

## BAB 3

### METODOLOGI PENELITIAN

#### 3.1 Metode Penelitian

Metode penelitian merupakan cara ilmiah untuk mendapatkan data dengan tujuan dan kegunaan tertentu (Sugiyono, 2017: 3). Dalam penelitian ini penulisan menggunakan pendekatan kuantitatif dengan metode penelitian eksperimen.

Penelitian kuantitatif merupakan pendekatan penelitian yang berfokus pada pengukuran secara objektif serta analisis data melalui prosedur statistik, yang diperoleh dari instrumen seperti survei, angket, maupun eksperimen. Pendekatan ini umumnya digunakan untuk menguji hipotesis atau teori yang telah ada dengan cara mengidentifikasi dan mengukur variabel-variabel tertentu, kemudian menganalisis hubungan maupun pola yang muncul di antara variabel-variabel tersebut. Data dalam penelitian kuantitatif biasanya berbentuk angka dan dianalisis menggunakan teknik statistik guna menguji kebenaran hipotesis yang diajukan. Selain itu, penelitian kuantitatif umumnya melibatkan sampel yang mewakili populasi yang lebih luas, sehingga hasil penelitiannya dapat digeneralisasikan (Hidayat dkk., 2024: 4).

Metode eksperimen merupakan jenis penelitian yang bertujuan untuk membuktikan adanya pengaruh suatu perlakuan terhadap hasil atau akibat yang ditimbulkan. Menurut Arikunto dalam Arib dkk (2024: 5498), penelitian eksperimen dilakukan dengan cara sengaja menciptakan atau menimbulkan suatu kejadian maupun kondisi tertentu untuk kemudian mengamati akibat yang muncul. Penelitian eksperimen dapat dipahami sebagai metode untuk mengkaji hubungan sebab-akibat antara dua faktor. Dalam konteks pendidikan, penelitian eksperimen digunakan untuk mengetahui pengaruh suatu tindakan pembelajaran terhadap perilaku peserta didik, serta menguji hipotesis mengenai ada atau tidaknya pengaruh suatu perlakuan dibandingkan dengan perlakuan lainnya. Perlakuan dalam penelitian eksperimen dikenal dengan istilah *treatment*, yaitu segala bentuk tindakan atau variasi yang diuji untuk mengetahui pengaruhnya.

Metode eksperimental yang diterapkan dalam penelitian ini adalah Quasi-Eksperimen menggunakan desain *Nonequivalent control group design*. Desain ini merupakan rancangan penelitian yang dilakukan dengan memberikan perlakuan kepada kelompok eksperimen dan menyediakan kelompok kontrol sebagai kelompok pembanding. Dalam desain ini, penentuan kelompok eksperimen maupun kelompok kontrol tidak dilakukan secara acak (Monalisa dkk., 2025: 42).

### 3.2 Desain Penelitian

Penelitian ini menggunakan desain *Nonequivalent control group design*. Desain ini merupakan rancangan penelitian yang melibatkan dua kelompok, yaitu kelompok eksperimen dan kelompok kontrol. Sebelum perlakuan diberikan, kedua kelompok terlebih dahulu menjalani *pretest* untuk menilai kondisi awal. Selanjutnya, perlakuan hanya diterapkan pada kelompok eksperimen dengan metode pembelajaran Numbered Head Together (NHT), sedangkan kelompok kontrol tidak diberikan perlakuan dengan metode pembelajaran konvensional. Pada tahap akhir, dilakukan *posttest* terhadap kedua kelompok yang akan dianalisis untuk menentukan apakah terdapat perbedaan yang signifikan antara kedua kelompok tersebut. mengukur dampak dari perlakuan yang telah diberikan.

**Tabel 3.1 Desain Penelitian**

| Kelas      | <i>Pretest</i> | Perlakuan                    | <i>Posttest</i> |
|------------|----------------|------------------------------|-----------------|
| Eksperimen | O <sub>1</sub> | Numbered Head Together (NHT) | O <sub>2</sub>  |
| Kontrol    | O <sub>3</sub> | -                            | O <sub>4</sub>  |

**Keterangan:**

- O<sub>1</sub>: *Pretest* pada kelompok eksperimen
- O<sub>2</sub> : *Posttest* pada kelompok eksperimen
- O<sub>3</sub> : *Pretest* pada kelompok kontrol
- O<sub>4</sub> : *Posttest* pada kelompok kontrol

### 3.3 Variabel Penelitian

Variabel penelitian merupakan segala sesuatu yang ditetapkan oleh peneliti untuk dikaji guna memperoleh informasi yang relevan mengenai objek tersebut, sehingga dapat ditarik suatu kesimpulan berdasarkan hasil analisis yang dilakukan

(Sugiyono 2017: 60). Dalam penelitian ini terdapat dua jenis variabel, yaitu variabel independen (variable bebas) dan variabel dependen (variable terikat). Variabel independen merupakan variabel yang berperan sebagai faktor yang memengaruhi atau menjadi penyebab terjadinya perubahan pada variabel lainnya. Sebaliknya, variabel dependen (Y) adalah variabel yang dipengaruhi atau muncul sebagai akibat dari adanya variabel independen. Pada penelitian ini, variabel independen (X) adalah Numbered Head Together (NHT), sedangkan variabel dependen (Y) adalah keterampilan berpikir kritis.

### 3.4 Populasi dan Sampel

Populasi merupakan keseluruhan wilayah generalisasi yang terdiri atas objek atau subjek yang memiliki karakteristik dan jumlah tertentu, yang telah ditetapkan oleh peneliti untuk diteliti serta dijadikan dasar dalam penarikan kesimpulan penelitian (Siyoto & Sodik, 2015: 63). Populasi dalam penelitian ini yaitu peserta didik kelas X MAN 3 Kabupaten Tasikmalaya yang menunjukkan kemampuan berpikir kritis rendah.

**Tabel 3.2 Jumlah Populasi Penelitian**

| No            | Kelas | Jumlah Peserta Didik |
|---------------|-------|----------------------|
| 1             | X-1   | 32                   |
| 2             | X-2   | 30                   |
| 3             | X-3   | 32                   |
| 4             | X-4   | 32                   |
| 5             | X-10  | 32                   |
| 6             | X-11  | 32                   |
| 7             | X-12  | 32                   |
| <b>Jumlah</b> |       | 222                  |

Sumber: Tata Usaha MAN 3 Kabupaten Tasikmalaya Tahun 2025/2026

Sampel merupakan bagian dari populasi yang memiliki karakteristik tertentu dan diambil melalui prosedur yang telah ditetapkan, sehingga mampu mewakili keseluruhan populasi. Apabila ukuran populasi terlalu besar dan peneliti tidak memungkinkan untuk meneliti seluruh anggotanya karena keterbatasan sumber daya seperti waktu, tenaga, dan biaya, maka peneliti dapat menggunakan sebagian dari populasi tersebut sebagai sampel. Oleh karena itu, sampel yang

digunakan harus benar-benar representatif agar dapat menggambarkan kondisi populasi secara akurat (Siyoto & Sodik, 2015: 64).

Pada penelitian ini, sampel yang digunakan terdiri atas kelas X-11 sebagai kelompok eksperimen dan kelas X-12 sebagai kelompok kontrol. Teknik pengambilan sampel yang diterapkan adalah *purposive sampling*, karena pemilihan kelas dilakukan secara sengaja berdasarkan pertimbangan tertentu yang sesuai dengan tujuan penelitian. Kelas X-11 dan X-12 ditetapkan sebagai kelompok kontrol dan eksperimen karena menunjukkan kemampuan berpikir kritis yang relatif lebih rendah dibandingkan dengan kelas yang lain.

**Tabel 3.3 Jumlah Sampel Penelitian**

| No.           | Kelas                          | Jenis Kelamin |    | Total |
|---------------|--------------------------------|---------------|----|-------|
|               |                                | L             | P  |       |
| 1             | Kelas Eksperimen<br>Kelas X-11 | 12            | 20 | 32    |
| 2             | Kelas Kontrol<br>Kelas X-12    | 14            | 18 | 32    |
| <b>Jumlah</b> |                                |               |    | 64    |

Sumber: Tata Usaha MAN 3 Kabupaten Tasikmalaya Tahun 2025/2026

### 3.5 Teknik Pengumpulan Data

Teknik pengumpulan data merupakan salah satu tahapan yang sangat penting dalam suatu penelitian, karena melalui proses inilah diperoleh informasi yang dibutuhkan untuk menjawab rumusan masalah atau menguji hipotesis. Kegiatan pengumpulan data harus dilakukan secara sistematis dan cermat agar data yang dihasilkan memiliki tingkat validitas dan reliabilitas yang tinggi. Oleh karena itu, pemilihan metode serta instrumen yang tepat menjadi faktor penting yang menentukan kualitas data yang diperoleh (Pugu dkk., 2024: 17). Pada penelitian ini teknik pengumpulan data yang digunakan yaitu:

#### 3.5.1 Tes

Tes merupakan metode pengumpulan data yang menggunakan instrumen berupa serangkaian soal atau tugas yang dirancang untuk mengukur kemampuan, pengetahuan, kecerdasan, kompetensi, maupun potensi yang dimiliki oleh individu

ataupun kelompok (Seplyana, 2025: 115). Dalam penelitian ini, instrumen yang digunakan berupa lembar tes yang terdiri atas soal *pretest* dan *posttest*. Soal *pretest* berfungsi untuk mengukur hasil belajar peserta didik pada kelas eksperimen dan kelas kontrol sebelum diberikan perlakuan tertentu, sedangkan soal *posttest* digunakan untuk menilai hasil belajar kedua kelas setelah perlakuan diberikan. Jenis pertanyaan yang digunakan dalam tes ini berbentuk pertanyaan terbuka, di mana peserta didik diminta memberikan jawaban dalam bentuk uraian. Tujuan penggunaan pertanyaan terbuka adalah untuk menggali pendapat serta ide peserta didik secara lebih mendalam. Melalui jawaban uraian tersebut, peneliti dapat mengamati dan menilai kemampuan berpikir kritis peserta didik. Setiap jawaban yang diberikan kemudian dinilai berdasarkan rubrik penilaian kemampuan berpikir kritis yang telah disusun.

Instrumen penelitian yang telah disusun selanjutnya divalidasi oleh para ahli untuk memastikan kesesuaian dan kelayakannya. Setelah memperoleh persetujuan dari para ahli, instrumen tersebut kemudian diuji cobakan. Tahapan berikutnya setelah uji coba dilakukan adalah pengujian validitas, reliabilitas, tingkat kesukaran, serta daya pembeda instrumen guna memastikan kualitas dan keandalan alat ukur yang digunakan dalam penelitian.

### **3.6 Instrumen Penelitian**

Instrumen yang digunakan dalam penelitian ini berupa tes uraian yang terdiri atas 5 butir soal. Instrumen tersebut disusun dengan tujuan untuk mengukur kemampuan berpikir kritis peserta didik pada kelas X-11 dan kelas X-12. Materi yang diujikan dalam tes ini berkaitan dengan topik Proses masuknya Hindu-Budha ke Indonesia. Setiap butir soal memiliki bobot penilaian yang berbeda tergantung jawaban yang diberikan oleh peserta didik.

Penyusunan instrumen penelitian ini didasarkan pada variabel serta indikator yang telah ditetapkan. Mengacu pada pendapat Ennis, kemampuan berpikir kritis mencakup indikator utama, yaitu *Elementary Clarification*, *Basic Support*, *Inference*, *Advanced Clarification*, *Strategy and Tactics*. Berdasarkan

unsur-unsur tersebut, disusun kisi-kisi tes kemampuan berpikir kritis sebagaimana disajikan berikut ini.

**Tabel 3.4 Kisi-kisi Instrumen Kemampuan Berpikir Kritis**

| No | Indikator Berpikir Kritis (Ennis)                                   | Indikator Soal                                                                              | No Soal |
|----|---------------------------------------------------------------------|---------------------------------------------------------------------------------------------|---------|
| 1  | Memberikan penjelasan sederhana ( <i>Elementary Clarification</i> ) | Menganalisis latar belakang dan bukti pendukung teori Brahmana                              | 1       |
| 2  | Membangun keterampilan dasar ( <i>Basic Support</i> )               | Menganalisis kelebihan dan kelemahan teori Kesatria berdasarkan bukti sejarah               | 2       |
| 3  | Menarik kesimpulan ( <i>Inference</i> )                             | Menyimpulkan keterbatasan beberapa teori masuknya Hindu-Buddha                              | 3       |
| 4  | Memberikan penjelasan lanjut ( <i>Advanced Clarification</i> )      | Menyusun argumentasi historis terkait perdebatan teori Arus Balik                           | 4       |
| 5  | Mengatur strategi dan taktik ( <i>Strategy and Tactics</i> )        | Menentukan dan mengombinasikan teori paling kuat untuk menjelaskan proses awal Hindu-Buddha | 5       |

**Tabel 3.5 Rubrik Penilaian**

| No | Indikator Berpikir Kritis (Ennis)                                   | Kriteria Penilaian                                                                     | Skor |
|----|---------------------------------------------------------------------|----------------------------------------------------------------------------------------|------|
| 1  | Memberikan penjelasan sederhana ( <i>Elementary Clarification</i> ) | Sangat Baik : Peserta didik menjawab dengan benar, lengkap, jelas, dan dapat dipahami. | 4    |
|    |                                                                     | Baik: Peserta didik menjawab dengan benar, kurang lengkap, jelas dan dapat dipahami.   | 3    |

|   |                                                          |                                                                                                   |   |
|---|----------------------------------------------------------|---------------------------------------------------------------------------------------------------|---|
|   |                                                          | Cukup:<br>Peserta didik menjawab dengan benar, kurang lengkap, kurang jelas, dan kurang dipahami. | 2 |
|   |                                                          | Kurang:<br>Peserta didik menjawab kurang sesuai, tidak lengkap, tidak jelas, dan kurang dipahami. | 1 |
| 2 | Membangun keterampilan dasar<br>( <i>Basic Support</i> ) | Sangat Baik :<br>Peserta didik menjawab dengan benar, lengkap, jelas, dan dapat dipahami.         | 4 |
|   |                                                          | Baik:<br>Peserta didik menjawab dengan benar, kurang lengkap, jelas dan dapat dipahami.           | 3 |
|   |                                                          | Cukup:<br>Peserta didik menjawab dengan benar, kurang lengkap, kurang jelas, dan kurang dipahami. | 2 |
|   |                                                          | Kurang:<br>Peserta didik menjawab kurang sesuai, tidak lengkap, tidak jelas, dan kurang dipahami. | 1 |
| 3 | Menarik kesimpulan<br>( <i>Inference</i> )               | Sangat Baik :<br>Peserta didik menjawab dengan benar, lengkap, jelas, dan dapat dipahami.         | 4 |
|   |                                                          | Baik:<br>Peserta didik menjawab dengan benar, kurang lengkap, jelas dan dapat dipahami.           | 3 |
|   |                                                          | Cukup:<br>Peserta didik menjawab dengan benar, kurang lengkap, kurang jelas, dan kurang dipahami. | 2 |
|   |                                                          | Kurang:<br>Peserta didik menjawab kurang sesuai, tidak lengkap, tidak jelas, dan kurang dipahami. | 1 |
| 4 | Memberikan penjelasan lanjut                             | Sangat Baik :<br>Peserta didik menjawab dengan benar, lengkap, jelas, dan dapat dipahami.         | 4 |

|   |                                                              |                                                                                                   |   |
|---|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------|---|
|   | <i>(Advanced Clarification)</i>                              | Baik:<br>Peserta didik menjawab dengan benar, kurang lengkap, jelas dan dapat dipahami.           | 3 |
|   |                                                              | Cukup:<br>Peserta didik menjawab dengan benar, kurang lengkap, kurang jelas, dan kurang dipahami. | 2 |
|   |                                                              | Kurang:<br>Peserta didik menjawab kurang sesuai, tidak lengkap, tidak jelas, dan kurang dipahami. | 1 |
| 5 | Mengatur strategi dan taktik ( <i>Strategy and Tactics</i> ) | Sangat Baik :<br>Peserta didik menjawab dengan benar, lengkap, jelas, dan dapat dipahami.         | 4 |
|   |                                                              | Baik:<br>Peserta didik menjawab dengan benar, kurang lengkap, jelas dan dapat dipahami.           | 3 |
|   |                                                              | Cukup:<br>Peserta didik menjawab dengan benar, kurang lengkap, kurang jelas, dan kurang dipahami. | 2 |
|   |                                                              | Kurang:<br>Peserta didik menjawab kurang sesuai, tidak lengkap, tidak jelas, dan kurang dipahami. | 1 |

Penilaian skor uraian:

Nilai = (Skor Perolehan / Skor Total)  $\times$  100

Kriteria Nilai:

A : 80–100 = Baik Sekali

B : 70–79 = Baik

C : 60–69 = Cukup

D : 0–59 = Kurang

Setelah menentukan instrumen yang akan digunakan untuk mengukur kemampuan berpikir kritis peserta didik, perlu dilakukan uji kelayakan yaitu uji validitas, reliabilitas, tingkat kesukaran, serta daya pembeda instrumen guna memastikan kualitas dan keandalan alat ukur yang digunakan dalam penelitian.



### 1. Uji Validitas Instrumen

Uji validitas merupakan prosedur yang digunakan untuk menilai apakah suatu alat ukur tersebut valid atau tidak valid. Pengujian validitas tidak hanya bertujuan untuk menilai ketepatan suatu tes, tetapi juga digunakan untuk menilai keakuratan instrumen penelitian. Sebuah instrumen dikatakan valid apabila mampu mengungkapkan data yang relevan dengan permasalahan yang diteliti secara tepat dan benar, serta mencerminkan situasi dan kondisi yang sesungguhnya (Saputri dkk., 2023: 2989).

Terdapat dua jenis pengujian terhadap instrumen penelitian, yaitu: (1) uji validitas isi, yang dilakukan dengan meminta pertimbangan dari para ahli untuk menilai kesesuaian antara indikator dan konsep yang diukur, (2) uji validitas konstruk, yang diterapkan pada instrumen berbentuk tes dengan cara membandingkan isi instrumen dengan materi pelajaran yang telah diberikan (Sugiyono, 2017: 177-183). Dalam penelitian ini, analisis validitas isi dilakukan oleh validator ahli. Peneliti memberikan kisi-kisi, butir instrumen, serta rubrik penskoran kepada para validator ahli untuk memperoleh masukan. Masukan yang diharapkan mencakup kesesuaian instrumen dengan indikator berpikir kritis, keselarasan antara indikator dan butir pernyataan, ketepatan isi pernyataan, serta kejelasan kalimat pada setiap butir. Berdasarkan saran dan pertimbangan tersebut, validator ahli menentukan apakah instrumen dapat digunakan tanpa revisi, perlu digunakan setelah dilakukan perbaikan, atau tidak layak digunakan sehingga harus disusun ulang secara keseluruhan. Proses penilaian dilakukan dengan menggunakan skala Likert sesuai dengan kriteria yang disajikan pada tabel berikut:

**Tabel 3.6 Kriteria Penilaian Butir Instrumen Oleh Validator**

| Nilai | Deskripsi    |
|-------|--------------|
| 1     | Tidak Valid  |
| 2     | Kurang Valid |
| 3     | Cukup Valid  |
| 4     | Valid        |
| 5     | Sangat Valid |

(Retnawati dalam Utami dkk., 2024: 63)

Proses uji validitas bisa dilakukan menggunakan perangkat lunak pengolahan data, seperti *IBM SPSS* versi 31 dan Microsoft Excel. Uji validitas instrumen yang digunakan dalam penelitian ini adalah uji V-Aiken, yang didasarkan pada hasil penilaian para ahli sebanyak dua orang. Penilaian tersebut mengacu pada tingkat kesepakatan para ahli terhadap kesesuaian setiap butir instrumen dengan indikator yang hendak diukur. Dalam penelitian ini, yang dimaksud dengan butir adalah soal tes yang digunakan sebagai instrumen penelitian. Pengujian validitas isi menggunakan uji Aiken (1985) dengan rumus sebagai berikut:

$$V = \frac{\sum S}{n(C - 1)}$$

$$S = R - L_0$$

$V$  = indeks Aiken

$S$  = skor yang diberikan oleh rater (penilai) dikurangi dengan skor paling rendah

$R$  = skor yang diberikan oleh rater

$L_0$  = skor penilaian terendah (1)

$C$  = skor penilaian tertinggi (4)

$n$  = jumlah rater

Jika indeks Aiken kurang dari 0,4 maka dikatakan validitasnya rendah, indeks Aiken diantara 0,4-0,8 dikatakan validitasnya sedang dan jika lebih dari 0,8 dikatakan tinggi (Utami dkk., 2024:63).

Analisis data dalam perhitungan validitas tes berpikir kritis oleh para ahli menggunakan bantuan aplikasi Microsoft Excel dengan hasil analisis uji validitas isi sebagai berikut:

**Tabel 3.7 Hasil Uji Validitas Tes Oleh Para Ahli**

| Butir | Penilai |    | S1 | S2 | $\sum S$ | $n(C - 1)$ | V    | Ket    |
|-------|---------|----|----|----|----------|------------|------|--------|
|       | I       | II |    |    |          |            |      |        |
| 1     | 3       | 5  | 2  | 4  | 6        | 4          | 1,5  | Tinggi |
| 2     | 3       | 5  | 2  | 4  | 6        | 4          | 1,5  | Tinggi |
| 3     | 3       | 5  | 2  | 4  | 6        | 4          | 1,5  | Tinggi |
| 4     | 3       | 5  | 2  | 4  | 6        | 4          | 1,5  | Tinggi |
| 5     | 4       | 5  | 3  | 4  | 7        | 4          | 1,75 | Tinggi |

|    |   |   |   |   |   |   |      |        |
|----|---|---|---|---|---|---|------|--------|
| 6  | 4 | 4 | 3 | 3 | 6 | 4 | 1,5  | Tinggi |
| 7  | 4 | 5 | 3 | 4 | 7 | 4 | 1,75 | Tinggi |
| 8  | 3 | 5 | 2 | 4 | 6 | 4 | 1,5  | Tinggi |
| 9  | 4 | 5 | 3 | 4 | 7 | 4 | 1,75 | Tinggi |
| 10 | 3 | 5 | 2 | 4 | 6 | 4 | 1,5  | Tinggi |
| 11 | 5 | 5 | 4 | 4 | 8 | 4 | 2    | Tinggi |
| 12 | 4 | 5 | 3 | 4 | 7 | 4 | 1,75 | Tinggi |

**Tabel 3.8 Kesimpulan Hasil Uji Validitas Tes Oleh Para Ahli**

| Butir      | Penilai |    | S1 | S2 | $\sum s$ | V     | Ket    |
|------------|---------|----|----|----|----------|-------|--------|
|            | I       | II |    |    |          |       |        |
| Butir 1-12 | 43      | 59 | 31 | 47 | 78       | 1,625 | Tinggi |

Berdasarkan tabel hasil uji validitas isi diatas dapat disimpulkan bahwa instrumen tes yang digunakan dalam penelitian ini memiliki tingkat validitas yang tinggi. Oleh karena itu, instrumen tes tersebut layak digunakan sebagai alat ukur untuk menilai kemampuan berpikir kritis peserta didik.

Pengujian validitas konstruk dilakukan dengan uji coba instrumen berpikir kritis peserta didik dilakukan pada kelas yang berada di luar sampel penelitian tetapi masih dalam populasi yang sama, tujuannya untuk menguji validitas instrumen sebelum diberikan ke sampel sebenarnya agar tidak terjadi bias dan memastikan instrumen siap digunakan untuk mengukur apa yang seharusnya diukur. Kelas yang dipilih untuk pelaksanaan uji validitas instrumen adalah kelas X-7, dengan pertimbangan bahwa kelas tersebut telah mempelajari materi yang akan diujikan. Pengujian validitas konstruk dilakukan menggunakan korelasi (*Pearson Correlation*) dengan bantuan program *software IBM SPSS Statistic 31*. Kriteria pengambilan keputusan yang digunakan adalah apabila nilai  $r$  hitung  $> r$  tabel dan signifikansi  $< 0,05$ , maka butir soal dinyatakan valid. Nilai  $r$  tabel yang digunakan dalam penelitian ini adalah 0,349 ( $df = 32 - 2 = 30$ ). Berdasarkan hasil output SPSS, diperoleh data uji validitas sebagai berikut:

**Tabel 3.9 Uji Validitas Tes Berpikir Kritis Peserta Didik**

| Butir Soal | r hitung (Pearson Correlation) | r tabel | Keterangan |
|------------|--------------------------------|---------|------------|
| Soal 1     | 0,728                          | 0,349   | Valid      |
| Soal 2     | 0,728                          | 0,349   | Valid      |
| Soal 3     | 0,396                          | 0,349   | Valid      |
| Soal 4     | 0,522                          | 0,349   | Valid      |
| Soal 5     | 0,507                          | 0,349   | Valid      |

Berdasarkan hasil validitas instrumen tes yang dilakukan peserta didik kelas X-7 di atas dapat ditarik kesimpulan bahwa instrumen penelitian soal 1 sampai soal 5 dinyatakan valid karena memiliki nilai  $r$  hitung  $>$   $r$  tabel (0,349) dan nilai signifikansi  $<$  0,05.

## 2. Uji Reliabilitas Instrumen

Uji reliabilitas digunakan untuk mengetahui tingkat konsistensi suatu alat ukur, yakni apakah instrumen yang digunakan dapat memberikan hasil yang stabil ketika dilakukan pengujian secara berulang. Sebuah alat ukur dikatakan reliabel apabila hasil yang diperoleh dari beberapa kali pengujian menunjukkan nilai yang relatif sama dan konsisten. Hasil pengukuran yang dinilai sama harus memiliki alat serta cara mengukur yang sama dengan peranan penting saat proses penilaian pada waktu yang bersamaan (Forester dkk., 2024: 1814).

Sebuah instrumen dinyatakan reliabel apabila proses pengambilan data yang dilakukan secara berulang mampu menghasilkan nilai dan hasil yang konsisten. Sebaliknya, apabila Instrumen tidak reliabel, maka pengambilan data yang dilakukan berulang kali akan menghasilkan nilai yang berbeda dari pengujian sebelumnya. Dengan demikian, uji reliabilitas pada suatu instrumen penelitian berfungsi untuk menilai sejauh mana instrument yang digunakan dalam proses pengumpulan data mampu memberikan hasil yang stabil dan dapat dipercaya.

Pada uji reliabilitas yang dijelaskan tersebut dapat menggunakan Alpha Cronbach. Dalam uji reliabilitas, suatu instrumen dinyatakan reliabel apabila nilai *Cronbach's Alpha* yang diperoleh lebih dari 0,60. Hal ini menunjukkan bahwa variabel atau alat ukur yang digunakan memiliki tingkat konsistensi yang baik

dalam proses pengukuran. Uji reliabilitas dapat diartikan sebagai suatu prosedur untuk menilai sejauh mana alat ukur memiliki konsistensi dalam proses pengambilan, pengumpulan, dan hasil akhir data. Instrumen dikatakan reliabel apabila alat, bahan, serta data yang diuji menunjukkan tingkat konsistensi yang sama ketika dilakukan pengujian secara berulang tanpa adanya perubahan yang berarti. Uji *Cronbach's Alpha* digunakan untuk mengukur tingkat reliabilitas suatu instrumen yang skornya bukan 1 atau 0 (Forester dkk., 2024: 1815). Pada pengukuran datanya *cronbach's alpha* menggunakan rumus sebagai berikut:

$$r_{11} = \frac{k}{(k-1)} \left[ 1 - \frac{\sum \sigma_b^2}{\sigma_t^2} \right]$$

$r_{11}$  = koefisiensi reliabilitas instrument ( total tes)

$k$  = jumlah butir pertanyaan yang sah

$\sum \sigma_b^2$  = jumlah varian butir

$\sigma_t^2$  = varian skor total

Rentang nilai Cronbach's Alpha digunakan untuk menentukan tingkat reliabilitas suatu instrumen penelitian. Nilai alpha < 0,50 menunjukkan reliabilitas rendah, nilai antara 0,50-0,70 menunjukkan reliabilitas sedang atau moderat, nilai > 0,70 menunjukkan reliabilitas yang memadai (*sufficient reliability*), nilai > 0,80 menunjukkan reliabilitas yang kuat, sedangkan nilai > 0,90 menunjukkan reliabilitas yang sangat tinggi atau sempurna. Semakin kecil nilai alpha, semakin banyak item yang tidak konsisten atau tidak reliabel. Suatu instrumen dapat dikatakan reliabel apabila nilai *Cronbach's Alpha* lebih dari 0,60. Dengan demikian, kriteria pengambilan keputusan dalam uji reliabilitas apabila nilai *Cronbach's Alpha* > 0,60, maka item pertanyaan dalam kuesioner dinyatakan dapat diandalkan (*reliable*), sedangkan apabila *Cronbach's Alpha* < 0,60, maka item pertanyaan tersebut dianggap tidak dapat diandalkan (*not reliable*) (Forester dkk., 2024: 1817).

Hasil analisis data dan pengujian reliabilitas dengan berbantuan *software IBM SPSS Statistic* versi 31, adapun hasil dari *Reliability Statistics* bisa dilihat dari nilai *Cronbach's Alpha*:

**Tabel 3.10 Nilai Reliability Statistics**

| <i>Cronbach's Alpha</i> | N of Items |
|-------------------------|------------|
| 0,895                   | 7          |

**Tabel 3.11 Reliability Statistics per Item**

| Butir Soal | Cronbach's Alpha if Item Deleted | Keterangan |
|------------|----------------------------------|------------|
| Soal 1     | 0,878                            | Reliabel   |
| Soal 2     | 0,874                            | Reliabel   |
| Soal 3     | 0,883                            | Reliabel   |
| Soal 4     | 0,887                            | Reliabel   |
| Soal 5     | 0,890                            | Reliabel   |

Data hasil dari uji Reliabilitas Statistik di atas menunjukkan seluruh butir tes instrumen berpikir kritis dinyatakan reliabel karena nilai *Cronbach's Alpha if Item Deleted* lebih rendah dari *Cronbach's Alpha* keseluruhan (0,895). Instrumen secara keseluruhan reliabel  $\alpha = 0,895 > 0,80$  dengan pengambilan keputusan dari nilai *Cronbach's Alpha*  $> 0,60$  dinyatakan dapat diandalkan (*reliable*).

### 3. Tingkat Kesukaran

Tingkat kesukaran butir soal merupakan perbandingan antara jumlah peserta tes yang menjawab benar pada suatu butir soal dengan jumlah seluruh peserta tes. Ukuran ini dinyatakan dalam bentuk Indeks Kesukaran Soal (IKS) dengan rentang nilai antara 0,00 hingga 1,00. Semakin tinggi nilai indeks kesukaran yang diperoleh dari hasil perhitungan, menunjukkan bahwa butir soal tersebut semakin mudah. Perhitungan indeks kesukaran dilakukan untuk setiap butir soal secara terpisah. Secara prinsip, nilai rata-rata skor yang diperoleh peserta pada suatu butir soal mencerminkan tingkat kesukaran dari butir soal tersebut (Munip dalam Dianova & Anwar, 2024: 6). Berikut adalah rumus yang digunakan:

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

Keterangan:

$P$  = Indeks kesukaran

$B$  = Jumlah jawaban benar

$JS$  = Jumlah seluruh peserta tes

Soal yang dikatakan baik ketika soal yang diujikan tidak terlalu mudah tidak terlalu sukar. Dalam menguji tingkat kesukaran dalam penelitian ini menggunakan bantuan *software IBM SPSS Statistic* versi 31, dengan kriteria pengambilan keputusan sebagai berikut:

**Tabel 3.12 Kriteria Pengambilan Keputusan Tingkat Kesukaran**

| Rentang Nilai IKS | Kategori |
|-------------------|----------|
| 0,00-0,30         | Sukar    |
| 0,31-0,70         | Sedang   |
| 0,71-1,00         | Mudah    |

( Qadir dkk., 2024: 1455)

**Tabel 3.13 Tingkat Kesukaran Soal Tes Berpikir Kritis**

| Butir Soal | Mean (Output SPSS) | Tingkat Kesulitan |
|------------|--------------------|-------------------|
| Soal 1     | 3,06               | Mudah             |
| Soal 2     | 2,81               | Mudah             |
| Soal 3     | 2,38               | Sedang            |
| Soal 4     | 2,16               | Sedang            |
| Soal 5     | 1,81               | Sedang            |

Berdasarkan data Tingkat Kesukaran di atas, soal yang diujikan tergolong mudah dan sedang, Rumus  $P = (\text{Mean} / \text{Skor Maksimum (4)}) \times 1,00$ . Berdasarkan kriteria pengambilan keputusan di atas, 2 butir soal dengan kategori mudah dan 5 butir soal dengan kategori sedang.

#### 4. Daya Pembeda

Analisis daya pembeda merupakan suatu metode untuk menilai sejauh mana suatu butir soal mampu membedakan antara peserta didik yang berkemampuan tinggi (kelompok atas) dan peserta didik yang berkemampuan rendah (kelompok bawah). Kemampuan pembeda ini dapat diketahui melalui nilai Indeks Daya Beda Butir (IDB). Nilai indeks tersebut umumnya dinyatakan dalam bentuk proporsi. Semakin tinggi nilai indeks daya pembeda, semakin baik kemampuan butir soal tersebut dalam membedakan peserta didik yang memiliki kemampuan tinggi

dengan yang berkemampuan rendah (Munip dalam Dianova & Anwar, 2024: 6).

Berikut adalah rumus yang digunakan dalam menganalisis daya pembeda:

$$DB = P_T - P_R$$

atau

$$DB = \frac{\sum T_B}{\sum T} - \frac{\sum R_B}{\sum R}$$

Keterangan:

$P_T$  = proporsi peserta didik pada kelompok berkemampuan tinggi yang menjawab benar suatu butir soal.

$P_R$  = proporsi peserta didik pada kelompok berkemampuan rendah yang menjawab benar butir soal.

$\sum T_B$  = jumlah peserta dari kelompok berkemampuan tinggi yang menjawab benar.

$\sum T$  = jumlah keseluruhan peserta dalam kelompok berkemampuan tinggi.

$\sum R_B$  = jumlah peserta dari kelompok berkemampuan rendah yang menjawab benar.

$\sum R$  = jumlah keseluruhan peserta dalam kelompok berkemampuan rendah.

Dalam pengujian Daya Pembeda dilakukan dengan bantuan *software IBM SPSS Statistic* versi 31. Dengan kriteria pengambilan keputusan sebagai berikut:

**Tabel 3.14 kriteria Pengambilan Keputusan Daya Pembeda**

| Koefisien Korelasi | Kategori Daya Beda                    |
|--------------------|---------------------------------------|
| 0,40-1,00          | Soal diterima/baik                    |
| 0,30-0,39          | Soal diterima tetapi perlu diperbaiki |
| 0,20-0,29          | Soal diperbaiki                       |
| 0,00-0,19          | Soal tidak dipakai/dibuang            |

( Naga dalam Qadir dkk., 2024: 1461)

Hasil data uji daya pembeda untuk instrumen berpikir kritis peserta didik dengan menggunakan bantuan *software IBM SPSS Statistic* versi 31 sebagai berikut:



**Tabel 3.15 Uji Daya Pembeda**

| Butir Soal | r hitung (Corrected Item-Total Correlation) | Kategori           |
|------------|---------------------------------------------|--------------------|
| Soal 1     | 0,708                                       | Soal diterima/baik |
| Soal 2     | 0,749                                       | Soal diterima/baik |
| Soal 3     | 0,672                                       | Soal diterima/baik |
| Soal 4     | 0,635                                       | Soal diterima/baik |
| Soal 5     | 0,599                                       | Soal diterima/baik |

Berdasarkan hasil analisis daya pembeda menggunakan *Corrected Item-Total Correlation* dengan kriteria keputusan daya pembeda memiliki kategori baik dengan nilai  $r > 0,40$ . Dengan demikian, seluruh item dinyatakan diterima dan memiliki kemampuan yang baik dalam membedakan peserta didik berkemampuan tinggi dan rendah.

### 3.7 Teknik Analisis Data

Teknik analisis data dalam penelitian kuantitatif dilakukan dengan menggunakan metode statistik. Secara umum, terdapat dua jenis statistik yang digunakan dalam analisis penelitian, yaitu statistik deskriptif dan statistik inferensial. Statistik deskriptif berfungsi untuk mendeskripsikan atau menggambarkan data yang diperoleh dari sampel tanpa bermaksud menarik kesimpulan yang berlaku bagi populasi secara umum. Sebaliknya, statistik inferensial digunakan untuk menganalisis data sampel dengan tujuan agar hasil analisis dapat digeneralisasikan terhadap populasi (Sugiyono, 2017: 207-209).

Statistik inferensial dibedakan menjadi dua jenis, yakni statistik parametrik dan statistik nonparametrik. Perbedaan antara keduanya terletak pada jenis data yang dianalisis. Apabila data yang digunakan berskala nominal atau ordinal (data kualitatif), maka analisis dilakukan dengan menggunakan statistik nonparametrik. Sementara itu, apabila data yang diperoleh berupa data interval atau rasio (data kuantitatif), maka analisis dilakukan dengan teknik statistik parametrik. Pemilihan antara statistik parametrik dan nonparametrik bergantung pada jenis data serta asumsi yang menyertainya. Dalam statistik parametrik, terdapat asumsi utama yang harus dipenuhi, yaitu data harus berdistribusi normal dan memiliki varians yang

homogen (Sugiyono, 2017: 210-211). Oleh karena itu, diperlukan uji normalitas dan uji homogenitas sebelum dilakukan analisis lebih lanjut. Apabila hasil pengujian menunjukkan bahwa data tidak berdistribusi normal dan tidak homogen, maka teknik analisis yang tepat adalah statistik nonparametrik, misalnya dengan menggunakan uji *Mann-Whitney*. Hal ini disebabkan karena statistik nonparametrik tidak memerlukan pemenuhan asumsi distribusi normal dan homogenitas data. Adapun teknik analisis data pada penelitian ini yaitu:

### 3.7.1 Uji Normalitas

Uji normalitas merupakan prosedur analisis statistik yang bertujuan untuk mengetahui apakah data yang diperoleh dari suatu populasi memiliki distribusi yang normal. Distribusi normal merupakan salah satu asumsi dasar yang harus dipenuhi dalam penerapan berbagai metode analisis statistik parametrik, seperti *uji-t*. Distribusi normal memiliki bentuk simetris menyerupai lonceng (*bell-shaped curve*), di mana sebagian besar data berkumpul di sekitar nilai rata-rata (mean), dan data secara bertahap menurun ke arah kiri dan kanan. Pelaksanaan uji normalitas memiliki beberapa tujuan utama yaitu, menentukan apakah data memenuhi asumsi untuk dilakukan analisis parametrik, menetapkan jenis metode statistik yang tepat digunakan, baik parametrik maupun non-parametrik dan menjamin keabsahan (validitas) dan ketepatan (akurasi) hasil analisis data yang diperoleh (Arifin, 2025: 177-178).

Dalam penelitian ini, peneliti menggunakan metode *Shapiro-Wilk* karena jumlah sampel yang digunakan kurang dari 50 responden. Uji *Shapiro-Wilk* merupakan metode pengujian normalitas yang memanfaatkan data mentah yang belum diolah atau dikelompokkan dalam tabel distribusi frekuensi, yang kemudian dikonversi ke dalam perhitungan *Shapiro-Wilk*. Metode ini juga dikenal sebagai W-Test dan dinilai efektif serta banyak digunakan dalam penelitian ilmiah (Ramadhani & Bina, 2021: 196).

Adapun beberapa persyaratan yang perlu dipenuhi dalam penerapan metode *Shapiro-Wilk* yaitu, data berskala interval atau rasio (kuantitatif), data masih berupa

data mentah, belum disusun dalam tabel distribusi frekuensi serta sampel penelitian diambil secara acak (random sampling).

Rumus uji normalitas data menggunakan metode *Shapiro-Wilk* (W-Test) sebagai berikut:

$$T_3 = \frac{1}{D} \left[ \sum_{i=1}^k a_1 (X_{n-i+1} - X_i) \right]^2$$

Di mana:

$D$  = berdasarkan rumus dibawah

$a_1$  = koefisien tes Shapiro-Whilk

$X_{n-i+1}$  = angka ke pada data

$X_i$  = Angka ke- $i$  pada data

$$D = \sum_{i=1}^n (X_i - \bar{X})^2$$

Di mana:

$X_i$  = angka ke- $i$  pada data

$\bar{X}$  = rata-rata data

$$G = b_n + c_n + \ln \left( \frac{T_3 - d_n}{1 - T_3} \right)$$

Di mana:

$G$  = identik dengan nilai  $z$  distribusi normal

$T_3$  = konversi statistik Shapiro Whilk pendekatan distribusi normal.

Interpretasi hasil uji dilakukan dengan membandingkan nilai signifikansi ( $p$ ) terhadap batas signifikansi 5% ( $\alpha = 0,05$ ), yaitu:

- Jika nilai  $p > 0,05$ , maka  $H_0$  (hipotesis nol) diterima, dan  $H_a$  (hipotesis alternatif) ditolak, artinya data berdistribusi normal.
- Jika nilai  $p < 0,05$ , maka  $H_0$  (hipotesis nol) ditolak dan  $H_a$  (hipotesis alternatif) diterima, yang berarti data tidak berdistribusi normal.

### 3.7.2 Uji Homogenitas

Uji homogenitas bertujuan untuk mengetahui apakah varians dari beberapa populasi memiliki kesamaan atau tidak. Pengujian ini merupakan salah satu prasyarat analisis statistik parametrik, seperti uji *Independent Sample t-test*. Dalam

analisis varians, terdapat asumsi bahwa varians antar populasi yang dibandingkan harus sama atau homogen. Uji kesamaan dua varians dilakukan untuk menguji apakah data yang digunakan bersifat homogen, dengan cara membandingkan besarnya varians antara dua kelompok atau lebih. Apabila varians kedua kelompok data tersebut relatif sama, maka data dianggap homogen dan uji homogenitas tidak perlu dilakukan kembali. Uji homogenitas hanya dapat dilakukan apabila data berasal dari populasi yang berdistribusi normal. Tujuan utama dari uji ini adalah memastikan bahwa perbedaan yang muncul dalam pengujian statistik parametrik seperti uji t, benar-benar disebabkan oleh adanya perbedaan antar kelompok, bukan karena adanya perbedaan varians dalam kelompok itu sendiri. Sebelum melakukan perbandingan antara dua kelompok atau lebih, pengujian homogenitas variansi sangat penting dilakukan agar perbedaan yang ditemukan tidak disebabkan oleh ketidakhomogenan data dasar (Usmadi, 2020: 51). Beberapa metode yang dapat digunakan untuk menguji homogenitas variansi antara lain uji *Hartley*, uji *Cochran*, uji *Levene*, dan uji *Bartlett*.

Menurut Usmadi (2020: 54), Uji *Levene* (Levene 1960) merupakan metode yang digunakan untuk menguji kesamaan varians dari beberapa populasi. Peneliti menggunakan uji Levene dengan rumus sebagai berikut:

$$W = \frac{(n - k) \sum_{i=1}^k n_i (\bar{Z}_i - \bar{Z}_{..})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2}$$

Dimana:

$n$  = jumlah perlakuan

$k$  = banyak kelompok

$Z_{ij} = Y_{ij} - Y_{i.}$

$Y_{i.}$  = rata rata dari kelompok ke  $-i$

$Z_i$  = rata rata kelompok dari  $Z_i$

$Z_{..}$  = rata rata menyeluruh dari  $Z_{ij}$

Jika  $W > F(\alpha; k - 1, n - k)$  Artinya, jika nilai hitung  $W$  lebih besar dari nilai kritis distribusi  $F$  pada taraf signifikansi tertentu  $\alpha = 0,05$  maka varians antar kelompok dianggap tidak homogen.

Keputusan uji *Levene* ditentukan berdasarkan nilai signifikansi (Sig.) dengan kriteria:

- a. Jika  $\text{Sig.} \geq 0,05$ , maka data yang dimiliki homogen secara univariat ( $H_0$  diterima).
- b. Jika  $\text{Sig.} < 0,05$ , maka data yang dimiliki tidak homogen secara univariat ( $H_0$  ditolak).

### 3.7.3 Uji Hipotesis

Setelah dilakukan uji normalitas dan uji homogenitas, tahap selanjutnya adalah pengujian hipotesis. Langkah awal dalam proses pengujian hipotesis yaitu menetapkan terlebih dahulu hipotesis statistik yang akan diuji.

- a. Hipotesis nol ( $H_0$ ) merupakan hipotesis yang menyatakan bahwa tidak terdapat pengaruh antara variabel independen (X) dan variabel dependen (Y).
- b. Hipotesis alternatif ( $H_a$ ) merupakan hipotesis yang menyatakan bahwa terdapat pengaruh antara variabel independen (X) dan variabel dependen (Y).

Salah satu metode yang digunakan dalam pengujian hipotesis adalah uji t (*sample t-test*). Uji t pada penelitian quasi-eksperimen digunakan untuk mengetahui apakah suatu pendekatan, model, metode, atau strategi pembelajaran yang diterapkan oleh guru berpengaruh secara signifikan atau tidak terhadap suatu aspek atau variabel tertentu (Isnawan, 2020: 26).

Uji t (*T-Test*) merupakan metode statistik yang digunakan untuk menguji perbedaan rata-rata antara dua kelompok sampel, baik yang berasal dari subjek yang sama maupun dari subjek yang berbeda setelah diberikan perlakuan tertentu. Dalam perangkat lunak SPSS, uji t terdiri atas dua jenis, yaitu *Independent Samples t-Test* dan *Paired Samples t-Test*. Uji *Independent Samples t-Test* diterapkan ketika peneliti ingin membandingkan rata-rata antara dua kelompok yang berbeda, sedangkan *Paired Samples t-Test* digunakan pada kelompok yang sama tetapi dalam dua kondisi atau waktu yang berbeda, misalnya sebelum dan sesudah

perlakuan, atau untuk menilai tanggapan yang berbeda dari kelompok yang sama terhadap pertanyaan yang berbeda (Pallant dalam Akbar dkk., 2023: 433).

Menurut Kim (2019: 1-4), Uji hipotesis dalam penelitian ini dilakukan dengan menggunakan uji t dua sampel independen (*Independent Samples t-Test*). Uji ini bertujuan untuk mengetahui apakah terdapat perbedaan rata-rata antara dua kelompok yang tidak saling berhubungan, yaitu antara kelas eksperimen dan kelas kontrol. Sebelum dilakukan pengujian, terlebih dahulu data diuji normalitas dan homogenitasnya untuk memastikan terpenuhinya asumsi uji parametrik. Apabila data berdistribusi normal dan homogen, maka digunakan *Student's t-test (Equal Variances Assumed)*. Namun, apabila data berdistribusi normal tetapi tidak homogen, maka digunakan *Welch's t-test (Equal Variances Not Assumed)*.

a. *Student's t-test*

Dalam *Student's t-test*, varians populasi diasumsikan sama. Oleh karena itu, hanya membutuhkan satu estimasi varians umum untuk 2 kelompok. Estimasi varians umum dihitung sebagai varians gabungan, rata-rata tertimbang dari 2 varians sampel sebagai berikut:

$$s_p^2 = \frac{(n_1 - 1)}{(n_1 - 1) + (n_2 - 1)} s_1^2 + \frac{(n_2 - 1)}{(n_1 - 1) + (n_2 - 1)} s_2^2$$

Dimana  $s_p^2$  dan  $s_2^2$  adalah varians sampel.

*Statistik uji- t* yang dihasilkan adalah suatu bentuk yang menunjukkan varians populasi,  $\sigma_1^2$  dan  $\sigma_2^2$ , dipertukarkan dengan estimasi varians umum,  $s_p^2$  nilai df diberikan sebagai  $n_1 + n_2 - 2$  untuk statistik uji- t .

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{\bar{X}_1 - \bar{X}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)}$$

b. *Welch's t-test*

Sebenarnya, terdapat banyak kasus di mana varians yang sama tidak dapat diasumsikan. Meskipun kecil kemungkinannya untuk mengasumsikan varians yang sama, tetapi tetap membandingkan rata-rata dua kelompok independen dengan melakukan uji-*t welch*. Uji-*t welch* lebih

reliabel ketika kedua sampel memiliki varians yang tidak sama atau ukuran sampel yang tidak sama. Jadi perlu mempertahankan asumsi normalitas. Karena varians populasi tidak sama, maka harus memperkirakannya secara terpisah dengan 2 varians sampel  $s_1^2$  dan  $s_2^2$  Sebagai hasilnya, bentuk statistik uji- $t$  diberikan sebagai berikut:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim tv$$

Di mana  $v$  adalah derajat kebebasan Satterthwaite.

$$v = \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{\frac{\frac{s_1^2}{n_1}}{(n_1 - 1)} + \frac{\frac{s_2^2}{n_2}}{(n_2 - 1)}}$$

Dari hasil perhitungan menggunakan rumus  $t$ -test, akan diperoleh nilai  $t$  hitung yang kemudian dibandingkan dengan  $t$  tabel pada taraf signifikansi 0,05 (5%). Kriteria pengambilan keputusan terhadap hipotesis adalah sebagai berikut:

- a. Jika nilai  $t$  hitung  $< t$  tabel atau nilai Sig. (2-tailed)  $\geq 0,05$ , maka  $H_0$  diterima dan  $H_a$  ditolak, yang berarti tidak terdapat pengaruh atau perbedaan signifikan antara variabel yang diuji.
- b. Jika nilai  $t$  hitung  $> t$  tabel atau nilai Sig. (2-tailed)  $< 0,05$ , maka  $H_0$  ditolak dan  $H_a$  diterima, yang berarti terdapat pengaruh atau perbedaan signifikan antara variabel yang diuji.

#### 3.7.4 Uji *N-Gain*

*N-Gain* (*Normalized Gain*) merupakan metode analisis yang digunakan untuk menilai efektivitas suatu proses pembelajaran dengan cara membandingkan peningkatan skor antara hasil pretest (sebelum pembelajaran) dan posttest (setelah pembelajaran). Melalui perhitungan ini, peneliti dapat mengetahui tingkat peningkatan pemahaman peserta didik terhadap materi yang dipelajari. Metode *N-Gain* memiliki beberapa keunggulan. Pertama, prosedurnya sederhana sehingga mudah diterapkan dalam berbagai penelitian pembelajaran. Kedua, *N-Gain*

menghasilkan nilai baku yang memungkinkan perbandingan antarkelompok atau antarmetode pembelajaran secara lebih objektif. Selain itu, metode ini mampu mengatasi keterbatasan pengukuran yang hanya menggunakan selisih skor mentah, yang sering kali kurang menggambarkan efektivitas pembelajaran secara menyeluruh (Coletta & Steinert, 2020: 1).

Menurut Nissen dkk., (2018: 1-3) Pada era modern, *N-Gain* banyak dimanfaatkan dalam berbagai bidang pendidikan. Dalam penelitian, metode ini digunakan untuk menilai efektivitas model pembelajaran, media, maupun teknologi pendidikan. Di tingkat kurikulum, *N-Gain* berperan dalam mengevaluasi keberhasilan penerapan kurikulum baru. Selain itu, metode ini juga digunakan dalam pengembangan instrumen asesmen untuk menilai kualitas alat ukur atau program pelatihan. Dengan kemampuannya memberikan informasi yang berbasis data, *N-Gain* tetap relevan sebagai alat evaluasi dalam proses pembelajaran. Fungsi lainnya adalah untuk menilai tingkat pemahaman siswa, yakni melihat sejauh mana peningkatan pemahaman setelah memperoleh intervensi pembelajaran.

Adapun rumus *N-Gain* dapat dituliskan sebagai berikut:

$$N - Gain = \frac{\text{Nilai Posttest} - \text{Nilai Pretest}}{\text{Nilai Maksimum} - \text{Nilai Pretest}}$$

Kategori umum *N-Gain* yang digunakan berasal dari penelitian Richard R. Hake (1998), diklasifikasikan ke dalam tiga kategori:

- a. Kategori tinggi jika  $N\text{-Gain} > 70$
- b. Kategori sedang jika  $30 \leq N\text{-Gain} \leq 70$
- c. Kategori rendah jika  $N\text{-Gain} < 30$

Penentuan kategori ini diperkenalkan dalam artikel Hake yang berjudul "Keterlibatan interaktif vs metode tradisional: Survei data uji mekanika enam ribu siswa untuk kursus pengantar fisika" (1998). Dalam penelitian tersebut, *N-Gain* digunakan untuk membandingkan efektivitas berbagai pendekatan pembelajaran, khususnya dalam bidang fisika. Hingga saat ini, klasifikasi tersebut tetap banyak digunakan karena sederhana serta memberikan interpretasi yang jelas mengenai tingkat efektivitas suatu proses pembelajaran.



### 3.8 Langkah-langkah Penelitian

Langkah-langkah dalam penelitian adalah serangkaian proses yang dimulai ketika seorang peneliti mengidentifikasi sebuah masalah. Setelah itu, peneliti akan berupaya menyelesaikan masalah tersebut, dan di akhir proses, ia akan membuat kesimpulan untuk menentukan apakah hasil penelitiannya berhasil memecahkan masalah yang ada atau tidak (Simamarta dkk., 2023:5). Adapun langkah langkah dalam penelitian ini yaitu:

1. Tahap Persiapan
  - a. Obsevasi dan perumusan masalah, diawali dengan kegiatan observasi terhadap fenomena yang terjadi di lapangan untuk mengidentifikasi dan merumuskan permasalahan yang akan menjadi fokus penelitian.
  - b. Menentukan landasan teori dan merumuskan hipotesis, setelah permasalahan dirumuskan, peneliti menentukan landasan teori yang relevan sebagai dasar konseptual penelitian serta merumuskan hipotesis yang akan diuji.
  - c. Menentukan prosedur penelitian, peneliti kemudian menyusun prosedur penelitian yang akan dilaksanakan sesuai dengan metode eksperimen yang dipilih.
  - d. Menyiapkan rencana penelitian, pada tahap ini, peneliti menyusun rencana penelitian yang mencakup metode penelitian, desain, variabel, tata cara pelaksanaan, serta instrumen yang akan digunakan dalam proses penelitian.
  - e. Pelaksanaan awal penelitian, setelah rencana penelitian tersusun secara sistematis, peneliti mulai melaksanakan penelitian sesuai dengan rancangan yang telah dibuat.
2. Tahap Pelaksanaan
  - a. Uji instrumen penelitian, pada tahap ini dilakukan pengujian instrumen penelitian untuk memastikan setiap butir soal atau instrumen yang digunakan memiliki validitas dan reliabilitas yang memadai.

- b. Pelaksanaan *pretest*, dilakukan baik pada kelas eksperimen maupun kelas kontrol untuk mengetahui kemampuan awal peserta didik sebelum diberikan perlakuan.
  - c. Pemberian perlakuan (*treatment*), peneliti memberikan perlakuan berupa penerapan metode Numbered Head Together (NHT) pada kelas eksperimen, sementara kelas kontrol tidak diberikan perlakuan tersebut.
  - d. Observasi pelaksanaan perlakuan, selama proses perlakuan berlangsung, peneliti melakukan pengamatan terhadap pelaksanaan metode Numbered Head Together (NHT) untuk memastikan prosedur berjalan sesuai rencana.
  - e. Pelaksanaan *posttest* setelah perlakuan diberikan, peneliti melaksanakan *posttest* pada kelas eksperimen dan kelas kontrol untuk mengetahui perubahan nilai peserta didik setelah perlakuan.
3. Tahap Pelaporan
- a. Pengolahan dan analisis data setelah seluruh rangkaian penelitian di lapangan selesai dan data terkumpul, peneliti melakukan pengolahan serta analisis data guna memperoleh kesimpulan dari hasil eksperimen yang telah dilaksanakan.
  - b. Pembahasan hasil penelitian data yang telah dianalisis kemudian dibahas secara mendalam untuk memberikan penjelasan dan interpretasi terhadap hasil penelitian serta menjawab rumusan masalah dan hipotesis.
  - c. Penyusunan laporan penelitian tahap akhir penelitian adalah penyusunan laporan hasil penelitian yang memuat seluruh proses, temuan, serta kesimpulan yang diperoleh dari pelaksanaan penelitian.

### 3.9 Waktu dan Tempat Penelitian

#### 3.9.1 Waktu Penelitian

**Tabel 3.16 Waktu Penelitian**

| No | Kegiatan                                                                                                                                 | Bulan |     |     |     |     |     |
|----|------------------------------------------------------------------------------------------------------------------------------------------|-------|-----|-----|-----|-----|-----|
|    |                                                                                                                                          | Sep   | Okt | Nov | Des | Jan | Feb |
| 1  | Tahap Perencanaan:<br>a. Menyusun dan mengajukan judul<br>b. Mengajukan proposal penelitian<br>c. Membuat dan memperoleh izin penelitian |       |     |     |     |     |     |
| 2  | Tahap pelaksanaan:<br>a. Mengumpulkan data penelitian<br>b. Mengolah data penelitian                                                     |       |     |     |     |     |     |
| 3  | Tahap akhir:<br>a. Menyusun laporan akhir                                                                                                |       |     |     |     |     |     |

#### 3.9.2 Tempat Penelitian

Tempat penelitian ini akan dilaksanakan di MAN 3 Tasikmalaya. Jl. Raya Panumbangan No.33 Desa Pakemitan Kec. Ciawi Kab. Tasikmalaya Prov. Jawa Barat. Sekolah ini dipilih sebagai lokasi karena representatif untuk mengumpulkan data yang sesuai dengan populasi dan sampel yang telah ditetapkan.