

The Effectiveness of Edpuzzle-Based Interactive Video on Fifth Grade Students' Learning Outcomes in Fact and Opinion Materials

Millah Fithriyani

Universitas Sultan Ageng Tirtayasa Serang
piscas.mee@gmail.com

Lukman Nulhakim

Universitas Sultan Ageng Tirtayasa Serang
lukman.nulhakim@untirta.ac.id

Ratna Sari Dewi

Universitas Sultan Ageng Tirtayasa Serang
ratna@untirta.ac.id

Corresponding Author: Millah Fithriyani

Article history: Received: Desember 03, 2025 | Revised: Januari 20, 2026 | Available Online: March 27, 2026

Abstract

Distinguishing fact from opinion is an essential information literacy skill for elementary students, yet achievement in Indonesia remains below minimum standards because instruction relies predominantly on text-based media. This study developed, validated, and tested the initial effectiveness of an Edpuzzle-based interactive video for fact-and-opinion instruction in Grade 5. Development followed the ADDIE model and applied principles of generative processing from the *Cognitive Theory of Multimedia Learning* through embedded questions and formative feedback. Three experts assessed content validity using Aiken's V. Instrument quality was evaluated through item validity, KR-20 reliability, difficulty index, and discrimination index. Effectiveness was tested using a one-group pretest–posttest design ($n = 35$) with a paired-samples t-test, Cohen's d , and normalized gain (N-Gain). Content validity yielded Aiken's $V = 0.793$, KR-20 reliability = 0.880, and a student response rate of 90.89%. Posttest scores ($M = 83.83$) were significantly higher than pretest scores ($M = 62.37$): $t(34) = 26.80$, $p < .001$, Cohen's $d = 2.14$, N-Gain = 0.60 (moderate gain). The video medium is effective as an initial instructional intervention for Grade 5 fact-and-opinion learning. Future studies should adopt quasi-experimental designs with a control group, incorporate performance-based assessment, and include.

Keywords: interactive video; Edpuzzle; facts and opinions; multimedia learning.

Introduction

The ability to distinguish factual statements from opinions is now a fundamental information literacy competency, not merely a language skill. The proliferation of unverified digital information requires children to evaluate claims independently, often without adult guidance.¹ Recent reviews indicate that children's exposure to misleading content begins at an early age. However, educational interventions that systematically train claim evaluation at the elementary school level remain very limited.²

From a developmental perspective, ages 10–11 represent a critical window for building text evaluation skills. Beal et al. demonstrate that children's ability to evaluate problematic texts and monitor their own comprehension develops rapidly between 3rd and 6th grade, and that these skills are responsive to structured comprehension-monitoring training.³ At the neurological level, inhibition, working memory, and cognitive flexibility mediate expository text comprehension in 4th-grade students,⁴ These executive functions also support children's differentiation between informative and narrative discourse a foundation for higher-order critical reading.⁵ The PIRLS framework, in turn, places the evaluation of content and textual elements as the highest level of competence in reading comprehension.⁶

One instructional design framework relevant to the learning of complex cognitive skills is the Cognitive Theory of Multimedia Learning (CTML), developed by Mayer over four decades.⁷ CTML is based on three assumptions: (a) separate processing of information through verbal and visual channels, (b) limited processing capacity, and (c) the necessity of active processing by the learner. From this third assumption emerges the

¹ F S Adler Berg et al., "Children and Digital Misinformation: A Scoping Review," *Global Studies of Childhood*, 2025, <https://doi.org/10.1177/20436106251398608>.

² Angela Blakston, Simon Chambers, and Tanya Notley, "Young People, Algorithms and News: Exploring the Relationship between Algorithmic Literacy and News Literacy," *Journal of Youth Studies*, 2025, 1–17.

³ C R Beal, A C Garrod, and G J Bonitatibus, "Fostering Children's Revision Skills Through Training in Comprehension Monitoring," *Journal of Educational Psychology* 82, no. 2 (1990): 275–80, <https://doi.org/10.1037/0022-0663.82.2.275>.

⁴ J.-P. Escobar and V Espinoza, "Direct and Indirect Effects of Inhibition, Working Memory and Cognitive Flexibility on Reading Comprehension of Narrative and Expository Texts: Same or Different Effects?," *Reading and Writing Quarterly* 41, no. 2 (2025): 176–92, <https://doi.org/10.1080/10573569.2024.2400993>.

⁵ Rachel M Best, Randy G Floyd, and Danielle S McNamara, "Differential Competencies Contributing to Children's Comprehension of Narrative and Expository Texts," *Reading Psychology* 29, no. 2 (2008): 137–64.

⁶ S Santos et al., "An Application of the Rasch Model to Reading Comprehension Measurement," *Psicologia: Reflexao e Critica* 29, no. 1 (2016), <https://doi.org/10.1186/s41155-016-0044-6>.

⁷ R E Mayer, "The Past, Present, and Future of the Cognitive Theory of Multimedia Learning," *Educational Psychology Review* 36, no. 1 (2024), <https://doi.org/10.1007/s10648-023-09842-1>.

principle of generative processing formulated by Fiorella and Mayer.⁸ According to this principle, meaningful learning does not occur through passive reception, but rather through generative cognitive activities such as summarizing, explaining, mapping, and answering questions that promote the integration of new information with prior knowledge. A systematic review by Noetel et al. confirms the consistent positive effects of generative principles on learning outcomes across various instructional contexts.⁹

Cognitive Load Theory (CLT) formulates design principles to minimize extrinsic cognitive load while optimally managing essential load,¹⁰ In the context of instructional videos, the combination of CTML and CLT yields concrete design principles: video segmentation, the insertion of reflective questions, the provision of feedback, and the activation of metacognitive monitoring.¹¹ Interactive videos based on embedded questions differ from passive videos in three ways: questions actively promote meaning construction, feedback corrects misconceptions before they take root, and segmentation prevents cognitive overload.¹²

Incorporating questions into pre-class videos in a flipped classroom model enhances learning effectiveness compared to videos without questions. Wang documented that instructional videos with feedback mechanisms improve learning performance and intrinsic motivation among pre-service teachers. In the context of elementary schools, Anastasiades et al. reported that interactive videos support students' independent learning when combined with an online platform.¹³ In the same study, participants rated embedded

⁸ Logan Fiorella and Richard E Mayer, *Learning as a Generative Activity* (Cambridge university press, 2015).

⁹ M Noetel et al., "Multimedia Design for Learning: An Overview of Reviews With Meta-Meta-Analysis," *Review of Educational Research* 92, no. 3 (2022): 413–54, <https://doi.org/10.3102/00346543211052329>.

¹⁰ Carolina Robledo-Castro, Piedad Rocio Lerma Castaño, and Luis Gerardo Pachón Ospina, "Programas de Rehabilitación de Las Funciones Ejecutivas En La Aduldez Joven y Media Basados En Sistemas Computacionales: Una Revisión de Alcance," *Revista de Investigación e Innovación En Ciencias de La Salud* 5, no. 1 (2023): 231–72.

¹¹ R Deng, S Feng, and S Shen, "Improving the Effectiveness of Video-Based Flipped Classrooms with Question-Embedding," *Education and Information Technologies* 29, no. 10 (2024): 12677–702, <https://doi.org/10.1007/s10639-023-12303-5>.

¹² H Zheng et al., "The Combination of Segmentation and Self-Explanation to Enhance Video-Based Learning," *Active Learning in Higher Education* 25, no. 2 (2024): 285–302, <https://doi.org/10.1177/14697874221126920>.

¹³ Panagiotos S. Anastasiades et al., "Interactive Videoconferencing for Collaborative Learning at a Distance in the School of 21st Century: A Case Study in Elementary Schools in Greece," *Computers and Education* 54, no. 2 (2010): 321–39, <https://doi.org/10.1016/j.compedu.2009.08.016>.

questions as meaningful, though they noted the potential for distraction if interactive elements appeared too frequently.¹⁴

Edpuzzle is a web-based platform that allows educators to embed multiple-choice questions, open-ended questions, notes, and audio comments into videos and provides an analytics dashboard to monitor student learning progress. This platform simultaneously applies three CTML principles: the *segmenting* principle through automatic pauses at question points, the *generative processing* principle through embedded questions, and the feedback principle through instant responses, Vinothini Ramasany¹⁵ and Cáceres,¹⁶ reports that Edpuzzle enhances engagement and learning outcomes across various educational levels.

The existing Edpuzzle literature has three limitations that restrict its applicability to elementary schools in developing countries. *First*, the majority of studies were conducted at the upper secondary to higher education levels; empirical evidence on elementary school students, particularly fifth graders, remains very limited. *Second*, existing studies generally use pre-existing media without development and validation procedures aligned with student characteristics and the national curriculum. *Third*, existing studies are dominated by science and physical education; language learning with a focus on critical literacy particularly fact-opinion competencies is scarcely represented.

This gap is exacerbated by the structural conditions of elementary education in Indonesia. Suwanto identified limitations hindering the implementation of digital literacy in elementary schools in Yogyakarta, namely, infrastructure and teachers' pedagogical capacity.¹⁷ Existing practices tend to focus on technical ICT skills, with a very limited

¹⁴ Anastasiades et al., "Interactive Videoconferencing for Collaborative Learning at a Distance in the School of 21st Century: A Case Study in Elementary Schools in Greece."

¹⁵ Vinothini Ramasany, Norah Md Noor, and Norasykin Mohd Zaid, "Effects of Learning Using Edpuzzle Interactive Video Application on Students' Interest, Engagement and Achievement in Science Subjects," *Innovative Teaching and Learning Journal* 6, no. 2 (2022): 59–72.

¹⁶ S Q Cáceres, "Motion Graphics and Edpuzzle as Teaching Resources in a Flipped Classroom: An Experience in the Teaching-Learning Process of a Theoretical Subject in Industrial Design," *International Journal of Technologies in Learning* 28, no. 2 (2021): 63–77, <https://doi.org/10.18848/2327-0144/CGP/V28I02/63-77>.

¹⁷ D H Suwanto, B Setiawan, and S Machmiyah, "Developing Digital Literacy Practices in Yogyakarta Elementary Schools," *Electronic Journal of E-Learning* 20, no. 2 (2022): 101–11, <https://doi.org/10.34190/ejel.20.2.2602>.

emphasis on information evaluation.¹⁸ At the same time, research by Dewanda¹⁹ and Zainil²⁰ indicates that the adoption of educational technology platforms in Indonesia is increasing. However, this increase has not been matched by the development of pedagogically structured and empirically validated media.

An exploratory study in Grade 5 at SDN Ciwedus, Kasemen District, Serang City, revealed that these structural conditions are concretely manifested at the classroom level. Observations of two Indonesian language learning sessions revealed four indicators: *first*, instructional materials rely on static content without interactive components; *second*, student participation in class discussions is low; *third*, student scores on fact-and-opinion material fall below the Minimum Achievement Criteria (KKM); and *fourth*, students fail to apply the fact-opinion concept to contextual texts consistently. The first two indicators reflect classroom media and climate conditions that do not encourage active cognitive engagement; the last two indicators reflect the impact on student understanding. This pattern suggests that the problem lies not in students' abilities but in the design of the available learning experiences. This situation can be addressed through interventions based on interactive media.

The principle of *generative processing* in CTML has been applied to various interactive video platforms and proven effective at the middle and high school levels. However, no studies have reported on the development and validation of Edpuzzle-based media for critical literacy in elementary schools, particularly in developing countries. This gap is most evident in the facts-and-opinions competency material that underpins students' ability to evaluate digital information critically.

This study contributes in three dimensions. First, instructional design contribution: this study produced Edpuzzle-based interactive video media developed through the ADDIE model and validated using Aiken's V coefficient by a panel of experts.²¹ *Second*,

¹⁸ Suwanto, Setiawan, and Machmiah, "Developing Digital Literacy Practices in Yogyakarta Elementary Schools."

¹⁹ F Dewanda, A I Suroso, and N Hasanah, "Students' Adoption of Educational Technology Platforms for Nonformal Education in Indonesia," *Information Development*, 2025, <https://doi.org/10.1177/02666669241312161>.

²⁰ M Zainil et al., "A Needs Analysis on the Utilization of Learning Management Systems as Blended Learning Media in Elementary School," *Journal of Education and E-Learning Research* 11, no. 1 (2024): 56–65, <https://doi.org/10.20448/jeelr.v11i1.5310>.

²¹ M Xue et al., "Does a Video-Based and 3D Animation Hybrid Learning System Improve Teaching Outcomes in Orthopedic Surgery? A Randomized Controlled Trial," *Journal of Surgical Education* 81, no. 9 (2024): 1305–19, <https://doi.org/10.1016/j.jsurg.2024.05.015>.

the empirical-contextual contribution: this study documents changes in fifth-grade students' learning outcomes using a pretest–posttest design and N-Gain and Cohen's *d* analyses, providing preliminary evidence of the feasibility of interactive media-based interventions in environments with limited digital facilities. *Third*, contribution to the Indonesian literature: this study provides a standardized, documented, and replicable procedure for developing Indonesian language media that researchers and practitioners can use in similar contexts.

This study is guided by three research questions: (1) What is the procedure for developing Edpuzzle-based interactive video media for fact and opinion materials in fifth-grade elementary school through the ADDIE model stages? (2) What is the level of content validity and media feasibility of the developed media based on assessments by subject matter experts, media experts, and learning experts? (3) How do changes in student learning outcomes on factual and opinion-based content occur after media use, as measured by pretest–posttest scores, N-Gain, and Cohen's *d*?

Research Methodology

This study employs a Research and Development (R&D) approach using the ADDIE model.²² The iterative structure of this model allows for validation and revision at each phase, a characteristic aligned with media development principles based on the Cognitive Theory of Multimedia Learning.²³ The effectiveness of the media was tested using a one-group pretest–posttest design ($O_1 \rightarrow X \rightarrow O_2$), in which *X* comprised a learning session with the Edpuzzle interactive video. This design was chosen because no equivalent comparison class was available at the research site; consequently, the findings are susceptible to bias from history, maturation, and testing factors.

The primary participants were 35 fifth-grade students (ages 10–11) at a public elementary school in Banten Province, selected *purposively* based on two criteria: *first*, active participation in the Indonesian Language subject during the current semester; *second*, the identification of learning gaps regarding fact-opinion material through the classroom teacher's initial assessment. A small-group pilot study involved 10 students from a parallel class outside the main sample. The study was conducted from October

²² Robert Maribe Branch and İlhan Varank, *Instructional Design: The ADDIE Approach*, vol. 722 (Springer, 2009).

²³ Richard E Mayer, *The Cambridge Handbook of Multimedia Learning* (Cambridge university press, 2005).

2025 to January 2026, following approval from the school principal and the institutional ethics committee. Written *informed consent* was obtained from the participants' parents/guardians and verbal *assent* from the students, with the right to withdraw without consequences and data anonymity guaranteed throughout the study.

Development followed the five ADDIE phases sequentially with cross-phase feedback. *First, Analyze.* Needs analysis combined semi-structured interviews with classroom teachers, observation of two learning sessions, and analysis of previous semester learning outcome documents. This triangulation identified five learning gaps: low motivation, difficulty distinguishing facts from opinions, lack of interactive media, low classroom participation, and learning outcomes below the minimum competency standard (KKM). The characteristics of learners at the concrete operational stage informed the selection of visual-interactive media as the intervention method.²⁴

Second, Design. This phase produced three *deliverables*: a segmented video *storyboard* at conceptual transition points according to the *segmenting* principle; a question scheme embedded at strategic *cognitive checkpoints* to trigger generative processing; and a supporting instrument grid covering three cognitive levels: identification, discrimination, and contextual application. *Third, Develop.* The video is produced in Canva with narrative-visual coordination according to *modality* principles, then integrated into Edpuzzle with multiple-choice questions and formative feedback at predetermined points. Validation was conducted by three experts in educational media, elementary school Indonesian language materials, and language, within their respective domains of specialization. Post-validation revisions included adding usage instructions, highlighting learning objectives at the beginning of the video, and adjusting vocabulary in accordance with PUEBI.

Fourth, Implementation. A small-group pilot test was conducted to identify usability issues before implementation with a larger group. The main session consists of six measurable phases: a 15-item online *pretest* (20 minutes), a procedural introduction (5 minutes), individual video viewing with automatic pauses at each embedded question (30–35 minutes), a guided discussion on the most difficult items based on *the* Edpuzzle *dashboard* (15 minutes), a parallel *posttest* (20 minutes), and completion of a response

²⁴ C R Beal, "The Development of Text Evaluation and Revision Skills," *Child Development* 61, no. 1 (1990): 247–58, <https://doi.org/10.1111/j.1467-8624.1990.tb02776.x>.

questionnaire (10 minutes). Classroom teachers served as facilitators rather than content deliverers to ensure consistency of treatment among participants. *Fifth, Evaluate.* Formative evaluation guided iterative revisions in each phase; summative evaluation at the end of implementation assessed fulfillment of product success criteria.

Four instruments were used, each with a classified function. The learning outcome test consists of a 15-item multiple-choice test that is paralleled between *the pretest* and *posttest* in terms of indicator coverage and difficulty level, but uses different text contexts to minimize the memory effect. The expert validation sheet, using a 5-point Likert scale, covers three domains: media, content, and language. The student response questionnaire measured four dimensions of user acceptance: ease of use, content clarity, usefulness, and motivation for reuse. Observation sheets and semi-structured interview guides (for classroom teachers and a subsample of five students) generated qualitative data for interpretive triangulation in the discussion, rather than as primary evidence of effectiveness.

Before being used in the main study, the test was pilot-tested on a sample outside the research sample and analyzed in terms of four aspects: item validity (*Product Moment* correlation, $\alpha = 0.05$), reliability (KR-20 with a minimum threshold of 0.70),²⁵ level of difficulty, and discriminative power.²⁶

Expert validation scores were analyzed using Aiken's V coefficient at a minimum threshold of 0.60.²⁷ Questionnaire response scores were converted to percentages using the Riduwan scale,²⁸ with practicality set at $\geq 75\%$. Meanwhile, effectiveness was analyzed through four sequential procedures. The Lilliefors test ($\alpha = 0.05$) was used to assess the normality of the distribution as a prerequisite for parametric testing. A paired *t-test* ($\alpha = 0.05$) was then used to assess the significance of the mean difference, appropriate for *within-subjects* data in a single-group design. The effect size was calculated using Cohen's *d* with *the pooled standard deviation*:

²⁵ Arikunto Suharsimi, "Prosedur Penelitian Suatu Pendekatan Praktik," *Jakarta: Rineka Cipta* 134 (2006): 252.

²⁶ Rostina Sundayana, "Statistika Penelitian Pendidikan (Alfabeta)," 2016.

²⁷ Lewis R Aiken, "Three Coefficients for Analyzing the Reliability and Validity of Ratings," *Educational and Psychological Measurement* 45, no. 1 (1985): 131–42.

²⁸ Riduwan, *Pengantar Statistika Sosial* (Bandung: Alfabeta, 2016).

$$d = \frac{M_{posttest} - M_{pretest}}{\sqrt{(SD_{pretest}^2 + SD_{posttest}^2)/2}}$$

Classification follows Cohen (1988), as expanded by Sawilowsky (2009): *small* (≥ 0.2), *medium* (≥ 0.5), *large* (≥ 0.8), and *very large* (≥ 1.2). Cohen's d is reported as a supplement for methodological transparency in the *within-subjects* design. Finally, individual *Normalized Gain* was calculated using Hake's formula.²⁹ with classifications of low ($g < 0.30$), moderate ($0.30 \leq g < 0.70$), and high ($g \geq 0.70$)

$$g = \frac{posttest - pretest}{100 - pretest}$$

A product is deemed successful if it meets three criteria simultaneously: valid (Aiken's $V \geq 0.60$; KR-20 ≥ 0.70), practical (average student response $\geq 75\%$), and effective (significant paired *t-test* at $p < 0.05$, *N-Gain* ≥ 0.30 , and Cohen's $d \geq 0.8$).

Results and Discussion

Media development through the five ADDIE phases produced a product ready for pilot testing in Grade 5. The analysis phase identified five learning gaps: low motivation, difficulty distinguishing facts from opinions, lack of interactive media, low participation, and learning outcomes below the minimum passing score. These five gaps served as the empirical basis for formulating operational learning objectives. The design phase placed embedded questions at strategic *cognitive checkpoints* in accordance with the CTML *generative processing* principle.³⁰ The development phase included video production using Canva, integration into Edpuzzle with formative feedback, and validation by a panel of three experts. Post-validation revisions included three improvements: adding a user guide page, including learning objectives at the beginning of the video, and adjusting vocabulary in accordance with PUEBI.

The three expert validators rated the prototype using a 1–5 Likert scale; the results are summarized in Table 1.

²⁹ Richard Hake, "R.(1999). Analyzing Change/Gain Scores. AREA-D American Education Research Association's Devison. D," *Measurement and Reasearch Methodology* 1, no. 4 (1999): 48–56.

³⁰ Mayer, "The Past, Present, and Future of the Cognitive Theory of Multimedia Learning."

Table 1.
Expert Panel Validation Results

Validator	Score (%)	Category	V Aiken	Category V
Media Expert	90	Highly Recommended	0.875	High
Content Expert	76	Good	0.741	Moderate
Language Expert	81	Highly Suitable	0.763	Moderate
Average	82.33	Very Good	0.793	Moderate–High

All validators met the “Acceptable” threshold, and the overall Aiken’s V (0.793) exceeded the 0.60 criterion (Aiken, 1985). However, the 14-point difference between media experts (90%) and content experts (76%) warrants further elaboration.

Table 2.
Test Instrument Quality Analysis

Indicator	Result	Criteria	Status
Item validity (Pearson, $\alpha = 0.05$)	15 out of 15 valid	calculated $r >$ table r	Retained
Reliability (KR-20)	0.880	≥ 0.70	High
Difficulty level (easy/moderate/difficult)	5/5/5	Balanced proportion	Met
Discriminant power ≥ 0.30	12/15	≥ 0.30	Acceptable

Three items (numbers 10, 12, and 14) showed marginal discriminant validity below the threshold ($D = 0.28, 0.27,$ and 0.26). Still, they were retained to maintain construct coverage, with a note for revision in the next iteration of the instrument.

Table 3.
Descriptive Statistics for the Pretest and Posttest (N = 35)

Statistics	Pretest	Posttest	Difference
Mean (M)	62.37	83.83	+21.46
Standard deviation (SD)	10.74	9.30	
Minimum	40	67	
Maximum	80	100	

The *pretest* mean ($M = 62.37$) was below the passing score (70), confirming the initial knowledge deficit identified during the analysis phase; the *posttest* mean ($M = 83.83$) exceeded the passing score by 13.83 points. The decrease in standard deviation from 10.74 to 9.30 indicates that the improvement was evenly distributed across all participants, rather than driven by a small group of high-performing students. The Liliefors test confirmed the normality of the distribution of both scores (L -calculated $<$ L -table, $\alpha = 0.05$).

Table 4.
Paired t-Test and Effect Size

<i>t</i>	<i>df</i>	<i>t</i> -table ($\alpha = 0.05$)	<i>p</i>	Cohen's <i>d</i> (pooled) ^a	Cohen's <i>d</i> (<i>paired</i>) ^b	Decision
----------	-----------	-------------------------------------	----------	--	---	----------

26.80	34	2.032	< 0.001	2.14	4.53	H ₀ is rejected
-------	----	-------	------------	------	------	----------------------------

^a $d = 21.46 / \sqrt{((10.74^2 + 9.30^2)/2)} = 21.46 / 10.05 = 2.14$. ^b $d < z < /sub > = 21.46 / 4.74 = 4.53$.

The difference in scores is statistically significant, $t(34) = 26.80$; $p < 0.001$. Cohen's $d = 2.14$ falls into the "huge" category according to Sawilowsky (2009). Cohen's $d < z < /sub > = 4.53$ is reported as a complementary metric for *within-subjects* designs; since it is calculated using the standard deviation of paired differences (rather than *the pooled SD*), the two metrics are not on the same interpretive scale and cannot be directly compared.

Table 5.
Distribution of Individual *N-Gain*

Category	Criteria	Number	%
High	$g \geq 0.70$	8	22.9
Moderate	$0.30 \leq g < 0.70$	27	77.1
Low	$g < 0.30$	0	0
Total		35	100

The average *N-Gain* of 0.60 falls in the moderate category; no participants fall in the low category. The pattern of "moderately distributed, with no low scores" serves as the starting point for interpretation in §2 Discussion.

The average student response rate of 90.89% (Very High category) exceeded the 75% practicality threshold by a substantial margin, a finding consistent with patterns of user acceptance of interactive videos in an international context.

Table 6.
Summary of Product Success Criteria Compliance

Criteria	Indicator	Target	Result	Status
Valid	V Aiken (overall)	≥ 0.60	0.793	✓
	Reliability (KR-20)	≥ 0.70	0.880	✓
Practical	Student response	$\geq 75\%$	90.89%	✓
Effective	Paired <i>t-test</i>	$p < 0.05$	< 0.001	✓
	<i>N-Gain</i>	≥ 0.30	0.60	✓
	Cohen's <i>d</i>	≥ 0.8	2.14	✓

The product meets all six indicators of the three success criteria simultaneously.

Discussion

Differential Validity of Multimedia in the Cognitive Theory of Multimedia Learning

Confirmed content validity (V Aiken = 0.793) establishes the general validity of the media. Still, a 14-point difference between media experts' and content experts' scores

indicates that this validity is not uniform across all evaluation dimensions. Content experts specifically identified limitations in the contextual examples for evaluative reasoning related to the task type that forms the core of the fact-opinion competency.

CTML asserts that the effectiveness of multimedia does not depend on the aesthetics of the presentation, but rather on facilitating generative processing by providing context that enables learners to construct meaning actively.³¹ Noetel et al., in a review of ten meta-analyses of multimedia design principles, concluded that the principles most influential on learning outcomes are those that promote meaning construction, not those that enhance visual appeal.³² High media expert scores reflect success in the technical dimensions of segmentation, coherence, and visual cues; moderate content expert scores indicate that the generative dimensions have not been fully realized. La Torre and Désiron's analysis of instructional videos used by classroom teachers documents a persistent gap between multimedia design principles and their actual implementation in schools, suggesting that pedagogical alignment requires sustained content-expert involvement throughout development rather than only at the final validation stage.³³

Impact of Design Limitations and Moderate Gains

High statistical significance, $t(34) = 26.80$, $p < 0.001$, and Cohen's $d = 2.14$ (classified as "huge" according to Sawilowsky) reflect a substantial effect size in the conditions studied. However, the *one-group pretest–posttest* design requires careful interpretation: *testing* effects, maturation, and *demand characteristics* cannot be ruled out, and part of the observed effect size likely reflects these factors rather than the media's impact alone.³⁴ Reporting two effect size indices enhances interpretive transparency: $d_{z} = 4.53$ reflects the consistency of change within paired samples—i.e., the low variability of score differences among participants—while $d = 2.14$ reflects the magnitude of change relative to population variability.³⁵

³¹ Mayer, "The Past, Present, and Future of the Cognitive Theory of Multimedia Learning."

³² Noetel et al., "Multimedia Design for Learning: An Overview of Reviews With Meta-Analysis."

³³ S La Torre and J C Désiron, "From Research to Practice: Are Multimedia Principles Present in Instructional Videos Used by Teachers in Science and History?," *Technology, Knowledge and Learning* 29, no. 4 (2024): 1993–2016, <https://doi.org/10.1007/s10758-024-09753-2>.

³⁴ Shlomo S Sawilowsky, "New Effect Size Rules of Thumb," *Journal of Modern Applied Statistical Methods* 8, no. 2 (2009): 26.

³⁵ Christine M Anderson-Cook, "Experimental and Quasi-Experimental Designs for Generalized Causal Inference" (Taylor & Francis, 2005).

An N-Gain of 0.60 complements the interpretation of Cohen's *d* with a more contextual perspective. This value indicates that the intervention achieved approximately 60% of the available potential for improvement, consistent with the moderate category under Hake's criteria. Comparable patterns have been documented at the elementary level: Su and Chiu, examining sixth-grade pupils' use of Edpuzzle in Taiwan, reported strong engagement-driven outcomes that nonetheless fell short of transformative gains, supporting the inference that single-session interactive video interventions for elementary learners typically yield moderate rather than high learning gains.³⁶ Combined with the high student response rate (90.89%), these findings suggest that the medium successfully fosters active engagement but is not yet sufficient to produce deeper evaluative transfer.

Generative processing and retrieval practice in the effectiveness of instructional videos based on the Cognitive Theory of Multimedia Learning

The observed improvement patterns have convergent theoretical explanations from two perspectives. From the CTML perspective, questions inserted at cognitive checkpoints activate what Fiorella and Mayer call generative processing — the activity of selecting relevant information, organizing it into a coherent structure, and integrating it with prior knowledge.³⁷ Recent experimental evidence from Zeithofer and colleagues confirms that interactive educational videos equipped with feedback features reduce extraneous cognitive load and enhance knowledge acquisition relative to passive video viewing, providing direct empirical support for the embedded-question architecture used in the present study.³⁸

From a cognitive psychology perspective, each response to embedded questions triggers *retrieval practice*, *strengthening* memory encoding even when the initial response is incorrect and corrected through feedback. The combination of video segmentation with generative activities surpasses the effects of each strategy individually. The decrease in standard deviation from *pretest* to *posttest* (10.74 → 9.30) in this study reflects a tighter

³⁶ C.-Y. Su and C.-H. Chiu, "Perceived Enjoyment and Attractiveness Influence Taiwanese Elementary School Students' Intention to Use Interactive Video Learning," *International Journal of Human-Computer Interaction* 37, no. 6 (2021): 574–83, <https://doi.org/10.1080/10447318.2020.1841423>.

³⁷ Fiorella and Mayer, *Learning as a Generative Activity*.

³⁸ Vincent Dusanek and Ingo Kollar, "Interactive Educational Videos to Learn about Psychological Theories: Effects of Learner Control and Feedback Features on Extraneous Cognitive Load and Knowledge Acquisition," *Psychology Learning & Teaching* 25, no. 1 (2026): 35–57.

post-intervention distribution consistent with the theoretical prediction that repeated *retrieval* reduces interindividual variability on retention tasks.

The high student response rate (90.89%) complements cognitive mechanisms with an affective dimension. Adiy Tweissi identified that questions provide meaningful input and timely feedback, creating two opportunities for interaction among elementary school students.³⁹ The Edpuzzle analytics dashboard used in this study reinforces these opportunities for interaction by allowing teachers to identify points of difficulty in real time.

The moderate N-Gain category calls for an explanation of why the intervention's effects were not fully optimal. The explanation most consistent with the data points to the type of processing that the question format triggers. CTML distinguishes between learning for retention (remembering information) and learning for transfer (applying information to new situations), with transfer requiring deeper generative processing. Multiple-choice questions effectively trigger retention processing but are less effective at fostering the evaluative reasoning required for authentic fact-opinion tasks.

Within the PIRLS framework, fact-opinion competence resides at the highest level of reading comprehension evaluating content and textual elements which requires evaluative consideration of claims rather than mere identification of information.⁴⁰ Beal et al. demonstrate that text-evaluation skills in elementary school students develop through structured comprehension-monitoring exercises.

Contributions to the Effectiveness of Edpuzzle and Elementary School Students' Critical Skills

Within the PIRLS framework, fact-opinion competence resides at the highest level of reading comprehension evaluating content and textual elements which requires evaluative consideration of claims rather than mere identification of information,⁴¹ This study extends that foundation to the elementary school level a gap that had not been systematically addressed previously. The documented effectiveness of Edpuzzle among

³⁹ Adiy Tweissi, "The Effects of Embedded Questions Strategy in Video among Graduate Students at a Middle Eastern University" (Ohio University, 2016).

⁴⁰ Joyce Wangui Gikandi, Donna Morrow, and Niki E Davis, "Online Formative Assessment in Higher Education: A Review of the Literature," *Computers & Education* 57, no. 4 (2011): 2333–51.

⁴¹ Ina V S Mullis and Michael O Martin, *PIRLS 2021 Assessment Frameworks*. (ERIC, 2019).

adult learners can be extended to students aged 10–11, provided that evaluative transfer is assessed with questions that more strongly encourage generative processing.

This study contributes to three interrelated bodies of literature. In the Edpuzzle literature, previous empirical evidence has been concentrated on higher education and professional training contexts including business pedagogy (Mischel) and biochemistry laboratory instruction with comparatively little attention to elementary classrooms.⁴² In the literature on children's critical reading, this study complements recent action research by Po and colleagues, who reported statistically significant gains in fact-and-opinion comprehension among Filipino third-grade pupils following a tiered reading intervention ($t = 22.88, p < .05$).⁴³ The replication of this pattern in the Indonesian context using video-based media reinforces the claim that text-evaluation competencies among elementary school students respond to structured, cross-modal interventions across diverse Southeast Asian classrooms. In the context of the spread of digital misinformation, developing fact-opinion competencies starting in elementary school is an urgent educational priority.

Practical Implications

For classroom teachers, this validated media serves as an accessible and student-friendly alternative to text-based media. Edpuzzle's analytics dashboard, which tracks item-level responses and viewing behavior, can serve as a formative diagnostic tool to identify concepts that require further guidance. However, this media is best positioned as an initial step, not the entirety of the learning process. Evaluative transfer that is not yet optimal in a single session needs to be supported by follow-up activities: analysis of authentic news texts, discussions of ambiguous examples, and writing assignments that require justification of classifications.

For instructional media developers, the difference in scores between media experts and subject-matter experts provides concrete guidance for design improvements. Subsequent iterations should: (a) include questions requiring evaluation of authentic text excerpts, not merely identification of definitions; (b) add open-ended question formats

⁴² Leann J Mischel, "Watch and Learn? Using EDpuzzle to Enhance the Use of Online Videos," *Management Teaching Review* 4, no. 3 (2019): 283–89.

⁴³ Lovelyn Po et al., "Leveraging Tiered Instruction Strategy to Enhance Reading Comprehension and Quality Education: An Action Research among Early Elementary Learners in the Philippines," *Recoletos Multidisciplinary Research Journal* 13, no. 2 (2025): 35–42.

requiring written explanations; and (c) apply video segmentation at points of conceptual confusion, not at uniform time intervals.

For curriculum designers, these findings reinforce the argument that digital literacy and language literacy in elementary school are interdependent and need to be integrated. The competency to distinguish facts from opinions is not merely a unit in the Indonesian language curriculum, but a foundation for students' ability to critically evaluate information amidst the ever-expanding flow of digital content.

Limitations and Future Directions

Three limitations constrain the scope of the claims. First, the *one-group pretest–posttest* design does not allow for causal inferences because the effects of *testing*, *maturation*, and *history* cannot be ruled out. The appropriate claim is that media is associated with a practically and statistically significant improvement under the conditions studied, not that media causes such improvement. Second, the study involved only 35 students from a single class in one school during a single learning session; long-term retention and transfer to new texts were not assessed. Third, the 15-item multiple-choice instrument is more sensitive to *recognition* and *recall* than to *evaluative reasoning*. Hence, the reported gains are likely higher than those measured by *performance assessment*.

Four research agendas emerge from these limitations. *First*, a quasi-experimental design with a comparison group is needed to strengthen causal inferences. *Second*, the effects of embedded multiple-choice and open-ended questions on evaluative transfer need to be compared. *Third*, retention should be assessed via a *delayed posttest* at least two weeks after the intervention, accompanied by a *performance assessment* using authentic texts. *Fourth*, the role of the Edpuzzle analytics dashboard as a pedagogical decision-making tool for elementary school teachers in Indonesia requires empirical examination.

Conclusion

This study developed, validated, and tested the initial effectiveness of Edpuzzle-based interactive video media for teaching facts and opinions in fifth-grade elementary school classes using the ADDIE model. All three research objectives were met. The media was developed through five systematic stages, integrating CTML generative processing

principles into Edpuzzle's affordances. Content validity met the established threshold (V Aiken = 0.793), and instrument reliability was classified as high (KR-20 = 0.880). The use of the media was associated with a statistically significant improvement in learning outcomes, $t(34) = 26.80$, $p < 0.001$, and was practically meaningful, with Cohen's $d = 2.14$ (very large effect), and N-Gain = 0.60 (moderate category). A student response rate of 90.89% confirms the media's practicality. These findings address a gap in the literature: validated R&D studies on Edpuzzle-based media for fact-opinion competencies in elementary schools, particularly in developing countries, were previously unavailable.

Theoretically, this study provides preliminary evidence that the generative processing principles of CTML can be applied to 10–11-year-old students in environments with limited digital resources. In practice, this validated media can be adopted by teachers as an alternative to static media, provided it is positioned as an initial stage, supplemented by post-video activities to support evaluative transfer. Three limitations constrain the claims: a single-group design without a control group limits causal inferences; a single-school sample limits generalizability; and the dominance of multiple-choice formats measures retention rather than evaluative reasoning. Future research should adopt a quasi-experimental design with a comparison group, include delayed retention tests and performance assessments based on authentic texts, and examine the interaction between embedded question types (multiple-choice vs. open-ended) and the depth of critical literacy learning.

Bibliography

- Adler Berg, F S, T E Lundtofte, L Heiselberg, and L Frischlich. "Children and Digital Misinformation: A Scoping Review." *Global Studies of Childhood*, 2025. <https://doi.org/10.1177/20436106251398608>.
- Aiken, Lewis R. "Three Coefficients for Analyzing the Reliability and Validity of Ratings." *Educational and Psychological Measurement* 45, no. 1 (1985): 131–42.
- Anastasiades, Panagiotis S., George Filippousis, Labros Karvunis, Spiros Siakas, Arsitokritos Tomazinakis, Panagiota Giza, and Hellen Mastoraki. "Interactive Videoconferencing for Collaborative Learning at a Distance in the School of 21st Century: A Case Study in Elementary Schools in Greece." *Computers and Education* 54, no. 2 (2010): 321–39. <https://doi.org/10.1016/j.compedu.2009.08.016>.
- Anderson-Cook, Christine M. "Experimental and Quasi-Experimental Designs for Generalized Causal Inference." Taylor & Francis, 2005.
- Beal, C R. "The Development of Text Evaluation and Revision Skills." *Child Development* 61, no. 1 (1990): 247–58. <https://doi.org/10.1111/j.1467-8624.1990.tb02776.x>.
- Beal, C R, A C Garrod, and G J Bonitatibus. "Fostering Children's Revision Skills Through Training in Comprehension Monitoring." *Journal of Educational Psychology* 82, no. 2 (1990): 275–80. <https://doi.org/10.1037/0022-0663.82.2.275>.
- Best, Rachel M, Randy G Floyd, and Danielle S McNamara. "Differential Competencies Contributing to Children's Comprehension of Narrative and Expository Texts." *Reading Psychology* 29, no. 2 (2008): 137–64.
- Blakston, Angela, Simon Chambers, and Tanya Notley. "Young People, Algorithms and News: Exploring the Relationship between Algorithmic Literacy and News Literacy." *Journal of Youth Studies*, 2025, 1–17.
- Branch, Robert Maribe, and İlhan Varank. *Instructional Design: The ADDIE Approach*. Vol. 722. Springer, 2009.
- Cáceres, S Q. "Motion Graphics and Edpuzzle as Teaching Resources in a Flipped Classroom: An Experience in the Teaching-Learning Process of a Theoretical Subject in Industrial Design." *International Journal of Technologies in Learning* 28, no. 2 (2021): 63–77. <https://doi.org/10.18848/2327-0144/CGP/V28I02/63-77>.
- Deng, R, S Feng, and S Shen. "Improving the Effectiveness of Video-Based Flipped Classrooms with Question-Embedding." *Education and Information Technologies* 29, no. 10 (2024): 12677–702. <https://doi.org/10.1007/s10639-023-12303-5>.
- Dewanda, F, A I Suroso, and N Hasanah. "Students' Adoption of Educational Technology Platforms for Nonformal Education in Indonesia." *Information Development*, 2025. <https://doi.org/10.1177/02666669241312161>.
- Dusanek, Vincent, and Ingo Kollar. "Interactive Educational Videos to Learn about Psychological Theories: Effects of Learner Control and Feedback Features on Extraneous Cognitive Load and Knowledge Acquisition." *Psychology Learning & Teaching* 25, no. 1 (2026): 35–57.

- Escobar, J.-P., and V Espinoza. "Direct and Indirect Effects of Inhibition, Working Memory and Cognitive Flexibility on Reading Comprehension of Narrative and Expository Texts: Same or Different Effects?" *Reading and Writing Quarterly* 41, no. 2 (2025): 176–92. <https://doi.org/10.1080/10573569.2024.2400993>.
- Fiorella, Logan, and Richard E Mayer. *Learning as a Generative Activity*. Cambridge university press, 2015.
- Gikandi, Joyce Wangui, Donna Morrow, and Niki E Davis. "Online Formative Assessment in Higher Education: A Review of the Literature." *Computers & Education* 57, no. 4 (2011): 2333–51.
- Hake, Richard. "R.(1999). Analyzing Change/Gain Scores. AREA-D American Education Research Association's Devison. D." *Measurement and Reasearch Methodology* 1, no. 4 (1999): 48–56.
- Mayer, R E. "The Past, Present, and Future of the Cognitive Theory of Multimedia Learning." *Educational Psychology Review* 36, no. 1 (2024). <https://doi.org/10.1007/s10648-023-09842-1>.
- Mayer, Richard E. *The Cambridge Handbook of Multimedia Learning*. Cambridge university press, 2005.
- Mischel, Leann J. "Watch and Learn? Using EDpuzzle to Enhance the Use of Online Videos." *Management Teaching Review* 4, no. 3 (2019): 283–89.
- Mullis, Ina V S, and Michael O Martin. *PIRLS 2021 Assessment Frameworks*. ERIC, 2019.
- Noetel, M, S Griffith, O Delaney, N R Harris, T Sanders, P Parker, B del Pozo Cruz, and C Lonsdale. "Multimedia Design for Learning: An Overview of Reviews With Meta-Meta-Analysis." *Review of Educational Research* 92, no. 3 (2022): 413–54. <https://doi.org/10.3102/00346543211052329>.
- Po, Lovelyn, Dame Francis Louise Ersando, Doreen Homecillo, Flordelis Quitara, Maria Jessca Amaquin, Abegail Quezon, Marie Zar Balanay, and Richard Mascardo. "Leveraging Tiered Instruction Strategy to Enhance Reading Comprehension and Quality Education: An Action Research among Early Elementary Learners in the Philippines." *Recoletos Multidisciplinary Research Journal* 13, no. 2 (2025): 35–42.
- Ramasany, Vinohthini, Norah Md Noor, and Norasykin Mohd Zaid. "Effects of Learning Using Edpuzzle Interactive Video Application on Students' Interest, Engagement and Achievement in Science Subjects." *Innovative Teaching and Learning Journal* 6, no. 2 (2022): 59–72.
- Riduwan. *Pengantar Statistika Sosial*. Bandung: Alfabeta, 2016.
- Robledo-Castro, Carolina, Piedad Rocio Lerma Castaño, and Luis Gerardo Pachón Ospina. "Programas de Rehabilitación de Las Funciones Ejecutivas En La Adultez Joven y Media Basados En Sistemas Computacionales: Una Revisión de Alcance." *Revista de Investigación e Innovación En Ciencias de La Salud* 5, no. 1 (2023): 231–72.
- Santos, S, I Cadime, F L Viana, G Prieto, S Chaves-Sousa, A G Spinillo, and I Ribeiro. "An Application of the Rasch Model to Reading Comprehension Measurement."

- Psicologia: Reflexao e Critica* 29, no. 1 (2016). <https://doi.org/10.1186/s41155-016-0044-6>.
- Sawilowsky, Shlomo S. “New Effect Size Rules of Thumb.” *Journal of Modern Applied Statistical Methods* 8, no. 2 (2009): 26.
- Su, C.-Y., and C.-H. Chiu. “Perceived Enjoyment and Attractiveness Influence Taiwanese Elementary School Students’ Intention to Use Interactive Video Learning.” *International Journal of Human-Computer Interaction* 37, no. 6 (2021): 574–83. <https://doi.org/10.1080/10447318.2020.1841423>.
- Suharsimi, Arikunto. “Prosedur Penelitian Suatu Pendekatan Praktik.” *Jakarta: Rineka Cipta* 134 (2006): 252.
- Sundayana, Rostina. “Statistika Penelitian Pendidikan (Alfabeta),” 2016.
- Suwarto, D H, B Setiawan, and S Machmiyah. “Developing Digital Literacy Practices in Yogyakarta Elementary Schools.” *Electronic Journal of E-Learning* 20, no. 2 (2022): 101–11. <https://doi.org/10.34190/ejel.20.2.2602>.
- Torre, S La, and J C Désiron. “From Research to Practice: Are Multimedia Principles Present in Instructional Videos Used by Teachers in Science and History?” *Technology, Knowledge and Learning* 29, no. 4 (2024): 1993–2016. <https://doi.org/10.1007/s10758-024-09753-2>.
- Tweissi, Adiy. “The Effects of Embedded Questions Strategy in Video among Graduate Students at a Middle Eastern University.” Ohio University, 2016.
- Xue, M, P Liu, J Zhang, Y Sun, Y Fang, J Yang, C Zeng, et al. “Does a Video-Based and 3D Animation Hybrid Learning System Improve Teaching Outcomes in Orthopedic Surgery? A Randomized Controlled Trial.” *Journal of Surgical Education* 81, no. 9 (2024): 1305–19. <https://doi.org/10.1016/j.jsurg.2024.05.015>.
- Zainil, M, Y Helsa, C Sutarsih, and S Nisa. “A Needs Analysis on the Utilization of Learning Management Systems as Blended Learning Media in Elementary School.” *Journal of Education and E-Learning Research* 11, no. 1 (2024): 56–65. <https://doi.org/10.20448/jeelr.v11i1.5310>.
- Zheng, H, R M Branch, L Ding, D Kim, E Jung, Z Lu, T Li, Z Pan, and M Yoon. “The Combination of Segmentation and Self-Explanation to Enhance Video-Based Learning.” *Active Learning in Higher Education* 25, no. 2 (2024): 285–302. <https://doi.org/10.1177/14697874221126920>.