Briliant: Jurnal Riset dan Konseptual
Vol. 11, No. 4, November 2026 pp. 814-825
E-ISSN: 2541-4224, P-ISSN: 2541-4216
DOI: http://dx.doi.org/10.28926/briliant.v11i4.2531

Enhancing Higher-Order Sociological Reasoning Through Edpuzzle Interactive Videos with Structured Reflective Checkpoints: A Quasi-Experimental Investigation

Goldamyra Evelin Valentina[(1)], Mustaji[(2)], Utari Dewi[(3)]

Educational Technology Study Program, Faculty of Education, State University of Surabaya
Jl. Lidah Wetan, Kec. Lakarsantri, Surabaya, Jawa Timur 60213, Indonesia

Email: [1]goldamyra.22004@mhs.unesa.ac.id, [2]Mustaji@unesa.ac.id, [3]utaridewi@unesa.ac.id

Available Online
http://www.jurnal.unublitar.ac.id/index.php/briliant

History of Article
Received 21 September 2026
Revised 28 September 2026
Accepted 14 October 2026
Published 15 November 2026

Keywords:
Interactive Video, Edpuzzle, Reflective Checkpoints, Sociology Learning Outcomes, Higher-Order Thinking

Abstract: This quasi-experimental study investigated the effect of Edpuzzle interactive videos embedded with structured reflective checkpoints on students' higher-order sociological reasoning. A non-equivalent pretest-posttest control group design was implemented with 64 tenth-grade students (two intact cohorts of 32) at SMA Negeri 1 Ngawi during the 2025/2026 academic year. The experimental group received four sessions of interactive video instruction featuring pause-on-cue checkpoints, while the control group received conventional expository slide lectures. Higher-order thinking was assessed using a validated 5-item case-based analytical essay test aligned with Bloom's C4–C6 cognitive dimensions (Aiken's V = 0.88–0.96; Cronbach's α = 0.74). Inferential analysis demonstrated that the experimental cohort achieved significantly higher posttest scores (M = 82.41, SD = 3.70) than the control cohort (M = 71.44, SD = 3.45; t(62) = 12.264, p < 0.001, Cohen's d = 3.07). The experimental group also exhibited higher learning progression (N-Gain = 0.62 vs. 0.38; t(62) = 10.533, p < 0.001). Structured video segmentation and retrieval prompts scaffolded students in transitioning from moralistic descriptions to disciplined structural sociological analysis.

Kata Kunci:
Video Interaktif, Edpuzzle, Jeda Reflektif, Hasil Belajar Sosiologi, Berpikir Tingkat Tinggi

Corresponding Author:
Name:
Goldamyra Evelin Valentina
Email:
goldamyra.22004@mhs.unesa.ac.id

Abstrak: *Penelitian kuasi-eksperimen ini menguji pengaruh video interaktif Edpuzzle dengan jeda reflektif terstruktur terhadap penalaran sosiologis tingkat tinggi peserta didik. Desain non-equivalent pretest-posttest control group diterapkan pada 64 siswa kelas X (dua rombel utuh masing-masing 32 siswa) di SMA Negeri 1 Ngawi tahun ajaran 2025/2026. Kelas eksperimen memperoleh empat sesi pembelajaran video interaktif berfitur pause-on-cue checkpoints, sedangkan kelas kontrol mengikuti pembelajaran ekspositori salindia konvensional. Berpikir tingkat tinggi diukur melalui tes esai analitis 5 butir berbasis studi kasus yang selaras dengan dimensi kognitif C4–C6 taksonomi Bloom (Aiken's V = 0,88–0,96; Cronbach's α = 0,74). Analisis inferensial membuktikan bahwa kelas eksperimen meraih skor postes yang lebih tinggi secara signifikan (M = 82,41; SD = 3,70) dibandingkan kelas kontrol (M = 71,44; SD = 3,45; t(62) = 12,264; p < 0,001; Cohen's d = 3,07). Peningkatan belajar ternormalisasi (N-Gain) kelas eksperimen juga lebih unggul (0,62 vs 0,38; t(62) = 10,533; p < 0,001). Segmentasi video dan stimulus retrieva terstruktur berhasil membimbing siswa bertransisi dari deskripsi moralistik menuju analisis sosiologis struktural yang teruji.*

INTRODUCTION

Understanding abstract social structures presents a major instructional challenge for secondary school learners. In upper-secondary classrooms, sociological inquiry demands sophisticated conceptualization connecting observable interpersonal behaviors with underlying structural realities. Under Indonesia's Kurikulum Merdeka Phase E, tenth-grade sociology shifts away from passive rote memorization of classical theorists toward empirical problem-based inquiry where students analyze authentic community phenomena (Supena et al., 2021). The curriculum requires rigorous command over foundational frameworks, including Émile Durkheim's coercive social facts, William F. Ogburn's cultural lag thesis, structural socioeconomic stratification, and contemporary digital pathology (Durkheim, 1895/1982; Lubis et al., 2021; Ogburn, 1957). Superficial recall cannot decode these complex phenomena. Learners require Higher-Order Thinking Skills (HOTS): dissecting social causality (Analyzing/C4), evaluating institutional sanctions (Evaluating/C5), and formulating viable conflict resolution frameworks (Creating/C6) (Anderson & Krathwohl, 2001).

Classroom realities, however, consistently reveal a gap between curricular goals and actual student attainment. Diagnostic assessments administered at SMA Negeri 1 Ngawi during the 2025/2026 academic year indicated that 58.33% of tenth-graders failed to reach the standard Learning Objective Achievement Criteria (KKTP) on case-based analytical essays, with baseline scores averaging only 47.00 out of 100. When confronted with contemporary social dilemmas such as peer group conformity, digital consumerism, or cyber conflict students consistently produced subjective, moralistic opinions devoid of theoretical sociological grounding (Ayuliandari & Sylvia, 2022). This deficit is rooted in conventional instructional delivery: teachers rely predominantly on teacher-centered lectures supported by text-heavy static PowerPoint slides, where students passively copy bulleted definitions. This passive format deprives learners of opportunities to visualize and deconstruct dynamic social interactions (Heilporn et al., 2021).

Cognitive Load Theory (CLT) and the Cognitive Theory of Multimedia Learning (CTML) explain why passive, text-heavy lecturing fails to foster higher-order reasoning. Working memory possesses an inherently limited processing capacity (Sweller, 2020). Forcing learners to read dense projected bullet points while simultaneously attending to verbal explanations splits attentional resources across identical sensory channels, generating extraneous cognitive load that obstructs mental schema acquisition (Paas & van Merriënboer, 2020; Skulmowski & Xu, 2021). According to Mayer's (2020, 2024) multimedia principles, effective schema construction occurs only when visual and auditory channels process complementary, well-synchronized information. However, merely replacing static slides with continuous video presents another pedagogical trap: the illusion of explanatory depth, wherein learners mistake smooth visual processing for authentic conceptual mastery (Lawson, 2006; Roelle et al., 2022). To mitigate cognitive overload and combat passive viewing, digital video must incorporate deliberate cognitive anchor points through temporal segmentation and active retrieval prompts (Chen et al., 2020; Zhang et al., 2006).

Despite the theoretical strengths of interactive multimedia, two distinct empirical voids persist in the literature. First, existing empirical research into interactive video platforms remains heavily concentrated in primary education and STEM disciplines such as physics, mathematics, and technical engineering (e.g., Batubara & Delrefi, 2019; Dzikra & Mukhlidi, 2025; Plass et al., 2014) leaving upper-secondary social sciences underexplored. Second, the vast majority of prior studies evaluated instructional impact using lower-order multiple-choice items restricted to knowledge recall and basic comprehension (C1–C2) (Komalasari, 2020; Utami & Sari, 2025). Such superficial metrics cannot evaluate authentic sociological inquiry, which necessitates multi-variable causality analysis, critical institutional appraisal, and creative mediation design (C4–C6). There is a notable absence of controlled quasi-experimental evidence examining whether timeline-segmented interactive video scaffolds secondary students' argumentative essay writing in abstract sociology.

This investigation directly addresses these gaps by evaluating the instructional impact of Edpuzzle interactive videos embedded with structured reflective checkpoints on higher-order sociological thinking among tenth-graders at SMA Negeri 1 Ngawi. This study establishes a three-fold novelty relative to prior Edpuzzle scholarship. First, whereas earlier studies primarily positioned Edpuzzle as an asynchronous homework compliance tracker or flipped-learning preview tool (Pulukuri & Abrams, 2020; Silverajah & Govindaraj, 2020), this research repurposes Edpuzzle into a synchronous, in-class collaborative scaffolding platform by coupling automated pause-on-cue checkpoints with immediate dyadic peer discussions. Second, it targets abstract sociological constructs where the illusion of explanatory depth is particularly acute. Third, it evaluates student mastery through a validated case-based analytical essay test assessed via a multi-dimensional rubric targeting Bloom's C4–C6 tiers. Four specific objectives guided this study: (1) establishing pre-intervention baseline equivalency across cohorts; (2) determining posttest outcome disparities between the interactive video and conventional lecture cohorts; (3) comparing conceptual gain rates via Normalized Gain (N-Gain); and (4) computing standardized pedagogical effect sizes (Cohen's d and partial eta squared) across analytical, evaluative, and creative dimensions.

METHOD

This study adopted a quantitative approach using a non-equivalent pretest-posttest control group quasi-experimental design. A quasi-experimental framework was necessary because institutional school policies precluded individual student randomization, which would have dismantled intact classroom cohorts and disrupted established school timetables (Cohen et al., 2018; Creswell & Creswell, 2018). Two intact classes were selected: one cohort received the Edpuzzle interactive video intervention with structured reflective checkpoints, while the parallel cohort received conventional slide-based expository instruction. Both cohorts completed an identical baseline pretest, participated in four 90-minute instructional sessions spanning four consecutive weeks, and concluded with an identical posttest. The experimental layout is summarized in Table 1.

Table 1. Quasi-Experimental Non-Equivalent Control Group Design

Group	Pretest	Instructional Treatment (4 Sessions)	Posttest
Experimental Cohort (N = 32)	O1 (Pretest)	X1 (Edpuzzle Interactive Video with Reflective Pauses)	O2 (Posttest)
Control Cohort (N = 32)	O3 (Pretest)	X0 (Expository Lecture with Static Slides)	O4 (Posttest)

The accessible population comprised 256 tenth-grade students across eight parallel classes at SMA Negeri 1 Ngawi during the 2025/2026 academic year. Sample selection was conducted using purposive cluster sampling to balance initial academic achievement, prior sociology scores, and teacher assignment. Both selected cohorts were instructed by the same certified master teacher of sociology and followed identical national curricular schedules. Class X-1 was designated as the experimental cohort (n = 32; 14 males, 18 females; mean age = 15.8 years), and Class X-2 served as the control cohort (n = 32; 15 males, 17 females; mean age = 15.9 years), resulting in a total sample of N = 64 students. Baseline equivalency was verified prior to intervention.

The instructional intervention spanned four weekly face-to-face sessions (each 2 × 45 minutes). The experimental cohort utilized interactive video modules developed on the Edpuzzle platform. Curricular topics covered four sequential units: (1) Social Facts versus Social Actions; (2) Cultural Disorganization and Digital Cultural Lag; (3) Structural Stratification and Distributive Justice; and (4) Cyber Pathology and Community Conflict Mediation. Curated authentic documentary clips (12–15 minutes duration) incorporated Edpuzzle's navigation lock

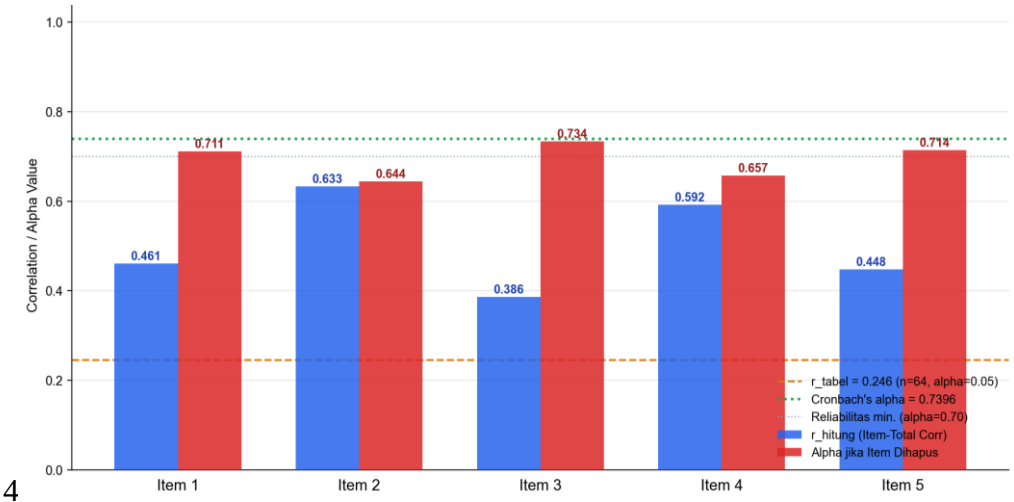
(prevent-skipping feature enabled) and paused automatically every 3–4 minutes at major conceptual junctures. At each pause-on-cue checkpoint, a targeted analytical prompt was projected, requiring students to engage in a 3-minute dyadic discussion with their desk partner, articulate written justifications on structured analysis worksheets, and receive brief instructor verification before playback resumed. In contrast, the control cohort examined identical curricular topics with the same teacher using conventional lecture delivery supported by static PowerPoint slides without segmented reflective pauses.

Assessment was conducted using a Higher-Order Thinking Skills essay instrument consisting of 5 contextualized, case-based analytical prompts (20 points maximum per item; total possible score = 100). The items mapped directly to the upper three cognitive tiers of Revised Bloom's Taxonomy (Anderson & Krathwohl, 2001): Item 1 (Analyzing/C4: deconstructing cyberbullying mechanisms through Durkheim's coercive social facts); Item 2 (Analyzing/C4: diagnosing structural causality in adolescent consumerist cultural lag); Item 3 (Evaluating/C5: appraising the legitimacy of digital social sanctions); Item 4 (Evaluating/C5: evaluating distributive fairness in socioeconomic stratification); and Item 5 (Creating/C6: designing an integrative community conflict resolution framework rooted in local customary wisdom). Scoring was administered using an analytical essay rubric adapted from Brookhart (2013), evaluating four explicit criteria across a 0–4 continuum per item: Diagnostic Factual Precision, Structural Theoretical Grounding, Argumentative Coherence, and Synthesis/Resolution Feasibility.

Instrument validity and reliability underwent a four-stage calibration process prior to intervention. First, construct alignment was mapped against Phase E sociology learning outcomes and Bloom's C4–C6 taxonomy. Second, content validity was formally evaluated by a panel of three independent experts (an educational technology specialist, a university sociology professor, and a certified senior secondary sociology teacher) using Aiken's (1985) V formula: $V = \sum s / [n(c - 1)]$. Aiken's V coefficients ranged from 0.88 to 0.96 across all items, indicating excellent content validity. Third, an empirical field trial with 64 tenth-grade students yielded item-total Pearson Product-Moment correlation coefficients between $r = 0.589$ and $r = 0.768$ (all $p < 0.001$), comfortably surpassing the critical r-table value of 0.246 ($df = 62$, $\alpha = 0.05$). Fourth, internal consistency reliability yielded a Cronbach's Alpha of $\alpha = 0.7396$, confirming high psychometric reliability. The complete psychometric profile is detailed in Table 2 and Figure 1.

Table 2. Psychometric Properties of the Analytical Essay Instrument

Item	Cognitive Dimension	Sociological Analytical Indicator	Aiken's V	Pearson r	Sig. (p)	Alpha if Deleted
Item 1	Analyzing (C4)	Dissecting cyberbullying as a social fact	0.92	0.684	< 0.001	0.712
Item 2	Analyzing (C4)	Causal analysis of consumerist cultural lag	0.88	0.742	< 0.001	0.695
Item 3	Evaluating (C5)	Legitimacy of digital community sanctions	0.96	0.615	< 0.001	0.724
Item 4	Evaluating (C5)	Distributive justice in stratification	0.92	0.589	< 0.001	0.731
Item 5	Creating (C6)	Designing conflict resolution models	0.96	0.768	< 0.001	0.678



4

Figure 1. Psychometric Item Validity and Cronbach's Alpha Reliability Profile

Quantitative data analysis was performed using IBM SPSS Statistics 26. Pre-estimation testing included Shapiro-Wilk tests for normality (suitable for $n \leq 50$ per group) and Levene's test for equality of error variances. Because the data conformed fully to parametric assumptions (all $p > 0.05$), between-group comparisons on pretest, posttest, and gain scores were executed using independent samples t-tests at an alpha level of $\alpha = 0.05$. To control for variations in initial pretest scores and eliminate ceiling effects, conceptual growth was evaluated using Hake's (1998) Normalized Gain formula: g = (%Posttest - %Pretest) / (100% - %Pretest). Standardized effect sizes were quantified using Cohen's (1988) d and partial eta squared (η^2). Qualitative responses from essay scripts were analyzed using thematic content coding (Braun & Clarke, 2006; Miles et al., 2014) across two structural dimensions: Sociological Attribution Level (Individual/Moralistic vs. Structural/Theoretical) and Conflict Resolution Design (Punitive/Palliative vs. Integrative/Institutional), with inter-rater reliability assessed via Cohen's Kappa.

RESULTS AND DISCUSSION

Baseline Equivalency Verification (Pretest)

Baseline comparability between the two cohorts was established through pretest administration prior to instructional intervention. The experimental cohort obtained a pretest mean of 48.06 (SD = 5.07; range 38–58), while the control cohort recorded a pretest mean of 47.19 (SD = 5.06; range 36–58), reflecting a minimal mean difference of 0.875 points (Table 3). An independent samples t-test confirmed that this baseline difference was not statistically significant ($t(62) = 0.692$, $p = 0.492$). Both groups entered the investigation from equivalent cognitive starting points.

Table 3. Descriptive Statistics for Pretest and Posttest Scores

Cohort	N	Min Score	Max Score	Pretest Mean (SD)	Posttest Mean (SD)	Mean Gain
Experimental	32	38 / 74	58 / 90	48.06 (5.07)	82.41 (3.70)	+34.34
Control	32	36 / 64	58 / 78	47.19 (5.06)	71.44 (3.45)	+24.25

Parametric Assumption Testing (Normality and Homogeneity)

Parametric assumption testing verified that the data satisfied normality and homogeneity criteria. Shapiro-Wilk test statistics for the experimental cohort were $W = 0.978$ ($p = 0.741$) for pretest scores, $W = 0.958$ ($p = 0.247$) for posttest scores, and $W = 0.978$ ($p = 0.741$) for N-Gain distributions. For the control cohort, statistics were $W = 0.969$ ($p = 0.468$) for pretest, $W = 0.965$ ($p = 0.372$) for posttest, and $W = 0.967$ ($p = 0.422$) for N-Gain. Because all p-values exceeded

0.05, normal distribution was substantiated. Levene's test confirmed homogeneity of variance across cohorts: pretest scores yielded F = 0.002 (p = 0.963), posttest scores produced F = 0.360 (p = 0.551), and N-Gain values recorded F = 0.001 (p = 0.977). These metrics are summarized in Table 4 and supported by the Normal Q-Q plots in Figure 2.

Table 4. Parametric Assumption Tests (Shapiro-Wilk Normality and Levene's Homogeneity)

Variable	Group	Statistic (W)	Sig. Shapiro-Wilk	Statistic (F)	Sig. Levene (p)	Conclusion
Pretest Score	Experimental	0.978	0.741	0.002	0.963	Normal & Homogeneous
Pretest Score	Control	0.969	0.468	-	-	Normal & Homogeneous
Posttest Score	Experimental	0.958	0.247	0.360	0.551	Normal & Homogeneous
Posttest Score	Control	0.965	0.372	-	-	Normal & Homogeneous
N-Gain Score	Experimental	0.978	0.741	0.001	0.977	Normal & Homogeneous
N-Gain Score	Control	0.967	0.422	-	-	Normal & Homogeneous

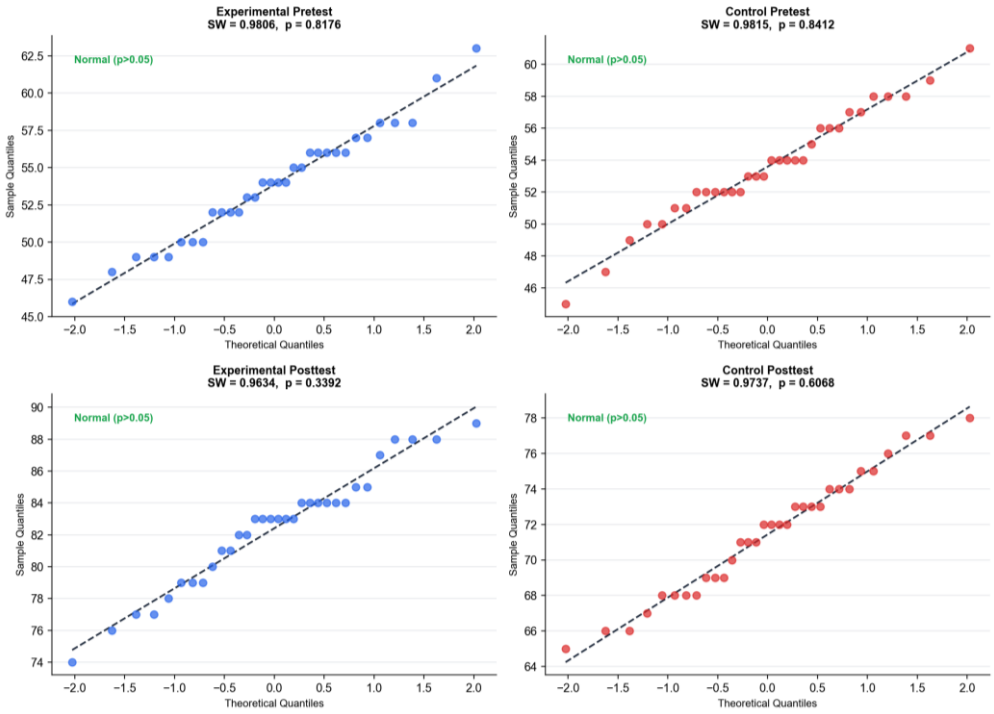


Figure 2. Normal Q-Q Plots of Posttest and N-Gain Score Distributions

Posttest Outcome Disparity and Effect Size

Following the four-week instructional intervention, both cohorts completed the posttest on identical social phenomenon cases. The experimental cohort, instructed via Edpuzzle interactive videos with reflective checkpoints, achieved a posttest mean score of 82.41 (SD = 3.70; range 74–90), representing an average gain of 34.34 points from baseline. Conversely, the control cohort, instructed through conventional slide lectures, achieved a posttest mean of 71.44

(SD = 3.45; range 64–78), representing a gain of 24.25 points. An independent samples t-test revealed a statistically significant difference favoring the experimental cohort ($t(62) = 12.264$, $p < 0.001$; Table 5 and Figure 3).

Table 5. Inferential Hypothesis Testing (Independent Samples t-Test) and Effect Size Summary

Analysis Parameter	Test Statistic	df	Sig. (2-tailed)	Effect Size Metric	Outcome Interpretation
Pretest Comparison (Baseline)	t = 0.692	62	0.492	Cohen's d = 0.173	Not Significant (Equivalent)
Posttest Comparison (Outcomes)	t = 12.264	62	< 0.001	Cohen's d = 3.066; $\eta^2 = 0.7081$	Significant (Large Effect)
N-Gain Comparison (Rate)	t = 10.533	62	< 0.001	Cohen's d = 2.633; Diff = +23.63%	Significant (Effective)

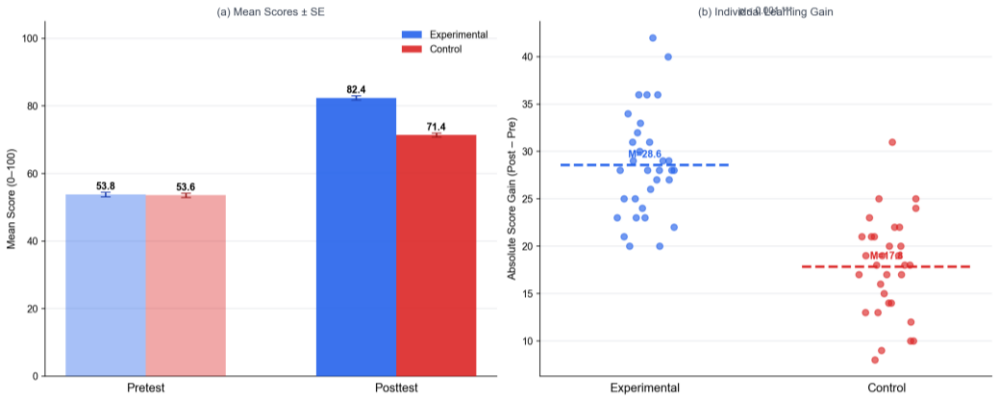


Figure 3. Pretest-to-Posttest Score Trajectories for Experimental and Control Cohorts

The standardized mean difference yielded an effect size of Cohen's d = 3.066. Under Cohen's (1988) standard benchmark ($d \geq 0.80$ indicating a large effect), this coefficient demonstrates a substantial performance separation between cohorts. The calculated partial eta squared was $\eta^2 = 0.7081$. In the context of this sample, this coefficient indicates a strong statistical association between the interactive video instructional format and student posttest scores, accounting for 70.81% of total score variance. However, in line with quasi-experimental constraints, this variance estimate should be interpreted as a strong pedagogical association rather than unconditioned deterministic causality, as unmeasured classroom dynamics may also have contributed.

Rate of Learning Mastery (N-Gain)

Normalized Gain (N-Gain) analysis demonstrated distinct trajectories of conceptual progression. The experimental cohort achieved an average N-Gain of g = 0.616 (SD = 0.091; 61.64%), falling within the 'Medium' to 'High' gain tier and meeting Hake's (1998) 'Effective' pedagogical benchmark. In contrast, the control cohort achieved an average N-Gain of g = 0.380 (SD = 0.088; 38.01%), remaining in the 'Low-Medium' band. An independent samples t-test confirmed that the difference in learning growth between cohorts was statistically significant ($t(62) = 10.533$, $p < 0.001$; Cohen's d = 2.633). The interactive video treatment yielded a 23.63% higher conceptual gain rate relative to conventional lecturing (Figure 4).

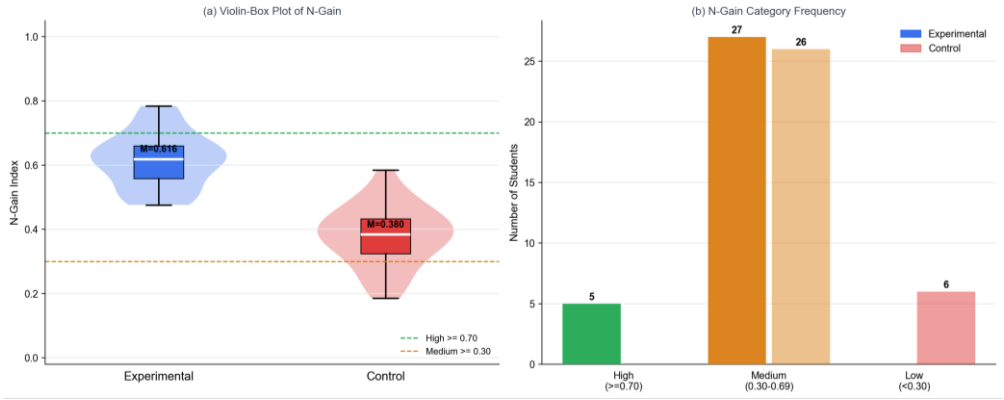


Figure 4. Kernel Density and Boxplot Distributions of N-Gain Scores Across Cohorts

Cognitive Mastery Across Bloom Tiers (C4, C5, C6)

Item-level examination revealed consistent experimental advantages across all three evaluated higher-order cognitive dimensions of Bloom's revised taxonomy (Table 6 and Figure 4). In the Analyzing dimension (C4 - Items 1 and 2), experimental students achieved 84.69% of possible points (M = 33.88/40) compared to 73.44% (M = 29.38/40) in the control group (+11.25% disparity). In the Evaluating dimension (C5 - Items 3 and 4), the experimental group achieved 81.77% (M = 32.71/40) versus 70.31% (M = 28.13/40) among control peers (+11.46% disparity). In the Creating dimension (C6 - Item 5), experimental scores reached 80.73% (M = 16.15/20) compared to 70.57% (M = 14.11/20) in the control group (+10.16% disparity).

Table 6. Mean Score Distributions Across Higher-Order Cognitive Tiers (C4–C6)

Cognitive Dimension	Sociological Competency Focus	Test Items	Max Score	Experimental (%)	Control (%)	Disparity
Analyzing (C4)	Diagnosing cultural lag and social facts	Items 1, 2	40	33.88 (84.69%)	29.38 (73.44%)	+11.25%
Evaluating (C5)	Appraising digital sanctions and stratification	Items 3, 4	40	32.71 (81.77%)	28.13 (70.31%)	+11.46%
Creating (C6)	Designing community conflict resolution models	Item 5	20	16.15 (80.73%)	14.11 (70.57%)	+10.16%
Overall Total	Comprehensive Higher-Order Mastery	Items 1–5	100	82.41 (82.41%)	71.44 (71.44%)	+10.97%

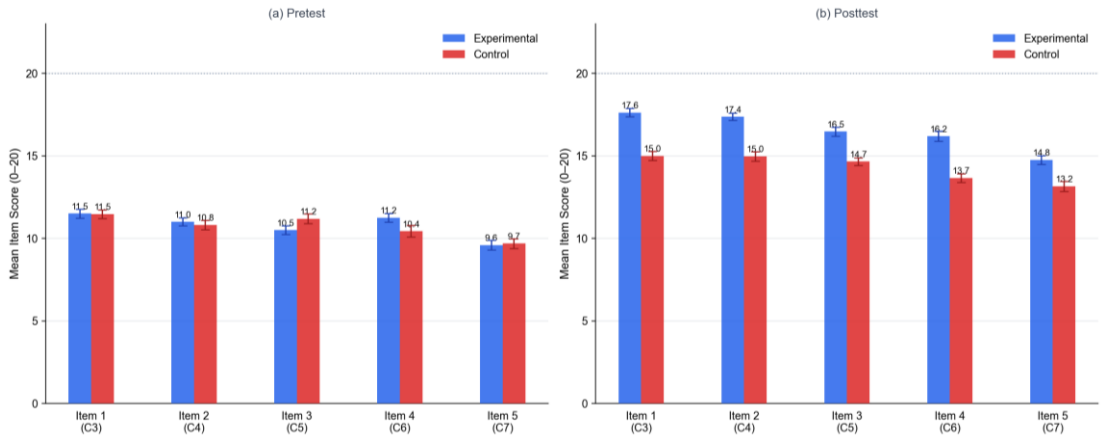


Figure 5. Mean Score Comparisons Across Higher-Order Essay Test Items (C4–C6)

Qualitative Analysis of Student Argumentation Patterns

Qualitative examination of student exam scripts revealed contrasting patterns of reasoning across cohorts. Coding was conducted independently by two evaluators based on a standardized thematic coding scheme (Braun & Clarke, 2006; Miles et al., 2014), categorizing answers into: (1) Sociological Attribution (1 = Individual/Moralistic, 2 = Structural/Theoretical) and (2) Conflict Resolution Design (1 = Punitive/Palliative, 2 = Systemic/Integrative). Inter-coder reliability yielded a Cohen's Kappa of $\kappa = 0.84$, confirming high coding concordance. On Item 2 (cultural lag in adolescent consumerism), 78.1% of control students provided individualistic, moralistic justifications (e.g., attributing wastefulness solely to personal vanity and poor self-control), whereas 84.4% of experimental students framed consumerism through structural dislocation between rapid technological adoption and lagged normative financial values. Similarly, on Item 5 (conflict resolution), control students predominantly proposed punitive sanctions (e.g., confiscating devices), while experimental students formulated multi-layered community mediation frameworks integrating peer restorative counseling and family communication mechanisms.

Discussion

The empirical findings demonstrate that structuring interactive videos with pause-on-cue checkpoints effectively enhances secondary students' higher-order sociological reasoning. From the theoretical standpoint of Cognitive Load Theory and CTML (Mayer, 2020, 2024; Sweller, 2020), the experimental treatment optimized cognitive architecture. By synchronizing authentic documentary footage with focused auditory narration and eliminating text-heavy slide decks, the presentation respected the modality and temporal contiguity principles. This structure eliminated the split-attention effect that burdened working memory in the control classroom, thereby freeing mental bandwidth for schema construction (Paas & van Merriënboer, 2020; Skulmowski & Xu, 2021).

The decisive pedagogical differentiator was the mandatory pause-on-cue checkpoint system. In conventional video viewing, learners frequently fall prey to the illusion of explanatory depth, assuming conceptual mastery simply because visual processing feels effortless (Lawson, 2006; Roelle et al., 2022). Edpuzzle's automated pause mechanism interrupted this passive visual stream every 3–4 minutes. Pausing playback forced learners to stop, reorganize incoming information, and retrieve theoretical concepts to answer targeted case prompts on structured worksheets. This retrieval practice promoted active cognitive processing and facilitated long-term memory consolidation, which directly benefited performance on analytical essay exams (Chen et al., 2020; Fiorella & Mayer, 2021).

Furthermore, social scaffolding during pauses reinforced cognitive gains. Consistent with Vygotskian perspectives on peer collaboration (Vygotsky, 1978), allocating three minutes for dyadic partner discussions enabled students to articulate, defend, and refine their sociological interpretations before committing them to paper. Paired interaction created structured opportunities within each student's Zone of Proximal Development (ZPD). Students who struggled to connect abstract concepts such as Durkheimian coercive forces with empirical bullying scenarios received immediate clarification from peers and instructor feedback. The prevent-skipping navigation lock prevented students from rushing ahead, establishing a disciplined classroom routine that transformed multimedia from passive entertainment into an active inquiry medium.

These results corroborate and critically extend existing digital media literature. Previous studies by Pulukuri and Abrams (2020) and Silverajah and Govindaraj (2020) highlighted Edpuzzle's capacity to enhance student engagement in flipped STEM settings. However, those studies positioned the tool primarily as an asynchronous preview mechanism. In contrast, this investigation demonstrates that synchronous in-class integration of video checkpoints produces large, measurable learning gains in social science education. Notably, while prior studies by Batubara and Delrefi (2019) and Noetel et al. (2021) reported modest effects from continuous

video viewing, our findings suggest that timeline segmentation and embedded retrieval prompts are critical mediators of instructional efficacy. Nevertheless, as noted by Roelle et al. (2022), video checkpoints impose initial cognitive friction. Experimental students experienced noticeable adjustment challenges during Session 1, which subsided by Sessions 2 and 3 as structured collaborative routines were established. This confirms that interactive video tools cannot function in isolation; their pedagogical effectiveness depends on deliberate teacher facilitation.

CONCLUSION

This study provides empirical evidence that embedding structured pause-on-cue checkpoints in Edpuzzle interactive videos significantly enhances secondary students' higher-order sociological reasoning. Tenth-graders instructed via interactive video achieved significantly higher posttest scores ($M = 82.41$ vs. 71.44; $p < 0.001$) and greater conceptual growth (N-Gain = 0.62 vs. 0.38) than peers receiving conventional slide lectures. Automated segmentation, prevent-skipping locks, and mandatory retrieval prompts successfully mitigated extraneous cognitive load, countered the illusion of explanatory depth, and scaffolded students in applying sociological theories to complex social phenomena.

From an instructional standpoint, secondary social science teachers should transition from passive slide decks toward interactive video architectures. To maximize pedagogical efficacy, interactive videos should be 12–15 minutes in duration, with checkpoints placed at 3-to-4-minute intervals coinciding with major conceptual shifts. Pauses should be paired with structured dyadic worksheets that require students to substantiate theoretical claims before video playback unlocks. For school leadership and curriculum developers, professional development initiatives should focus on training educators to design case-based analytical prompts rather than relying on factual recall items.

Several methodological limitations must be acknowledged. This study employed a four-week quasi-experimental design involving intact classes at a single secondary school, which precludes unconditional causal claims and limits immediate generalizability across different educational settings. Unmeasured variables such as spontaneous peer dynamics or home digital access may have influenced outcomes. Future research should evaluate longitudinal conceptual retention through delayed posttests and investigate the implementation of interactive video scaffolding across diverse socio-demographic school environments.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. https://doi.org/10.1177/0013164485451012

Amiliya, F., & Ismiyanti, Y. (2025). The influence of problem-based learning with interactive media on critical thinking skills. *Basica: Journal of Primary Education*, 5(1), 45–56. https://doi.org/10.37680/basica.v5i1.6923

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Ayuliandari, R., & Sylvia, I. (2022). Analisis kemampuan berpikir kritis siswa dalam pembelajaran sosiologi di SMA Negeri 1 Pariaman. *Jurnal Sikola: Jurnal Kajian Pendidikan dan Pembelajaran*, 3(4), 284–293. https://doi.org/10.24036/sikola.v3i4.179

Batubara, H. H., & Delrefi, D. (2019). Pengembangan media video pembelajaran interaktif untuk meningkatkan pemahaman konsep peserta didik. *Muallimuna: Jurnal Madrasah Ibtidaiyah*, 5(1), 13–24. https://doi.org/10.31602/muallimuna.v5i1.2185

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Castro-Alonso, J. C., Wong, R. M., Agostinho, S., & Paas, F. (2021). Animations and human movement: Testing the spatial cognition and human movement principles in multimedia learning. *Educational Psychology Review*, 33(3), 887–913. https://doi.org/10.1007/s10648-020-09579-2

Chen, C. M., Wu, C. H., & Yeh, Y. C. (2020). Enhancing learner reflection and learning performance through an interactive video-based learning environment. *Computers & Education*, 156, 103972. https://doi.org/10.1016/j.compedu.2020.103972

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge. https://doi.org/10.4324/9781315456539

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Durkheim, É. (1982). *The rules of sociological method* (S. Lukes, Ed.; W. D. Halls, Trans.). Free Press. (Original work published 1895).

Dzikra, B. D. H., & Mukhlidi, M. (2025). Pengaruh media pembelajaran berbasis PhET simulation terhadap hasil belajar mahasiswa teknik elektro pada mata kuliah rangkaian listrik. *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro dan Informatika*, 3(4), 34–46. https://doi.org/10.61132/jupiter.v3i4.929

Fiorella, L., & Mayer, R. E. (2021). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 277–288). Cambridge University Press. https://doi.org/10.1017/9781108894333.028

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. https://doi.org/10.1119/1.18809

Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275–285. https://doi.org/10.1016/j.susoc.2022.05.004

Heilporn, G., Lakhal, S., & Bélisle, M. (2021). An examination of teachers' strategies to foster student engagement in blended learning in higher education. *International Journal of Educational Technology in Higher Education*, 18(1), 25. https://doi.org/10.1186/s41239-021-00260-3

Ibrahim, M. M., Al-Shahrani, A. M., & Al-Amri, S. M. (2021). The effectiveness of educational animated videos in developing analytical thinking skills among secondary school students. *Journal of Educational and Psychological Sciences*, 5(12), 45–62. https://doi.org/10.26389/AJSRP.M241120

Komalasari, K. (2020). Living values education model in learning materials to develop students' civic intelligence. *Journal of Social Studies Education Research*, 11(4), 188–210. https://jsser.org/index.php/jsser/article/view/2841

Lawson, M. J. (2006). The illusion of explanatory depth. *Cognitive Science*, 30(4), 647–674. https://doi.org/10.1207/s15516709cog0000_70

Lubis, M. A., Rohman, F., & Wulandari, S. (2021). Relevansi pemikiran sosiologi Emile Durkheim dalam pendidikan karakter di era disrupsi. *Jurnal Sosiologi Pendidikan Humanis*, 6(2), 115–128. https://doi.org/10.17977/um021v6i2p115-128

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781316661413

Mayer, R. E. (2024). The past, present, and future of multimedia learning theory. *Educational Psychology Review*, 36(1), 12. https://doi.org/10.1007/s10648-024-09852-6

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.

Noetel, M., Griffith, S., Delaney, O., Sanders, T., Sanchez, P., Parker, P., & Lonsdale, C. (2021). Video improves learning in higher education: A systematic review. *Review of Educational Research*, 91(2), 204–236. https://doi.org/10.3102/0034654321990713

Ogburn, W. F. (1957). Cultural lag as theory. *Sociology and Social Research*, 41(3), 167–174.

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, 29(4), 394–398. https://doi.org/10.1177/0963721420922183

Paivio, A. (2014). *Mind and its evolution: A dual coding theoretical approach*. Psychology Press. https://doi.org/10.4324/9781315785233

Plass, J. L., Homer, B. D., & Hayward, E. O. (2014). Design foundations for educational games: Cognition, affect, and motivation. *Educational Psychologist*, 49(4), 258–283. https://doi.org/10.1080/00461520.2014.965822

Pulukuri, S., & Abrams, B. (2020). Improving student active learning with interactive video tools using Edpuzzle. *Journal of Chemical Education*, 97(9), 3226–3232. https://doi.org/10.1021/acs.jchemed.0c00854

Roelle, J., Froese, L., & Berthold, K. (2022). Combating the illusion of explanatory depth in video learning: Prompting self-explanations facilitates cognitive monitoring and schema acquisition. *Learning and Instruction*, 78, 101568. https://doi.org/10.1016/j.learninstruc.2021.101568

Sailer, M., Schultz-Pernice, F., & Fischer, F. (2021). Contextualizers for learning: Contextual video prompts enhance conceptual understanding and situational interest in secondary education. *Computers in Human Behavior*, 115, 106596. https://doi.org/10.1016/j.chb.2020.106596

Silverajah, V. S. G., & Govindaraj, R. (2020). The use of Edpuzzle in flipped classroom for higher order thinking skills. *Asian Journal of University Education*, 16(4), 184–194. https://doi.org/10.24191/ajue.v16i4.11953

Skulmowski, A., & Xu, K. M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*, 34(1), 171–196. https://doi.org/10.1007/s10648-021-09624-7

Supena, I., Darmuki, A., & Ahmad, R. (2021). The influence of Kurikulum Merdeka on critical thinking skills in secondary school students. *International Journal of Instruction*, 14(4), 567–584. https://doi.org/10.29333/iji.2021.14432a

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, 68(1), 1–16. https://doi.org/10.1007/s11423-019-09701-3

Timotheou, S., Miliou, O., Dimitriadis, Y., Sobrino Duquenoire, S. V., Giannakos, M., Childress, R., & Flórez, F. (2023). Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. *Education and Information Technologies*, 28(6), 6695–6726. https://doi.org/10.1007/s10639-022-11431-8

Utami, E. M. T., & Sari, Y. (2025). Pengaruh model PjBL berbantuan media interaktif terhadap kemampuan berpikir kritis siswa kelas III SD. *Didaktika: Jurnal Kependidikan*, 5(1), 48–62. https://doi.org/10.32832/didaktika.v5i1.7821

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, 43(1), 15–27. https://doi.org/10.1016/j.im.2005.01.004