

The Influence Of The *DeepThink* Learning Model On Students' Critical Thinking Skills In History At Sma Negeri 1 Parakan

Pengaruh Model Pembelajaran *DeepThink* Terhadap Keterampilan Berpikir Kritis Siswa Dalam Mata Pelajaran Sejarah Di Sma Negeri 1 Parakan

Choirizki Amanda^{1*}, Ba'in², Fendi Aditiya³

^{1,2,3} Jurusan Sejarah, Fakultas Ilmu Sosial dan Ilmu Politik, Universitas Negeri Semarang

(*) Corresponding Author:

choirizkiamaanda@students.unnes.ac.id

HP : 085643219215

How to Cite: Amanda. C., Bain., Aditiya. F. (2026). The Influence Of The *DeepThink* Learning Model On Students' Critical Thinking Skills In History At Sma Negeri 1 Parakan. *Santhet*, 2(2), 1-5. doi: 10.36526/js.v3i2.9289

Received : 30-07-2026

Revised : 31-08-2026

Accepted : 30-09-2026

Keywords:

DeepThink Model,
Critical Thinking Skills,
History Learning,
HOTS.

Abstract

This study aimed to investigate the effectiveness was to examine how the *DeepThink* learning model affects students' critical thinking skills at SMA Negeri 1 Parakan. This study focused on students' critical thinking skills, which are largely influenced by surface learning and a lack of critical thinking skills. The research method used a quasi-experimental approach with a Nonequivalent Control Group design. The study sample included two classes, selected through intact group sampling. The experimental class used the *DeepThink* model, while the control class used conventional methods. The data collection instrument consisted of 20 multiple-choice questions that fall into the HOTS category and align with the critical thinking indicators outlined by Facione (2015). Data were analyzed using the Shapiro-Wilk normality test, Levene's homogeneity test, Wilcoxon Signed-Rank test, Mann-Whitney U test, and N-Gain test. The Wilcoxon test showed a significant increase in the experimental class ($Z = -2.103$; $p = 0.035$) and the control class ($Z = -3.972$; $p = 0.000$). The Mann-Whitney U test showed a significance level of 0.128 ($p > 0.05$), meaning there was no significant difference between the *Posttest* results of the experimental and control classes. The N-Gain analysis showed that the experimental class experienced greater improvement than the control class. Therefore, the use of the *DeepThink* model can help improve students' critical thinking skills, although the level of difference is not yet very significant.

PENDAHULUAN

Pembelajaran sejarah di tingkat sekolah menengah tidak lagi cukup berorientasi pada penguasaan fakta, tokoh, tanggal, dan kronologi peristiwa. Siswa perlu mampu memahami informasi secara konseptual, menganalisis bukti, mengevaluasi berbagai perspektif, menarik kesimpulan, dan menjelaskan hubungan sebab-akibat secara logis. Kebutuhan tersebut menempatkan critical thinking sebagai salah satu kompetensi penting dalam pembelajaran abad ke-21. (Facione, 2015) menjelaskan bahwa berpikir kritis mencakup interpretation, analysis, evaluation, inference, explanation, dan self-regulation. Dalam pembelajaran sejarah, kemampuan tersebut diperlukan agar siswa tidak hanya mengingat peristiwa, tetapi mampu memahami dan menilai informasi sejarah secara kritis. Kebutuhan terhadap keterampilan berpikir kritis sejalan dengan arah pembelajaran dalam Kurikulum Merdeka yang menekankan pengembangan kompetensi dan pemahaman siswa (Kemdikbudristek, 2022). Salah satu pendekatan yang relevan adalah deep learning, yaitu pembelajaran yang mendorong keterlibatan kognitif dan pemahaman yang lebih mendalam dibandingkan surface learning (J. Biggs et al., 2022). Dalam konteks pendidikan Indonesia, pendekatan deep learning diarahkan pada pengalaman belajar yang bermakna, sadar, dan menyenangkan (Khasanah & dkk., 2025; Mustapa et al., 2025). Sejumlah

penelitian juga menunjukkan relevansinya dengan pengembangan literasi sejarah, berpikir kritis, dan berpikir historis siswa (Fitroh, 2025; Hasibuan, 2025; Suryana & Jono, 2025).

Tuntutan tersebut belum sepenuhnya tercermin dalam praktik pembelajaran sejarah. Observasi awal di SMA Negeri 1 Parakan menunjukkan bahwa sebagian siswa masih cenderung menghafalkan fakta dan kronologi sejarah serta belum optimal dalam menganalisis sumber sejarah. Partisipasi siswa dalam mengkritisi informasi juga masih terbatas, sementara hasil penilaian menunjukkan kemampuan kognitif pada tingkat C4–C6 belum optimal dan belum sepenuhnya memenuhi Kriteria Ketercapaian Tujuan Pembelajaran (KKTP). Kondisi ini menunjukkan adanya kesenjangan antara tuntutan pembelajaran sejarah yang berorientasi pada keterampilan berpikir tingkat tinggi dan praktik pembelajaran yang masih cenderung berorientasi pada hafalan. Salah satu model yang dirancang untuk menjawab kebutuhan tersebut adalah *DeepThink*. (Basri et al., 2024) mengembangkan model *DeepThink* sebagai model pembelajaran sejarah yang berorientasi pada pengembangan keterampilan berpikir kritis melalui tujuh tahapan, yaitu pembukaan, apersepsi, analisis, inferensi, interpretasi dan penjelasan, evaluasi dan regulasi diri, serta penutupan. Sintaks tersebut memiliki keterkaitan dengan indikator berpikir kritis (Facione, 2015), karena siswa diarahkan untuk menginterpretasikan informasi, menganalisis bukti, menarik inferensi, mengevaluasi informasi, menjelaskan hasil pemikiran, dan melakukan regulasi diri. Kesesuaian teoritis tersebut belum cukup untuk memastikan efektivitas *DeepThink* dalam pembelajaran nyata. Penelitian (Basri et al., 2024) memberikan dasar pengembangan model, tetapi bukti empiris mengenai penerapannya pada pembelajaran sejarah di tingkat SMA masih perlu diperluas. Penelitian (Andriani et al., 2021) menunjukkan bahwa inovasi pembelajaran sejarah dapat meningkatkan kemampuan berpikir kritis siswa, sementara penelitian mengenai deep learning lebih banyak membahas relevansi dan penerapannya secara umum (Nurjanah & Suryadi, 2025; Waruwu & Setiawati, 2025; Wathon, 2024). Dengan demikian, masih terdapat research gap berupa keterbatasan bukti empiris mengenai penerapan *DeepThink* secara langsung dalam pembelajaran sejarah SMA dan perbandingannya dengan pembelajaran konvensional.

Kebaruan penelitian ini terletak pada pengujian empiris model *DeepThink* dalam pembelajaran sejarah pada jenjang SMA melalui desain quasi-experimental dengan kelompok eksperimen dan kelompok kontrol. Penelitian ini menganalisis perubahan keterampilan berpikir kritis siswa sebelum dan sesudah pembelajaran serta membandingkan hasil kedua kelompok berdasarkan indikator berpikir kritis Facione. Dengan demikian, penelitian ini diharapkan dapat memberikan bukti empiris mengenai penerapan *DeepThink* sekaligus memperkaya kajian pembelajaran sejarah yang berorientasi pada keterampilan berpikir tingkat tinggi. Penelitian ini bertujuan untuk menguji pengaruh model pembelajaran *DeepThink* terhadap keterampilan berpikir kritis siswa dalam mata pelajaran sejarah di SMA Negeri 1 Parakan. Penelitian ini menganalisis perubahan kemampuan berpikir kritis pada kelompok eksperimen dan kontrol, membandingkan hasil kedua kelompok setelah pembelajaran, serta mengidentifikasi tingkat peningkatan melalui analisis N-Gain.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan desain *quasi-experimental* berbentuk *Nonequivalent Control Group Design*. Penelitian dilaksanakan pada siswa kelas X SMA Negeri 1 Parakan tahun ajaran 2025/2026. Sampel dipilih menggunakan teknik intact group sampling, yaitu pemilihan kelompok berdasarkan kelas yang telah terbentuk tanpa mengacak siswa secara individual. Teknik tersebut digunakan untuk mempertahankan struktur kelas dan tidak mengganggu proses pembelajaran yang telah berlangsung. Berdasarkan pertimbangan karakteristik akademik dan hasil *Pretest*, satu kelas ditetapkan sebagai kelompok eksperimen dan satu kelas sebagai kelompok kontrol. Kelompok eksperimen memperoleh pembelajaran menggunakan model *DeepThink*, sedangkan kelompok kontrol memperoleh pembelajaran konvensional. Kedua kelompok mempelajari materi yang sama, yaitu Kerajaan Mataram Islam.

Data hasil penelitian menunjukkan bahwa jumlah data yang dianalisis berbeda pada setiap pengukuran, yaitu 30 siswa pada *Pretest* dan 34 siswa pada *Posttest* kelompok eksperimen serta 31 siswa pada *Pretest* dan 33 siswa pada *Posttest* kelompok kontrol. Instrumen penelitian berupa tes tertulis untuk mengukur keterampilan berpikir kritis siswa. Tes disusun dalam bentuk pilihan ganda berbasis higher-order thinking skills (HOTS) dan mengacu pada enam indikator berpikir kritis menurut (Facione, 2015), yaitu interpretation, analysis, evaluation, inference, explanation, dan self-regulation. Setiap butir dirancang untuk mengukur kemampuan siswa dalam memahami dan menafsirkan informasi, menganalisis hubungan atau bukti, mengevaluasi informasi, menarik kesimpulan, menjelaskan hasil pemikiran, serta merefleksikan proses berpikir. Dalam naskah penelitian, instrumen utama disebut terdiri atas 20 butir soal yang digunakan untuk *Pretest* dan *Posttest*. Instrumen diuji melalui validitas isi dan validitas empiris. Validitas isi dilakukan melalui penilaian ahli untuk memastikan kesesuaian butir soal dengan indikator keterampilan berpikir kritis. Selanjutnya, validitas empiris diuji menggunakan korelasi Pearson Product Moment dengan melihat hubungan antara skor setiap butir dan skor total. Berdasarkan hasil pengujian, seluruh 20 butir soal memiliki nilai signifikansi kurang dari 0,05 sehingga dinyatakan valid. Nilai koefisien korelasi setiap butir berada pada rentang 0,325–0,747.

Reliabilitas instrumen diuji menggunakan koefisien Cronbach's Alpha. Hasil pengujian menunjukkan nilai Cronbach's Alpha sebesar 0,745. Nilai tersebut berada pada rentang $0,60 < r_{11} \leq 0,80$ yang dikategorikan memiliki tingkat reliabilitas tinggi berdasarkan klasifikasi yang digunakan dalam penelitian. Dengan demikian, instrumen memiliki konsistensi internal yang memadai untuk digunakan dalam pengukuran keterampilan berpikir kritis siswa. Pengumpulan data dilakukan melalui tiga tahap, yaitu pra-lapangan, lapangan, dan pasca-lapangan. Pada tahap pra-lapangan dilakukan observasi awal dan penentuan kelas eksperimen serta kelas kontrol, kemudian kedua kelompok diberikan *Pretest*. Pada tahap lapangan, kelompok eksperimen memperoleh pembelajaran menggunakan model *DeepThink*, sedangkan kelompok kontrol memperoleh pembelajaran konvensional yang berpusat pada guru. Kedua kelompok menerima materi pembelajaran yang sama. Setelah perlakuan selesai, kedua kelompok diberikan *Posttest* untuk mengetahui perubahan keterampilan berpikir kritis siswa. Analisis deskriptif digunakan untuk menggambarkan nilai minimum, maksimum, rerata, dan standar deviasi hasil *Pretest* dan *Posttest*. Selanjutnya, uji normalitas dilakukan menggunakan Shapiro–Wilk, sedangkan uji homogenitas menggunakan Levene. Hasil uji normalitas menunjukkan bahwa data tidak berdistribusi normal sehingga pengujian hipotesis dilakukan menggunakan statistik nonparametrik. Uji Wilcoxon Signed-Rank digunakan untuk mengetahui perbedaan skor *Pretest* dan *Posttest* dalam masing-masing kelompok, sedangkan Mann–Whitney U digunakan untuk membandingkan hasil *Posttest* antara kelompok eksperimen dan kelompok kontrol. Selain itu, peningkatan keterampilan berpikir kritis dianalisis menggunakan N-Gain dengan kategori rendah apabila $g < 0,30$, sedang apabila $0,30 \leq g < 0,70$, dan tinggi apabila $g \geq 0,70$. Alur analisis penelitian dilakukan mulai dari pengumpulan data *Pretest* dan *Posttest*, analisis deskriptif, pengujian normalitas dan homogenitas, pemilihan statistik nonparametrik, pengujian Wilcoxon dan Mann–Whitney U, hingga analisis N-Gain untuk mengetahui tingkat peningkatan keterampilan berpikir kritis siswa.

HASIL DAN PEMBAHASAN

Tabel 1. Hasil Uji Validitas Instrumen Soal *Pretest-Posttest*

No	Butir Soal	r Hitung	Sig.	Keputusan
1	Soal_1	0,441	0,000	Valid
2	Soal_2	0,493	0,000	Valid
3	Soal_3	0,497	0,000	Valid
4	Soal_4	0,343	0,006	Valid
5	Soal_5	0,399	0,001	Valid
6	Soal_6	0,325	0,009	Valid
7	Soal_7	0,747	0,000	Valid
8	Soal_8	0,696	0,000	Valid
9	Soal_9	0,353	0,005	Valid
10	Soal_10	0,455	0,000	Valid
11	Soal_11	0,733	0,000	Valid
12	Soal_12	0,444	0,000	Valid
13	Soal_13	0,569	0,000	Valid
14	Soal_14	0,583	0,000	Valid
15	Soal_15	0,733	0,000	Valid
16	Soal_16	0,649	0,000	Valid
17	Soal_17	0,583	0,000	Valid
18	Soal_18	0,709	0,000	Valid
19	Soal_19	0,626	0,000	Valid
20	Soal_20	0,543	0,000	Valid

Sumber: Data primer diolah, 2026

Selain menguji validitas, instrumen penelitian juga diperiksa reliabilitasnya menggunakan koefisien Cronbach Alpha untuk mengetahui seberapa konsisten instrumen tersebut dalam mengukur variabel penelitian.

Gambar 1. Uji Reliabilitas

<i>Reliability Statistics</i>	
<i>Cronbach's Alpha</i>	<i>N of Items</i>
.745	26

Hasil pengujian yang ditunjukkan pada gambar 1. menunjukkan bahwa nilai Cronbach Alpha adalah 0,745. Skor tersebut termasuk dalam kategori reliabilitas yang dapat diterima, yang berarti bahwa instrumen tersebut memiliki tingkat konsistensi internal yang baik dan stabil, sehingga cocok untuk digunakan dalam penelitian. Penentuan tingkat reliabilitas dilakukan dengan membandingkan nilai koefisien yang diperoleh dengan kriteria klasifikasi reliabilitas (Andriani et al., 2021; Emerentiana, 2024).

Tabel 2. Klasifikasi Tingkat Kriteria Alpha Cronbach

Rentang Nilai Koefisien (r11)	Tingkat Reliabilitas
0,80 < r11 1,00	Sangat Tinggi
0,60 < r11 0,80	Tinggi
0,40 < r11 0,60	Sedang
0,20 < r11 0,40	Rendah

-1,00 $r_{11} < 0,20$

Sangat Rendah (Tidak Reliabel)

Berdasarkan tabel tersebut, dapat disimpulkan bahwa semua item dalam instrumen penelitian ini valid dan reliabel. Oleh karena itu, instrumen tersebut cocok untuk digunakan sebagai alat pengumpulan data dalam penelitian untuk mengukur kemampuan berpikir kritis siswa secara akurat dan dapat diandalkan. Penelitian ini dilakukan dalam tiga tahap, yaitu pra-lapangan, lapangan, dan pasca lapangan. Tahap pra-lapangan merupakan tahap observasi untuk pemilihan kelas yang akan menerima eksperimen (kelas eksperimen) dan kelas kontrol, ditahap ini juga dilakukan *Pretest* pada kedua kelas tersebut. Dalam tahap lapangan, kelas eksperimen menerima pengajaran menggunakan model *DeepThink*, sedangkan kelas kontrol menggunakan pembelajaran yang berpusat pada guru atau konvensional (ceramah). Pengukuran dilakukan selama dua kali, yaitu sebelum perlakuan (*Pretest*) dan setelah perlakuan (*Posttest*).

Tabel 3. Klasifikasi Perolehan Nilai *Gain Score*

Nilai N-Gain	Kategori
$g \geq 0,70$	Tinggi
$0,30 \leq g < 0,70$	Sedang
$g < 0,30$	Rendah

(Hake, 1999)

Analisis Deskriptif

Analisis deskriptif dilakukan untuk memberikan gambaran umum tentang kemampuan berpikir kritis siswa sebelum dan sesudah menerima perlakuan pembelajaran. Analisis ini bertujuan untuk memahami data penelitian dengan melihat nilai minimum, maksimum, rata-rata, dan standar deviasi untuk kelompok eksperimen dan kelompok kontrol. Dengan analisis deskriptif, peneliti dapat melihat dengan jelas bagaimana kemampuan berpikir kritis siswa berkembang sebelum melakukan pengujian statistik lebih lanjut. Ringkasan statistik deskriptif untuk kedua kelompok penelitian ditunjukkan pada Tabel 4.

Tabel 4. Statistik Deskriptif *Pretest* dan *Posttest* Kelas Eksperimen dan Kontrol

Kelompok	Data	N	Min	Maks	Rerata	StD
Eksperimen	<i>Pretest</i>	30	35	100	84,83	19,41
Eksperimen	<i>Posttest</i>	34	55	100	92,65	8,19
Kontrol	<i>Pretest</i>	31	10	90	66,77	22,90
Kontrol	<i>Posttest</i>	33	65	100	88,94	10,21

Sumber: Data primer diolah, 2026

Berdasarkan Tabel 4, terlihat bahwa pada tahap *Pretest*, kelas eksperimen memiliki nilai rata-rata 84,83 dengan simpangan baku 19,41, sedangkan kelas kontrol memiliki nilai rata-rata 66,77 dengan simpangan baku 22,90. Hasil ini menunjukkan bahwa kemampuan berpikir kritis awal siswa di kelas eksperimen umumnya lebih tinggi daripada siswa di kelas kontrol. Perbedaan kemampuan awal ini kemudian didukung oleh hasil uji

homogenitas *Pretest*, yang menunjukkan perbedaan signifikan antar kelompok ($p = 0,002 < 0,05$). Namun, varians kedua kelompok dianggap homogen dengan nilai F sebesar 1,235 dan signifikansi 0,271. Kondisi ini menunjukkan bahwa kedua kelompok memiliki distribusi data yang serupa meskipun terdapat perbedaan kemampuan awal rata-rata mereka. Setelah diberikan perlakuan pengajaran, kedua kelompok menunjukkan peningkatan pada hasil *Posttest* mereka. Kelas eksperimen memiliki skor rata-rata *Posttest* sebesar 92,65 dengan standar deviasi 8,19, sedangkan kelas kontrol memiliki skor rata-rata 88,94 dengan standar deviasi 10,21. Dimana terjadi peningkatan rata-rata pada kedua kelompok.

Hasil penelitian menunjukkan bahwa proses pembelajaran yang digunakan mampu meningkatkan kemampuan berpikir kritis siswa. Namun, kelas eksperimen menunjukkan hasil yang lebih baik daripada kelas kontrol. Selain itu, standar deviasi skor *Posttest* untuk kedua kelompok menurun dibandingkan dengan *Pretest*, yang menunjukkan bahwa kemampuan siswa setelah proses pembelajaran menjadi lebih konsisten. Hal ini menunjukkan bahwa metode pengajaran yang digunakan di kelas eksperimen memiliki efek positif dalam meningkatkan kemampuan berpikir kritis siswa secara lebih efektif daripada pengajaran di kelas kontrol.

Uji Normalitas

Uji normalitas dilakukan untuk mengetahui apakah data penelitian berdistribusi normal atau tidak. Dalam penelitian ini, digunakan metode *Shapiro-Wilk* untuk menguji normalitas karena ukuran sampel relatif kecil, dan metode ini lebih akurat untuk memeriksa distribusi data pada sampel terbatas. Pengujian dilakukan pada data *Pretest* dan *Posttest* untuk kelas eksperimen dan kelas kontrol. Hasil pengujian menunjukkan bahwa semua data penelitian menerima nilai signifikansi (Sig.) 0,000 atau kurang dari 0,05 ($p < 0,05$). Hal ini dapat disimpulkan bahwa data *Pretest* dan *Posttest* untuk kedua kelompok tidak berdistribusi normal. Pada kelas eksperimen, baik data *Pretest* maupun *Posttest* menunjukkan nilai signifikansi di bawah 0,05, yang berarti bahwa keduanya tidak memenuhi asumsi normalitas. Hasil yang sama diamati pada kelompok kontrol, di mana data *Pretest* dan *Posttest* menunjukkan distribusi abnormal. Kondisi ini menunjukkan bahwa data penelitian cenderung tidak merata dan tidak berdistribusi normal.

Akibatnya, penelitian ini tidak dapat dianalisis menggunakan uji parametrik seperti uji t . Berdasarkan hasil uji normalitas, langkah selanjutnya dalam pengujian hipotesis dalam penelitian ini dilakukan dengan menggunakan statistik nonparametrik, *Wilcoxon Signed-Rank Test* dan Uji *Mann-Whitney U*. Menurut Sugiyono (2019, hlm.166), jika data penelitian tidak memenuhi asumsi normalitas, lebih baik menggunakan teknik statistik nonparametrik untuk analisis data. Dengan demikian, hasil uji akan tetap valid dan dapat dibenarkan secara ilmiah. Oleh karena itu, penggunaan uji nonparametrik dalam penelitian ini telah mengikuti prosedur analisis statistik yang benar.

Uji Homogenitas

Uji ini bertujuan untuk memeriksa apakah kedua kelompok penelitian memiliki tingkat keragaman data yang serupa baik sebelum maupun sesudah perlakuan diberikan. Hasil uji homogenitas pada data *Pretest* menunjukkan nilai signifikansi sebesar 0,271, yang lebih besar dari 0,05 ($p > 0,05$). Hal ini berarti varians data *Pretest* baik untuk kelas eksperimen maupun kelas kontrol dianggap homogen. Hal ini menunjukkan bahwa kedua

kelompok memiliki tingkat keragaman keterampilan awal yang serupa sebelum proses pembelajaran dimulai. Namun, hasil perbedaan rata-rata *Pretest* menunjukkan nilai signifikansi sebesar 0,002, yang lebih kecil dari 0,05 ($p < 0,05$). Hal ini berarti terdapat perbedaan yang signifikan dalam kemampuan awal antara kelas eksperimen dan kelas kontrol sebelum perlakuan diberikan. Sementara itu, hasil uji homogenitas pada data *Posttest* menunjukkan nilai signifikansi sebesar 0,013, yang lebih kecil dari 0,05 ($p < 0,05$). Hal ini berarti tidak terdapat perbedaan yang signifikan antara hasil tes akhir kelas eksperimen dan kelas kontrol. Dengan demikian, penggunaan model pembelajaran *DeepThink* belum menunjukkan dampak yang signifikan terhadap kemampuan berpikir kritis siswa.

Uji Wilcoxon Signed-Rank Test

Uji *Wilcoxon Signed-Rank Test* dilakukan untuk melihat apakah ada peningkatan yang signifikan dalam keterampilan berpikir kritis siswa sebelum dan sesudah perlakuan yang diberikan kepada setiap kelompok penelitian. Pengujian ini dilakukan karena data penelitian tidak memenuhi asumsi normalitas, sehingga analisisnya menggunakan statistik nonparametrik. Uji *Wilcoxon Signed-Rank* adalah digunakan untuk menguji perbedaan antara dua kondisi pengukuran pada sampel yang sama, terutama saat data tidak berdistribusi normal. Dalam penelitian ini, uji *Wilcoxon* digunakan untuk membandingkan hasil *Pretest* dan *Posttest* pada kelas eksperimen dan kelas kontrol. Tujuannya adalah untuk mengetahui apakah ada peningkatan keterampilan berpikir kritis siswa setelah proses pembelajaran dilakukan.

Tabel 5. Hasil Uji Wilcoxon Signed-Rank Kelas Eksperimen dan Kelas Kontrol

Kelas	Z	Asymp. Sig. (2-tailed)	Keterangan
Eksperimen	-2,103	0,035	Terdapat peningkatan signifikan
Kontrol	-3,972	0,000	Terdapat peningkatan signifikan

Sumber: Data primer diolah, 2026

Dari hasil pengujian yang ada pada **Tabel 5.** terlihat bahwa kelas eksperimen mendapatkan nilai Z sebesar -2,103 dengan nilai Asymp. Sig. (2- tailed) sebesar 0,035. Nilai signifikansi itu lebih kecil dari 0,05 ($p < 0,05$), jadi bisa disimpulkan bahwa terdapat perbedaan yang signifikan antara hasil *Pretest* dan *Posttest* di kelas eksperimen. Hasil ini menunjukkan bahwa setelah menggunakan model pembelajaran *DeepThink*, kemampuan berpikir siswa meningkat secara signifikan. Peningkatan ini menunjukkan bahwa model pembelajaran *DeepThink* dapat membantu siswa dalam meningkatkan kemampuan siswa untuk menganalisis, mempelajari, dan menyelesaikan masalah dengan cara yang lebih kritis selama proses pembelajaran. Sementara itu, di kelas kontrol, nilai Z yang didapat adalah -3,972 dengan nilai Asymp. sig. (2- tailed) sebesar 0,000. Nilai itu juga lebih kecil dari 0,05 ($p < 0,05$), yang berarti terdapat perbedaan yang signifikan antara nilai *Pretest* dan *Posttest* pada kelas kontrol. Dengan demikian, kelas kontrol juga menunjukkan peningkatan dalam keterampilan berpikir kritis setelah proses pembelajaran berlangsung. Peningkatan

yang terlihat pada kelas kontrol menunjukkan bahwa metode pembelajaran konvensional yang digunakan tetap berperan dalam meningkatkan kemampuan berpikir kritis siswa, meskipun tidak menggunakan model pembelajaran *DeepThink*.

Hasil dari uji *Wilcoxon Signed-Rank Test* menunjukkan bahwa kedua kelompok yang diteliti mengalami peningkatan yang signifikan dalam keterampilan berpikir kritis setelah proses pembelajaran. Namun, peningkatan yang terjadi di kelas eksperimen menunjukkan bahwa penggunaan model pembelajaran *DeepThink* memberikan dampak positif pada proses belajar dan membantu siswa mengembangkan kemampuan berpikir kritis mereka. Hasil ini menunjukkan bahwa penggunaan model pembelajaran yang berfokus pada aktivitas berpikir mendalam dapat membantu siswa menjadi lebih aktif dalam memahami, menganalisis, dan menyelesaikan materi pelajaran dengan cara yang kritis.

Uji Mann-Whitney U Test

Uji *Mann-Whitney U Test* dilakukan untuk melihat apakah terdapat perbedaan yang signifikan antara hasil *Posttest* kelas eksperimen dan kelas kontrol setelah mendapatkan perlakuan. Uji ini digunakan sebagai alternatif statistik nonparametrik untuk dua sampel yang tidak saling bergantung karena data penelitian tidak berdistribusi normal. Uji *Mann-Whitney U* digunakan untuk membandingkan hasil akhir dari dua kelompok penelitian. Tujuannya adalah untuk mengetahui bagaimana model pembelajaran yang diterapkan dapat mempengaruhi keterampilan berpikir kritis siswa. Hasil dari pengujian *Mann-Whitney U Test* ditampilkan dalam Tabel 6.

Tabel 6. Hasil Uji Mann-Whitney U Posttest Kelas Eksperimen dan Kelas Kontrol

Uji	Sig.	Keputusan	Keterangan
Mann-Whitney U (<i>Posttest</i>)	0,128	H_0 diterima	Tidak ada perbedaan signifikan

Sumber: Data primer diolah, 2026

Hasil uji *Mann-Whitney U* menunjukkan nilai signifikansi sebesar 0,128. Nilai ini lebih besar dari 0,05 ($p > 0,05$), sehingga kita menerima H_0 dan menolak H_1 . Oleh karena itu, dapat disimpulkan bahwa tidak ada perbedaan statistik yang signifikan antara hasil *Posttest* kelas eksperimen dan kelas kontrol. Hasil ini menunjukkan bahwa kemampuan berpikir kritis siswa di kedua kelompok berada pada tingkat yang hampir sama setelah proses pengajaran dan pembelajaran berlangsung. Namun, hasil ini perlu dipahami dengan cermat dan tidak langsung menunjukkan bahwa model pembelajaran *DeepThink* tidak berkinerja baik. Hal ini terjadi karena hasil uji *Wilcoxon Signed-Rank* sebelumnya menunjukkan bahwa baik kelas eksperimen maupun kelas kontrol mengalami peningkatan yang signifikan dalam kemampuan berpikir kritis mereka setelah proses pembelajaran.

Dengan kata lain, kedua cara belajar yang digunakan sama-sama membantu meningkatkan kemampuan berpikir kritis siswa. Tidak ada perbedaan yang berarti antara kelompok dapat dipengaruhi oleh beberapa hal, seperti perbedaan kemampuan awal siswa, durasi perlakuan yang cukup singkat, dan juga pengaruh pembelajaran di kelas kontrol yang dapat membantu meningkatkan keterampilan berpikir kritis siswa. Rata-rata hasil *Posttest* dari kedua kelompok yang sama-sama tinggi menunjukkan bahwa siswa di kedua kelas telah menunjukkan perkembangan kemampuan berpikir kritis yang baik setelah proses pembelajaran. Secara umum, hasil dari uji *Mann-Whitney U Test* menunjukkan bahwa

model pembelajaran *DeepThink* belum bisa memberikan perbedaan pengaruh yang nyata jika dibandingkan dengan pembelajaran di kelas kontrol. Walaupun demikian, model pembelajaran *DeepThink* masih menunjukkan potensi yang baik dalam meningkatkan kemampuan berpikir siswa, terutama terlihat dari peningkatan yang signifikan pada kelas eksperimen selama proses belajar.

Analisis N-Gain

Analisis N- Gain dilakukan untuk melihat seberapa efektif peningkatan keterampilan berpikir kritis siswa setelah menerima perlakuan pembelajaran. Perhitungan N-Gain digunakan untuk melihat seberapa besar peningkatan hasil belajar antara nilai *Pretest* dan *Posttest* dari setiap siswa. Dalam penelitian ini, analisis N-Gain mengacu pada rumus yang dikemukakan oleh Richard Hake pada tahun 1999. Menurut Hake, efektivitas peningkatan hasil belajar bagian menjadi tiga kategori. Kategori rendah jika nilai N-Gain kurang dari 0,30, kategori sedang jika nilai N-Gain berada di antara 0,30 hingga 0,70, dan kategori tinggi jika nilai N-Gain lebih dari 0,70. Dengan analisis ini, peneliti dapat memahami seberapa besar pembelajaran yang dilakukan dapat meningkatkan kemampuan berpikir kritis siswa di kelas eksperimen dan kelas kontrol.

Tabel 7. Ringkasan Analisis N-Gain Kelas Eksperimen dan Kelas Kontrol

Statistik	Kelas Eksperimen	Kelas Kontrol
N (valid)	19	30
N-Gain Minimum	-1,00	-1,50
N-Gain Maksimum	1,00	1,00
Rerata N-Gain	0,38	0,47
Kategori Dominan	Tinggi	Sedang-Tinggi

Sumber: Data primer diolah, 2026

Berdasarkan analisis yang ada pada **Tabel 7.** dapat dilihat bahwa rata-rata N-Gain pada kelas eksperimen adalah 0,38, yang termasuk dalam kategori sedang. Sementara itu, kelas kontrol mendapatkan rata-rata N-Gain sebesar 0,47, yang juga masuk dalam kategori sedang. Hasilnya menunjukkan bahwa kelompok kedua menunjukkan peningkatan dalam keterampilan berpikir kritis setelah proses pembelajaran berlangsung. Walaupun rata-rata N-Gain di kelas kontrol sedikit lebih tinggi dibandingkan di kelas eksperimen, perbedaan tersebut tidak secara langsung menunjukkan bahwa pembelajaran di kelas kontrol lebih efektif dibandingkan dengan model *DeepThink*. Kondisi ini bisa dipengaruhi oleh kemampuan awal siswa di kelas kontrol yang lebih rendah jika dibandingkan dengan kelas eksperimen, sehingga kemungkinan peningkatan skor di kelas kontrol menjadi lebih besar. Selain melihat rata-rata N-Gain, peningkatan penyebaran per individu di kedua kelompok juga menunjukkan ciri-ciri yang berbeda. Pada kelas eksperimen, sebagian besar siswa menunjukkan peningkatan yang lebih stabil pada kategori tinggi jika dibandingkan dengan kelas kontrol. Hal ini menunjukkan bahwa penggunaan model pembelajaran *DeepThink* dapat membantu siswa dalam meningkatkan kemampuan berpikir kritis mereka dengan lebih merata. Sementara itu, di kelas kontrol, peningkatan yang terjadi biasanya berada dalam kategori sedang hingga tinggi, tetapi distribusinya tidak mengancam kelas eksperimen.

Analisis ini juga menunjukkan bahwa ada beberapa siswa yang mendapatkan skor N-Gain negatif, baik di kelas eksperimen maupun di kelas kontrol. Skor negatif ini menunjukkan bahwa ada penurunan nilai *Posttest* jika dibandingkan dengan *Pretest* pada beberapa orang. Kondisi ini mungkin dipengaruhi oleh faktor-faktor nonakademik yang berada di luar kontrol penelitian, seperti motivasi belajar siswa, seberapa fokus siswa saat mengerjakan ujian, keadaan fisik dan mental, serta kondisi lingkungan belajar. Jadi, hasil N-Gain tidak hanya dipengaruhi oleh model pembelajaran yang digunakan, tetapi juga oleh berbagai faktor dari dalam diri siswa dan faktor dari luar yang mempengaruhi proses belajar mereka. Analisis hasil N-Gain menunjukkan bahwa kedua kelompok mengalami peningkatan dalam keterampilan berpikir kritis dengan kategori yang sedang. Meskipun demikian, model pembelajaran *DeepThink* masih menunjukkan potensi yang baik dalam meningkatkan kemampuan berpikir kritis siswa, terutama terlihat dari konsistensi peningkatan individu yang berada pada kategori tinggi di kelas eksperimen. Dengan demikian, penggunaan model pembelajaran *DeepThink* dapat dianggap sebagai salah satu pilihan dalam pembelajaran yang dapat membantu siswa mengembangkan keterampilan berpikir kritis siswa selama proses belajar.

PEMBAHASAN

Penelitian ini menjelaskan tentang seberapa efektif model pembelajaran *DeepThink* dalam pembelajaran sejarah, khususnya dalam meningkatkan kemampuan berpikir kritis siswa. Berdasarkan data penelitian, baik kelompok eksperimen maupun kelompok kontrol menunjukkan peningkatan yang signifikan dalam kemampuan berpikir kritis menurut hasil uji *Wilcoxon*. Hasil tersebut menunjukkan bahwa proses pembelajaran selama penelitian, baik menggunakan model *DeepThink* maupun metode pembelajaran konvensional, dapat membantu siswa meningkatkan kemampuan berpikir kritis mereka. Dengan kata lain, kegiatan pembelajaran yang dilakukan selama penelitian berhasil mendorong pertumbuhan kemampuan kognitif siswa dalam memahami, menganalisis, dan berbagi materi sejarah. Peningkatan yang terlihat pada kelas eksperimen selaras dengan landasan teoritis model *DeepThink* yang dikembangkan oleh (Basri et al., 2024). Model pembelajaran ini diciptakan untuk membantu siswa berpikir lebih dalam dengan melalui beberapa langkah, seperti interpretasi, analisis, evaluasi, inferensi, eksplanasi, dan regulasi diri (Facione, 2015). Langkah-langkah tersebut memerlukan keterlibatan keterampilan berpikir tingkat tinggi yang terkait dengan taksonomi Bloom. Dalam proses pembelajaran, siswa tidak hanya diminta untuk menghafal informasi sejarah, tetapi mereka juga didorong untuk memahami makna peristiwa, mengevaluasi keakuratan informasi, menganalisis sebab dan akibat, dan membentuk opini berdasarkan fakta yang mereka kumpulkan. Melalui proses ini, siswa menjadi lebih terlibat dalam mengembangkan pemahaman mereka dan merefleksikan pelajaran sejarah.

Dengan demikian, model *DeepThink* dapat membantu menciptakan proses pembelajaran sejarah yang lebih aktif dan bermakna. Temuan penelitian ini juga didukung oleh studi (Fitroh, 2025), yang menyatakan bahwa pembelajaran mendalam dalam mata pelajaran sejarah dapat membantu meningkatkan kemampuan berpikir kritis, keterampilan memecahkan masalah, dan kepekaan yang tinggi. Selain itu, penggunaan model *DeepThink* juga sejalan dengan arah kebijakan Kurikulum Merdeka, yang menekankan gagasan pembelajaran mendalam. Menurut (Basri et al., 2024; Suryana & Jono, 2025), pembelajaran

mendalam dalam Kurikulum Merdeka memiliki tiga pilar utama diantaranya yaitu *mindful learning*, *meaningful learning*, dan *joyful learning*. Ketiga pilar ini diharapkan dapat diintegrasikan di setiap tahap pendidikan sehingga siswa dapat belajar lebih sadar, memahami materi secara bermakna, dan memiliki pengalaman belajar yang menyenangkan. Penggunaan model pembelajaran inovatif dalam pendidikan sejarah untuk meningkatkan kemampuan berpikir kritis juga didukung oleh penelitian yang dilakukan (Andriani et al., 2021). Penelitian tersebut menemukan bahwa kemampuan berpikir kritis siswa meningkat sebesar 26% melalui penggunaan metode pengajaran sejarah yang inovatif. Hasil ini menunjukkan bahwa penggunaan metode pengajaran inovatif secara konsisten memiliki efek positif pada pengembangan keterampilan berpikir siswa. Oleh karena itu, penggunaan model *DeepThink* dapat menjadi salah satu cara untuk mengajarkan sejarah yang membantu meningkatkan keterampilan berpikir kritis siswa di sekolah.

Hasil uji *Mann-Whitney U* menunjukkan bahwa tidak ada perbedaan signifikan antara hasil *Posttest* kelas eksperimen dan kelas kontrol, hal ini tertulis dengan nilai signifikansi 0,128 ($p > 0,05$). Hasil tersebut harus dipahami dengan hati-hati dan tidak langsung menunjukkan bahwa model *DeepThink* dianggap tidak efektif. Terdapat beberapa keterbatasan dalam penelitian yang mungkin dapat memengaruhi hasil pengujian. Pertama, terdapat perbedaan yang cukup besar dalam kemampuan awal antara kelas eksperimen dan kelas kontrol sebelum perlakuan diberikan, seperti yang ditunjukkan oleh nilai signifikansi *Pretest* sebesar 0,002. Perbedaan kemampuan awal ini dapat memengaruhi pemahaman kita terhadap hasil *Posttest* antara kedua kelompok. Kedua, jumlah sampel yang relatif kecil dalam penelitian ini membatasi kekuatan statistik penelitian. Situasi ini membuat perbedaan pengaruh antar kelompok sulit diamati secara statistik. Ketiga, durasi perlakuan yang diberikan dalam penelitian ini masih cukup singkat, yang berarti model *DeepThink* belum sepenuhnya menunjukkan dampaknya pada kemampuan berpikir kritis siswa. Hasil penelitian menunjukkan bahwa model pembelajaran *DeepThink* memiliki potensi positif untuk meningkatkan kemampuan berpikir kritis siswa dalam pelajaran sejarah. Meskipun belum menunjukkan perbedaan yang signifikan dibandingkan dengan kelompok kontrol, model ini tetap dapat membantu siswa berpikir lebih dalam, reflektif, dan aktif selama proses pembelajaran. Oleh karena itu, model *DeepThink* dapat dianggap sebagai salah satu model pembelajaran yang efektif dan model pembelajaran inovatif yang relevan untuk mendukung pendekatan yang lebih kritis dan bermakna dalam pengajaran sejarah di era Kurikulum Merdeka.

PENUTUP

Penelitian ini menunjukkan bahwa penerapan model pembelajaran *DeepThink* meningkatkan keterampilan berpikir kritis siswa secara signifikan pada kelompok eksperimen ($Z = -2,103$; $p = 0,035$), tetapi tidak terdapat perbedaan yang signifikan antara kelompok eksperimen dan kontrol berdasarkan uji *Mann-Whitney U* ($p = 0,128$), sehingga *DeepThink* belum dapat dinyatakan lebih unggul secara statistik dibandingkan pembelajaran konvensional. Rata-rata N-Gain pada kelompok eksperimen sebesar 0,38 dan kontrol sebesar 0,47, keduanya berada pada kategori sedang. Temuan ini memberikan implikasi bahwa *DeepThink* berpotensi menjadi alternatif pembelajaran sejarah yang mendorong keterlibatan kognitif dan berpikir kritis, sekaligus memberikan kontribusi empiris mengenai penerapannya pada jenjang SMA. Namun, hasil penelitian perlu dipertimbangkan dengan adanya keterbatasan berupa perbedaan kemampuan

awal antarkelompok, jumlah sampel yang terbatas, dan durasi perlakuan yang relatif singkat. Penelitian selanjutnya disarankan menggunakan sampel yang lebih besar, durasi penerapan yang lebih panjang, pengendalian kemampuan awal yang lebih kuat, serta analisis berdasarkan setiap indikator berpikir kritis Facione untuk memperoleh bukti yang lebih komprehensif.

DAFTAR PUSTAKA

- Andriani, S. P., Agung, D. A. G., & Subekti, A. (2021). Pengaruh media pembelajaran sejarah film Guru Bangsa Tjokroaminoto terhadap kemampuan berpikir kritis siswa kelas X MIPA A2 SMA Negeri 3 Malang. *Historiography: Journal of Indonesian History and Education*, 1(4), 419–429.
- Basri, I., Fatimah, S., & Hastuti, H. (2024). Learning history with *DeepThink*: A model to train critical thinking skills. *Al-Ishlah: Jurnal Pendidikan*, 16(2), 1041–1051. <https://doi.org/10.35445/alishlah.v16i2.4915>
- Biggs, J. B. (1987). *Student approaches to learning and studying: Research monograph*. <https://eric.ed.gov/?id=ED308201>
- Biggs, J., Tang, C., & Kennedy, G. (2022). *Teaching for Quality Learning at University* (5th ed.). McGraw-Hill Education (UK).
- Emerentiana, P. (2024). Pengembangan buku panduan penilaian karakter kurikulum merdeka pada mata pelajaran IPA sekolah dasar. *Biogenerasi: Jurnal Pendidikan Biologi*, 10(1). <https://ejournal.my.id/biogenerasi>
- Facione, P. A. (2015). *Critical Thinking: What It Is and Why It Counts*. Measured Reasons LLC.
- Fitroh, I. (2025). Deep learning: Strategi inovatif dalam penguatan literasi sejarah peserta didik. *Jurnal Penelitian Ilmu Pendidikan Indonesia*, 4(3), 1973–1979. <https://jpcion.org/index.php/jpi>
- Hasibuan, F. J. (2025). The Influence of Deep Learning on Students' Critical and Historical Thinking Skills. *Napak Tilas: Jurnal Pendidikan Sejarah*, 1, 31–37.
- Kemdikbudristek. (2022). *Panduan Implementasi Kurikulum Merdeka*.
- Khasanah, U., & dkk. (2025). *Deep Learning dalam Pendidikan: Pendekatan Pembelajaran Bermakna, Sadar, dan Menyenangkan* (M. Hasan (ed.)). Tahta Media Group.
- Maszeri, & others. (2025). *Pendekatan Deep Learning dalam Pendidikan: Pendekatan Pembelajaran Bermakna, Sadar, dan Menyenangkan*.
- Mustapa, A., Ramadhani, K., Dewi, L. P., Oktarina, N., & Widodo, J. (2025). Implementasi Pendekatan Pembelajaran Kurikulum Merdeka: Understanding by Design, Berdiferensiasi, dan Deep Learning. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 10, 427–441.
- Nurjanah, S., & Suryadi, A. (2025). Analisis kesiapan guru dalam menerapkan pendekatan deep learning pada pembelajaran sejarah kelas X SMA Sint Louis. *JKIP: Jurnal Kajian Ilmu Pendidikan*, 6(3), 943–953. <http://journal.al-matani.com/index.php/jkip/index>
- Putri, R., Ardiansyah, S. S., Kurnia, H., Sari, M. I., Fierna, M., & Putri, L. J. (2022). *Penerapan Pendekatan Deep Learning dalam Pendidikan di Indonesia* (Vol. 2).
- Ramadhani, M. I., Mappedasse, M. Y., & Sanatang. (2026). Pengaruh penggunaan educational games sebagai alat pembelajaran terhadap hasil belajar siswa pada pembelajaran matematika di sekolah menengah pertama. *Pinisi Journal of Education*, 6(2), 65–79.
- Raup, A., Ridwan, W., Khoeriyah, Y., Supiana, & Zaqiah, Y. Q. (2022). *Deep Learning dan Penerapannya dalam Pembelajaran*. <http://jiip.stkipyapisdompu.ac.id>
- Sugiyono. (2013). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Suryana, A., & Jono. (2025). Paradigma Deep Learning dalam Pembelajaran Sejarah Abad 21: Upaya Membangun Kemampuan Berpikir Kritis Siswa. *J-KIP (Jurnal Keguruan Dan Ilmu Pendidikan)*, 6(2), 530–535. <https://doi.org/10.25157/jkip.v6i2.18872>
- Waruwu, D. E. R., & Setiawati, E. (2025). Integrasi Kurikulum Pendekatan Deep Learning dalam Pendidikan: Strategi dan Tantangan. *Jurnal Sosialita: Jurnal Kependidikan Dan Ilmu Sosial*, 20(1), 69–80. <https://doi.org/10.31316/js.v20i1.7663>



Wathon, A. (2024). Kesesuaian Kurikulum Merdeka dengan Kurikulum Pendekatan Deep Learning. *Arzusin*, 4(6), 1280–1300. <https://doi.org/10.58578/arzusin.v4i6.4442>