

BAB 3

METODOLOGI PENELITIAN

3.1 Metode Penelitian

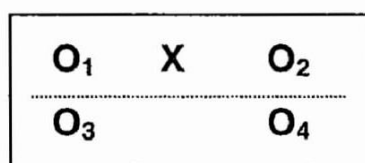
Metode penelitian adalah seperangkat prosedur terstruktur yang digunakan dalam pelaksanaan suatu kajian ilmiah. Dalam penelitian ini digunakan pendekatan kuantitatif sebagai dasar analisisnya. Menurut Danuri & Maisaroh (2019: 207) penelitian kuantitatif merupakan bentuk penelitian yang bersifat deduktif, karena berangkat dari teori yang sudah ada dan ditujukan untuk menguji kebenaran teori tersebut. Pendekatan peneliti ini berlandaskan paradigma positivisme, yang berfokus pada pengkajian populasi atau sampel tertentu. Sampel penelitian dipilih menggunakan Teknik pengambilan sampel yang didasarkan pada pertimbangan serta kriteria khusus yang ditetapkan peneliti guna memperoleh data yang relevan dengan tujuan penelitian. Pengumpulan data dilaksanakan dengan menggunakan instrumen yang telah terstandarisasi, sedangkan analisis data dilakukan secara kuantitatif melalui pengujian statistic untuk menguji kebenaran hipotesis yang telah dirumuskan sebelumnya.

Melalui pendekatan kuantitatif, proses pengolahan data pada setiap tahapan penelitian direpresentasikan dalam bentuk data numerik, mulai dari proses pengumpulan hingga interpretasi data dan pada penarikan kesimpulan serta penyusunan laporan. Penelitian ini menggunakan metode kuasi eksperimen dengan desain *Nonequivalent Control Group Design*. Dalam penelitian ini, metode kuasi eksperimen digunakan untuk mengkaji dan membuktikan adanya pengaruh perlakuan tertentu terhadap variabel lain melalui perbandingan hasil yang diperoleh selama proses penelitian untuk mencari hasil mana yang lebih baik dari penggunaan media Quizalize terhadap kemampuan berpikir sejarah pada siswa di kelas XI SMA Negeri 8 Tasikmalaya.

3.2 Desain Penelitian

Desain penelitian merupakan kerangka konseptual yang disusun secara sistematis pedoman dalam proses pengumpulan data yang dibutuhkan sebagai dasar pengolahan informasi guna menjawab pertanyaan penelitian (Abdullah, 2015: 19). Penelitian ini menggunakan desain kuasi eksperimen sebagai rancangan penelitian

yang digunakan dalam pelaksanaan penelitian. Pemilihan desain kuasi eksperimen dipilih karena sampelnya sudah terbentuk secara organik (*convenient*), serupa dengan kelas yang disusun oleh sekolah, sehingga tidak memungkinkan digunakannya metode pemilihan acak untuk pemilihan peserta (*non-randomly assignment*). Peneliti menggunakan desain kuasi eksperimen dengan rancangan *pretest-posttest with non-equivalent control group*. Gambar berikut memberikan rincian tentang metodologi penelitian:



Gambar 3.1 Desain Kuasi Eksperimen

Sumber: Sugiyono (2019: 120)

Keterangan:

- O₁ : *Pretest* pada kelas eksperimen
- O₃ : *Pretest* pada kelas kontrol
- X : Perlakuan (Media Quizalize)
- O₂ : *Posttest* pada kelas eksperimen
- O₄ : *Posttest* pada kelas kontrol

Skema penelitian pada gambar tersebut menunjukkan bahwa kelompok kontrol tidak diberikan perlakuan khusus selama proses penelitian, sedangkan kelompok eksperimen memperoleh perlakuan (X) berupa penerapan media Quizalize dalam kegiatan pembelajaran. Kedua kelompok penelitian menjalani *pretest* dan *posttest* sebagai instrumen pengukuran variabel dependen, seperti kemampuan berpikir Sejarah, guna mengetahui perubahan yang terjadi sebelum dan sesudah pemberian perlakuan.

3.3 Variabel Penelitian

Variabel penelitian pada dasarnya, merupakan unsur yang dapat diukur dan diamati, memiliki karakteristik tertentu, serta menunjukkan variasi. Variabel ini dimanfaatkan untuk mengetahui adanya hubungan atau pengaruh antar variabel. Melalui penentuan variabel, peneliti dapat mengumpulkan data yang selanjutnya dianalisis guna menjawab rumusan masalah yang telah ditetapkan. Sejalan dengan

Sugiyono (2019: 38), bahwasannya variabel penelitian dapat diartikan sebagai objek atau karakteristik yang menjadi fokus penelitian, yang dapat diamati, diukur, dan memiliki kemungkinan mengalami perubahan atau variasi nilai.

Penelitian ini melibatkan variabel independent dan variabel dependen. Variabel dependen merupakan variabel yang berfungsi sebagai faktor penyebab atau pemberi pengaruh terhadap variabel dependen, baik secara positif maupun negatif. Di sisi lain variabel terikat disebut juga variabel dependen, merupakan fokus utama yang menjadi perhatian peneliti dan menerima pengaruh dari variabel bebas. Dalam penelitian ini, variabel bebas (X) yang digunakan adalah media Quizalize, sedangkan variabel terikat (Y) yang dianalisis adalah kemampuan berpikir sejarah, siswa kelas XI 7 SMAN 8 Tasikmalaya.

3.4 Populasi dan Sampel

3.4.1 Populasi

Sugiyono (2019: 126) mengemukakan populasi diartikan sebagai seluruh elemen yang menjadi sasaran penelitian dan menjadi dasar dalam melakukan generalisasi hasil penelitian. Elemen populasi terdiri atas seluruh subjek yang diukur sebagai unit penelitian. Populasi dalam penelitian ini adalah seluruh kelas XI SMAN 8 Tasikmalaya tahun ajaran 2025/2026. Sebanyak 4 kelas dengan jumlah peserta 139 orang.

Tabel 3.1 Populasi Siswa Kelas XI SMA Negeri 8 Tasikmalaya

No	Kelas	Jumlah Siswa
1.	XI 3	35 Orang
2.	XI 4	35 Orang
3.	XI 6	34 Orang
4.	XI 7	33 Orang
Jumlah		139

Sumber : Staff Tata Usaha (TU) SMA Negeri 8 Tasikmalaya

3.4.2 Sampel

Sampel penelitian digunakan sebagai sumber data dalam penelitian ini berasal dari sebagian anggota populasi yang mempunyai karakteristik tertentu sesuai dengan tujuan penelitian (Sugiyono, 2019: 127). Penetapan sampel penelitian

dilakukan melalui perencanaan yang matang dengan memperhatikan tujuan dan kebutuhan penelitian yang ingin dicapai. Pemilihan sampel dalam penelitian ini dilakukan menerapkan teknik non-probability sampling, yaitu pemilihan sampel berdasarkan pertimbangan tertentu yang sesuai dengan tujuan penelitian, Pendekatan ini memungkinkan peneliti untuk menentukan sampel secara cermat sehingga karakteristik sampel yang dipilih sesuai dengan tujuan penelitian.

Berikut ini adalah contoh sampel yang di gunakan dalam penelitian ini:

Tabel 3.2 Sampel Penelitian

No	Kelas	Jumlah Siswa	Keterangan
1.	XI 6	34 Orang	Kelas Kontrol
2.	XI 7	33 Orang	Kelas Eksperimen
Jumlah		67	

Sumber : Staff Tata Usaha (TU) SMA Negeri 8 Tasikmalaya

3.5 Teknik Pengumpulan Data

3.5.1 Tes

Tes merupakan salah satu teknik pengumpulan data yang menggunakan instrumen berupa soal atau tugas tertentu untuk memperoleh informasi mengenai kemampuan, pengetahuan, kompetensi, kecerdasan, dan potensi peserta yang menjadi objek penelitian (Seplyana, dkk., 2025: 115). Penelitian ini memanfaatkan lembar tes sebagai instrumen pengumpulan data yang terdiri atas *pretest* dan *posttest*. *Pretest* dilaksanakan untuk mengidentifikasi tingkat hasil belajar awal siswa pada kelas eksperimen maupun kelas kontrol sebelum pemberian perlakuan sedangkan *posttest* bertujuan untuk mengukur hasil belajar siswa setelah perlakuan dilaksanakan. Tes yang digunakan berupa pertanyaan tertutup dengan bentuk pilihan ganda. Berdasarkan jawaban siswa, peneliti dapat menganalisis dan menilai kemampuan berpikir sejarah siswa.

Instrumen penelitian yang telah dikembangkan kemudian menjalani proses validasi oleh para ahli untuk memperoleh penilaian mengenai kesesuaian isi dan kelayakannya sebagai instrumen penelitian. Setelah dinyatakan layak, instrumen tersebut terlebih dahulu diujicobakan sebelum digunakan dalam penelitian. Selanjutnya, instrumen dianalisis melalui uji validitas, reliabilitas, tingkat

kesukaran dan uji daya beda untuk mengetahui kualitas serta keandalannya dalam mengukur variabel penelitian.

3.6 Instrumen Penelitian

Sugiyono (2019: 156) mendefinisikan instrumen penelitian adalah suatu alat yang digunakan mengukur fenomena alam maupun sosial yang yang diamati. Instrumen penelitian dapat didefinisikan juga sebagai alat atau metode untuk mengumpulkan atau menghimpun data serta informasi dari suatu fenomena atau permasalahan yang sedang diamati atau diteliti.

Penyusunan instrumen penelitian ini didasarkan pada varaiabel serta indikator yang telah ditetapkan. Mengacu pada pendapat Wineburg, kemampuan berpikir sejarah mencakup indikator utama, yaitu berpikir sesuai dengan konteksnya, urutan waktu dan keterhubungan peristiwa, keterampilan menelusuri hubungan sebab dan akibat, penafsiran serta bukti, kecakapan dalam melakukan penalaran kritis, dan keterampilan menyampaikan informasi. Berdasarkan indikator tersebut, disusun kisi-kisi tes kemampuan berpikir sejarah sebagaimna disajikan berikut ini.

Tabel 3.3 Kisi-Kisi Tes Kemampuan Berpikir Sejarah

Indikator yang diukur	Sub Indikator	No Soal	Level Kognitif	Jumlah Soal
Berpikir sesuai konteks sejarah	Menganalisis tujuan pembagian wilayah pemerintahan militer Jepang berdasarkan konteks Perang Dunia II	1	C4	5
	Menganalisis latar belakang perlawanan Aceh terhadap seikerei	4	C4	
	Menilai dampak sistem ekonomi perang Jepang bagi petani	14	C4	
	Mengaitkan pelatihan militer pemuda dengan kebutuhan perang Jepang	19	C4	
	Menganalisis pembentukan Jawa Hokokai dikaitkan dengan posisi Jepang dalam perang	31	C5	
Urutan waktu dan	Menafsirkan kebijakan budaya Jepang (waktu, bendera, lagu)	2	C4	4

Indikator yang diukur	Sub Indikator	No Soal	Level Kognitif	Jumlah Soal
keterhubungan peristiwa	sebagai upaya penanaman loyalitas			
	Menganalisis perubahan sistem romusha dari sukarela ke paksa	15	C4	
	Menyimpulkan nilai pendidikan masa Jepang bagi masa positif kemerdekaan	27	C5	
	Menyimpulkan pola perlawanan rakyat Aceh berdasarkan urutan kejadian	32	C5	
Keterampilan menelusuri hubungan sebab akibat	Menganalisis faktor kegagalan Gerakan Tiga A	3	C4	7
	Menilai faktor pemicu perlawanan rakyat Singapura	5	C4	
	Menilai dampak organisasi semi-militer Jepang terhadap perjuangan kemerdekaan	7	C4	
	Menganalisis dampak sosial romusha terhadap masyarakat desa	10	C4	
	Menganalisis latar belakang pembentukan Heiho	11	C4	
	Menjelaskan penyebab krisis pangan di Jawa pada masa Jepang	17	C4	
	Menyimpulkan dampak logis romusha terhadap kehidupan desa	23	C5	
Penafsiran serta bukti	Menilai perlakuan terhadap anggota Heiho berdasarkan tujuan Jepang	12	C4	5
	Menilai kebijakan penyerahan hasil panen dari sudut pandang petani	16	C4	
	Menyimpulkan arah kebijakan ekonomi Jepang berdasarkan eksploitasi wilayah	21	C5	
	Menafsirkan makna istilah "Prajurit Ekonomi" sebagai propaganda	22	C5	
	Menyimpulkan kebijakan Jepang terhadap MIAI	33	C5	

Indikator yang diukur	Sub Indikator	No Soal	Level Kognitif	Jumlah Soal
	berdasarkan perubahan sikap tahun 1942-1943			
Kecakapan dalam melakukan penalaran kritis	Mengevaluasi pembentukan Putera sebagai refleksi kegagalan propaganda Jepang	6	C4	10
	Menganalisis strategi tokoh nasional bekerja sama dengan Jepang	8	C4	
	Menganalisis fungsi tokoh nasional dalam Barisan Pelopor di bawah kontrol Jepang	9	C4	
	Membandingkan kedudukan dan tugas Heiho dan PETA	13	C4	
	Menganalisis tujuan pembubaran dan pembentukan organisasi politik Jepang	20	C4	
	Mengevaluasi pembukaan lahan baru sebagai dampak eksploitasi	24	C5	
	Menilai dampak positif kebijakan budaya Jepang bagi bangsa Indonesia	26	C5	
	Menyimpulkan keterkaitan kebijakan pemerintahan, organisasi massa, dan semi-militer Jepang	28	C5	
	Menyimpulkan tujuan utama kebijakan politik, militer, dan budaya Jepang	29	C5	
	Menyimpulkan pelajaran sejarah dari penderitaan masa pendudukan Jepang	30	C5	
Keterampilan menyampaikan informasi	Menyimpulkan tujuan utama mobilisasi rakyat oleh Jepang	18	C4	4
	Menyimpulkan pelajaran sejarah dari kebijakan Jepang bagi generasi sekarang	25	C5	
	Menyimpulkan secara komprehensif jalannya perlawanan faktor Pang Suma, strategi gerilya, serta perlawanan berdasarkan bukti sejarah	34	C6	

Indikator yang diukur	Sub Indikator	No Soal	Level Kognitif	Jumlah Soal
	Merumuskan pelajaran sejarah dari perlawanan Pang Suma dengan mengaitkan konteks lokal, strategi perlawanan, dan kelemahan yang dimanfaatkan Jepang	35	C6	

Tabel 3.3 menunjukkan kisi-kisi instrumen tes kemampuan berpikir sejarah yang disusun berdasarkan 6 indikator, yaitu berpikir sesuai konteks sejarah, keterhubungan waktu dan peristiwa, keterampilan menelusuri hubungan sebab akibat, penafsiran serta bukti, kecakapan dalam melakukan penalaran kritis, dan keterampilan menyampaikan informasi. Setiap indikator dijabarkan ke dalam beberapa subindikator yang disesuaikan dengan materi Masuknya Jepang dan Jatuhnya Hindia Belanda serta dikembangkan menjadi butir soal berdasarkan tingkat kognitif. Secara keseluruhan, instrumen terdiri atas 35 butir soal yang mencakup level kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta). Penyusunan kisi-kisi ini bertujuan agar instrumen mampu mengukur kemampuan berpikir sejarah peserta didik secara komprehensif sesuai dengan indikator yang telah ditetapkan.

Tabel 3.4 Rubrik Penilain

No	Indikator	Kriteria Penilain	Skor
1	Berpikir sesuai dengan konteks sejarah	Benar	1
		Salah	0
2.	Urutan waktu dan keterhubungan peristiwa	Benar	1
		Salah	0
3.	Keterampilan menelusuri hubungan sebab- akibat	Benar	1
		Salah	0
4.	Penafsiran serta penggunaan bukti sejarah	Benar	1
		Salah	0
5.	Kecakapan dalam melakukan penalaran kritis	Benar	1
		Salah	0
6.	Keterampilan menyampaikan informasi sejarah	Benar	1
		Salah	0

$$\text{Skor Nilai} = \frac{\text{Jumlah skor perolehan}}{\text{Jumlah skor maksimum}} \times 100$$

Berdasarkan Tabel 3.4, rubrik penilaian yang digunakan dalam penelitian ini mengacu pada enam indikator kemampuan berpikir sejarah, yaitu berpikir sesuai dengan konteks sejarah, urutan waktu dan keterhubungan peristiwa, keterampilan menelusuri hubungan sebab-akibat, penafsiran serta penggunaan bukti sejarah, kecakapan dalam melakukan penalaran kritis, dan keterampilan menyampaikan informasi sejarah. Penilaian dilakukan menggunakan skor dikotomis, yakni penilaian yang dilakukan dengan memberikan skor 1 pada jawaban yang benar dan skor 0 pada jawaban yang salah. Rubrik ini digunakan untuk memudahkan proses pengukuran dan pengolahan data kemampuan berpikir sejarah peserta didik secara objektif.

Setelah instrumen pengukuran kemampuan berpikir sejarah siswa ditetapkan, perlu dilakukan serangkaian uji kelayakan, yang diawali dengan uji validitas, reliabilitas, uji tingkat kesukaran, dan uji pembeda guna memastikan kualitas dan keandalan alat ukur yang digunakan dalam penelitian.

3.6.1 Uji Validitas

Uji validitas dilakukan untuk menilai sejauh mana instrumen penelitian mampu mengukur aspek yang seharusnya diukur. Kimberlin dan Winsterstein dalam Ramadhani & Bina (2021: 129) suatu instrumen dinilai valid apabila butir-butir pertanyaan dalam kuesioner atau tes mampu menggambarkan dengan tepat aspek yang ingin diukur. Tujuan uji validitas adalah memastikan bahwa alat yang digunakan benar-benar tepat untuk mengukur kemampuan belajar siswa. Dalam pelaksanaan penelitian atau penilaian, pengujian instrumen sangat penting agar dapat dipastikan bahwa alat tersebut mampu menilai kemampuan yang dituju secara akurat.

1) Validitas isi

Validitas isi adalah bentuk validitas yang diperoleh melalui penilaian terhadap kesesuaian dan relevansi materi dalam instrumen, yang dinilai secara logis oleh pihak yang berkompeten atau melalui proses penilaian ahli (*expert judgement*). Validitas isi instrumen pada penelitian ini ditetapkan berdasarkan kesepakatan para ahli. Penggunaan kesepakatan ahli dalam menetapkan validitas isi dilakukan karena suatu instrumen, baik tes maupun non-tes, dianggap tepat apabila para ahli

meyakini bahwa instrumen tersebut benar-benar mampu mengukur kemampuan yang dimaksud. Indeks validitas Aiken dimanfaatkan untuk menilai sejauh mana para ahli mencapai kesepakatan tersebut.

Penilaian dari beberapa ahli tersebut selanjutnya dianalisis menggunakan uji validitas Aiken's V dengan menerapkan rumus berikut.

$$V = \frac{\sum s}{[n(c - 1)]}$$

Gambar 3.2 Rumus Uji Validitas Aiken's

Sumber : Azwar dalam (Alpindo 2020: 4)

Keterangan;

- s : r - I_o
- I_o : Angka penilaian validitas yang terendah
- c : Angka penilaian validitas tertinggi
- r : Angka yang diberikan oleh seorang penilai
- n : Jumlah Penilai

Uji validitas isi dilakukan terhadap 35 soal tes pilihan ganda dengan melibatkan dua orang ahli, yaitu Ibu Rofiatul Azizah, M.Pd. (Penilai I) dan Ibu Sri Haryani, S.Pd. (Penilai II). Penilai I merupakan dosen aktif Program Studi Pendidikan Sejarah, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Siliwangi, sedangkan Penilai II merupakan guru mata pelajaran sejarah di SMAN 8 Tasikmalaya. Masing-masing butir soal dinilai menggunakan skala 1 sampai 5 yang menunjukkan tingkat kesesuaian soal dengan indikator kemampuan berpikir sejarah yang diukur. Perhitungan validitas isi dilakukan menggunakan rumus V-Aiken, dimana skor dari masing-masing penilai dikonversi menjadi nilai $s = r - I_o$ dengan $I_o = 1$, kemudian dijumlahkan dan dihitung nilai V berdasarkan jumlah kategori skala dan jumlah penilai. Berikut merupakan hasil analisis validitas isi menggunakan rumus V-Aiken.

Tabel 3.5 Hasil Uji Validitas Isi

Butir Soal	Penilai		S1	S2	ΣS	n (C-1)	V	KET
	I	II						
1	5	4	4	3	7	8	0,875	SANGAT VALID
2	5	5	4	4	8	8	1	SANGAT VALID
3	4	5	3	4	7	8	0,875	SANGAT VALID
4	5	5	4	4	8	8	1	SANGAT VALID
5	5	4	4	3	7	8	0,875	SANGAT VALID
6	4	5	3	4	7	8	0,875	SANGAT VALID
7	5	5	4	4	8	8	1	SANGAT VALID
8	4	5	3	4	7	8	0,875	SANGAT VALID
9	4	4	3	3	6	8	0,75	VALID
10	5	5	4	4	8	8	1	SANGAT VALID
11	5	5	4	4	8	8	1	SANGAT VALID
12	5	5	4	4	8	8	1	SANGAT VALID

Tabel 3.6 Kesimpulan Hasil Uji Validitas Tes Oleh Para Ahli

Butir Soal	Penilai		SI	S2	ΣS	n (C-1)	V	KET
	I	II						
1-12	56	57	44	45	89	96	0,927083	TINGGI

Berdasarkan hasil perhitungan menggunakan rumus Aiken's V terhadap butir soal instrumen yang telah divalidasi, seluruh butir soal menunjukkan tingkat validitas yang baik. Dari hasil tersebut terdapat 5 butir soal yang memiliki nilai $V \geq 0,875$ termasuk dalam kategori sangat valid, 6 butir soal dengan nilai $V = 1$ yang juga termasuk dalam kategori sangat valid, serta 1 butir soal yang memiliki nilai V sebesar 0,75 dan berada pada kateori valid. Berdasarkan hasil pengujian, seluruh butir soal dinyatakan telah sesuai dengan indikator kemampuan berpikir sejarah oleh karena, itu instrumen penelitian dapat dinyatakan layak dan dapat digunakan sebagai alat pengumpul data dalam penelitian.

2) Validitas Konstruk

Uji validitas konstruk dilaksanakan setelah instrumen dinyatakan layak untuk diujicobakan berdasarkan hasil penilaian para ahli. Keputusan mengenai validitas ditentukan melalui perbandingan nilai hasil perhitungan. Suatu butir pertanyaan dinyatakan valid apabila nilai $r_{hitung} > r_{tabel}$. Suatu butir pertanyaan dinyatakan tidak memenuhi kriteria validitas apabila nilai $r_{hitung} < r_{tabel}$. Perbandingan nilai korelasi dengan r_{tabel} dilakukan menggunakan derajat kebebasan $N-2$ serta tingkat signifikansi 5% atau 0,05. Rumus berikut dipakai peneliti untuk melakukan pengujian validitas instrumen.

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

Gambar 3.3 Rumus Uji Validitas

Sumber : Sugiyono (2019: 246)

Keterangan:

- r : Koefisien korelasi *Pearson*
- $\sum xy$: Jumlah hasil kali skor x dan y
- $\sum x$: Jumlah skor x
- $\sum y$: Jumlah skor y
- $\sum x^2$: Jumlah kuadrat skor x
- $\sum y^2$: Jumlah kuadrat skor y
- N : Jumlah peserta

Berikut adalah hasil uji validitas konstruk yang dilakukan oleh peneliti pada kelas XI 4 :

Tabel 3.7 Hasil Uji Validitas Konstruk

No	r- hitung	r- tabel	Keterangan
1	0,396	0,374	Valid
2	0,207	0,374	Tidak Valid
3	0,196	0,374	Tidak Vilad
4	0,53	0,374	Tidak Valid
5	0,582	0,374	Valid
6	0,359	0,374	Tidak Valid
7	0,805	0,374	Valid
8	0,505	0,347	Valid
9	0,467	0,347	Valid
10	0,613	0,347	Valid

No	r- hitung	r- tabel	Keterangan
11	0,660	0,347	Valid
12	0,653	0,347	Valid
13	0,280	0,347	Tidak Valid
14	0,805	0,347	Valid
15	0,570	0,347	Valid
16	0,805	0,347	Valid
17	0,178	0,347	Tidak Valid
18	0,539	0,347	Valid
19	0,596	0,347	Valid
20	0,521	0,347	Valid
21	0,818	0,347	Valid
22	0,660	0,347	Valid
23	0,634	0,347	Valid
24	0,626	0,347	Valid
25	0,218	0,347	Tidak Valid
26	0,395	0,347	Valid
27	0,354	0,347	Tidak Valid
28	0,483	0,347	Valid
29	0,470	0,347	Valid
30	0,457	0,347	Valid
31	0,488	0,347	Valid
32	0,081	0,347	Tidak Valid
33	0,311	0,347	Tidak Valid
34	0,359	0,347	Tidak Valid
35	0,493	0,347	Valid

Berdasarkan uji validitas konstruk yang dilakukan pada 35 soal dengan melibatkan 30 siswa kelas XI 4 SMA Negeri 8 Tasikmalaya dengan menggunakan *software IBM SPSS Statistic 29* menghasilkan temuan bahwa 11 butir pertanyaan dinyatakan tidak valid. Oleh sebab itu, butir yang tidak memenuhi syarat tersebut yaitu 2, 3, 4, 6, 13, 17, 25, 27, 32, 33, 34. Pada soal yang lain, 24 butir soal lainnya dinyatakan valid dan layak digunakan dalam instrumen tes penelitian.

3.6.2 Uji Reliabilitas

Uji reliabilitas adalah proses evaluasi terhadap instrumen penelitian untuk memastikan bahwa data yang diperoleh bersifat stabil dan konsisten, meskipun digunakan dalam pengukuran berulang. Uji reliabilitas dilakukan setelah instrumen terbukti valid. Sebuah instrumen dapat dikategorikan memiliki reliabilitas yang memadai apabila mampu menghasilkan hasil yang konsisten, siapa pun yang

menggunakannya, serta di mana pun dan kapan pun pengukuran dilakukan (Ramdhani & Bina, 2021: 135).

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 instrumen yang digunakan dalam proses pengumpulan data menghasilkan data yang stabil dan memiliki tingkat keandalan yang baik.

Uji reliabilitas dengan teknik *Cronbach alpha* diterapkan dalam penelitian ini untuk menilai konsistensi instrumen. Uji reliabilitas dilaksanakan apabila nilai koefisien yang diperoleh mencapai batas minimum 0,6. Instrumen dinyatakan reliabel apabila nilai r_{hitung} (*Cronbach's Alpha*) melebihi angka 0,6. Sebaliknya, apabila nilai r_{hitung} berada di bawah 0,6, maka instrumen tersebut dianggap tidak memenuhi standar keandalan. Berikut rumus realibilitas *Cronbach alpha*:

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

Gambar 3.4 Rumus Uji Reliabilitas Cronbach Alpha

Keterangan:

- r_{11} : Reliabilitas Instrumen
- k : Banyak butir pertanyaan
- $\sum \sigma_b^2$: Variansi butir
- σ_t^2 : Variansi total

Tabel 3.8 jumlah sampel (Responden)

Case Processing Summary		N	%
Cases	Valid	30	100.0
	Excluded ^a	0	.0
	Total	30	100.0

a. Listwise deletion based on all variables in the procedure.

Tabel tersebut menunjukkan bahwa jumlah responden dalam uji coba instrumen terdapat 30 responden.

Tabel 3.9 Hasil Uji reliabilitas

Reliability Statistics	
Cronbach's Alpha	N of Items
.888	24

Berdasarkan data Reliabilitas Statistics, diperoleh nilai Cronbach's Alpha sebesar 0,888 pada 24 butir soal, yang memperlihatkan bahwa instrumen memenuhi kriteria reliabilitas tinggi karena berada di atas minimum 0,70. Hal ini berarti seluruh butir soal menunjukkan keseragaman antarbutir yang baik serta mampu mengukur variabel penelitian secara stabil serta saling berkaitan satu sama lain. Dengan demikian, instrumen tersebut dapat dipercaya sebagai alat pengumpulan data penelitian sebab menghasilkan pengukuran yang konsisten dan tidak bersifat acak.

Tabel 3.10 Hasil Uji Reliabilitas per Butir Soal

Butir Soal	Cronbach's Alpha If item Deleted	Keterangan
Soal 1	0,888	Reliabel
Soal 2	0,881	Reliabel
Soal 3	0,879	Reliabel
Soal 4	0,886	Reliabel
Soal 5	0,885	Reliabel
Soal 6	0,880	Reliabel
Soal 7	0,881	Reliabel
Soal 8	0,880	Reliabel
Soal 9	0,903	Reliabel
Soal 10	0,879	Reliabel
Soal 11	0,882	Reliabel
Soal 12	0,879	Reliabel

Butir Soal	Cronbach's Alpha If item Deleted	Keterangan
Soal 13	0,882	Reliabel
Soal 14	0,885	Reliabel
Soal 15	0,877	Reliabel
Soal 16	0,879	Reliabel
Soal 17	0,882	Reliabel
Soal 18	0,880	Reliabel
Soal 19	0,888	Reliabel
Soal 20	0,885	Reliabel
Soal 21	0,887	Reliabel
Soal 22	0,885	Reliabel
Soal 23	0,886	Reliabel
Soal 24	0,884	Reliabel

Berdasarkan tabel hasil uji reliabilitas per butir soal, diketahui bahwa nilai Cronbach's Alpha if Item Deleted pada 24 butir soal berada pada rentang 0,877 sampai 0,903. Seluruh nilai tersebut menunjukkan angka yang lebih besar dari 0,70, sehingga setiap butir soal termasuk dalam kategori reliabel. Meskipun seluruh butir soal dinyatakan reliabel, pada penelitian ini 4 soal, yaitu nomor 2, 4, 5, dan 6, tidak digunakan pada tahap pelaksanaan penelitian. Pengurangan butir soal dilakukan sebagai bentuk penyesuaian terhadap alokasi waktu pembelajaran sejarah yang terbatas, sehingga pelaksanaan *pretest* dan *posttest* dapat berlangsung secara efektif tanpa mengurangi ketercapaian indikator kemampuan berpikir sejarah yang diukur. Pada tabel terlihat bahwa seluruh nilai tetap berada pada kategori tinggi, sehingga dapat diartikan bahwa seluruh butir soal memberikan kontribusi yang baik terhadap reliabilitas instrumen. Dengan demikian, semua butir soal dinilai layak untuk digunakan dalam penelitian.

3.6.3 Tingkat Kesukaran

Tingkat kesukaran butir soal adalah ukuran yang digunakan untuk mengetahui seberapa mudah atau sulit suatu soal berdasarkan proporsi peserta tes yang menjawab benar terhadap jumlah seluruh peserta tes. Tingkat kesukaran butir soal diukur menggunakan Indeks Kesukaran Soal (IKS) yang memiliki rentang nilai dari 0,00 sampai 1,00. Nilai indeks yang mendekati 1 menunjukkan bahwa soal cenderung mudah, sedangkan nilai yang mendekati 0 menunjukkan bahwa soal cenderung sulit. Secara umum, tingkat kesukaran butir soal merupakan ukuran yang didasarkan pada rata-rata skor yang diperoleh peserta tes pada butir soal tersebut

(Munip dalam Dianova and 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

Dalam pengujian tingkat kesukaran pada penelitian ini dilakukan dengan bantuan *software IBM SPSS Statistic* versi 29, dengan kriteria penentuan keputusan sebagai berikut:

Tabel 3.11 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 3.12 Tingkat Kesukaran Soal Tes Kemampuan Berpikir Sejarah

Butir Soal	Mean	Tingkat Kesukaran
Soal 1	0,80	Mudah
Soal 2	0,70	Sedang
Soal 3	0,93	Mudah
Soal 4	0,63	Sedang
Soal 5	0,83	Mudah
Soal 6	0,80	Mudah
Soal 7	0,77	Mudah
Soal 8	0,80	Mudah
Soal 9	0,80	Mudah
Soal 10	0,93	Mudah
Soal 11	0,87	Mudah
Soal 12	0,93	Mudah
Soal 13	0,83	Mudah
Soal 14	0,77	Mudah
Soal 15	0,90	Mudah
Soal 16	0,77	Mudah
Soal 17	0,93	Mudah
Soal 18	0,80	Mudah
Soal 19	0,63	Sedang
Soal 20	0,63	Sedang

Butir Soal	Mean	Tingkat Kesukaran
Soal 21	0,43	Sedang
Soal 22	0,60	Sedang
Soal 23	0,53	Sedang
Soal 24	0,80	Mudah

Hasil pengujian tingkat kesukaran pada 24 butir soal menunjukkan bahwa sebagian besar soal memiliki tingkat kesukaran yang rendah. Dari 24 soal yang dianalisis, 16 butir soal termasuk kategori mudah dan 8 butir berkategori sedang. Butir soal yang tergolong sedang terdapat pada nomor 2, 4, 19, 20, 21, 22, dan 23. Adapun soal kategori mudah terdapat pada nomor 1, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, dan 24.

Dominannya butir soal berkategori mudah menunjukkan bahwa sebagian besar peserta didik mampu menjawab soal dengan benar. Meskipun demikian, keberadaan beberapa butir soal berkategori sedang tetap penting karena dapat memberikan variasi tingkat kesulitan dan membantu mengukur kemampuan peserta didik secara lebih optimal. Dengan demikian, keseluruhan butir soal memiliki tingkat kesukaran yang masih sesuai untuk digunakan sebagai instrumen penelitian dalam mengukur kemampuan berpikir sejarah peserta didik.

3.6.4 Daya Pembeda

Analisis daya pembeda digunakan untuk menilai tingkat kemampuan suatu butir soal dalam membedakan siswa yang memiliki penguasaan materi tinggi dengan siswa yang penguasaan materinya rendah. Besarnya daya beda soal dapat dilihat dari nilai Indeks Daya Beda (IDB) yang dimiliki oleh butir tersebut. Indeks daya beda pada umumnya dinyatakan dalam bentuk proporsi yang menggambarkan kemampuan soal dalam membedakan peserta tes. Semakin besar indeks daya beda yang dimiliki suatu butir soal, semakin baik kemampuan soal tersebut dalam mengidentifikasi perbedaan antar kelompok siswa pandai dan kelompok siswa yang kurang pandai (Munip dalam Dianova and 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 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 Statistis versi 29. Dengan kriteria pengambilan keputusan sebagai berikut:

Tabel 3.13 Kriteria Pengambilan Keputusan Daya Pembeda

Koefisien Korelasi	Kategori Daya Beda
0,40-1,00	Sangat baik
0,30-0,39	Baik
0,20-0,29	Cukup
0,00-0,19	Jelek

(Naga dalam Qadir dkk., 2024: 1461)

Hasil data uji daya pembeda untuk instrumen kemampuan berpikir Sejarah deangan menggunakan bantuan software IBM SPSS Statistic versi 29 sebagai berikut:

Tabel 3.14 Uji Daya Beda

Butir Soal	r hitung (Corrected Item-Total Correlation)	Kategori
Soal 1	0,282	Cukup
Soal 2	0,551	Baik
Soal 3	0,820	Sangat baik
Soal 4	0,393	Cukup
Soal 5	0,436	Baik
Soal 6	0,598	Baik
Soal 7	0,581	Baik
Soal 8	0,616	Baik
Soal 9	0,354	Cukup
Soal 10	0,820	Sangat baik
Soal 11	0,545	Baik
Soal 12	0,820	Sangat baik
Soal 13	0,530	Baik
Soal 14	0,412	Baik
Soal 15	0,840	Sangat baik

Butir Soal	r hitung (Corrected Item-Total Correlation)	Kategori
Soal 16	0,632	Baik
Soal 17	0,622	Baik
Soal 18	0,598	Baik
Soal 19	0,320	Cukup
Soal 20	0,423	Baik
Soal 21	0,386	Cukup
Soal 22	0,434	Baik
Soal 23	0,420	Baik
Soal 24	0,456	Baik

Analisis daya pembeda terhadap 24 butir soal menghasilkan temuan bahwa 4 butir soal tergolong sangat baik, 15 butir berkategori baik, dan 5 butir soal cukup. Butir soal yang menunjukkan kemampuan membedakan siswa secara sangat baik adalah soal nomor 3, 10, 12, dan 15. Soal berkategori baik terdapat pada nomor 2, 5, 6, 7, 8, 11, 13, 14, 16, 17, 18, 20, 22, 23, dan 24. Adapun soal berkategori cukup terdapat pada nomor 1, 4, 9, 19, dan 21.

Hasil tersebut menunjukkan bahwa sebagian besar butir soal memiliki daya pembeda yang baik, sehingga mampu membedakan peserta didik yang memiliki tingkat penguasaan materi tinggi dengan peserta didik yang memiliki tingkat penguasaan materi lebih rendah. Selain itu, beberapa butir soal yang termasuk dalam kategori sangat baik menunjukkan kemampuan yang sangat optimal dalam membedakan kemampuan peserta didik. Adapun butir soal yang berada pada kategori cukup masih memenuhi kriteria daya pembeda sehingga tetap layak digunakan sebagai instrumen penelitian.

3.7 Teknik Analisis Data

Analisis data adalah proses mengolah informasi untuk menemukan jawaban atas rumusan masalah penelitian, menilai kesesuaiannya dengan tujuan yang telah ditetapkan, serta melihat faktor-faktor yang dapat memengaruhi hasil penelitian (Balaka, 2022: 3). Analisis data kuantitatif dalam penelitian ini dilakukan menggunakan teknik statistik dengan bantuan aplikasi IBM SPSS statistic 29 untuk menguji instrumen tes yang digunakan. Untuk mendapatkan hasil uji instrumen tersebut, peneliti melaksanakan beberapa tahapan pengujian sebagai berikut:

3.7.1 Uji Normalitas

Uji normalitas merupakan tahapan penting dalam analisis statistik inferensial yang bertujuan memastikan bahwa data memenuhi asumsi distribusi normal serta keseragaman varians sebelum dilakukan pengujian hipotesis. Pada penelitian ini, uji normalitas dilakukan dengan menggunakan metode Shapiro-Wilk. Proses pengujian Shapiro-Wilk dilakukan dengan bantuan aplikasi IBM SPSS statistic 29. Nilai signifikansi yang lebih besar dari 0,05 menunjukkan bahwa data berdistribusi normal, sedangkan nilai signifikansi yang lebih kecil dari 0,05 mengindikasikan bahwa data tidak berdistribusi normal.

3.7.2 Uji Homogenitas

Uji homogenitas dilakukan untuk memastikan apakah kelompok perlakuan dan kelompok kontrol memiliki variasi atau karakteristik yang sebanding. Apabila tidak terdapat perbedaan yang signifikan pada varians kedua kelompok, maka sampel dianggap homogen, untuk menentukan apakah data bersifat homogen atau tidak, penelitian ini menggunakan uji Levene. Uji Levene merupakan prosedur statistik yang digunakan untuk menilai keseragaman atau homogenitas varians antar beberapa populasi atau kelompok yang dibandingkan (Sianturi, 2022: 392). Proses pengujiannya dilakukan dengan bantuan aplikasi IBM SPSS Statistics 29. Uji homogenitas melalui SPSS dilakukan dengan acuan bahwa data dinyatakan homogen apabila nilai signifikansinya melebihi 0,05. Jika nilai signifikansi berada di bawah 0,05, berarti data kemungkinan tidak seragam. Jika hasil pengujian menunjukkan bahwa data homogen, maka langkah berikutnya adalah melanjutkan ke tahap uji hipotesis.

3.7.3 Uji Hipotesis

Tujuan uji hipotesis adalah menilai kebenaran atau ketepatan suatu pernyataan atau dugaan yang diajukan dalam penelitian. Tahap ini dilaksanakan untuk memberikan dasar yang kuat dalam proses pengumpulan data serta memperoleh bukti empiris penelitian. Setelah data berhasil dihimpun, pengujian hipotesis dilakukan menggunakan uji Independent Sample T-test. Peneliti menggunakan program IBM SPSS Statistics 29 untuk menguji hipotesis. Dasar penarikan kesimpulan adalah sebagai berikut jika nilai Sig. (2-tailed) lebih besar

dari 0,05, maka H_0 diterima dan rata-rata kedua kelompok dianggap sama. Sebaliknya, jika nilai Sig. (2-tailed) kurang dari 0,05, maka H_0 ditolak dan rata-rata kedua kelompok dinyatakan berbeda.

3.7.4 Uji N-Gain

Uji Normalitas Gain/N-Gain digunakan untuk mengidentifikasi pola peningkatan hasil belajar melalui perbandingan skor sebelum dan sesudah penerapan suatu metode pembelajaran. Analisis data N-Gain dilakukan setelah nilai pretest dan posttest diperoleh serta dianalisis. Data N-Gain dimanfaatkan untuk menilai sejauh mana efektivitas perlakuan yang diberikan. Nilai N-Gain diperoleh dengan menerapkan rumus berikut:

$$N\text{-Gain} = \frac{S_{posttest} - S_{pretest}}{S_{max} - S_{pretest}}$$

Keterangan:

N-Gain : nilai uji normalitas
 $S_{posttest}$: skor *posttest*
 $S_{pretest}$: skor *pretest*
 S_{max} : skor maksimal

3.8 Langkah-Langkah Penelitian

Langkah-langkah penelitian mencakup berbagai proses yang ditempuh peneliti untuk memperoleh solusi atas permasalahan yang diteliti. Adapun tahapan-tahapan yang dilakukan peneliti dalam pelaksanaan penelitian adalah sebagai berikut:

3.8.1 Tahap Penyusunan

Pada tahap ini peneliti melakukan observasi ke sekolah yang akan menjadi tempat penelitian untuk mendapatkan informasi mengenai kondisi kelas yang akan menjadi sampel penelitian. Selanjutnya, peneliti mengajukan judul penelitian kepada dosen pembimbing 1 dan dosen pembimbing 2, setelah memperoleh persetujuan dari kedua dosen pembimbing, peneliti menyusun proposal penelitian. Selanjutnya, proposal yang telah direvisi berdasarkan masukan dari dosen pembimbing 1 dan dosen pembimbing 2 diajukan untuk diuji kelayakannya

melalui sidang proposal. Sesudah proposal mendapatkan persetujuan dalam sidang proposal, peneliti melaksanakan tahap persiapan yang mendukung pelaksanaan penelitian, termasuk pengembangan instrumen penelitian dan pengujian instrumen.

3.8.2 Tahap Pelaksanaan

Pada tahap pelaksanaan, peneliti akan mengunjungi kelas XI SMA Negeri 8 Tasikmalaya untuk memantau siswa yang sedang belajar menggunakan media Quizalize. peneliti melaksanakan kegiatan pembelajaran pada dua kelompok, yaitu kelas eksperimen dan kelas kontrol. Kelas eksperimen diberikan perlakuan menggunakan media pembelajaran Quizalize, sedangkan kelas kontrol mengikuti pembelajaran konvensional tanpa penggunaan Quizalize. Sebelum perlakuan diberikan, kedua kelompok diberikan *pretest* untuk mengukur kemampuan awal berpikir sejarah. Setelah pembelajaran selesai, kedua kelompok kembali diberikan *posttest* untuk mengetahui perubahan kemampuan berpikir sejarah setelah perlakuan. Seluruh proses pelaksanaan dilakukan dengan pengawasan peneliti untuk memastikan kesesuaian prosedur penelitian. Data hasil *pretest* dan *posttest* kemudian dikumpulkan untuk dianalisis secara kuantitatif. Tahap berikutnya adalah melakukan pengolahan dan analisis terhadap data yang telah terkumpul. Pengolahan data bertujuan untuk mempermudah peneliti dalam memperoleh informasi yang relevan serta memastikan bahwa data yang digunakan akurat dan dapat dipertanggungjawabkan.

3.8.3 Tahap Akhir

Tahap terakhir yang dilakukan peneliti yaitu mengolah seluruh data yang terkumpul selama proses pengambilan data berlangsung. Laporan yang disusun memuat data yang telah tervalidasi dan diperoleh dari hasil analisis data kuantitatif yang dilakukan oleh peneliti.

3.9 Waktu dan Tempat Penelitian

3.9.1 Waktu Penelitian

Penelitian ini dilaksanakan pada periode September 2025 hingga Mei 2026 dengan durasi pelaksanaan selama sembilan bulan. Adapun rincian waktu setiap tahapan kegiatan peneliti dijabarkan sebagai berikut

Tabel 3.15 Waktu Penelitian

No	Jenis Kegiatan	Bulan							
		Sept	Nov	Des	Jan	Feb	Mar	Apr	Mei
1.	Tahap Penyusunan								
	a. Pengajuan tema dan penyusunan proposal								
	b. Pengajuan proposal								
	c. Seminar proposal								
	d. Perizinan Penelitian								
	e. Penyusunan instrumen								
2.	Tahap Pelaksanaan								
	a. Uji coba instrumen								
	b. Pengumpulan data								
	c. Pengelolaan data								
3.	Tahap Akhir								
	a. Penyusunan laporan akhir								

3.9.2 Tempat Penelitian

SMA Negeri 8 Tasikmalaya yang terletak di JL. Mulyasari No. 03, Mulyasari, Kec. Tamansari, Kab. Tasikmalaya, Jawa Barat 46196.