Vol. 4 No. 3 September 2026, pp. 1266-1283
INTERNATIONAL JOURNAL OF
POST AXIAL
FUTURISTIC TEACHING AND LEARNING
https://journal.amorfati.id/index.php/postaxial | | ISSN 3025-7549

The Use of Gemini AI as a Learning Tool for Poster Creation in the Topic of Technological and Scientific Development for Grade XII Students at SMA Al-Hikmah Surabaya

Muhammad Yazid Al Busthomi[1a]*, Ajitama Putra Nugraha[2b], Muhibbul Anwar[2c]
[1]Universitas Sebelas Maret, Indonesia.
[2]Universitas Negeri Surabaya, Indonesia.

[a]muhammadyazidalbusthomi@gmail.com , [b]ajitama.23006@mhs.unesa.ac.id,
[c]24041344066@mhs.unesa.ac.id

Article History:
Received:
12-08-2026
Revised:
22-08-2026
Accepted:
04-09-2026

Keywords:
Gemini AI; history education; digital posters; AI literacy; information curation

***Correspondence Address:**
muhammadyazidalbusthomi@gmail.com

Abstract:
This study analyzes how 25 Grade XII students used Gemini AI to create posters on the development of science and technology. A descriptive qualitative design was employed with 25 students at Al-Hikmah Senior High School, Surabaya. The data comprised 25 prompts submitted to Gemini and students' responses to ten open-ended questions after the activity. The data were examined through reflexive thematic analysis, complemented by a descriptive reading of prompt length and components. The findings show that students used Gemini as an integrated production environment to locate information, formulate titles and copy, generate visual ideas, and revise designs. Their practices nevertheless formed a spectrum ranging from near-complete delegation to collaboration in which students acted as directors and editors. All 25 students reported encountering limitations in either the poster activity or their previous use of AI. Task-related difficulties included formulating or repeating prompts (15 students), visual mismatches or errors in image-embedded text (11 students), and overly long, general, excessive, or questionable content (7 students). Students reported using Gemini to search for or organize information (21/25), formulate titles or poster copy (22/25), generate visual ideas or images (24/25), and obtain feedback or revise designs (17/25). Twenty-one students described selecting, shortening, combining, or editing AI-generated materials before incorporating them into their posters. This study contributes the concept of dual learning: students learned the history of technological development while developing AI literacy through prompting, curation, verification, and creative decision-making.

Introduction (مقدمة)

Generative artificial intelligence is changing students' relationships with information, media, and the production of academic work. Rather than merely retrieving existing documents through search engines, students can ask a system to generate summaries, text structures, illustrations, and visual designs from natural-language instructions. This shift makes generative AI relevant to learning because it shortens the distance between an idea and an artifact. At the same time, it relocates cognitive effort from searching toward formulating requests, selecting outputs, checking their accuracy, and editing them. Recent scholarship identifies these capacities as elements of AI literacy, which includes knowledge of AI, the ability to use it, critical evaluation of its outputs, and ethical awareness of its consequences (Ng et al., 2021; Wang et al., 2023; Biagini, 2025).

This shift is particularly important in secondary education. For several years, research on AI in education has developed more extensively in higher education, while empirical evidence of its actual use in school classrooms remains comparatively limited and fragmented (Zawacki-Richter et al., 2019; Bond et al., 2024; Lee & Kwon, 2024). The gap is not simply a matter of the number of studies. Secondary students are forming epistemic habits: how they define questions, accept or reject claims, and distinguish persuasive-sounding answers from accountable information. AI integration at the K–12 level must therefore cultivate critical judgment, protect student agency, and develop ethical understanding rather than merely improve task-completion efficiency (Akgun & Greenhow, 2022; Holmes et al., 2022; Walter, 2024).

In history education, the issue is especially acute because historical knowledge cannot be reduced to a collection of chronological facts. Learning history requires students to relate change and continuity, evaluate evidence, understand causation, and situate present developments within longer temporal trajectories. Historical consciousness is built when learners make meaning from relationships among the past, present, and possible futures, not when they simply memorize lists of inventions or dates (Popa, 2022). Generative systems can summarize and visualize sequences of technological development, yet their linguistic fluency may conceal oversimplification, presentism, or factual inaccuracy. The educational value of AI in history should consequently be assessed through the reasoning that accompanies media production, not merely through the speed or appearance of the final product.

The development of science and technology offers a distinctive setting for examining AI. Gemini is both a learning medium and an example of the contemporary technology under study. Students can use it to organize a timeline from radio and computers to the internet, smartphones, and AI. This dual relationship makes the topic concrete, but may also present AI as the culmination of technological history and obscure the social, political, economic, and scientific processes shaping change. Generative AI therefore requires designs that preserve efficiency without removing student thinking (Lim et al., 2023; Farrokhnia et al., 2024; Kasneci et al., 2023).

Posters require students to transform information into multimodal communication by choosing facts, condensing sentences, creating hierarchy, relating text and images, and adapting a message to an audience. Limited visual space forces prioritization. Combining linguistic, visual, and spatial resources can broaden engagement and enable representational decisions (Jiang, 2018; Zhang et al., 2023). Quality therefore depends on connecting meaning, form, tools, and communicative purpose, creating additional instructional demands for teachers (Jiang et al., 2022).

Research on digital posters suggests that poster tasks can strengthen conceptual understanding, creativity, and communication when students participate actively in selecting and presenting information (Ahmad, 2019; Wisanti et al., 2024). In history education, e-posters have also been reported to support historical awareness by visualizing content and making it readily accessible (Sumantri et al., 2024). These studies, however, generally treat the poster as a medium or learning product. Far fewer studies examine how generative AI changes the cognitive division of labor behind the poster: which tasks are delegated to AI, which decisions remain with students, how outputs are checked, and when students recognize that an answer or image should be rejected.

The division of cognitive labor matters because generative AI can simultaneously find ideas, write text, provide feedback, and generate images. This diversity enables personalization and follow-up dialogue, but also allows students to transfer a chain of decisions to one system (Su & Yang, 2023; Tlili et al., 2023). Meaningful learning requires a dialogic arrangement in which AI proposes possibilities while students question, compare, and remain accountable. Such interaction can promote critical questioning and knowledge co-construction when teachers deliberately design assessment and reflection (Tang et al., 2024).

Prompt formulation is part of this process because a prompt represents purpose, prior knowledge, task boundaries, visual preferences, and success criteria. "Make a technology poster" leaves more decisions to the system than a request specifying audience, focus, chronology, dimensions, color, information hierarchy, illustration style, and exclusions. Detail may improve relevance but cannot guarantee factual or aesthetic quality. Prompting proficiency must be paired with output evaluation and strategic revision (Walter, 2024; Kasneci et al., 2023).

Evaluation is essential because probabilistic generation does not guarantee correspondence with evidence. Outputs may contain plausible but unsupported data, dates, or causal relationships. Such hallucinations can be expressed fluently and remain difficult to detect without background knowledge (Ji et al., 2023). Accuracy, bias, privacy, dependency, and integrity consequently require pedagogical mediation (Yan et al., 2024; Holmes et al., 2022). Asking for "reliable sources" is insufficient unless students verify their existence, relevance, and support for each claim.

History education provides a strong foundation for verification. Students often judge credibility from appearance or read one source vertically, whereas fact-checkers identify publishers, examine evidence, and compare independent reports (McGrew et al., 2018; Breakstone et al., 2021). Explicit lateral-reading instruction improves digital evaluation when practiced with authentic sources (McGrew, 2020; McGrew & Breakstone, 2023). AI adds a synthesized text that may conceal the origins of individual claims, making source verification integral to poster assignments.

Creativity and ownership present another concern. Generative AI can expand initial options and support students who struggle to begin, yet similar recommendations may make outputs visually or conceptually alike. AI assistance can improve individual work, especially for creators with lower initial ability, while reducing collective diversity (Doshi & Hauser, 2024). School assessment should therefore examine the decision trail: what originated in the student's purpose, what AI suggested, what was rejected, and how the poster was revised.

Studies of user perceptions in education show that students generally value AI's speed, accessibility, and support for idea generation, while remaining concerned about reliability, dependency, and effects on learning (Chan & Hu, 2023). A SWOT analysis of ChatGPT similarly indicates that benefits such as rapid feedback and personalized support can become weaknesses when outputs are accepted without examination or when the teacher's role is reduced to supervising a tool (Farrokhnia et al., 2024). This position is relevant to Gemini because the principal pedagogical issues are not determined by a system's brand but by how its generative capacities are arranged within activities, reference sources, assessment rubrics, and user responsibilities.

AI literacy extends beyond operating an interface to conceptual understanding, effective use, critical evaluation, communication with systems, and ethical consideration (Ng et al., 2021; Biagini, 2025). Competence measures distinguish understanding, application, evaluation, and ethical awareness; image generation alone does not constitute mature literacy (Wang et al., 2023). In K–12 classrooms, concept introduction and project creation still dominate, while critical evaluation and attention to social consequences require further development (Lee & Kwon, 2024).

On this basis, the use of Gemini in poster assignments should be analyzed as human–AI orchestration. Orchestration refers to how students divide, sequence, and supervise work among themselves, AI, and other sources. Students may act as operators who provide a single prompt, curators who choose among several outputs, or co-creators who establish goals, engage in dialogue, verify facts, and edit results. This spectrum avoids a simplistic dichotomy between work "made by students" and work "made by AI." Assessment instead centers on the quality and accountability of decisions, in line with calls for AI in education to be developed with ethics, transparency, collaboration, and methodological rigor (Bond et al., 2024; Akgun & Greenhow, 2022).

Despite growing research on generative AI in education, at least four gaps remain. First, much of the existing evidence has been produced in higher education, language learning, or STEM contexts rather than secondary history classrooms. Second, studies of digital posters generally examine learning outcomes or final products without investigating the AI-assisted production process. Third, studies of students' perceptions do not adequately explain how students formulate prompts, evaluate generated claims, revise unsuitable outputs, and transfer selected materials into a poster. Fourth, limited attention has been given to the distribution of decision-making between students and generative AI. Consequently, students' actual roles as directors, curators, verifiers, and editors remain insufficiently documented.

The Indonesian context is also significant. Research on e-posters has demonstrated the value of visual media for strengthening conceptual understanding and communication, including in history education, but generative AI as part of the production process has received less attention (Wisanti et al., 2024; Sumantri et al., 2024). International research on generative AI remains concentrated in higher education and in language or STEM fields, whereas its use to study the history of technology in secondary education involves different epistemic demands (Zawacki-Richter et al., 2019; Lee & Kwon, 2024). A study grounded in student voices is necessary to understand actual practices rather than only conceptually proposed possibilities.

This study addresses that gap through a qualitative analysis of prompts and reflective responses from Grade XII students at Al-Hikmah Senior High School, Surabaya, after they created posters on science and technology development using Gemini AI. Its purpose is neither to test a causal increase in achievement nor to assess the model's technical sophistication. The study examines how students used Gemini at different stages, how they selected and verified outputs, what they understood about science and technology development, how they maintained creativity, and what difficulties arose. This process-centered focus is important because educational technology adoption can appear successful in a final product while concealing the epistemic decisions that determine learning quality.

The research question was: How do Grade XII students use Gemini AI in the process of creating posters on the development of science and technology? Four analytical concerns elaborated the question: Gemini's functions within the production workflow; patterns of prompting and revision; strategies of curation, verification, and creative control; and the relationship between the experience of using AI and students' understanding of science and technology development. The study aimed to produce a contextualized explanation and practical principles for designing AI-assisted history learning that remains centered on reasoning, sources, and student agency.

This study contributes a process-centered account of students' actual practices when using Gemini in a secondary history classroom. Its novelty lies in two related contributions. First, it introduces dual learning, defined as the simultaneous development of historical understanding and AI literacy through direct interaction with AI-generated information and visuals. Second, it maps students' changing roles as directors who define objectives, curators who select outputs, verifiers who examine claims against external sources, and editors who revise language and visual elements. This perspective extends previous studies that have focused primarily on perceptions, technological affordances, or the quality of final products. Empirically, the study maps variation in students' uses of Gemini in a history classroom. Conceptually, it proposes dual learning, a situation in which AI is both a learning medium and an object understood as part of the history of technology. Practically, it identifies the need for a process rubric that evaluates objectives, prompts, verification, editing, and reflection rather than only the poster's visual quality. From this perspective, AI is positioned neither as a shortcut nor as a singular threat but as an environment that must be governed through pedagogical design and human responsibility (Su & Yang, 2023; Tang et al., 2024; Walter, 2024).

Method (منهج)

This study employed a descriptive qualitative design bounded by one Gemini AI–assisted poster-making activity on the development of science and technology in a Grade XII class at Al-Hikmah Senior High School, Surabaya. This design was selected because the research question concerned how students interpreted and enacted a process rather than the measurement of a treatment effect. The analysis attended both to recurring patterns of meaning and to important variations among students. This orientation is consistent with reflexive thematic analysis, which treats themes as products of the researcher's interpretive engagement with data rather than as categories presumed to emerge automatically (Braun & Clarke, 2019; Braun & Clarke, 2022).

The learning activity required students to create a poster on the development of information technology or science and technology using Gemini. The activity was conducted in a two session lasting approximately 150 minutes on August 3, 2026. After completing the assignment, students submitted a form on August 10, 2026. It recorded the prompt they used, a link to the uploaded poster, their prior knowledge of Gemini, sources and idea-development practices used before Gemini, parts of the process supported by AI, reasons for selecting outputs, understanding of the subject after the activity, experiences with problematic outputs, verification strategies, efforts to maintain creativity, and difficulties encountered. This sequence enabled the researcher to examine experiences before, during, and after poster production without converting them into causal claims.

The corpus included all 25 available and complete responses from Grade XII students. Each data row contained a student identity, attendance number, one prompt, one upload link, and ten open-ended responses. Inclusion criteria were participation in the activity, submission of a prompt and assignment link, and completion of all open-ended questions. No response was excluded because all 25 rows met these criteria. Demographic information other than grade level and school was unavailable; accordingly, the study did not compare students by gender, age, academic ability, or technological experience.

The amount of data was determined by complete coverage of the available classroom case rather than by terminating recruitment at saturation. This distinction matters because qualitative sample adequacy should be considered in relation to the study's purpose, group specificity, data richness, and case boundaries rather than a universal numerical threshold (Vasileiou et al., 2018; Hennink & Kaiser, 2022). The corpus provided 250 open-ended answers and 25 prompts, which was sufficient to identify patterns and contrasting cases within the class but was not intended to represent all senior high school students.

Student names were removed from the analytical corpus and replaced with codes P01–P25. Quotations in the article use these codes. Spelling in translated quotations was lightly standardized for readability without altering the substance, key word choices, or the student's position toward AI. Upload links were retained only as metadata confirming completeness of data collection. Poster images were not coded because the image files were not embedded in the analyzed workbook. This boundary prevents unsupported claims about visual quality, factual accuracy within the posters, or correspondence between prompts and products.

The instrument was an open-ended questionnaire designed to capture six dimensions: prior knowledge; information-search and idea-development practices; Gemini's functions in the production process; selection and verification of outputs; understanding of the subject matter; and creative agency together with encountered difficulties. Questions requested narratives and examples rather than agreement with closed statements. This format allowed students to report practices not fully anticipated by the researcher, although the depth of each response remained dependent on their reflective capacity and willingness to write.

Prompt data were treated as process artifacts that complemented self-reports. Their inclusion made it possible to inspect whether students' descriptions of content, style, format, and visual requirements were visible in the instructions they submitted. Combining these two data types was not presented as perfect validation, but as limited source triangulation between what students requested from AI and how they described using it. Thematic analysis can integrate different textual data when units of analysis, production contexts, and interpretive limits are explained consistently (Kiger & Varpio, 2020; Naeem et al., 2023).

The analysis followed six stages adapted from reflexive thematic analysis. First, all responses were read repeatedly to understand tone, terminology, and contradictions within each case. Second, initial semantic codes were assigned to relevant segments, including information search, text composition, visual generation, prompt revision, selection, source checking, design control, chronological understanding, and awareness of limitations. Third, codes were grouped into candidate themes. Fourth, themes were reviewed against every response to locate cases that supported, extended, or challenged dominant patterns. Fifth, theme boundaries and names were refined. Sixth, themes were written as arguments that answered the research question (Braun & Clarke, 2019; Byrne, 2022; Kiger & Varpio, 2020).

The analytical process did not equate frequency with thematic importance. An experience reported by only one student was retained when it revealed substantive variation, such as an admission that AI produced the entire poster or an explanation linking science and technology development to the Second World War and the Cold War. This principle is consistent with the view that themes are constructed from patterns of meaning relevant to the question rather than from word counts alone (Braun & Clarke, 2022; Byrne, 2022). Frequencies were used descriptively to show the distribution of practices within the case, not to test hypotheses or generalize to a population.

Prompts were analyzed through two layers of reading. The first calculated word counts to describe variation in instructional density. Based on the observed distribution, prompts were classified as minimal (20 words or fewer), directive (21–60 words), and structured (more than 60 words). These boundaries were operational categories for this corpus, not universal standards of prompt quality. The second layer identified requested components, including topic, chronology, information type, dimensions, orientation, color, style, illustrations, legibility, audience, resolution, and design constraints. Prompt features were then compared with students' narratives of difficulties and revision strategies.

Initial codes were recorded in a case-by-code matrix and subsequently refined to make decisions traceable. The researcher's disciplinary background and involvement in the learning activity were treated as part of the interpretive process. Reflexive notes were maintained during coding to record assumptions, changes in code definitions, and decisions concerning theme boundaries. Interpretations were repeatedly compared with prompts, student responses, and divergent cases to prevent the analysis from merely confirming the researcher's initial expectations. The matrix supported consistency without turning the analysis into a mechanical procedure. Codebook development and a decision trail improve transparency, especially when data contain responses of varying length and overlapping content (Roberts et al., 2019; Naeem et al., 2023). Intercoder-agreement coefficients were not calculated because the analysis was conducted reflexively by the researcher and did not claim that one coding scheme represented the only objective account of the data. Debates on intercoder reliability indicate that such choices should be aligned with the analytical paradigm and reported openly (O'Connor & Joffe, 2020; Braun & Clarke, 2022).

Trustworthiness was supported through four procedures. First, every theme was rechecked against the complete set of cases, including divergent responses. Second, major claims were supported by direct quotations and descriptive frequencies. Third, prompts were compared with open-ended answers to reduce reliance on a single form of data. Fourth, claim boundaries were made explicit: the study did not analyze poster images, measure learning gains, or infer causal relationships. These procedures follow recommendations that thematic analysis should report links among data, codes, themes, and the researcher's interpretive position transparently (Roberts et al., 2019; Kiger & Varpio, 2020; Braun & Clarke, 2019).

Participants were protected during analysis and reporting by removing names, withholding personal links, and limiting quotations to information relevant to the research question. Because the raw data contain student identities, they are not included as a public appendix. Responses were used only for educational analysis, and no individual assessment of student competence was made. This article reports an anonymization process that can be inspected within the manuscript; The students were informed that their responses would be anonymized and used for research purposes. Participation in completing the reflective questionnaire was voluntary, and no information that could identify individuals was included in the analysis..

Discussion (مناقشة)

Result: Gemini as an Integrated Production Environment

The data show that Gemini was not restricted to a single stage of the assignment. Of the 25 students, 21 explicitly described using it to search for or organize information; 22 used it for titles, slogans, or poster copy; 24 incorporated a visual, image, or design function; and 17 referred to feedback or revision. These categories overlap because one student could use several functions. Students therefore experienced Gemini as an integrated production environment linking content search, linguistic packaging, and visual execution.

Table 1. Distribution of Major Codes for Gemini AI Use

Analytical code	n of 25	Indication in responses
Searching for or organizing information	21	Finding data, summarizing material, and organizing the development of technology
Poster titles, slogans, or copy	22	Creating titles, condensing sentences, and producing readable text
Visual ideas, design, or image generation	24	Determining color, style, illustration, layout, and generating images
Feedback or design revision	17	Changing backgrounds, improving composition, repeating prompts, and requesting alternatives
Selecting, editing, or not copying the entire output	21	Choosing relevant content, removing material, revising language, and determining the final product
Encountering an AI limitation or mismatch	25	Inaccurate information, overly general content, visual mismatch, and errors in image text

Source: Research data, 2026. Frequencies are based on descriptive coding and are not intended as inferential statistics.

Despite the breadth of functions, students' positions toward AI were not uniform. At one end of the spectrum, P01 stated, "For this assignment, Gemini AI did every part; I only typed the prompt." This statement indicates near-complete delegation: the student's contribution was concentrated in the initial command, while content and form decisions were assigned to the system. Several other responses indicated that Gemini could perform "the entire sequence," although these students still referred to selecting points or editing after an output was generated.

At the other end, students positioned AI as a tool under user control. P11 explained, "I treated Gemini AI as a discussion partner (copilot), not as the decision maker for the final product." P20 drew a more specific boundary: AI served as a visual-execution tool, whereas the concept, content sequence, and final design decisions remained the student's responsibility. Between these endpoints lay the most common pattern: Gemini generated initial materials, after which students selected, edited, combined, or transferred them to Canva. The label "using AI" therefore conceals considerable differences in cognitive division of labor.

Before the activity, 20 students named Google, websites, or articles as their routes to information. Some also used Canva, Pinterest, books, teacher materials, peers, ChatGPT, or other AI systems. After adopting Gemini, activities previously distributed across several applications were consolidated in one interface. This consolidation helps explain Gemini's appeal: students could move from an idea to text and visuals without changing the basic form of interaction. Yet the convenience also increases the risk that one system becomes the source, editor, and designer without external scrutiny.

These findings support the view that a poster assignment is a multimodal activity rather than a transfer of text onto an image plane. Students manage relationships among content, typography, color, illustration, and visual sequence, as emphasized in studies of digital multimodal composition (Jiang, 2018; Zhang et al., 2023). Digital posters can strengthen conceptual understanding and communication when students make representational decisions (Wisanti et al., 2024). The present study adds an important condition: these benefits depend on who makes the decisions, how the decisions are reviewed, and whether students can explain the reasons behind AI-informed choices.

Prompts Ranged from Minimal Commands to Structured Specifications

Prompt length ranged from 10 to 299 words, with a mean of 68.88 and a median of 31 words. No single type dominated the distribution. Eight prompts contained 20 words or fewer, nine contained 21–60 words, and eight exceeded 60 words. This variation reflects different ways of conceptualizing the task. Minimal prompts named only a topic and a broad expectation, while structured prompts included historical sequence, style, color, dimensions, audience, image components, text quality, and even elements to avoid.

Table 2. Operational Complexity of Student Prompts

Category	**Operational boundary**	**n**	**Dominant characteristics**
Minimal	≤20 words	8	A topic and general request; many decisions left to AI
Directive	21–60 words	9	Topic accompanied by color, mood, visual objects, or a content focus
Structured	>60 words	8	Detailed content, chronology, format, style, audience, visual limits, and print output

Source: Research data, 2026. Categories were created to interpret this corpus and are not universal prompt standards.

Minimal prompts included commands such as "Please create a poster about the development of information technology" or requests for a poster that was "interesting and funny." These commands were sufficient to initiate generation but did not specify periodization, scope, sources, visual hierarchy, or success criteria. AI consequently received substantial latitude to determine both content and style. This pattern accords with several students' recognition that outputs could easily shift away from the intended result when a prompt was "slightly off track" or insufficiently detailed.

Directive prompts added preferences such as a historical atmosphere, particular colors, a focus on mobile phones or high-speed networks, and images ranging from early computers to

AI. Students began to define a visual identity and content focus, although several aspects remained implicit. Structured prompts demonstrated an ability to design specifications. P04, for example, requested a vertical A4 poster, outlined stages from letters and telephones through AI, required positive and negative effects and a message about responsible use, and specified illustration type, legibility, Indonesian language, and print readiness. P19 used negative constraints, asking the system to avoid generic corporate graphics, excessive neon, complicated backgrounds, and an overly formal infographic style.

Prompt sophistication did not remove the need for revision. Fifteen students reported difficulty formulating or repeating prompts. Eleven encountered visual mismatch, generic design, an unsuitable background, or typographical errors within generated images. Seven faced answers that were too long, too general, excessive, or of questionable accuracy. Two reported network or upload limitations, while two reported no task-specific difficulty. P20 summarized a recurring image-generation problem: text embedded in an image was difficult to make "100% correctly spelled" and could contain glitches.

These data indicate that effective prompting is iterative. Students must not only add details but also diagnose the cause of failure: whether it lies in the content, relationship among elements, legibility, or the system's visual interpretation. The finding supports the inclusion of prompt engineering in AI literacy while showing that prompting must be understood alongside critical evaluation rather than as a one-shot technique (Walter, 2024; Kasneci et al., 2023). Prompt length is not synonymous with quality; what matters is precision of purpose and criteria, together with the ability to use an output as feedback for the next instruction.

Curation and Verification Emerged, but Their Depth Varied

Twenty-one students explicitly stated that they did not copy the entire output directly. They selected relevant information, removed material that had not been taught, shortened text, revised language, or adjusted the design. The most frequently reported criteria were thematic relevance, accuracy, clarity, readability, and compliance with assignment requirements. Most students thus recognized the need for human intervention after generation.

Negative experiences reinforced this awareness. All 25 students reported encountering at least one limitation of Gemini or another AI system, either during the poster activity or in previous use. Problems included information that differed from books, overly general answers, questionable dates or data, unsuitable formulas, misspelled words in images, overlapping composition, and visuals that did not follow the instructions. P12 wrote that even when AI was asked to provide complete content, the output remained generic and included "invented data." P09 referred to possible "AI hallucination," meaning data or dates that appeared convincing but did not match official sources when checked.

The most common verification route was the web, articles, journals, or trusted websites, which appeared in 21 cases. Twelve students named books, notes, presentations, or teacher-provided materials. Some compared outputs with another AI, while others asked Gemini to display sources. These strategies show an appropriate orientation but varying depth. Comparison with a textbook or official website offers a clearer evidentiary basis than simply "searching on Google." Agreement between two AI systems does not guarantee accuracy if both reproduce the same error. Asking the system for sources can provide a starting point, but students must still check the source's existence, authority, date, relevance, and actual support for the claim.

The important finding is not that students achieved perfect verification, but that productive distrust began to develop. Incorrect outputs prompted them to shift from recipients to examiners. P21 noted that AI could be a learning aid, "but the results still need to be checked and used wisely." At the same time, several answers suggested circular verification, such as asking AI to correct AI, accepting a correctness score as proof of accuracy, or assuming that including the words "credible and valid" in a prompt would guarantee a reliable output.

This pattern is consistent with evidence that students require explicit strategies to evaluate online information (McGrew et al., 2018; Breakstone et al., 2021). Lateral reading requires users to leave a source and compare independent evidence rather than judge the fluency of an answer (McGrew, 2020; McGrew & Breakstone, 2023). The need is even greater with generative AI because system-generated synthesis can break the visible connection between claims and their sources. The risk of natural-language hallucination strengthens the case for claim-by-claim verification, especially for dates, sequences of invention, and causal relationships (Ji et al., 2023).

Dual Learning Occurred in the History of Technology and AI Literacy

Students' reports of what they understood most clearly followed two paths. The first was chronological understanding: movement from simple communication tools toward radio, computers, the internet, smartphones, cloud technologies, and AI. Several students emphasized the speed of change and the conveniences that technology brought to communication, education, the economy, employment, and daily life. P08 stated that the activity helped them understand the journey of technology "from simple communication tools to modern technologies such as the internet, smartphones, and AI."

The second path concerned AI as a contemporary technology. Students described AI's ability to summarize information, generate images, process multiple modalities, and accelerate work. Poster production therefore became a direct demonstration of science and technology development: students used a recent technology to represent the history of technology. P11 described the activity as "direct experience, not merely theory from a textbook," because it demonstrated how AI could support creative work. P23 added that technological development produces not only new devices but also changes how humans obtain information, learn, communicate, and complete work.

Only a small number of responses developed deeper historical explanations. P05 connected science and technology development with the Second World War and United States–Soviet competition during the Cold War in space exploration, weaponry, and communication. Other responses more commonly presented a linear narrative in which technology "becomes more advanced over time." This pattern demonstrates achievement and limitation simultaneously. AI-assisted posters helped students recognize sequences and examples, but they did not automatically develop an understanding of causation, contestation, unequal consequences across groups, or relationships between change and continuity.

The learning was therefore dual but not always balanced. Students learned the content of technological history and, through failed outputs and revisions, learned about the nature of AI. They recognized that AI was fast and versatile but not consistently accurate. The experience extended the topic from "which technologies developed" to "how humans should relate to technology." Without strong historical questions, however, the novelty of AI can dominate the narrative and turn earlier developments into a simple staircase toward the present. Historical consciousness requires a more complex temporal interpretation than a list of linear progress (Popa, 2022); teachers should consequently add questions of cause, consequence, interest, and inequality to the poster task.

Creativity Was Maintained through Decisions, Not through the Absence of AI

Students generally did not define creativity as producing every component without assistance. They maintained it through decisions about concepts, colors, arrangements, images, language, and the final form. P07 stated that Gemini was used only to obtain ideas, while content, design choices, information sequence, images, and the final product were determined independently. P06 added ideas and examples relevant to personal experience. Other students transferred outputs to Canva for editing or used Gemini for visuals while developing text from books and web sources.

The data also reveal aesthetic tension. P17 complained that generated images were unsuitable, commonplace, and repetitive. P25 identified the main difficulty as finding a way to prevent the poster from "looking as if it was made by AI." This concern is more than a technical problem; it indicates recognition of a generic aesthetic that may weaken the work's identity. Structured prompts containing style prohibitions represented one attempt to preserve differentiation, but post-generation editing decisions remained necessary.

The finding supports an understanding of creativity as a traceable sequence of choices. AI assistance may widen an individual's possibilities while narrowing diversity when users accept similar initial suggestions (Doshi & Hauser, 2024). In the classroom, ownership is therefore more appropriately assessed through evidence of decision-making than through speculation about whether an image "looks AI-generated." Initial prompts, revised versions, source lists, reasons for rejecting outputs, and final reflections provide a fairer trail for evaluating creativity and student responsibility.

Discussion (مناقشة)

Discussion: From Tool Use to Human–AI Orchestration

The central finding is that Gemini transformed poster production from a linear workflow into the orchestration of several forms of work within one conversation. In a conventional sequence, students search for information, take notes, open a design application, select images, and assemble the poster. With Gemini, several of these stages occurred through repeated prompts and responses. This change explains the efficiency perceived by students, but efficiency cannot serve as the sole indicator of learning quality. AI literacy requires the capacity to understand, use, evaluate, and position a system ethically; operating an interface or successfully generating an image is only one component (Ng et al., 2021; Wang et al., 2023; Biagini, 2025).

The delegation–orchestration spectrum helps explain differences within the class. Under delegation, a student established a topic and accepted most system decisions. Under curation, a student selected outputs considered appropriate and discarded unnecessary parts. Under co-creation, a student formulated questions, engaged in follow-up dialogue, introduced external sources, edited content, and explained final decisions. These positions are not fixed labels for individuals; a student may delegate image generation but control the text, or do the reverse. The spectrum extends the framework of AI as a dialogic partner by showing that partnership quality is determined by the distribution of control at each stage (Su & Yang, 2023; Tlili et al., 2023; Tang et al., 2024).

This analysis also shows why assessing only the final product is inadequate. A polished poster may result from complete delegation, while a less visually perfect poster may reflect richer verification and revision. If a teacher assesses aesthetics alone, the system indirectly rewards the students who obtain the most polished images rather than those who demonstrate the strongest historical reasoning. The generative AI paradox lies in its ability to improve surface quality while potentially concealing reduced cognitive engagement (Lim et al., 2023; Farrokhnia et al., 2024). Assessment should therefore shift part of its emphasis from a "visible result" to "decisions that can be explained."

Dual Learning and the Risk of a Linear Progress Narrative

Dual learning is the conceptual contribution of this study. Students used a product of science and technology development to construct representations of that development. In the first loop, they learned subject matter: sequences of communication devices, computers, the internet, smartphones, and AI. In the second loop, they learned about technology through experience: outputs could help, fail, display bias, remain overly general, or require revision. The two loops reinforce one another when technical experience is connected to historical questions about changes in how humans work, learn, communicate, and produce knowledge.

This potential makes an AI-assisted poster more than a presentation medium. It can become a site where students negotiate time, cause, consequence, and representational choices. Previous studies suggest that e-posters can support the presentation of historical content and historical awareness, while digital poster tasks can strengthen concepts and communication (Sumantri et al., 2024; Wisanti et al., 2024). The present study adds that engagement with a generative process can produce literacy about the tool itself. Experience of AI does not automatically become historical consciousness, however. Teachers must turn the experience into an object of reflection, for example by asking students to compare communication technologies in two periods and explain who benefited, who was excluded, and what new problems emerged.

The prevalence of statements that “technology is becoming increasingly advanced and sophisticated” indicates the risk of a linear progress narrative. Such a narrative is easily visualized as a timeline but may overlook reversal, social cost, unequal access, shifts in power, and policy choices. Historical consciousness involves making meaning from change, not merely sequencing objects (Popa, 2022). Gemini tends to produce orderly, teleological structures when a prompt does not request conflict or alternative perspectives. Prompts for history content should consequently require analytical elements such as causal drivers, actors, positive and negative consequences, different group experiences, and supporting evidence for each period.

A poster also demands transduction, or the movement of meaning from extended prose to a visual-spatial form. Students must decide which information to preserve, condense, or remove. Multimodal composition can increase engagement and provide space for identity investment, but its success requires support for connecting modes coherently (Jiang, 2018; Jiang et al., 2022; Zhang et al., 2023). AI can suggest icons and layouts, yet it cannot know the classroom’s pedagogical purpose unless that purpose is specified and subsequently checked. The teacher’s role is to ensure that visual simplification does not become historical oversimplification.

Prompting as Epistemic Design, Not Merely a Technical Skill

Variation among prompts indicates different levels of readiness to translate objectives into instructions. Structured prompts reveal knowledge of format, audience, content, and aesthetics. Nevertheless, factual and visual errors in long prompts challenge the assumption that a longer prompt is invariably better. An effective prompt contains criteria that can be checked. A request to “use valid sources” is insufficient if it does not specify source types, a time range, or a requirement to cite references. “Make it attractive” is difficult to evaluate, whereas “use a five-stage timeline, no more than 35 words per stage, and distinguish facts from consequences” establishes a clearer basis for revision.

Prompt instruction should therefore begin with epistemic design. Students first formulate what they want to know and how they will determine whether an answer is adequate. They then establish content boundaries, output form, and standards of evidence. Prompt engineering detached from critical thinking risks increasing only the ability to obtain persuasive responses. Educational scholarship stresses that prompt literacy should be integrated with source evaluation and awareness of AI limitations (Walter, 2024; Kasneci et al., 2023). In history education, prompts should also request periodization, evidence, comparison, and explicit uncertainty where sources differ.

Students who repeated prompts after an unsuitable output demonstrated an early form of metacognitive regulation. They compared a result with their purpose, recognized a gap, and changed the instruction. This process can be made more explicit through a short revision log documenting the first prompt, the problem identified, the change introduced, and the reason for it. The log need not be lengthy, but it helps the teacher distinguish random repetition from diagnosis-based revision. It also turns AI failure into learning material rather than a disruption to be concealed.

Verification as the Core of AI Literacy in History Education

All students had encountered AI limitations, yet their checking strategies did not follow a shared standard. This finding indicates that experience of an error does not automatically produce rigorous verification. Students need a simple and repeatable protocol. Research on online reasoning shows that the questions “Who is behind this information?”, “What is the evidence?”, and “What do other sources say?” are more usable than long checklists that are rarely applied (McGrew et al., 2018; McGrew & Breakstone, 2023). Lateral reading is effective when students actually open other sources and evaluate their authority (McGrew, 2020; Breakstone et al., 2021).

For a Gemini-assisted task, this protocol can be operationalized through five actions. First, mark verifiable claims, including dates, inventors, sequences, figures, and causal relationships. Second, ask AI to identify sources as preliminary leads. Third, open the original or another trusted independent source. Fourth, determine whether the source supports the same claim rather than merely discussing a similar topic. Fifth, revise the poster and retain a record of the source. This protocol addresses two weaknesses evident in the data: trusting the word “valid” within a prompt and treating agreement among chatbots as evidence.

Hallucination means that verification must target individual claims rather than general confidence in the system. A model may be correct throughout most of a passage but wrong about one date that determines the chronology. Linguistic fluency also increases the likelihood that students will overlook an error (Ji et al., 2023). Ethical guidance for AI in education identifies transparency, accountability, and human oversight as foundational principles (Holmes et al., 2022; Yan et al., 2024). In classroom practice, these principles can be translated into requirements that students provide a source list and explain at least one correction made to an AI output.

Creative Agency, Authenticity, and Generic Aesthetics

Students’ concern that a poster should not look “AI-made” shows that they recognized authenticity as a social issue. Judging authenticity from appearance alone is nevertheless unreliable. Humans can produce a generic aesthetic through templates, while AI-generated visuals can be extensively edited. Process evidence is more accountable than detection by intuition. Prompts, sketches, sources, versions, and reflections demonstrate student contributions without requiring every pixel to have been produced manually.

Emphasis on decision-making also protects creativity from two extremes. The first assumes that any AI assistance eliminates creativity. The second equates the number of AI-generated variations with student creativity. Experimental evidence indicates that AI can raise individual quality while narrowing the diversity of collective output (Doshi & Hauser, 2024). Teachers can therefore ask students to develop two initial concepts before using AI, compare multiple outputs, and incorporate an element drawn from personal experience or interpretation. This approach preserves ideational support while preventing the first generated output from becoming the sole anchor.

Creative agency is also ethical. Students should know when an image or text was AI-generated, how factual sources were recorded, and which parts were edited. Transparency is not a punishment but a way of acknowledging the division of labor. Ethical frameworks for K–12 education emphasize the need to connect technical capability with fairness, privacy, accountability, and social consequences (Akgun & Greenhow, 2022; Holmes et al., 2022). In a poster assignment, a practical minimum is a short process note that explains the function of AI and identifies the non-AI sources used.

Pedagogical Implications: A Human–AI Curation Cycle

Based on the findings, this study proposes a five-stage human–AI curation cycle. The cycle is not specific to one platform and may be applied to Gemini or another generative system. Its purpose is to ensure that students remain the owners of learning objectives, examiners of claims, and decision makers, while AI supplies generative support. The design responds to the multidimensional nature of AI literacy and the need for dialogic learning in secondary education (Ng et al., 2021; Lee & Kwon, 2024; Tang et al., 2024).

Table 3. A Human–AI Curation Cycle for a History Poster Assignment

Stage	Student activity	Process evidence	Assessment focus
Establish a historical objective	Formulate a question, period, concept of change, and audience	Concept map or initial brief	Precision of focus and subject relevance
Engage in dialogue through prompts	Compose a prompt, request alternatives, and constrain the output	Initial prompt and one meaningful revision	Clarity of instructions and rationale for revision
Verify claims	Mark facts, open sources, and read laterally	Claim–source matrix	Accuracy, credibility, and evidentiary fit
Edit multimodally	Condense text, select visuals, organize hierarchy, and correct typographical errors	Draft and final poster	Text–visual coherence and readability
Reflect on contributions	Explain AI assistance, rejected elements, and personal decisions	A 150–200-word reflection	Agency, transparency, and ethical use

The cycle provides a basis for a rubric with four balanced domains: historical accuracy, verification quality, multimodal communication, and reflective agency. Visual quality remains important but does not dominate. Teachers can give greater weight to accuracy and sources because an attractive but incorrect poster can reinforce misconceptions. Feedback should also address relationships among decisions: why a particular image represents a period, why a claim was retained, and how a source supports it.

A further implication is the need for differentiation. Students with minimal and structured prompts were present in the same class. Prompt templates may support beginners, but they should not produce uniformity. Teachers can provide required components topic, period, audience, sources, and text limit alongside optional components that allow stylistic and creative choices. This approach offers structure without taking over the student's voice. Research on AI in education cautions that responsible implementation requires the integration of pedagogy, ethics, and disciplinary objectives rather than access to a tool alone (Bond et al., 2024; Biagini, 2025).

The research agenda should also move from cataloging benefits toward observing human–AI processes in classrooms. Recommendations for generative AI research emphasize designs that examine how goals, feedback, and learner regulation shape technological benefits rather than merely comparing the presence or absence of AI (Chiu, 2024). In history education, such an agenda could be realized through analysis of prompt changes, factual corrections, visual decisions, and the development of temporal reasoning across a sequence of assignments.

Conclusion (خاتمة)

This study shows that Grade XII students at Al-Hikmah Senior High School, Surabaya, used Gemini AI as a poster-production environment integrating information search, title and copy development, visual production, and design revision. The pattern of use was not uniform.

Some students delegated almost the entire process and limited their contribution to an initial prompt. Most adopted a curatorial pattern, receiving material from AI and then selecting, shortening, correcting, or transferring it to a design application. Others positioned Gemini as a copilot subordinate to student objectives, sources, and creative decisions.

The research objective is answered through four conclusions. First, Gemini accelerates multimodal production, but efficiency must be distinguished from depth of learning. Second, student prompts ranged from minimal commands to structured specifications; detail helped direct the system but did not guarantee accuracy or visual appropriateness. Third, every student had encountered an AI limitation, and most selected outputs and checked sources, although verification quality remained uneven. Fourth, the activity produced dual learning: students understood sequences of technological development while gaining concrete experience of the capacities and limitations of AI as a contemporary technology.

The study's contribution lies in demonstrating that the quality of AI-assisted learning is determined by the orchestration of decisions, not by the presence or absence of AI. Creativity remains when students establish objectives, select alternatives, inspect information, edit language and visuals, and accept responsibility for the final result. Conversely, an attractive poster does not necessarily indicate learning if every decision has been surrendered to a system.

Teachers should consequently assess process alongside product through a five-stage cycle: establish a historical objective, engage in dialogue through prompts, verify claims, edit multimodally, and reflect on human–AI contributions. Assignments should require the submitted prompt, a source matrix, evidence of one revision, the final poster, and a short reflection. With this design, Gemini can serve as a productive medium for history education without replacing student reasoning. The most defensible pedagogical position is neither absolute acceptance nor rejection of AI, but ensuring that the technology extends students' capacities to question, verify, create, and explain their decisions.

Acknowledgment (شكر وتقدير)

The author wishes to express profound gratitude to the principal, teachers, and students of Class XII-10 at SMA Al-Hikmah Surabaya for their cooperation and active participation during the data collection and research implementation phases. Furthermore, the author extends appreciation to the History Education Study Program at Universitas Sebelas Maret, as well as the History Education and Social Studies Education programs at Universitas Negeri Surabaya, for providing the institutional resources, collaborative spaces, and academic environment that made this research possible.

Bibliography ((مراجع

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. AI and Ethics, 2(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Ahmad, S. Z. (2019). Digital posters to engage EFL students and develop their reading comprehension. Journal of Education and Learning, 8(4), 169–184. https://doi.org/10.5539/jel.v8n4p169

Biagini, G. (2025). Towards an AI-literate future: A systematic literature review exploring education, ethics, and applications. International Journal of Artificial Intelligence in Education, 35, 2616–2666. https://doi.org/10.1007/s40593-025-00466-w

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., . . . Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. International Journal of Educational Technology in Higher Education, 21(1), 4. https://doi.org/10.1186/s41239-023-00436-z

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. Qualitative Psychology, 9(1), 3–26. https://doi.org/10.1037/qup0000196

Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M., . . . Saavedra, A. (2021). Students' civic online reasoning: A national portrait. Educational Researcher, 50(8), 505–515. https://doi.org/10.3102/0013189X211017495

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. Quality & Quantity, 56(3), 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. Computers and Education: Artificial Intelligence, 6, 100197. https://doi.org/10.1016/j.caeai.2023.100197

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. Science Advances, 10(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. Innovations in Education and Teaching International, 61(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine, 292, 114523. https://doi.org/10.1016/j.socscimed.2021.114523

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., . . . Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., . . . Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), 1–38. https://doi.org/10.1145/3571730

Jiang, L. (2018). Digital multimodal composing and investment change in learners' writing in English as a foreign language. Journal of Second Language Writing, 40, 60–72. https://doi.org/10.1016/j.jslw.2018.03.002

Jiang, L., Yu, S., & Zhao, Y. (2022). Incorporating digital multimodal composing through collaborative action research: Challenges and coping strategies. Technology, Pedagogy and Education, 31(1), 45–61. https://doi.org/10.1080/1475939X.2021.1978534

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. Medical Teacher, 42(8), 846–854. https://doi.org/10.1080/0142159X.2020.1755030

Lee, S. J., & Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes. Computers and Education: Artificial Intelligence, 6, 100211. https://doi.org/10.1016/j.caeai.2024.100211

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. The International Journal of Management Education, 21(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. Computers & Education, 145, 103711. https://doi.org/10.1016/j.compedu.2019.103711

McGrew, S., & Breakstone, J. (2023). Civic online reasoning across the curriculum: Developing and testing the efficacy of digital literacy lessons. AERA Open, 9, 1–14. https://doi.org/10.1177/23328584231176451

McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. Theory & Research in Social Education, 46(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. International Journal of Qualitative Methods, 22, 1–18. https://doi.org/10.1177/16094069231205789

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, 100041. https://doi.org/10.1016/j.caeai.2021.100041

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. International Journal of Qualitative Methods, 19, 1–13. https://doi.org/10.1177/1609406919899220

Popa, N. (2022). Operationalizing historical consciousness: A review and synthesis of the literature on meaning making in historical learning. Review of Educational Research, 92(2), 171–208. https://doi.org/10.3102/00346543211052333

Roberts, K., Dowell, A., & Nie, J.-B. (2019). Attempting rigour and replicability in thematic analysis of qualitative research data: A case study of codebook development. BMC Medical Research Methodology, 19(1), 66. https://doi.org/10.1186/s12874-019-0707-y

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. ECNU Review of Education, 6(3), 355–366. https://doi.org/10.1177/20965311231168423

Sumantri, P., Nababan, S. A., Sumanti, S. T., Tanjung, Y., Mulyani, F. F., & Jali, J. M. (2024). Effectiveness of use of interactive e-poster history learning media in increasing historical awareness. Jurnal Ilmu Pendidikan, 30(1), 41–48. https://doi.org/10.17977/um048v30i1p40-47

Tang, K.-S., Cooper, G., Rappa, N., Cooper, M., Sims, C., & Nonis, K. (2024). A dialogic approach to transform teaching, learning & assessment with generative AI in secondary education: A proof of concept. Pedagogies: An International Journal, 19(3), 493–503. https://doi.org/10.1080/1554480X.2024.2379774

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period. BMC Medical Research Methodology, 18(1), 148. https://doi.org/10.1186/s12874-018-0594-7

Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education, 21(1), 15. https://doi.org/10.1186/s41239-024-00448-3

Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. Behaviour & Information Technology, 42(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Wisanti, W., Indah, N. K., & Putri, E. K. (2024). Scientific digital poster assignments: Strengthen concepts, train creativity, and communication skills. International Journal of Evaluation and Research in Education, 13(2), 1035–1044. https://doi.org/10.11591/ijere.v13i2.25909

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., . . . Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. British Journal of Educational Technology, 55(1), 90–112. https://doi.org/10.1111/bjet.13370

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education Where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, M., Akoto, M., & Li, M. (2023). Digital multimodal composing in post-secondary L2 settings: A review of the empirical landscape. Computer Assisted Language Learning, 36(4), 694–721. https://doi.org/10.1080/09588221.2021.1942068