentasikan populasi.

Dalam penelitian ini, yang dimaksud populasi adalah seluruh siswa kelas X yang berjumlah sembilan kelas di salah satu SMA di kota Bandung. Sedangkan sampelnya adalah siswa di salah satu kelas X yang diambil dengan teknik *purposive sampling*, yaitu teknik penentuan sampel yang didasarkan atas pertimbangan-pertimbangan tertentu. Dalam hal ini pertimbangan tersebut didasarkan atas informasi dari guru bahwa setiap kelas memiliki karakteristik akademis yang berbeda sehingga tingkat penyerapan materi akan berbeda tiap

kelasnya meskipun diberikan perlakuan yang sama. Sesuai dengan rekomendasai guru bidang studi fisika yang mengajar di kelas X maka sampel penelitian yang digunakan adalah kelas X – 7 dengan jumlah siswa sebanyak 30 orang yang mengikuti kegiatan penelitian dari *pretest* sampai *posttest*.

D. Teknik Pengumpulan Data

Teknik pengumpulan data merupakan cara-cara yang digunakan untuk memperoleh data-data empiris yang digunakan untuk mencapai tujuan penelitian. Sedangkan alat yang digunakan untuk memperoleh data tersebut disebut instrumen penelitian. Dalam penelitian ini, teknik pengumpulan data meliputi tes keterampilan berpikir kritis dan observasi aktivitas guru dan siswa sesuai dengan sintaks model pembelajaran induktif.

1. Tes Keterampilan Berpikir Kritis Siswa

Tes keterampilan berpikir kritis digunakan untuk mengukur peningkatan keterampilan berpikir kritis siswa pada aspek yang diteliti. Tes yang digunakan berbentuk soal pilihan ganda dengan 5 (lima) pilihan yang terdiri dari tes awal (*pretest*) dan tes akhir (*posttest*). Soal *pretest* dan *posttest* menggunakan soal yang sama dengan anggapan bahwa dengan soal yang sama maka peningkatan keterampilan berpikir kritis akan lebih mudah terlihat dan terukur. Butir-butir soal dalam *pretest* dan *posttest* mencakup soal-soal sesuai dengan indikator keterampilan berpikir kritis yang akan diteliti. Untuk mengetahui bagaimana pengaruh penerapan model pembelajaran induktif terhadap peningkatan keterampilan berpikir kritis siswa, maka hasil *pretest*

dan *posttest* kelompok eksperimen pada tiap seri pembelajaran diolah dan dianalisis dengan menggunakan uji statistik.

2. Observasi Aktivitas Guru dan Siswa

Instrumen yang digunakan dalam mengobservasi aktivitas guru dan siswa berupa lembar observasi. Lembar observasi merupakan lembar yang diisi oleh observer selama pembelajaran berlangsung. Tujuannya untuk mengetahui efektivitas pembelajaran serta kekurangan dan kelebihan terhadap model pembelajaran yang diterapkan. Dalam penelitian ini observer adalah beberapa orang (minimal tiga orang) yang berada di dalam kelas selama proses pembelajaran selain guru dan siswa. Observer bertugas untuk mengamati aktivitas guru dan siswa selama proses pembelajaran sesuai dengan sintaks model pembelajaran induktif.

Observasi terhadap aktivitas guru dilakukan untuk mengetahui keterlaksanaan model pembelajaran induktif sesuai dengan aktivitas guru yang teramati. Sedangkan pengamatan terhadap siswa dilakukan untuk untuk mengetahui keterlaksanaan model pembelajaran induktif sesuai dengan aktivitas siswa yang teramati.

E. Prosedur Penelitian

Dalam penelitian ini meliputi tiga tahap yaitu tahap perencanaan, tahap pelaksanaan dan tahap akhir.

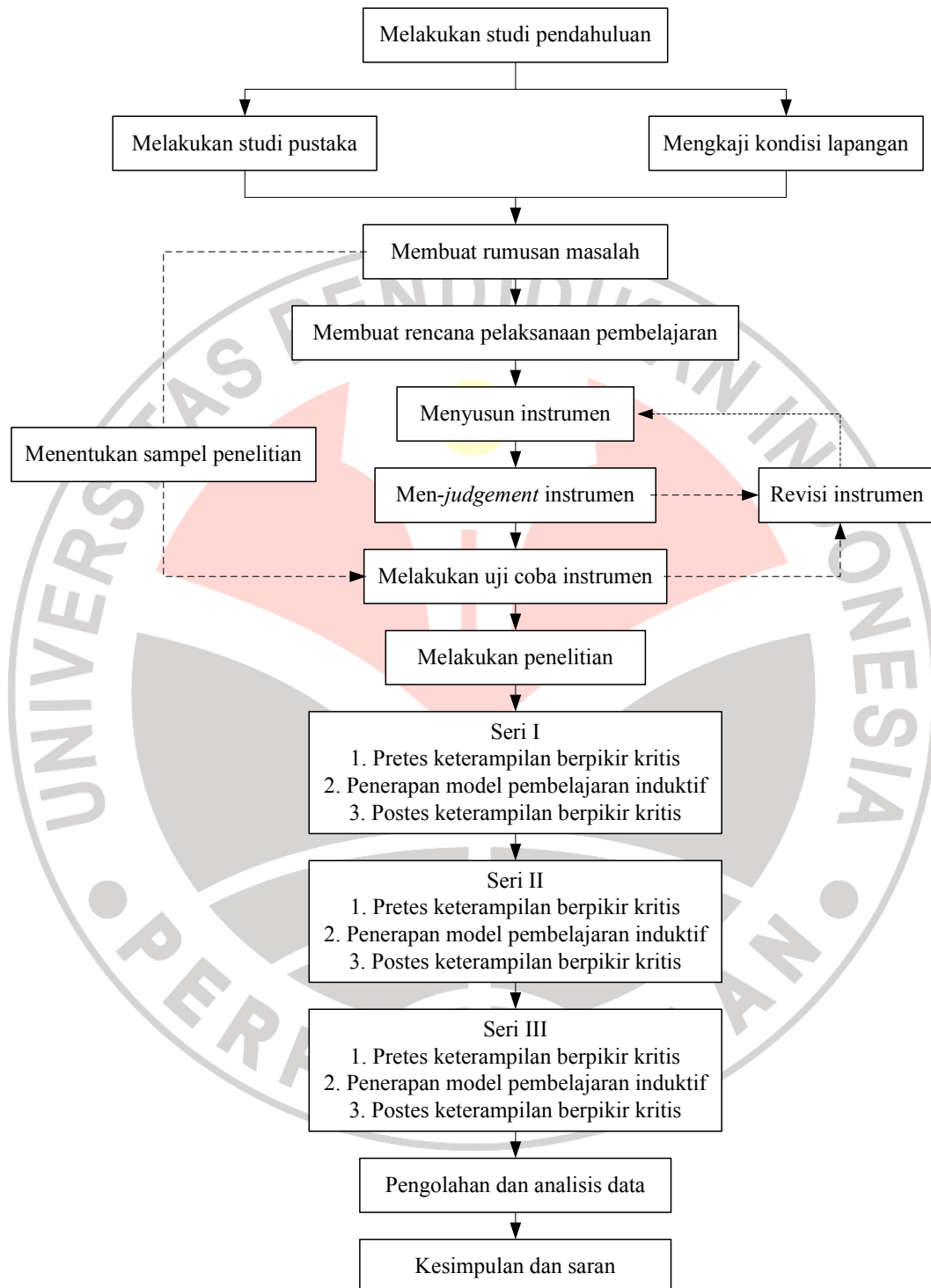
1. Tahap Perencanaan

- a. Telaah kompetensi mata pelajaran fisika untuk jenjang SMA.
- b. Menentukan sekolah yang akan dijadikan tempat penelitian, mengurus surat izin penelitian dan menghubungi pihak sekolah tempat penelitian akan dilaksanakan.
- c. Observasi awal yang dilakukan dengan cara pengamatan langsung mengenai proses pembelajaran di kelas serta wawancara dengan guru dan siswa mengenai pembelajaran yang biasa dilakukan.
- d. Studi literatur terhadap buku, artikel, dan jurnal dan laporan penelitian mengenai model pembelajaran induktif.
- e. Perumusan masalah penelitian
- f. Telaah kurikulum fisika SMA kelas X dan penentuan materi pembelajaran yang dijadikan materi pembelajaran dalam penelitian serta untuk mengetahui tujuan atau kompetensi dasar yang hendak dicapai.
- g. Menyusun perangkat pembelajaran yang meliputi rencana pelaksanaan pembelajaran (RPP) dan LKS.
- h. Menyusun instrumen penelitian yang meliputi tes keterampilan berpikir kritis dan lembar observasi keterlaksanaan model pembelajaran induktif oleh guru dan siswa.
- i. Men-*judgement* instrumen (tes) kepada dua orang dosen dan satu guru

mata pelajaran fisika yang ada di sekolah tempat penelitian akan dilaksanakan.

- j. Merevisi/memperbaiki instrumen.
 - k. Melakukan uji coba instrumen.
 - l. Menganalisis hasil uji coba instrumen yang meliputi validitas, tingkat kesukaran, daya pembeda dan reliabilitas sehingga layak dipakai untuk dijadikan sebagai instrumen penelitian.
2. Tahap Pelaksanaan
- a. Penentuan sampel penelitian (kelompok eksperimen).
 - b. Pelaksanaan tes awal (*pretest*) bagi kelompok eksperimen.
 - c. Memberikan perlakuan yaitu berupa penerapan model pembelajaran induktif pada kelompok eskperimen.
 - d. Selama proses pembelajaran berlangsung, observer melakukan observasi keterlaksanaan model pembelajaran induktif sesuai dengan aktivitas guru dan siswa yang teramati.
 - e. Pelaksanaan tes akhir (*posttest*) bagi kelompok eksperimen.
 - f. Mengulang pelaksanaan penelitian sebanyak tiga seri pembelajaran.
3. Tahap Akhir
- a. Mengolah data hasil *pretest*, *posttest* dan instrumen lainnya.
 - b. Menganalisis dan membahas temuan penelitian.
 - c. Menarik kesimpulan.

Untuk lebih jelasnya, alur penelitian yang telah dilakukan dapat digambarkan seperti pada gambar 3.1 berikut.



Gambar 3.1 Bagan Alur Penelitian

F. Uji Coba Instrumen

Sebelum soal *pretest* dan *posttest* digunakan pada kelas yang dijadikan sampel penelitian, terlebih dahulu soal ini diujicobakan di kelas lain yang bukan merupakan sampel penelitian. Analisis soal yang digunakan meliputi uji tingkat kesukaran butir soal, daya pembeda butir soal, uji validitas butir soal, dan uji reliabilitas.

1. Tingkat Kesukaran Butir Soal

Tingkat kesukaran suatu butir soal merupakan proporsi dari keseluruhan siswa yang menjawab benar pada butir soal tersebut. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Tingkat kesukaran dihitung dengan menggunakan rumus sebagai berikut:

$$P = \frac{B}{JS}$$

keterangan:

P = indeks kesukaran.

B = banyaknya siswa yang menjawab dengan benar pada suatu soal.

JS = jumlah seluruh siswa peserta tes.

Interpretasi dari nilai indeks kesukaran yang diperoleh adalah sebagai berikut:

Tabel 3. 2

Interpretasi Indeks Kesukaran (Arikunto, 2008: 210)

Nilai P	Kriteria
0.00 – 0.30	Sukar
0.31 – 0.70	Sedang
0.71 – 1.00	Mudah

Soal-soal yang dianggap baik adalah soal-soal sedang yaitu soal-soal yang mempunyai indeks kesukaran 0.30 sampai dengan 0.70.

2. Daya Pembeda Butir Soal

Daya pembeda merupakan keterampilan suatu soal untuk membedakan siswa yang berketerampilan tinggi dengan siswa yang berketerampilan rendah.

Untuk menghitung daya pembeda, digunakan rumus:

$$DP = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

keterangan:

DP = indeks daya pembeda butir soal.

J_A = banyaknya peserta kelompok atas.

J_B = banyaknya peserta kelompok bawah.

B_A = banyaknya peserta kelompok atas yang menjawab soal itu dengan benar.

B_B = banyaknya peserta kelompok bawah yang menjawab soal itu dengan benar.

Sedangkan interpretasi nilai daya pembeda adalah sebagai berikut:

Tabel 3.3

Interpretasi Daya Pembeda (Arikunto, 2008: 218)

Nilai <i>DP</i>	Kategori
< 0.00	Tidak baik
0.01 – 0.20	Jelek (<i>poor</i>)
0.21 – 0.40	Cukup (<i>satisfactory</i>)
0.41 – 0.70	Baik (<i>good</i>)
0.71 – 1.00	Baik sekali (<i>excellent</i>)

3. Validitas Soal

Validitas berhubungan dengan ketepatan atau kesahihan instrumen yaitu kesesuaian tujuan dengan alat ukur yang digunakan. Sebuah tes dikatakan memiliki validitas jika hasilnya sesuai dengan kriteria dalam arti memiliki kesejajaran antara hasil tes dengan kriteria. Teknik untuk mengetahui kesejajaran tersebut salah satunya dengan menggunakan rumus γ_{pbi} atau rumus korelasi poin biseral (Arikunto, 2008: 72), yaitu:

$$\gamma_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

keterangan:

γ_{pbi} = koefisien korelasi biseral.

M_p = rata-rata skor dari subjek yang menjawab betul untuk butir soal yang dicari validitasnya.

M_t = rata-rata skor total.

S_t = standar deviasi dari skor total.

p = proporsi siswa yang menjawab benar atau banyaknya siswa yang menjawab benar dibagi dengan jumlah seluruh siswa.

q = proporsi siswa yang menjawab salah ($q = 1 - p$).

Sedangkan interpretasi besarnya koefisien korelasi r_{xy} adalah sebagai berikut:

Tabel 3. 4

Interpretasi Validitas Butir Soal (Arikunto, 2008: 75)

Koefisien Korelasi	Kriteria
0.00 – 0.20	Sangat rendah
0.21 – 0.40	Rendah
0.41 – 0.60	Sedang
0.61 – 0.80	Tinggi
0.81 – 1.00	Sangat tinggi

4. Reliabilitas

Reliabilitas merupakan ukuran sejauh mana suatu tes dapat dipercaya untuk menghasilkan skor yang konsisten. Dalam penelitian ini teknik yang akan digunakan untuk menentukan reliabilitas tes adalah dengan menggunakan rumus K-R 20 dengan persamaan (Arikunto, 2008: 100), yaitu:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \Sigma pq}{S^2} \right)$$

keterangan:

- r_{11} = reliabilitas tes secara keseluruhan
- p = proporsi subjek yang menjawab item dengan benar
- q = proporsi subjek yang menjawab item dengan salah ($q = 1 - p$)
- Σpq = jumlah hasil perkalian antara p dan q
- n = banyaknya item
- S = standar deviasi dari tes

Sedangkan interpretasi besar koefisien korelasi adalah sebagai berikut:

Tabel 3.5
Interpretasi Reliabilitas Tes (Arikunto, 2008: 75)

Koefisien Korelasi	Kriteria
0.00 – 0.20	Sangat rendah
0.21 – 0.40	Rendah
0.41 – 0.60	Sedang
0.61 – 0.80	Tinggi
0.81 – 1.00	Sangat tinggi

G. Teknik Pengolahan Data

1. Analisis Skor *Pretest* dan *Posttest*

Dari skor *pretest* dan *posttest*, dihitung gain (G) dan gain yang dinormalisasi (<g>). Perbedaan skor *pretest* dan *posttest* (gain) diasumsikan sebagai efek dari perlakuan (Panggabean, 1996). Sedangkan gain yang dinormalisasi diinterpretasikan sebagai kriteria untuk menyatakan efektivitas pembelajaran serta menunjukkan besarnya peningkatan antara skor *posttest* dan *pretest*. Untuk menghitung gain dan gain yang dinormalisasi digunakan rumus sebagai berikut (Hake, 1999):

$$G = T_2 - T_1$$

$$\langle g \rangle = \frac{T_2 - T_1}{I_s - T_1}$$

keterangan:

G = gain

<g> = gain yang dinormalisasi

T₁ = skor *pretest*

T_2 = skor *posttest*

I_s = skor ideal

Interpretasi $\langle g \rangle$ yang diperoleh ditunjukkan pada tabel berikut:

Tabel 3.6
Interpretasi Gain yang Dinormalisasi $\langle g \rangle$

$\langle g \rangle$	Kriteria
0.71 – 1.00	Tinggi
0.41 – 0.70	Sedang
< 0.40	Rendah

2. Uji Hipotesis

Secara umum pengujian hipotesis statistik dapat dilakukan dengan dua cara yaitu uji statistik parametrik dan uji statistik non-parametrik. Untuk menentukan uji statistik mana yang dipakai, maka terlebih dahulu dilakukan uji normalitas dan uji homogenitas.

a. Uji Normalitas

Dalam penelitian ini, uji normalitas yang akan digunakan adalah uji *Chi-Kuadrat* (χ^2) dengan langkah-langkah sebagai berikut:

- 1) Menentukan banyak kelas (K) dengan rumus:

$$K = 1 + 3,3 \log n ; n \text{ adalah jumlah siswa}$$

- 2) Menentukan panjang kelas (P) dengan rumus:

$$P = \frac{R}{K} = \frac{\text{rentang}}{\text{banyak kelas}} ; R = \text{skor maksimum} - \text{skor minimum}$$

- 3) Menghitung rata-rata dan standar deviasi dari data yang akan diuji normalitasnya.

Untuk menghitung nilai rata-rata (*mean*) dari gain digunakan persamaan:

$$\bar{x} = \frac{\sum x_i}{n}$$

Sedangkan untuk menghitung besarnya standar deviasi dari gain digunakan persamaan:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n-1)}}$$

keterangan: \bar{x} = nilai rata-rata skor *pretest* atau *posttest*
 x_i = skor *pretest* atau *posttest* yang diperoleh siswa
 n = jumlah siswa
 S = standar deviasi

- 4) Menentukan nilai baku z dengan menggunakan persamaan :

$$z = \frac{bk - \bar{x}}{S} ; bk = \text{batas kelas}$$

- 5) Mencari luas daerah dibawah kurva normal (l) untuk setiap kelas interval

$$l = |l_1 - l_2|$$

keterangan: l = luas kelas interval

l_1 = luas daerah batas bawah kelas interval

$l_2 =$ luas daerah batas atas kelas interval

- 6) Mencari frekuensi observasi (O_i) dengan menghitung banyaknya respon yang termasuk pada interval yang telah ditentukan.
- 7) Mencari frekuensi harapan E_i

$$E_i = n \times l$$

- 8) Mencari harga *Chi-Kuadrat* (χ^2) dengan menggunakan persamaan :

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

keterangan: χ^2_{hitung} = chi kuadrat hasil perhitungan

O_i = frekuensi observasi

E_i = frekuensi yang diharapkan

- 9) Membandingkan harga χ^2_{hitung} dengan χ^2_{tabel}

Jika $\chi^2_{hitung} < \chi^2_{tabel}$, maka data berdistribusi normal, sedangkan jika

$\chi^2_{hitung} > \chi^2_{tabel}$, maka data tidak berdistribusi normal.

b. Uji Homogenitas

Uji Homogenitas dilakukan untuk memeriksa apakah skor-skor pada penelitian yang dilakukan mempunyai variansi yang homogen atau tidak untuk taraf signifikansi α . Untuk uji homogenitas ini digunakan persamaan berikut:

$$F = \frac{s^2_b}{s^2_k}$$

keterangan : $s^2_b =$ Varians yang lebih besar

$s^2_k =$ Varians yang lebih kecil

Setelah nilai F diperoleh melalui perhitungan, kemudian menentukan nilai F dari tabel distribusi frekuensi dengan derajat kebebasan sebesar $(dk) = n - 1$ pada taraf signifikansi α . Dari kedua nilai tersebut dibandingkan dengan interpretasi sebagai berikut:

Jika $F_{hitung} < F_{tabel}$, maka variansi sampel homogen.

Jika $F_{hitung} > F_{tabel}$, maka variansi sampel tidak homogen.

c. Uji Hipotesis

Setelah dilakukan uji normalitas dan uji homogenitas ternyata diperoleh bahwa data terdistribusi normal dan homogen, maka untuk menguji hipotesis dapat digunakan uji statistik parametrik yaitu *uji-t*. Akan tetapi apabila salah satu data tidak normal atau tidak homogen maka statistik yang digunakan adalah statistik non-parametrik yaitu dapat menggunakan uji Wilcoxon.

Langkah-langkah untuk menguji hipotesis dengan menggunakan uji- t adalah sebagai berikut:

- 1) Menghitung nilai t (untuk sampel besar $n \geq 30$) dengan menggunakan rumus (Luhut Panggabean, 1996: 102)

$$t = \frac{M_2 - M_1}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n}}}$$

keterangan:

M_1 = rata-rata skor *pretest*.

M_2 = rata-rata skor *posttest*.

s_1^2 = standar deviasi skor *pretest*.

s_2^2 = standar deviasi skor *posttest*.

n = jumlah siswa

- 2) Mencari nilai t pada tabel distribusi t untuk tes dua ekor dengan derajat kebebasan $dk = n - 1$ pada taraf signifikansi α .
- 3) Membandingkan nilai t ; jika $t_{hitung} < t_{tabel}$ maka disimpulkan bahwa H_0 diterima dan H_1 ditolak.

Sedangkan langkah-langkah untuk menguji hipotesis dengan menggunakan uji Wilcoxon adalah sebagai berikut:

- 1) Membuat daftar rank.
- 2) Menghitung nilai W yaitu bilangan terkecil dari jumlah rank positif atau jumlah rank negatif dari daftar rank yang telah dibuat.
- 3) Menentukan nilai W_{tabel} untuk jumlah sampel n pada taraf signifikansi α .
- 4) Membandingkan nilai W ; jika $W_{hitung} > W_{tabel}$ maka disimpulkan bahwa H_0 diterima dan H_1 ditolak.

H. Analisis Hasil Uji Coba Instrumen

Untuk mendapatkan instrumen yang benar-benar mengukur keterampilan berpikir kritis siswa, maka instrumen yang telah disusun terlebih dahulu di-*judgement* dan diujicobakan. *Judgement* instrumen dilakukan oleh dua orang dosen fisika dan satu orang guru bidang studi fisika. Berdasarkan hasil *judgement* itu kemudian instrumen diperbaiki untuk selanjutnya dilakukan uji coba. Lembar

judgement untuk masing-masing seri pembelajaran dapat dilihat pada lampiran B.1.c, lampiran B.2.c, dan lampiran B.3.c.

Uji coba dilaksanakan di sekolah yang sama pada jenjang kelas yang sama tetapi di kelas yang tidak diteliti. Data hasil uji coba instrumen kemudian dianalisis untuk mengetahui tingkat kesukaran butir soal, daya pembeda butir soal, validitas butir soal, dan reliabilitas tes. Soal-soal yang telah diujicoba dan dianalisis kemudian digunakan sebagai instrumen dalam penelitian yang dilakukan di kelas X, kelas yang berbeda dengan kelas yang dijadikan sampel uji coba instrumen. Kelas X yang dijadikan sebagai sampel penelitian adalah kelas X yang belum mendapatkan materi listrik dinamis, sedangkan kelas uji coba instrumen adalah kelas X yang telah mendapatkan materi listrik dinamis oleh guru bidang studi fisika di sekolah tersebut. Soal yang diujicobakan seluruhnya berjumlah dua puluh tujuh soal dengan masing-masing seri I, II, dan III berturut-turut adalah sembilan soal, sepuluh soal, dan delapan soal. Pengolahan data hasil uji coba instrumen dapat dilihat pada lampiran C, sedangkan rekapitulasi hasil analisis uji coba instrumen dapat dilihat dalam Tabel 3.7 berikut ini.

Tabel 3.7
Rekapitulasi Hasil Analisis Uji Coba Instrumen

Seri	No. Soal	Validitas		Tingkat Kesukaran		Daya Pembeda		Reliabilitas		Ket.
		Nilai	Kriteria	Nilai	Kriteri	Nilai	Kriteria	Nilai	Kriteria	
1	1	0.55	Sedang	0.61	Sedang	0.67	Baik	0.47	Cukup	Digunakan
	2	0.26	Rendah	0.88	Mudah	0.11	Jelek			Diperbaiki
	3	0.64	Tinggi	0.79	Mudah	0.33	Cukup			Digunakan
	4	0.76	Tinggi	0.91	Mudah	0.33	Cukup			Digunakan
	5	0.59	Sedang	0.45	Sedang	0.78	Sgt Baik			Digunakan
	6	0.60	Sedang	0.91	Mudah	0.44	Baik			Digunakan
	7	-0.37	Sgt rendah	0.12	Sukar	-0.22	Sgt Jelek			Dibuang
	8	0.44	Sedang	0.52	Sedang	0.89	Sgt Baik			Digunakan
	9	0.49	Sedang	0.88	Mudah	0.22	Cukup			Digunakan
2	1	0.21	Rendah	0.85	Mudah	0.11	Jelek	0.58	Cukup	Dibuang
	2	0.45	Sedang	0.85	Mudah	0.44	Baik			Digunakan
	3	0.62	Tinggi	0.56	Sedang	0.67	Baik			Digunakan
	4	0.29	Rendah	0.91	Mudah	0.11	Jelek			Diperbaiki
	5	0.60	Sedang	0.45	Sedang	0.67	Baik			Digunakan
	6	0.24	Rendah	0.09	Sukar	0.22	Cukup			Digunakan
	7	0.76	Tinggi	0.42	Sedang	0.89	Sgt Baik			Digunakan
	8	0.39	Rendah	0.48	Sedang	0.33	Cukup			Digunakan
	9	0.74	Tinggi	0.73	Sedang	0.89	Sgt Baik			Digunakan
	10	0.08	Sgt rendah	0.45	Sedang	0.33	Cukup			Dibuang
3	1	0.52	Sedang	0.64	Sedang	0.56	Baik	0.75	Tinggi	Digunakan
	2	0.67	Tinggi	0.85	Mudah	0.44	Baik			Digunakan
	3	0.72	Tinggi	0.69	Sedang	0.56	Baik			Digunakan
	4	0.08	Sgt rendah	0.09	Sukar	0.00	Jelek			Diperbaiki
	5	0.79	Tinggi	0.64	Sedang	0.67	Baik			Digunakan
	6	0.92	Sgt tinggi	0.67	Sedang	0.89	Sgt Baik			Diperbaiki
	7	0.32	Rendah	0.09	Sukar	0.11	Jelek			Diperbaiki
	8	0.76	Tinggi	0.91	Mudah	0.22	Cukup			Digunakan

1. Analisis Taraf Kesukaran Butir Soal

Berdasarkan analisis tingkat kesukaran butir soal yang telah dilakukan menunjukkan bahwa pada seri I terdapat lima soal atau 55.55% soal dengan kategori mudah, tiga soal atau 33.33% dengan kategori sedang dan satu soal atau

11.11% dengan kategori sukar. Pada seri II, tiga soal atau 30% termasuk dalam kategori mudah, enam soal atau 60% termasuk dalam kategori sedang, dan satu soal atau 10% termasuk dalam kategori sukar. Sedangkan pada seri III, empat soal atau 50% termasuk kategori mudah, dua soal atau 25% termasuk kategori sedang, dan dua soal atau 25% termasuk kategori sukar.

2. Analisis Daya Pembeda Butir Soal

Berdasarkan analisis tingkat kesukaran butir soal yang telah dilakukan menunjukkan bahwa pada seri I terdapat satu soal atau 11.11% yang termasuk dalam kategori sangat jelek, satu soal atau 11.11% termasuk dalam kategori jelek, tiga soal atau 33.33% kategori cukup, dua soal atau 22.22% kategori baik, dan dua soal atau 22.22% kategori sangat baik. Soal dengan daya pembeda sangat jelek yaitu soal nomor tujuh tidak digunakan dalam instrumen penelitian. Sedangkan soal dengan daya pembeda jelek tetap digunakan sebagai instrumen penelitian namun diperbaiki terlebih dahulu.

Pada seri II, terdapat dua soal atau 20% termasuk kategori jelek, tiga soal atau 30% termasuk kategori cukup, tiga soal atau 30% kategori baik, dan dua soal atau 20% kategori sangat baik. Soal-soal dengan kategori jelek adalah soal nomor satu dan empat. Untuk soal nomor satu keputusan yang diambil adalah soal ini tidak digunakan karena memiliki daya pembeda yang jelek serta validitasnya rendah. Sedangkan soal nomor empat diperbaiki untuk selanjutnya digunakan dalam instrumen penelitian. Untuk seri III, terdapat dua soal atau 25% termasuk kategori jelek, satu soal atau 12.5% kategori cukup, empat soal atau 50% kategori baik, dan satu soal atau 12.5% kategori sangat baik. Pada seri III ini, soal-soal

yang termasuk kategori jelek diperbaiki untuk selanjutnya digunakan dalam instrumen penelitian.

3. Analisis Validitas Butir Soal

Hasil pengolahan data untuk menghitung validitas soal dengan menggunakan persamaan korelasi *point-biserial* didapatkan bahwa pada seri I terdapat satu soal atau 11.11% yang termasuk dalam kategori sangat rendah, satu soal atau 11.11% termasuk dalam kategori rendah, lima soal atau 55.56% termasuk kategori sedang, dan dua soal atau 22.22% termasuk kategori tinggi. Soal yang mempunyai validitas sangat rendah adalah soal nomor tujuh. Soal ini tidak digunakan dalam instrumen penelitian karena selain mempunyai validitas soal yang sangat rendah juga mempunyai tingkat kesukaran yang tinggi dan daya pembeda yang sangat jelek. Sedangkan soal yang validitasnya rendah diperbaiki dan digunakan dalam instrumen penelitian.

Pada seri II, terdapat satu soal atau 10% termasuk dalam kategori sangat rendah, empat soal atau 40% kategori rendah, dua soal atau 20% kategori sedang, dan tiga soal atau 30% kategori tinggi. Sama seperti halnya pada seri I, soal dengan kategori sangat rendah yaitu soal nomor 10, tidak digunakan dalam instrumen penelitian. Soal-soal yang validitasnya rendah yaitu soal nomor 1, 4, 6, dan 8. Untuk soal nomor, keputusan yang diambil adalah soal tersebut tidak digunakan karena selain mempunyai validitas yang rendah, soal tersebut juga mempunyai daya pembeda yang jelek. Sedangkan soal lain yang validitasnya rendah tetap digunakan dalam instrumen penelitian namun sebelumnya dilakukan perbaikan dalam hal penulisan dan tata bahasa tanpa mengurangi konten soal

tersebut. Untuk seri III, terdapat satu soal 12.5% yang termasuk dalam kategori sangat rendah, satu soal atau 12.5% kategori rendah, satu soal atau 12.5% kategori sedang, empat soal atau 50% kategori tinggi dan satu soal atau 12.5 kategori sangat tinggi. Soal-soal yang memiliki validitas rendah dan sangat rendah diperbaiki untuk selanjutnya digunakan dalam instrumen penelitian. Sedangkan rincian mengenai pengolahan data untuk analisis validitas butir soal dapat dilihat pada lampiran C.1.

4. Analisis Reliabilitas Tes

Untuk menghitung reliabilitas tes masing-masing seri pembelajaran I, II, dan III digunakan rumus *K-R 20*. Berdasarkan perhitungan tersebut didapatkan bahwa nilai koefisien reliabilitas untuk seri I adalah 0.47 yang termasuk dalam kategori cukup. Sedangkan untuk seri II dan III masing-masing nilai koefisien reliabilitasnya adalah 0.58 dan 0.75 yang termasuk dalam kategori cukup dan tinggi. Rincian pengolahan data untuk analisis reliabilitas instrumen dapat dilihat pada lampiran C.1.

Berdasarkan hasil analisis uji coba instrumen didapatkan bahwa soal yang layak untuk digunakan sebagai instrumen penelitian adalah sebanyak 24 soal dan tiga soal lainnya tidak digunakan. Soal-soal tersebut terdistribusi dalam tiga seri pembelajaran I, II, dan III masing-masing sebanyak delapan soal. Soal-soal tersebut merupakan soal yang disusun sedemikian rupa sehingga dapat mengukur keterampilan berpikir kritis siswa pada aspek membuat generalisasi dan membuat hipotesis.