

BAB 3

METODOLOGI PENELITIAN

3.1 Metode Penelitian

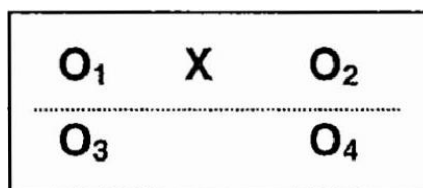
Metode penelitian yakni prosedur ilmiah dalam upaya memperoleh data dengan sistematis, terencana, serta terkontrol dengan tujuan serta kegunaan tertentu sesuai dengan fokus penelitian yang dilakukan (Ramdhan, 2021: 1). Pendekatan kuantitatif yaitu serangkaian langkah ilmiah yang dilakukan oleh peneliti dengan menggunakan pendekatan kuantitatif guna mendapatkan data yang mampu diukur secara numerik, kemudian dianalisis secara statistik guna menjawab atau memecahkan permasalahan penelitian (Muin, 2023: 3).

Penelitian ini menggunakan metode kuasi eksperimen, yaitu rancangan penelitian yang menggunakan kelompok pembandingan (kontrol), namun keberadaannya tidak sepenuhnya mengontrol variabel eksternal yang memengaruhi jalannya kelompok eksperimen. Desain *non equivalent control group* digunakan di mana kelompok eksperimen dan kelompok kontrol tidak dipilih secara random (Sugiyono, 2019: 118-120). Dalam penerapannya, kelas yang memperoleh *treatment* adalah kelas eksperimen dan kelas yang tidak memperoleh *treatment* adalah kelas kontrol (Abraham & Supriyati, 2022: 2480). Kelas eksperimen akan diterapkan media pembelajaran Flippity. Kelas kontrol akan diterapkan media pembelajaran yang biasa dimanfaatkan, kemudian peserta didik akan mengerjakan tes. Tes digunakan sebagai alat pengumpulan data yang diisi oleh peserta didik guna mengetahui kemampuan peserta didik dalam berpikir sejarah. Hasil tes pada kelas eksperimen kemudian akan dibandingkan dengan hasil tes dari kelas kontrol.

3.2 Desain Penelitian

Kajian ini menggunakan kuasi eksperimen dengan bentuk *pre-test* dan *post-test*. Desain ini memiliki 2 kelompok peserta didik dari SMAN 5 Tasikmalaya dipilih, yakni kelompok eksperimen yang akan mendapatkan *treatment* menggunakan media interaktif Flippity serta kelompok kontrol yang akan dilakukan pembelajaran tanpa memanfaatkan media interaktif.

Penggunaan *quasi experiment* dilakukan karena sampelnya menggunakan kelas yang telah terbentuk sebelumnya sesuai dengan kebijakan sekolah, oleh karena itu metode pemilihan acak tidak dapat digunakan. Peneliti menggunakan *quasi experiment* beserta rancangan *pre-test* serta *post-test* dengan kontrol yang tidak sebanding (*non equivalent control*).



Gambar 3.1 Desain Penelitian

(Sugiyono, 2019: 120)

Keterangan:

O^1 : *Pre-test* pada kelas eksperimen

O^3 : *Pre-test* pada kelas kontrol

X : Perlakuan (Media ajar Flippity)

O^2 : *Post-test* pada kelas eksperimen

O^4 : *Post-test* pada kelas kontrol

Pada Gambar 3.1 dijelaskan mengenai pembagian kelompok dalam penelitian, yakni kelompok yang mendapat *treatment* khusus serta kelompok yang tidak mendapat *treatment* khusus. Kelompok eksperimen memperoleh perlakuan (X) berupa penggunaan Flippity sebagai media pembelajaran, sedangkan kelompok kontrol tidak memperoleh *treatment* tersebut. Kedua kelompok tersebut menjalani *pre-test* serta *post-test* guna mengukur kemampuan berpikir sejarah peserta didik sebelum juga sesudah penerapan *treatment* khusus.

3.3 Variabel Penelitian

Variabel penelitian merupakan fokus utama atau hal yang diteliti dalam suatu penelitian. Variabel penelitian bisa berupa benda, manusia, atau sasaran penelitian yang menjadi fokus amatan utama peneliti untuk dikaji serta dipahami lebih dalam (Dartiningsih dkk., 2016: 132). Variabel penelitian terdiri dari beberapa macam, misalnya variabel penelitian independen serta variabel penelitian dependen. Variabel penelitian independen dikenal juga sebagai variabel bebas, yaitu variabel yang memberikan pengaruh atau yang memicu perubahan pada variabel dependen. Variabel dependen atau variabel terikat diposisikan sebagai unsur yang terdampak

dari keberadaan variabel bebas (Sugiyono, 2019: 68-69). Variabel independent yang peneliti gunakan yaitu penggunaan media flippity, sementara variabel dependen yang digunakan yakni kemampuan berpikir sejarah peserta didik kelas XI.11 SMAN 5 Tasikmalaya.

3.4 Populasi dan Sampel

3.4.1 Populasi

Populasi yakni seluruh kelompok yang meliputi objek atau subjek yang mempunyai jumlah serta kriteria khusus yang ditetapkan oleh peneliti untuk dikaji lebih lanjut serta dijadikan landasan dalam merumuskan simpulan penelitian (Sugiyono, 2019: 126). Peneliti menjadikan populasi sebagai dasar utama dalam menentukan ruang lingkup serta batasan penelitian. Populasi yang ditetapkan pada penelitian ini mencakup peserta didik kelas XI SMAN 5 Tasikmalaya yang sesuai dengan kriteria peneliti, sebagaimana ditampilkan pada tabel berikut:

Tabel 3.1 Populasi Kelas XI IPS

No	Kelas	Jumlah Peserta didik
1	XI - 9	33
2	XI - 11	28
3	XI - 12	28
Total Peserta didik		89

Sumber: Staff Tata Usaha SMAN 5 Tasikmalaya

3.4.2 Sampel

Prosedur penentuan sampel dalam penelitian menerapkan *non-probability sampling*, di mana populasi tidak memiliki probabilitas yang setara untuk terpilih sebagai sampel. Secara spesifik, peneliti menggunakan teknik *purposive sampling*, yakni pemilihan sampel yang didasarkan pada kriteria atau pertimbangan tertentu agar sesuai dengan tujuan penelitian (Septiani dkk., 2020: 135-136). Sampel dipilih berdasarkan kriteria kemampuan berpikir sejarah yang belum optimal yang dilihat dari pengamatan di kelas dan diperoleh dari pertimbangan guru sejarah kelas XI, yaitu kelas XI.11 yang merupakan kelas eksperimen serta kelas XI.12 yang merupakan kelas kontrol.

Tabel 3.2 Sampel Penelitian

No	Kelas	Jenis Kelamin		Jumlah
		L	P	
1	XI - 11	12	16	28
2	XI - 12	9	19	28
Total Peserta didik				56

Sumber: Staff Tata Usaha SMAN 5 Tasikmalaya

3.5 Teknik Pengumpulan Data

3.5.1 Tes

Prosedur pengumpulan data dalam penelitian ini dilakukan melalui instrumen tes, yang merupakan metode yang dilaksanakan melalui memberikan tes kognitif kepada partisipan yang berkaitan dengan fenomena penelitian. Instrumen yang berbentuk tes umumnya dimanfaatkan guna mengukur kemampuan dasar serta pencapaian prestasi belajar (Muin, 2023: 58).

3.6 Instrumen Penelitian

Instrumen penelitian merupakan panduan tertulis yang berisi pedoman wawancara, observasi, serta daftar pertanyaan yang disusun guna memperoleh informasi yang dibutuhkan dalam penelitian (Ovan & Saputra, 2020: 2). Instrumen penelitian yang akan diterapkan pada penelitian ini meliputi:

3.6.1 Tes

Tes didefinisikan sebagai instrumen sistematis dan terstandarisasi yang dirancang guna mengukur kemampuan peserta didik secara objektif melalui prosedur penelitian yang baku (Kurniawan, 2021: 6). Penelitian ini menerapkan tes tipe pilihan ganda yang mencakup dari 20 *item* yang wajib diselesaikan dengan memilih 1 dari 5 opsi jawaban yang tersedia pada setiap soal. Tes kemampuan berpikir sejarah meliputi beberapa indikator diantaranya berpikir kronologis, pemahaman sejarah, analisis dan interpretasi sejarah, kemampuan riset sejarah, serta analisis isu sejarah (Sukmawati dkk., 2025: 256-257). Kisi-kisi tes berpikir sejarah, diantaranya:

Tabel 3.3 Kisi-Kisi Tes Berpikir Sejarah

Indikator yang diukur	Sub Indikator	Nomor Soal	Level Kognitif	Jumlah Soal
Chronological Thinking (Berpikir Kronologis)	Mengidentifikasi urutan peristiwa sejarah.	1	C4	1
	Memahami hubungan sebab-akibat antar peristiwa.	2	C4	1
	Mengenali perubahan dan kesinambungan dalam sejarah.	3, 4	C4	2
Historical Comprehension (Pemahaman Sejarah)	Menafsirkan makna peristiwa dari perspektif masa itu.	5	C4	1
	Memahami konteks sosial, politik, ekonomi dari peristiwa sejarah.	6, 7	C4	2
	Menyimpulkan pesan atau pelajaran dari peristiwa sejarah.	8	C5	1
Historical Analysis and Interpretation (Analisis dan Interpretasi Sejarah)	Menganalisis keakuratan dan bias sumber sejarah.	9, 16	C4	2
		10	C5	1
	Menafsirkan peristiwa sejarah berdasarkan bukti.	11	C5	1
		12	C4	1
Historical Research Capabilities (Kemampuan Riset Sejarah)	Mengidentifikasi dan mengumpulkan sumber sejarah.	13, 14	C4	2
	Mengorganisasi informasi untuk membangun kesimpulan sejarah.	15	C4	1
Historical Issues (Analisis Isu Sejarah)	Menilai isu sejarah dari berbagai perspektif.	17, 18	C5	2
	Mengambil keputusan atau posisi berdasarkan bukti sejarah.	19	C4	1
		20	C5	1

3.6.2 Uji Validitas

Uji validitas adalah pengujian yang bertujuan guna mengetahui apakah suatu instrument penelitian dinyatakan valid atau tidak. Instrumen yang dimaksud dalam konteks ini berupa butir-butir pertanyaan yang terdapat dalam soal tes. Suatu tes dinyatakan valid jika setiap pertanyaan didalamnya mampu mengukur aspek yang hendak diteliti sesuai dengan tujuan penelitian (Janna & Herianto, 2021: 2).

1) Validitas Isi

Validitas isi bertujuan untuk memastikan bahwa instrumen penelitian telah relevan dengan variabel yang seharusnya diukur (Siregar, 2017: 46). Dalam proses penilaian validitas isi, para ahli diminta menilai sejauh mana setiap butir instrumen benar-benar merepresentasikan konstruk yang ingin diukur. Hasil penilaian para ahli tersebut lalu dianalisis dengan teknik statistik. Salah satu teknik yang umum digunakan adalah Aiken's V. Aiken mengembangkan rumus Aiken's V pada tahun 1985 guna menghitung koefisien validitas isi berdasarkan skor yang diberikan oleh sejumlah pakar terhadap tiap butir, dengan fokus pada tingkat keterwakilan item terhadap konstruk yang diukur (Widodo, 2022: 74). Indikator yang digunakan untuk menilai validitas isi meliputi relevansi antara indikator dengan butir instrumen, konsistensi butir soal terhadap aspek-aspek penelitian, kejelasan aspek kebahasaan atau visual dalam soal, kelayakan instrumen bagi subjek penelitian, ketepatan materi atau konsep yang diujikan (Novikasari, 2016: 2). Nilai indeks Aiken yang $< 0,4$ maka dikategorikan memiliki validitas rendah, jika terletak diantara $0,4 - 0,8$ dikategorikan memiliki validitas sedang, sementara apabila nilai indeks Aiken $> 0,8$ instrumen tersebut dianggap memiliki kualifikasi validitas yang tinggi (Utami dkk., 2024: 63). Berdasarkan hal tersebut maka instrumen penelitian ini dilakukan uji coba kepada kelas yang tidak termasuk kelas eksperimen serta kelas kontrol.

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

Gambar 3.2 Rumus Uji Validitas Aikens'V
(Utami dkk., 2024: 63)

Keterangan:

- V : Indeks aiken
- s : r - lo
- lo : Angka penilaian terendah (misalnya 1)
- c : Angka penilaian tertinggi (misalnya 4)
- r : Angka yang diberikan oleh penilai
- n : Jumlah penilai

Tabel 3.4 Hasil Uji Validitas Isi

Butir	Penilai		S1	S2	Σs	n(c-1)	V	Ket
	1	2						
Butir 1	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 2	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 3	5	5	4	4	8	8	1	SANGAT TINGGI
Butir 4	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 5	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 6	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 7	5	5	4	4	8	8	1	SANGAT TINGGI
Butir 8	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 9	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 10	5	5	4	4	8	8	1	SANGAT TINGGI
Butir 11	4	5	3	4	7	8	0.875	SANGAT TINGGI
Butir 12	5	5	4	4	8	8	1	SANGAT TINGGI

Tabel 3.5 Kesimpulan Hasil Uji Validitas

Butir	Penilai		S1	S2	Σs	n(c-1)	V	Ket
	1	2						
Butir 1-12	52	60	40	48	88	8	0.91667	SANGAT TINGGI

Sumber: Hasil Uji Validitas Menggunakan Excel

2) Validitas Konstruk

Validitas konstruk digunakan untuk menguji terkait beberapa aspek yang akan diukur benar-benar sesuai dengan konsep teoretis tertentu, kemudian dikonsultasikan kepada para ahli (Sugiyono, 2019: 179). Proses ini dilakukan terhadap setiap variabel dalam instrumen penelitian untuk memastikan keakuratan dan ketepatan alat ukur yang digunakan. Instrumen dinyatakan valid apabila nilai $r_{hitung} >$ dari r_{tabel} pada tingkat signifikansi 5% $\alpha = 0,05$. Instrumen dianggap

tidak valid jika $r_{hitung} < r_{tabel}$. Uji validitas dalam penelitian ini dilakukan dengan memanfaatkan program SPSS versi 29.0.

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

Gambar 3.3 Rumus Uji Validitas Konstruk
(Sugiyono, 2019: 246)

Keterangan (Siregar, 2017: 48):

n : Jumlah Responden

x : Skor variabel (jawaban responden)

Y : Skor total dari variabel (jawaban responden)

Tabel 3.6 Hasil Uji Validitas Konstruk

No Soal	r hitung	r tabel N = 31 A = 0,05	Kriteria r hitung > r tabel atau r hitung < r tabel	Validitas
1.	,433	0,355	,433 > 0,355	Valid
2.	,107	0,355	,107 < 0,355	Tidak Valid
3.	,466	0,355	,466 > 0,355	Valid
4.	,220	0,355	,220 < 0,355	Tidak Valid
5.	,539	0,355	,539 > 0,355	Valid
6.	,709	0,355	,709 > 0,355	Valid
7.	,829	0,355	,829 > 0,355	Valid
8.	,342	0,355	,342 < 0,355	Tidak Valid
9.	,722	0,355	,722 > 0,355	Valid
10.	,622	0,355	,622 > 0,355	Valid
11.	,256	0,355	,256 < 0,355	Tidak Valid
12.	,829	0,355	,829 > 0,355	Valid
13.	,065	0,355	,065 < 0,355	Tidak Valid
14.	,319	0,355	,319 < 0,355	Tidak Valid
15.	,663	0,355	,663 > 0,355	Valid
16.	,414	0,355	,414 > 0,355	valid
17.	,677	0,355	,677 > 0,355	Valid
18.	,724	0,355	,724 > 0,355	Valid
19.	,595	0,355	,595 > 0,355	Valid
20.	,532	0,355	,532 > 0,355	Valid
21.	,793	0,355	,793 > 0,355	Valid
22.	,709	0,355	,709 > 0,355	Valid
23.	,322	0,355	,322 < 0,355	Tidak Valid
24.	,810	0,355	,810 > 0,355	Valid
25.	,687	0,355	,687 > 0,355	Valid

No Soal	r hitung	r tabel N = 31 A = 0,05	Kriteria r hitung > r tabel atau r hitung < r tabel	Validitas
26.	,699	0,355	,699 > 0,355	Valid
27.	,678	0,355	,678 > 0,355	Valid

Sumber: Hasil Uji Validitas Menggunakan IBM SPSS Statistics 29

Hasil pengolahan data melalui IBM SPSS Statistics 29 menunjukkan bahwa dari total 27 soal yang diujicobakan, terdapat 20 *item* yang memenuhi syarat validitas. Sebaliknya *item* soal nomor 2, 4, 8, 11, 13, 14, dan 23 dianggap tidak valid. Dengan demikian, diperoleh 20 *item* soal yang dinyatakan layak dijadikan sebagai instrumen pengumpulan data dan diimplementasikan sebagai instrumen final yang didistribusikan kepada kelas eksperimen serta kelas kontrol.

3.6.3 Uji Reliabilitas

Uji reliabilitas bertujuan guna mengukur tingkat keandalan instrumen. Pengujian ini untuk memastikan bahwa alat ukur yang digunakan memberikan hasil yang tetap ajek meskipun dilakukan pengukuran secara berulang kali dengan menggunakan instrumen yang sama (Amanda dkk., 2019: 183). Instrumen dinyatakan reliabel jika memiliki data yang stabil atau konsisten saat digunakan pada jenis maupun subjek penelitian yang sama. Uji reliabilitas dilaksanakan dengan memanfaatkan rumus Cronbach's Alpha. Dasar yang digunakan dalam penentuan reliabilitas ini yakni jika Cronbach's Alpha > 0,6 atau nilai *r hitung* > *r tabel*, maka instrumen dianggap reliabel (Forester dkk, 2024: 1814). Instrumen dianggap tidak reliabel jika nilai *r hitung* < *r tabel*.

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

Gambar 3.4 Rumus Uji Reliabilitas
(Sugiyono, 2019: 187)

Keterangan (Siregar, 2017: 58):

σ_t^2 : Varians total

$\sum \sigma_b^2$: Jumlah varians butir

k : Jumlah butir pertanyaan

r_{11} : Koefisien reliabilitas instrumen

Tabel 3.7 Hasil Uji Reliabilitas

No	Scale Mean If Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Keterangan
1.	15.8182	45.966	.372	.907	Reliabel
2.	15.9697	48.280.	.042	.912	Reliabel
3.	15.7273	45.705	.407	.906	Reliabel
4.	15.5152	47.570	.157	.910	Reliabel
5.	15.5758	45.377	.488	.905	Reliabel
6.	15.4848	44.633	.676	.901	Reliabel
7.	15.5758	43.439	.806	.899	Reliabel
8.	15.6970	46.593	.276	.909	Reliabel
9.	15.7273	43.892	.684	.901	Reliabel
10.	15.6061	44.746	.576	.903	Reliabel
11.	15.9091	47.273	.190	.910	Reliabel
12.	15.5455	43.568	.806	.899	Reliabel
13.	15.8788	48.547	-.005	.914	Reliabel
14.	15.7576	46.752	.251	.909	Reliabel
15.	15.6667	44.354	.620	.902	Reliabel
16.	15.6667	46.104	.352	.907	Reliabel
17.	15.6061	44.371	.636	.902	Reliabel
18.	15.5152	44.383	..691	.901	Reliabel
19.	15.5758	45.002	.548	.903	Reliabel
20.	15.5455	45.506	.481	.905	Reliabel
21.	15.5152	43.945	.768	.900	Reliabel
22.	15.4848	44.633	.676	.901	Reliabel
23.	15.7879	46.735	.255	.909	Reliabel
24.	15.5455	43.693	.785	.899	Reliabel
25.	15.4545	44.943	.653	.902	Reliabel
26.	15.6667	44.104	.659	.901	Reliabel
27.	15.4848	44.820	.642	.902	Reliabel

Sumber: Hasil Uji Reliabilitas Menggunakan IBM SPSS Statistics 29

3.6.4 Uji Tingkat Kesukaran

Sudijono (2011: 370-372), mengungkapkan bahwa mutu suatu butir soal dapat ditentukan dari tingkat kesukarannya. Butir soal dapat dikatakan baik apabila tidak terlalu sulit dan juga tidak terlalu mudah, sehingga berada pada kategori kesukaran sedang atau cukup. Adapun rumus uji tingkat kesukaran yaitu:

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

Gambar 3.5 Rumus Uji Tingkat Kesukaran
(Sudijono, 2011: 372)

Keterangan:

P : Angka indeks kesukaran item

B : Banyaknya responden yang mampu menjawab benar

JS : Jumlah responden yang mengikuti tes

Tabel 3.8 Kriteria Penafsiran Uji Tingkat Kesukaran

Besarnya Angka Indeks Kesukaran	Keterangan
< 0.30	Sukar
0.30 – 0.70	Sedang
> 0.70	Mudah

Tabel 3.9 Hasil Uji Tingkat Kesukaran

No Soal	Mean	Keterangan
1	0.42	Sedang
2	0.27	Sukar
3	0.51	Sedang
4	0.72	Mudah
5	0.66	Sedang
6	0.75	Mudah
7	0.66	Sedang
8	0.54	Sedang
9	0.51	Sedang
10	0.63	Sedang
11	0.33	Sedang
12	0.69	Sedang

No Soal	Mean	Keterangan
13	0.36	Sedang
14	0.48	Sedang
15	0.57	Sedang
16	0.57	Sedang
17	0.63	Sedang
18	0.72	Mudah
19	0.66	Sedang
20	0.69	Sedang
21	0.72	Mudah
22	0.75	Mudah
23	0.45	Sedang
24	0.69	Sedang
25	0.78	Mudah
26	0.57	Sedang
27	0.75	Mudah

Sumber: Hasil Uji Tingkat Kesukaran Menggunakan IBM Statistics 29

Berdasarkan hasil analisis yang dilakukan terhadap 27 butir soal, diperoleh bahwa sebanyak 1 butir soal pada kategori sukar, 19 butir soal pada kategori sedang, serta 7 butir soal pada kategori mudah. Dengan demikian, sebagian besar butir soal berada pada tingkat sedang yang berarti memiliki kualitas yang baik dan layak digunakan dalam penelitian.

3.6.5 Uji Daya Beda

Daya pembeda suatu butir soal merupakan kemampuan item tersebut dalam membedakan peserta didik yang berkemampuan tinggi dengan yang berkemampuan rendah. Hal ini ditunjukkan dengan kondisi di mana sebagian besar peserta didik berkemampuan tinggi mampu menjawab soal dengan benar, sedangkan sebagian besar peserta didik berkemampuan rendah tidak dapat menjawab soal dengan benar. Daya pembeda dapat diketahui dengan melihat besar atau kecilnya nilai indeks diskriminasi. Indeks diskriminasi merupakan angka atau yang menunjukkan besar kecilnya angka indeks diskriminasi item (Sudijono, 2011:

385-387). Adapun rumus yang digunakan guna angka indeks diskriminasi item yaitu sebagai berikut:

$$D = P_A - P_B$$

Gambar 3.6 Rumus Uji Daya Beda
(Sudijono, 2011: 389)

Keterangan:

D : Diskriminasi power (angka indeks diskriminasi item)

P_A : Proporsi responden kelompok atas yang dapat menjawab dengan benar

P_B : Proporsi responden kelompok bawah yang dapat menjawab dengan benar

Tabel 3.10 Kriteria Penafsiran Uji Daya Beda

Besarnya Angka Indeks Diskriminasi	Keterangan
> 0.20	Jelek
$0.20 - 0.40$	Sedang
$0.40 - 0.70$	Baik
$0.70 - 1.00$	Sangat Baik

Tabel 3.11 Hasil Uji Daya Beda

No soal	r hitung (Correlated Item-Total Correlation)	Keterangan
1	0.372	Sedang
2	0.042	Baik
3	0.407	Baik
4	0.157	Jelek
5	0.488	Baik
6	0.676	Baik
7	0.806	Sangat Baik
8	0.276	Sedang
9	0.684	Baik
10	0.576	Baik
11	0.190	Jelek

No soal	r hitung (Correlated Item-Total Correlation)	Keterangan
12	0.806	Sangat Baik
13	-0.005	Jelek
14	0.251	Sedang
15	0.620	Baik
16	0.352	Sedang
17	0.636	Baik
18	0.691	Baik
19	0.548	Baik
20	0.481	Baik
21	0.768	Baik
22	0.676	Baik
23	0.255	Sedang
24	0.785	Baik
25	0.653	Baik
26	0.659	Baik
27	0.642	Baik

Sumber: Hasil Uji daya Beda Menggunakan IBM SPSS 29

Berdasarkan hasil analisis yang dilakukan terhadap 27 butir soal, diperoleh bahwa sebanyak 2 butir soal pada kategori sangat baik, 17 butir soal pada kategori baik, 5 butir soal pada kategori sedang, serta 3 butir soal pada kategori jelek. Dengan demikian, sebagian besar butir soal memiliki kualitas yang baik dan layak digunakan dalam penelitian.

3.7 Teknik Analisis Data

3.7.1 Uji Normalitas

Uji normalitas bertujuan memastikan bahwa data sampel berdistribusi normal sehingga layak digunakan dalam analisis statistik. Instrumen penelitian dianggap memenuhi syarat apabila data yang diperoleh menunjukkan pola distribusi normal.

Uji normalitas memanfaatkan uji *Shapiro-Wilk* sebab sampel tidak mencapai 100. Keputusan diambil berdasarkan pada tingkat signifikansi sebesar 5% atau 0.05.

Kriteria pengambilan keputusan dalam uji normalitas merujuk pada nilai signifikansi (nilai probabilitas), di mana nilai signifikansi $> 0,05$, mengindikasikan bahwa data memiliki sebaran normal, sedangkan nilai signifikansi $< 0,05$ menunjukkan sebaliknya. Hasil uji yang mengindikasikan bahwa data sebaran normal menjadi dasar untuk melakukan uji homogenitas pada tahap selanjutnya.

3.7.2 Uji Homogenitas

Tahap ini dilaksanakan untuk menguji kesamaan varians antar sampel guna memastikan bahwa data berasal dari populasi dengan karakteristik yang homogen. Melalui penerapan uji Levene, kriteria pengambilan keputusan ditetapkan berdasarkan nilai signifikansi, apabila diperoleh nilai signifikansi < 0.05 , maka dapat disimpulkan adanya perbedaan varians atau data dinyatakan tidak homogen. Namun, uji homogenitas tidak dilakukan karena pengujian tersebut merupakan prasyarat pada dalam analisis statistik parametrik (Usmadi, 2020: 51).

3.7.3 Uji Hipotesis

Uji hipotesis merupakan asumsi atau dugaan awal yang kebenarannya masih perlu dibuktikan (Siregar, 2017: 38). Penelitian ini menggunakan *Mann Whitney U test* guna membandingkan rata-rata antara dua kelompok yang tidak saling berhubungan dengan taraf signifikansi 0,05. Keputusan diambil berdasarkan nilai Sig. (2 tailed), yaitu H_0 diterima apabila nilai signifikansi $> 0,05$ yang menunjukkan rata-rata kedua kelompok sama, sedangkan H_0 ditolak jika nilai Sig. (2-tailed) $< 0,05$ yang menunjukkan adanya perbedaan rata-rata.

3.7.4 Uji N-Gain

Perhitungan N-Gain bertujuan guna mengetahui tingkat efektivitas penggunaan media atau metode pembelajaran dengan membandingkan hasil *pre-test* serta *post-test* agar dapat diketahui besarnya peningkatan kemampuan berpikir sejarah yang muncul pada kelas eksperimen serta kelas kontrol setelah diberikan perbedaan *treatment* (Condong dkk., 2026: 18).

3.8 Langkah-Langkah Penelitian

3.8.1 Tahap Perencanaan

Peneliti mencari dan menentukan fokus penelitian terhadap permasalahan yang akan dikaji terlebih dahulu pada tahap ini. Peneliti kemudian mengajukan topik penelitian kepada dosen pembimbing satu dan dua hingga memperoleh persetujuan. Peneliti selanjutnya menyusun proposal yang berisi konsep serta rencana pelaksanaan penelitian setelah topik disetujui. Proposal penelitian tersebut dipresentasikan dalam sidang proposal untuk memperoleh masukan dan persetujuan, kemudian peneliti mempersiapkan berbagai kebutuhan penelitian, termasuk menyusun instrumen pengumpulan data.

3.8.2 Tahap Pelaksanaan

Peneliti mulai melaksanakan kegiatan penelitian di sekolah pada tahap ini. Penelitian dimulai dengan memberikan *pre-test* kepada peserta didik, kemudian dilanjutkan dengan penerapan media pembelajaran Flippity di kelas XI.11 SMAN 5 Tasikmalaya. Peserta didik diberikan *post-test* setelah pembelajaran selesai. Data hasil tes kemudian dikumpulkan dan dianalisis dengan tujuan untuk melihat pengaruh penggunaan media Flippity terhadap kemampuan berpikir sejarah peserta didik.

3.8.3 Tahap Akhir

Peneliti melakukan penyusunan laporan penelitian sebagai tahap akhir dari rangkaian proses penelitian pada tahap ini. Peneliti menganalisis serta mengolah data yang telah diperoleh, kemudian menjawab pertanyaan yang dirumuskan dalam rumusan masalah. Peneliti menarik kesimpulan yang menggambarkan temuan utama dari penelitian yang telah dilakukan berdasarkan hasil analisis tersebut.

3.9 Waktu dan Tempat Penelitian

3.9.1 Waktu Penelitian

Studi ini berlangsung selama 7 bulan yakni dari bulan Oktober sampai April 2026. Rincian waktu kegiatan penelitiannya, sebagai berikut:

Tabel 3.12 Rincian Waktu Kegiatan Penelitian

No	Rincian Kegiatan	Bulan						
		Okt	Nov	Des	Jan	Feb	Mar	April
1	Tahap Pra-penelitian							
2	Tahap Pembuatan Proposal Penelitian							
3	Seminar Proposal Penelitian							
4	Mengurus Administrasi Penelitian							
5	Melaksanakan Observasi							
6	Menyusun Perangkat Penelitian							
7	Melaksanakan Penelitian							
8	Mengumpulkan Data							
9	Mengolah Data							
10	Ujian Skripsi							

3.9.2 Tempat Penelitian

Penelitian ini bertempat di SMAN 5 Tasikmalaya yang berlokasi di Jl. Tentara Pelajar No. 58, Kec. Tawang, Kota Tasikmalaya, Jawa Barat, 46113 dan akan dilaksanakan pada kelas XI.11 SMAN 5 Tasikmalaya tahu ajaran 2025/2026.