

To cite this article: Phan Thi Thu, Pham Van Gieng, Do Thi Ngoc Hien and Nguyen Thi Viet Ha (2026). GENERATIVE ARTIFICIAL INTELLIGENCE IN HISTORY EDUCATION: FROM CONTENT GENERATION TO SUPPORTING HISTORICAL THINKING COMPETENCE, International Journal of Education and Social Science Research (IJESSR) 9 (5): 377-395 Article No. 1302, Sub Id 1991

GENERATIVE ARTIFICIAL INTELLIGENCE IN HISTORY EDUCATION: FROM CONTENT GENERATION TO SUPPORTING HISTORICAL THINKING COMPETENCE

Phan Thi Thu¹, Pham Van Gieng², Do Thi Ngoc Hien³, Nguyen Thi Viet Ha⁴

¹Institute of Education Strategy and Policy

²Hanoi Pedagogical University

³Institute of Education Strategy and Policy

⁴Institute of Education Strategy and Policy

DOI : <https://doi.org/10.37500/IJESSR.2026.9521>

ABSTRACT

Generative artificial intelligence (GenAI) is changing the cognitive environment and creating new demands for History education. Undirected use of GenAI may increase dependence on AI outputs and reduce opportunities for independent thinking. Drawing on theories of historical thinking, cognition and AI in education, this article proposes the AI-Historical Thinking Competence Model (AI-HCTM). The study uses a conceptual approach, drawing on existing theories and research to develop the proposed model. The model shifts the focus from using GenAI to generate content to using it to support historical inquiry and reasoning, while learners retain an active role. AI-HCTM emphasizes three directions: prompting grounded in evidence and context; tasks that require verification of AI outputs against historical sources; and assessment of the reasoning process rather than only the final product. In the model, the pedagogical value of GenAI lies in its capacity to support learners in sourcing, contextualization, corroboration, information verification and evidence-based reasoning.

KEYWORDS: Generative artificial intelligence (GenAI), History learning, historical thinking, AI-HCTM, AI output verification

1. INTRODUCTION

Generative artificial intelligence is changing how learners' access, select, process and construct information. In History learning, this shift requires changes in the organization of learning activities. AI should not be used only as a tool for providing information; it should be organized as a means of supporting the development of historical thinking. As learners face an increasingly large, rapid and multidimensional flow of information, the central issue is no longer access to information but the capacity to navigate, appraise, verify and construct historical knowledge from it.

At the same time, GenAI creates a paradox for History education: the faster and easier it becomes to generate content, the greater the risk that learners' independent thinking processes will be reduced. Accepting highly coherent but unverified answers may create an "illusion of understanding," in which learners equate fluent expression with the validity of historical knowledge. In particular, if AI replaces sourcing, contextualization and corroboration, learners may have fewer opportunities to practise the cognitive operations that constitute historical thinking. The issue, therefore, is not to restrict AI's capacity to generate content, but to design its use in ways that require learners to examine, verify and reason.

This requirement calls for a different approach to the role of AI in History learning. Rather than treating AI mainly as a tool for producing learning products, it should be embedded in tasks that encourage "slow thinking," giving learners time to analyse evidence, recognize uncertainty, consider alternative interpretations and justify their conclusions. From this perspective, the pedagogical value of GenAI lies not in providing the fastest answer but in creating a dialogic, critical and inquiry-based environment in which learners engage more deeply in historical thinking.

Although current AI competency frameworks, such as UNESCO's AI Competency Framework for Students and AILFS 1.0, provide important guidance on AI competence, critical thinking and ethics in digital environments, they are largely cross-curricular and do not fully specify how AI can be integrated with discipline-specific forms of thinking. In History, the theoretical gap concerns how AI can be designed and used to support the distinctive cognitive practices of the discipline without replacing learner agency. Teaching students how to use AI tools is therefore not sufficient; a theoretical framework is needed to connect the characteristics of GenAI with historical epistemology, learners' cognitive activity and pedagogical task design.

On this basis, the article proposes the AI–Historical Thinking Competence Model (AI-HCTM), which seeks to shift the role of GenAI from a "content generation tool" to a "tool for supporting the development of historical thinking competence." The model clarifies the mechanisms through which GenAI can be integrated into the use, verification, corroboration and interpretation of historical sources. It also proposes directions for prompting, inquiry tasks and assessment so that AI supports thinking without replacing learners' historical judgment and reasoning.

2. THEORETICAL FOUNDATIONS

The development of technology-integrated learning tools for History needs to be grounded in both theories of historical thinking competence and scientific perspectives on cognition and experiential learning.

2.1. Historical thinking competence

2.1.1. Peter Seixas's Big Six model and the Vietnamese context

The Big Six historical thinking model proposed by Peter Seixas and Tom Morton identifies six core historical thinking concepts: historical significance, evidence, continuity and change, cause and consequence, historical perspectives and the ethical dimension (Seixas & Morton, 2013). The importance of the model lies not only in identifying the components of historical thinking but also in requiring learning activities that enable learners to ask questions, analyse evidence, explain events in historical context and make reasoned judgments. In this approach, learning History goes beyond receiving and reproducing information about the past and aims to develop the ability to reason and construct evidence-based historical explanations.

For the integration of GenAI into History learning, the Big Six provides a reference framework for identifying the cognitive activities that a learning tool should support. GenAI may be used to prompt questions about the historical significance of an event; help learners identify, analyse and compare evidence from historical sources; organize comparisons to identify continuity and change; suggest causal relationships; expose learners to different interpretations of the same event; or encourage discussion of ethical issues when evaluating human actions in historical contexts. However, these affordances have educational value only when they are translated into purposeful learning tasks with clear reasoning requirements and assessment criteria. A GenAI-generated answer cannot in itself be treated as historical evidence or as evidence of a learner's historical thinking competence.

The applicability of the Big Six to research on Vietnamese students' historical thinking has also been examined by Ninh et al. (2025). Their study surveyed 448 students from five upper-secondary schools in Hanoi and reported Cronbach's alpha coefficients ranging from 0.770 to 0.923 for scales corresponding to the six components. These results provide initial evidence of internal consistency in the sample and suggest the potential use of the Big Six framework in research on historical thinking in Vietnam. However, internal consistency does not mean that the full model has been validated as a measurement model, nor does it directly demonstrate the effectiveness of integrating GenAI into History learning.

Accordingly, this study uses the Big Six as a theoretical reference framework to identify the components of historical thinking that should be emphasized when designing GenAI-integrated learning tools. The framework guides the construction of learning tasks, learner-tool interaction and assessment requirements. The effectiveness of a tool should therefore not be judged simply by its ability to generate historical content, but by the extent to which it supports learners in performing historical thinking operations, especially using and verifying evidence, constructing arguments and producing justified interpretations. Its actual effectiveness still needs to be examined with appropriate empirical data.

2.1.2. Andreas Körber's competence model and Jörn Rüsen's historical consciousness

While Seixas and Morton's Big Six provides a framework for identifying components of historical thinking, Andreas Körber's model of historical competence and Jörn Rüsen's concept of historical consciousness help explain how learners use historical understanding to interpret the past, make sense of the present and orient themselves toward the future. Building on earlier work on historical competence, Körber (2011) emphasizes the relationship among questioning, the use of historical methods and orientation through historical understanding. Rüsen (2005), meanwhile, views historical consciousness as the interpretation of past experience to create meaning for present understanding and action while providing orientation for the future. These approaches complement the Big Six by emphasizing the use of historical knowledge to construct meaning rather than only the analysis of events and evidence.

For the integration of GenAI into History learning, these approaches suggest that learning tools should support learners' processes of inquiry, interpretation and orientation. GenAI can be used to suggest research questions, generate alternative explanations of an event, organize information in temporal and causal relationships or create situations in which learners consider links between past and present. However, GenAI's ability to generate answers or scenarios does not itself develop historical consciousness. The educational value of the tool depends on whether learners actively check information, evaluate the basis of explanations and construct their own historical arguments.

One useful idea for learning-task design is the "if thinking" approach discussed by Miki et al. (2015). This approach asks learners to consider how a change in a historical condition or factor might affect related factors, thereby clarifying causal relationships among historical events and conditions. Miki et al. used a learning support system to present hypothetical situations and evaluate responses; their study was not about GenAI. It is therefore used here as a suggestion for task-design principles rather than as direct evidence of GenAI effectiveness. Applying this principle, GenAI can help generate hypothetical situations and critical questions so that learners consider different possible developments. Learners must nevertheless use historical context, evidence and causal reasoning to assess the plausibility of such hypotheses; AI-generated scenarios must not be treated as events that actually occurred.

Thus, Körber's model of historical competence and Rüsen's concept of historical consciousness add a theoretical basis for designing GenAI-integrated learning tools to develop historical thinking competence. Such tools should enable learners to inquire, verify, interpret and evaluate information actively, while considering the significance of historical understanding for making sense of the present. Assessment should focus not on the quality of AI-generated text but on how learners use evidence, construct arguments and produce justified interpretations. This design direction requires empirical examination of learning processes and competence development.

2.2. Cognitive foundations and mechanisms of cognitive support from GenAI

The use of GenAI in History teaching should be considered in relation to the nature of cognition when learners encounter and interpret historical information. Learning History involves more than receiving and remembering information about the past. Learners must activate prior knowledge, formulate questions, evaluate information and evidence, place sources in context, compare sources and construct justified explanations. Research on historical reasoning shows that learners often have difficulty spontaneously using source information, evaluating credibility, corroborating sources and contextualizing documents; these operations need to be deliberately organized and supported through instruction (Wineburg, 2001; van Drie & van Boxtel, 2008).

From a cognitive perspective, GenAI can support learning by organizing and transforming information, generating questions, offering alternative explanations or providing initial feedback. However, the system's capacity to process and generate content does not mean that the corresponding cognitive activities have been performed by the learner. In History in particular, a coherent AI-generated interpretation of an event does not ensure that the interpretation is based on appropriate evidence, properly contextualized or verifiable. GenAI outputs should therefore be treated as inputs for learners' subsequent cognitive activity rather than as final learning outcomes.

This view is consistent with supported learning approaches in which learners are provided with tools and cognitive scaffolds to undertake tasks that would otherwise be difficult to perform independently. In the teaching of historical thinking, cognitive apprenticeship has been used to make expert practices observable, practicable and open to feedback, for example by modelling source reading, guiding learners in sourcing, contextualization and corroboration and gradually reducing support as learners become able to perform these operations independently. Within this logic, GenAI can be considered one component of a cognitively supportive environment, but it does not replace the teacher's role in task design or the learner's judgment.

GenAI can provide different forms of support depending on its position in the learning sequence. During preparation, it can help organize information and suggest questions; during inquiry, it can generate hypotheses or interpretations for learners to examine; and during reflection, it can pose critical questions or ask learners to explain the basis of their conclusions. This differs fundamentally from asking GenAI to produce a complete answer and then using that output directly. In the first case, AI output becomes an object of cognitive activity; in the second, AI may perform part of the cognitive work that the lesson is intended to develop in the learner.

This distinction is especially important in History because historical knowledge is interpretive and evidence-based. Learners need to identify who created a source, under what circumstances it was produced, what its information means in historical context and whether it is consistent with other sources. Wineburg and subsequent research have shown that sourcing, contextualization and corroboration are important operations in historical reasoning and that guided instruction can help

learners develop them. When GenAI is used in History learning, verifying AI output can therefore be designed as a cognitive task rather than treated merely as a technical requirement.

On this basis, the study adopts the principle that GenAI should support but not replace learners' core cognitive activity. The tool may perform or assist with technical operations such as organizing, transforming and presenting information, but activities central to historical thinking—defining the problem, evaluating evidence, verifying sources, contextualizing, comparing interpretations and constructing arguments—should remain with the learner. This principle is also consistent with a human-centred approach to GenAI, in which technology serves educational purposes while preserving learner agency.

From a tool-design perspective, this cognitive foundation shifts the focus from “AI generates content for learners” to “AI creates conditions and objects for learners to engage in cognitive activity.” A GenAI-integrated learning tool genuinely supports historical thinking only when it creates a sequence in which learners must receive input, ask questions, check and compare information, identify weaknesses, revise their understanding and construct evidence-based conclusions. This process provides the link between AI's generative capacity and the goal of developing historical thinking competence.

The cognitive foundation is therefore not used to claim that GenAI itself improves historical thinking competence. Instead, it specifies the pedagogical and cognitive conditions under which GenAI can become a learning support tool: technology provides inputs, suggestions, feedback or scaffolding; learners verify, interpret and reason; and teachers design tasks, regulate the level of support and assess the quality of the learning process. This provides a basis for considering experiential learning, in which interaction with problems, tools, feedback and reflection is organized into a purposeful learning cycle.

2.3. Experiential learning and learner motivation

In a conceptual analysis of experiential learning, Kong (2021) synthesizes the role of experiential learning in learner motivation and engagement; the study does not examine GenAI. Building on the principle that learning activities should centre on learners' experience and participation, the present study proposes placing GenAI within historical inquiry tasks in which learners ask questions, verify information and construct their own arguments rather than merely receive technological outputs.

3. RESEARCH METHODS

This conceptual study does not use empirical data to test the effectiveness of the proposed model. Its purpose is to synthesize and connect theoretical foundations in historical thinking, cognitive science and GenAI-integrated education and from these to construct a conceptual framework for using GenAI to support the development of historical thinking competence.

The sources analysed include foundational works on historical thinking and historical consciousness; studies of specific practices such as sourcing, contextualization and corroboration; and recent documents and studies on AI competence, AI use in education and issues such as hallucination, information verification and learner responsibility. Sources were selected for their direct relevance to the theoretical components required to construct the model. The study was not designed as a systematic review and therefore does not claim comprehensive coverage of the literature in the field.

The model was developed in three steps. First, the study identified core components of historical thinking from published theoretical frameworks. Second, the affordances and limitations of GenAI were compared with these cognitive operations to identify where the technology may support or risk replacing learners' thinking. Third, on the basis of this synthesis, the study proposed AI-HCTM and specified it through principles for task design, prompting techniques, AI-output verification activities and process-oriented assessment. Because AI-HCTM is a theoretical model, the proposed relationships should be treated as assumptions that require further empirical testing.

4. FINDINGS AND DISCUSSION

Integrating GenAI requires a qualitative change in the structure of historical cognition, shifting the focus from acquiring information to appraising and critically examining AI-generated outputs.

4.1. AI-supported Historical Thinking Competence Model (AI-HCTM)

The integration of generative artificial intelligence (GenAI) into History learning should be approached as a pedagogical process with a clear mechanism of action, rather than simply as the addition of a technological tool to existing learning activities. AI-HCTM is based on the assumption that the educational effectiveness of GenAI depends on the interaction among technological characteristics, learners' cognitive activity, access to and appraisal of historical sources, historical thinking operations and learning-task design. GenAI can rapidly generate text, support interactive dialogue and synthesize information, but it also has limitations such as hallucination, inherent bias and inaccurate citations (Do et al., 2025). These characteristics shift the cognitive demands placed on learners from searching for and reproducing information to appraising, verifying, comparing and critically examining AI outputs. In History learning, primary sources provide an important basis for checking AI-generated claims and for practising discipline-specific operations such as sourcing, contextualization and corroboration (Pratama et al., 2024). However, the effects of GenAI on cognition and learning outcomes are not determined by the technology itself; they depend substantially on how it is integrated into pedagogical design. The same digital tool can augment thinking but can also reduce opportunities for learners to perform cognitive operations themselves if it is used to replace the learning process (Do et al., 2025). AI-HCTM therefore identifies learning-task design as a key mediating mechanism through which technological affordances can become opportunities for competence development. On this basis, the model proposes five core principles:

First, preserve learner agency in Historical Inquiry. Learners retain the central role in defining problems, formulating questions and selecting historical sources for verification. GenAI mainly supports the processing and organization of information and does not replace learners' cognitive judgments and decisions.

Second, place Historical Reasoning at the centre. Assessment should shift from the completeness of the final product to the quality of the reasoning process, reflected in learners' ability to explain how claims were formed, select and use evidence and defend evidence-based historical conclusions.

Third, use GenAI to create conditions for Cognitive Conflict. Prompts designed around learning objectives can generate different or conflicting interpretations of the same historical issue, placing learners in situations where they must analyse them, compare them with historical sources, evaluate the basis of each interpretation and revise or form their own evidence-based conclusions.

Fourth, conduct Evidence-based Verification. GenAI responses should be treated as provisional claims whose accuracy has not been established. Learners must compare and cross-check these claims against primary sources and reliable historical scholarship, assess the alignment between claims and historical evidence and identify unsupported information, inaccuracies or interpretations that go beyond what the evidence allows.

Fifth, preserve Human Agency in historical interpretation. GenAI is used to support inquiry, analysis and critique, but learners remain responsible for evaluating evidence, constructing historical meaning, making ethical judgments and taking responsibility for their conclusions under teacher guidance. In summary, AI-HCTM emphasizes a pedagogical mediation mechanism in which GenAI provides technical affordances, historical sources provide the evidentiary reference base, learning tasks regulate interaction and learners remain the agents who construct historical arguments.

4.2. Restructuring historical cognition under the influence of GenAI

The emergence of GenAI changes not only how learners access information but also the possible sequence of cognitive activities in History learning. In a traditional learning environment, learners mainly receive information from textbooks, historical materials and teacher-selected sources, then analyse it and form conclusions. In a GenAI environment, learners can actively ask questions, request the generation or interpretation of information and continue to check and evaluate system outputs (Do et al., 2025). Cognitive activity therefore tends to shift from reception → processing → product creation toward inquiry → comparison → verification → interpretation → reasoning, with learners continually evaluating both the inputs and outputs of the technology.

This change is particularly important in History because historical thinking is not developed simply by receiving a body of information about the past. According to Seixas and Morton (2013), historical thinking includes cognitive operations such as using evidence, determining historical significance,

analysing cause and consequence, recognizing continuity and change and considering different historical perspectives. Wineburg (2001) likewise shows that understanding a historical source requires learners to move beyond its surface content and consider authorship, context of production, purpose and relationships with other evidence. In a GenAI environment, these operations become even more necessary because generated answers may be linguistically persuasive while still containing inaccurate information, unsupported interpretations or even unverifiable references (Miao et al., 2024). The pedagogical value of GenAI in History education therefore lies less in its capacity to produce quick answers than in its capacity to create an object that learners can question, verify and interpret. When appropriately organized, an AI-generated response can become an “object of cognition” through which students identify claims, trace the origins of information, compare them with historical sources, detect inaccuracies or bias and then revise and defend their own explanations. In this way, GenAI does not replace historical thinking operations but can make them more explicit, interactive and observable in the learning process.

From this perspective, learning with GenAI should shift the focus from AI-generated products to learners’ cognitive processes. A complete AI-generated answer cannot by itself be treated as evidence of historical thinking competence. Rather, competence is demonstrated through learners’ ability to identify historical problems, formulate questions, select and compare sources, assess the credibility of evidence, detect limitations in AI outputs, contextualize information and construct justified arguments. GenAI can thus function as an intermediary tool in the historical cognition cycle, supporting learners in performing and reflecting on thinking operations rather than performing those operations for them. The important effect of GenAI on History education is therefore not to reduce the need to remember and work with historical sources, but to shift cognitive activity from receiving information toward questioning, verification and the construction of historical meaning. This shift creates the possibility of transforming GenAI from a content generation tool into a tool for supporting historical thinking competence, provided that its use remains under learners’ cognitive control and is organized through evidence-based learning tasks.

The issue for History education is therefore not to replace existing modes of learning with GenAI, but to design activities in which the technology’s generative capacity supports rather than replaces learners’ historical thinking operations. Previous research on historical thinking and learning support systems indicates the value of organizing tasks around causality, sources and evidence (Miki et al., 2015; Rüsen, 2005). On this basis, the present study proposes treating GenAI as a generative dialogue tool: system outputs become objects for learners to question, compare, critique and revise on the basis of historical sources rather than final answers.

Accordingly, changes in the organization of History learning under GenAI can be summarized through the core criteria presented in Table 1.

Table 1. Shifts in History learning with GenAI integration (AI-HCTM)

Criterion	Traditional learning model	AI-supported Historical Thinking Competence Model (AI-HCTM)
Cognitive objective	Emphasizes the formation and consolidation of knowledge about events, people and dates; in some cases, this may lead to the reception of pre-existing, one-sided interpretations (Fikri et al., 2022).	Strengthens higher-order cognitive operations: contextualization, identifying the origin and purpose of sources (sourcing), corroboration, evaluating arguments and considering multiple perspectives (multiperspectivity) (Ninh et al., 2025).
Use of historical sources	Mainly uses materials selected and provided in advance in textbooks or by teachers (Fikri et al., 2022).	Combines GenAI inquiry with the use of historical sources and scholarship; cross-checks AI-generated information, compares claims with primary-source evidence and identifies unsupported statements or hallucinations.
Teacher role	Organizes, provides and guides the reception of knowledge; in transmission-oriented models, the teacher may remain central to the provision of information (Fikri et al., 2022).	Designs learning tasks, selects reference sources, develops criteria for process assessment, coordinates inquiry and verification and supports students in evaluating the quality of GenAI-generated information (Do et al., 2025).
Learner agency	Learners tend to receive, reproduce and apply selected knowledge; the degree of active participation depends on the teaching method and specific learning task (Ahmad, 2010).	Learners construct prompts, evaluate AI responses, compare them with primary-source evidence, revise their own interpretations and take academic responsibility for the historical arguments they construct (Miao et al., 2024).

This shift is intended to reduce the risk that learners will delegate cognitive operations to technology. In AI-HCTM, a GenAI response is treated as a claim to be verified, not as established knowledge. Learners practise three foundational historical thinking operations—sourcing, contextualization and corroboration—using reliable historical sources. Teachers shift from providing information to designing tasks, selecting reference sources and supporting appraisal, while learners remain responsible for the final argument. This structure clarifies the distinct roles of learners, teachers and GenAI without assuming that the mere presence of technology automatically produces learning gains.

4.3. Specific pedagogical strategies

4.3.1. Discipline-specific prompting for History

In a history learning environment that integrates GenAI, prompting should not be treated as a routine technical skill. General prompts that merely ask the system to summarize or reproduce information can encourage “cognitive offloading” and passive, one-sided reception of knowledge (Do et al., 2025). By contrast, when prompts are structured around the characteristic thinking operations of History, GenAI can support discipline-specific cognition. This approach directly draws on Wineburg’s (2001) model of reading historical sources through three core operations—sourcing, contextualization and corroboration—and is closely aligned with Seixas and Morton’s (2013) principle of justifying historical interpretations with evidence. The academic value of discipline-specific prompting therefore lies not in technological complexity but in the extent to which prompts restructure the cognitive demands of the task and move learners from extracting content toward practising historical reasoning.

To put this orientation into practice, the learning process can use a set of discipline-specific prompting techniques corresponding to core historical thinking operations:

Evidence-based Prompting: This technique sets clear constraints on the sources and scope of information used in GenAI outputs, requiring the system to remain within the provided historical material and avoid adding unsupported information beyond the reference text. A sample prompt instructs the AI: “Read the entire attached document carefully. Use only information that can be identified directly from this document in your answer. Do not add information from external sources unless the document provides a basis for doing so. For each claim, clearly distinguish: (1) information directly stated in the source; (2) an inference drawn from the source; and (3) information for which there is insufficient evidence to reach a conclusion. If the document does not provide enough evidence, state this clearly and indicate that additional reliable sources are needed.” The pedagogical purpose is to make the process of establishing historical claims more transparent. Learners must locate credible source texts and understand them sufficiently to distinguish raw source information, historical facts and evidence-based historical conclusions, thereby retaining responsibility for academic judgment rather than transferring it to technology.

Sourcing and Contextualization Prompting: This technique directs learners beyond surface content to examine why a source was created, when and where it was produced and the contextual conditions surrounding it. A sample prompt is: “Analyse this document as a historical source. Identify: (1) the source provided; (2) when and where it was created; (3) its intended audience; (4) the purpose of its creator; and (5) features of the historical context that may have influenced its content and presentation. For each claim, indicate the evidence in the document or state clearly when there is insufficient information to determine it.” This operation helps students understand that historical sources are not a “neutral store of facts” but situated products that reflect the perspectives of historical actors under specific conditions (Wineburg, 2001). Understanding context helps learners avoid presentism—the unreflective imposition of contemporary standards on the past—and gradually develop historically grounded empathy (Seixas & Morton, 2013).

Corroboration Prompting: To extend cross-checking from a single source to a network of sources, this technique asks AI to support comparison of agreements and contradictions among different materials. A structured prompt may read: “Identify the main claim of Source A and Source B in turn. Show where the two sources confirm, complement, differ from or contradict each other. For each difference, consider whether it may relate to the time of creation, the author’s position, purpose, intended audience or source type. Do not decide which source is correct simply because it matches an initial view; identify what evidence still needs to be checked.” The key is to move textual comparison from surface description to evaluation of relationships among sources and interpretations, requiring students to judge evidentiary value and choose explanations independently rather than fully delegating analysis to an algorithm.

Critical and Socratic-dialogue Prompting: GenAI can be used as a questioning dialogue partner that continuously challenges students' arguments. A sample prompt is: "Act as an academic critic. Do not reach a conclusion for me. For each claim I make about event X, ask in turn: (1) What is the evidence? (2) What source does the evidence come from? (3) What are the limitations of the source? (4) Is there another source that gives a different account? (5) Does my claim go beyond what the evidence allows?" Unlike asking AI to generate a complete interpretation, this form of interaction keeps academic responsibility and cognitive control with the learner and reduces the risk of one-way cognitive dependence on technology (Miao et al., 2024).

The difference between general and discipline-specific prompting can be illustrated by deconstructing the historical claim: "Ngo Quyen used a proactive military plan to defeat the Southern Han army." Rather than asking GenAI simply to confirm or explain the given claim, the teacher guides students through a sequence of prompts that reflects the logic of historical inquiry. First, students ask AI to extract and list the actual military actions of Ngo Quyen directly recorded in historical texts, such as sharpening the tops of wooden stakes and planting them in the bed of the Bach Dang River, in order to identify factual evidence. Next, a sourcing prompt asks AI to examine whether the phrase "proactive military plan" appears in a primary source or is a later interpretation that applies a modern concept. Students then ask AI to compare Vietnamese historical sources (Dai Viet Su Ky Toan Thu) with Chinese historical texts (Shiguo Chunqiu, Xin Wudai Shi and Zizhi Tongjian) to identify points of agreement, supplementation or contradiction concerning the course of the battle. Finally, students synthesize the evidentiary structure themselves and judge whether the claim of "proactiveness" is a well-established historical conclusion, a conditional conclusion or only a scholarly hypothesis.

In this sequence, GenAI is not used as the source or validator of the final answer. Instead, it helps make visible the structure of evidence and the reasoning steps leading to the historical claim. Students distinguish empirical data from interpretive inference and shift from receiving information to practising evidence-based historical reasoning (Do et al., 2025).

These discipline-specific prompting techniques should not exist as isolated "technology tips" or disconnected prompting tricks (Do et al., 2025). They need to be systematically integrated into a sequence of cognitive tasks corresponding to historical inquiry: defining a question, locating sources, extracting evidence, sourcing, contextualization, corroboration and constructing and justifying a final conclusion (Wineburg, 2001). In this sense, the prompt becomes an intermediary structure connecting GenAI's language-processing capacity with discipline-specific historical thinking operations. The pedagogical value of GenAI is then measured not by response speed or the polish of generated text, but by its capacity to activate, support and prompt reflection on learners' actual cognitive processes. This provides an important basis for shifting assessment in History education in the digital era from judging the correctness of the final product to evaluating learners' engagement in evidence analysis, source verification, multidirectional corroboration and academic responsibility for their own historical conclusions (Ninh et al., 2025).

4.3.2. Designing AI-output verification tasks to develop competence in using historical evidence

One limitation of GenAI in History learning is its capacity to generate inaccurate or difficult-to-verify information, dates, people or references (Do et al., 2025). Students should therefore be organized to treat AI outputs as claims that require verification rather than as inherently reliable sources of knowledge. Pedagogically, unverified outputs can become objects through which students practise using evidence by sourcing, contextualizing and corroborating sources (Wineburg, 2001). The learning task thus shifts from “receiving the AI answer” to examining the relationship among claims, evidence and historical reasoning.

Within AI-HCTM, verification tasks can be organized in three steps:

Step 1. Claim Identification. The teacher asks GenAI to generate an explanation, timeline or response about a historical issue for which sufficient sources are available for verification. Students identify claims that need to be checked, such as dates, people, figures, causal relationships or quotations.

Step 2. Evidence Verification. Students compare AI claims with historical sources and reliable references, focusing on three operations: sourcing, contextualization and corroboration (Wineburg, 2001). Verification should identify which claims are supported by evidence, which lack sufficient support and which conflict with the sources.

Step 3. Evidence-based Evaluation. Students reach conclusions about the reliability of each claim and use evidence to explain those conclusions. The final product is not simply a list of AI errors, but an argument showing why a claim should be accepted, revised or rejected on the basis of historical evidence.

From a pedagogical design perspective, this three-stage cycle can create controlled cognitive conflict: learners compare an apparently coherent output with historical sources and recognize the gap between persuasive expression and evidentiary support. The task then guides learners through a sequence of operations: receiving a claim → questioning it → tracing its source → evaluating credibility → corroborating across sources → revising the claim → justifying it with evidence. This is an AI-HCTM proposal based on human-centred principles and accountability in AI use (Miao et al., 2024). Its effectiveness for historical thinking competence still requires empirical testing.

4.3.3. Renewing assessment in AI-integrated learning environments

The emergence of GenAI calls for a review of how learning is assessed in History. When learners can use technological tools to rapidly produce well-structured, fluent and academically persuasive texts, the final product is no longer a sufficient indicator of students’ actual cognitive process (Do et al., 2025). A polished historical essay may reflect a learner’s own competence, but it may also result from excessive AI assistance. If assessment focuses only on the final static text, teachers may find it difficult to determine how far learners themselves selected evidence, verified sources, constructed claims and

reached conclusions. This uncertainty calls for adaptive assessment that focuses on the thinking process and restructures academic-integrity standards for the digital era (Do et al., 2025; USSH, 2025). Shifting from product assessment to assessment of the historical reasoning process: In a GenAI-supported learning environment, assessment design should make intermediate stages of historical cognition observable and assessable. In addition to the final written product, teachers can use multiple forms of evidence such as learning logs, process portfolios, argument maps, evidence-verification sheets, records of prompt interactions with AI or oral explanation tasks. For a GenAI-supported historical inquiry task, the process portfolio should make four core components visible:

Prompting history: Students should clarify the cognitive purpose of each prompt, explain why they changed wording or context and show how these changes affected the quality of system responses.

AI-output verification process: Students identify important historical claims generated by AI, retrieve primary sources for comparison and analyse the degree of alignment between claims and evidence.

Handling misinformation: Students actively identify factual errors, hallucinations or historically unsupported claims in AI responses and propose specific corrections (Do et al., 2025).

Final independent historical argument: Students present their own conclusions, make explicit the historical evidence used and justify the relevance or limitations of that evidence (Seixas & Morton, 2013).

This approach allows teachers to assess not only “what answer students give” but, more importantly, “how students arrived at that answer.” A correct conclusion copied directly from an AI output should not be evaluated as equivalent to a conclusion constructed through rigorous inquiry, verification and corroboration of primary historical sources. Assessment rubrics should therefore shift from the accuracy of the text alone to indicators of competence: the quality of inquiry questions; the ability to identify and appraise sources; alignment between evidence and claims; corroboration across sources; recognition of historical uncertainty; and the ability to account for the process of using AI (Seixas & Morton, 2013).

Using GenAI to design assessment situations involving historical judgment and ethics: An important extension of assessment in AI-integrated environments is to use the technology to design tasks that assess learners’ ability to address value conflicts and make historical moral judgments. Historical moral judgment should, however, be clearly distinguished from asking students for purely emotional or subjective moral judgments (Gibson, 2014). In the theoretical perspectives of Seixas and Morton (2013) and Rösen (2005), assessing the ethical dimension of History requires learners to make academically grounded judgments after considering the context, values, norms and limits of knowledge available to historical actors at the time, thereby avoiding the trap of presentism.

On this basis, GenAI can be used as a cognitive partner to simulate complex moral dilemmas, requiring students to perform a sequence of higher-order cognitive operations:

Identify the specific historical context and objective factors that constrained the choices available to historical actors (Gibson, 2014);

Identify conflicting values, interests and motivations among social groups within the same historical period (Seixas & Morton, 2013);

Use empirical historical evidence to explain why a historical decision may have been regarded as reasonable or controversial at the time (Gibson, 2014);

Distinguish historical explanation from subjective moral judgment based on present-day perspectives (Rüsen, 2005);

Construct a conditional moral judgment that makes explicit the value framework, evidentiary limits and degree of certainty of the conclusion (Seixas & Morton, 2013).

In assessment design, GenAI can generate or suggest historical situations involving multiple interpretations and conflicting perspectives, requiring students to make judgments and explain their choices. The object of assessment, however, is not the quality or persuasiveness of the AI-generated situation, but students' ability to reconstruct context, use evidence and build arguments to defend historical judgments.

Reframing academic integrity: From technology control to transparency and accountability

The spread of GenAI requires a reconsideration of academic integrity in educational settings (Do et al., 2025). Rather than asking only whether learners used AI, recent AI competency guidance also emphasizes the ability to evaluate outputs, use AI responsibly and maintain learner agency (Miao et al., 2024; USSH, 2025). Academic-integrity risks arise when learners conceal the extent of technological assistance, present AI-generated content as their independent work, use unverified data or include unverifiable references. Assessment in GenAI environments should therefore emphasize transparency in AI use and accountability for learning products.

Transparency in AI Co-creation: When submitting work produced with technological support, learners should clearly disclose the AI tool used, the purpose of use (for example, idea generation, translation, language editing or data analysis), prompt history and the actual extent of AI contribution to the final product (Miao et al., 2024; Do et al., 2025).

Evidentiary traceability and accountability: Historical facts, figures, dates, direct quotations and references included in student work should have clear origins and be independently verifiable (USSH, 2025). AI-provided information does not replace students' responsibility to verify historical sources. Fabricated references, DOI numbers or supposed facts generated through AI hallucination that cannot be verified against original sources should be removed from assessed work (USSH, 2025).

This principle of evidentiary traceability has particular methodological significance in History because source authenticity and the traceability of evidence are core components of historical argument, not merely formal technical requirements of citation conventions (Seixas & Morton, 2013). Ultimate academic responsibility for the integrity of historical knowledge lies with learners and AI users (Do et al., 2025; Miao et al., 2024). Renewing assessment in this direction helps maintain the academic rigour of History while preparing learners to use GenAI as a responsible cognitive support tool without transferring authority for validating historical knowledge to technology.

5. CONCLUSIONS AND IMPLICATIONS

5.1. Reframing the role of GenAI in History learning

The development of GenAI changes the conditions under which information is accessed, processed and created, requiring a reconsideration of how cognitive activity is organized in History education. As AI can support some content-generation and information-processing operations, learning should place greater emphasis on asking historical questions, sourcing and evaluating sources, verifying evidence, analysing perspectives and constructing justified interpretations.

Drawing on approaches to historical thinking, cognition and AI in education, this study proposes AI-HCTM as a theoretical framework for using GenAI to support the development of historical thinking. GenAI can support hypothesis generation, initiate dialogue and process initial information, while learners must directly perform core operations such as verifying historical sources, evaluating evidence, contextualizing, corroborating and constructing arguments. The educational value of GenAI therefore lies not in rapid content generation but in how the technology is integrated into cognitive tasks, systems of historical sources and learners' reasoning processes.

AI-HCTM proposes the following chain of relationships: GenAI characteristics → organization of cognitive activity → access to and verification of historical sources → historical thinking operations → task design → development of historical thinking competence. This theoretical framework requires further empirical testing and refinement.

5.2. Limitations and directions for future research

AI-HCTM is currently based mainly on theoretical synthesis and previous research on GenAI, cognition and historical thinking; its empirical effectiveness in developing historical thinking competence has therefore not yet been tested. The operation of the model may be influenced by the characteristics of the AI tool, the quality of prompts and tasks, the historical sources used, teachers' professional competence, students' digital competence and implementation conditions.

Future research should test the model in specific classroom contexts. Three directions may be prioritized. First, experimental or quasi-experimental designs can compare changes in historical thinking competence across different forms of learning organization. Second, assessment should prioritize direct performance through source-based tasks, argumentative writing and explanatory

products rather than relying only on self-report scales. Third, pre–post designs can use tasks that are structurally equivalent but differ in historical content; the choice of N-gain or other statistical tests should depend on the research design, data characteristics and relevant statistical assumptions.

In particular, data on interactions among students, historical sources and GenAI can be used to examine not only final outcomes but also changes in how students ask questions, verify information and construct historical arguments.

5.3. Pedagogical implications and recommendations for implementation

Implementing AI-HCTM in History teaching requires more than introducing AI tools into lessons. Three components should be prioritized:

First, develop a verified digital repository of historical sources. Primary and secondary historical sources with clear provenance should be systematically organized and digitized so that students can use them to compare with and verify AI outputs.

Second, develop a bank of tasks and prompts oriented toward historical thinking. Tasks should be linked to discipline-specific operations such as sourcing, contextualization, corroboration, evidence evaluation and historical reasoning rather than simply teaching students to use AI to generate content.

Third, renew assessment and safeguard academic integrity. Rubrics should assess the quality of reasoning, use of evidence and learners' ability to account for their interaction with AI, while clearly identifying which tasks AI may support and which cognitive operations learners must perform themselves.

In addition, teacher professional development should shift from training in tool use toward developing competence in pedagogical design with AI, particularly the design of cognitive tasks, selection of historical sources, organization of verification and assessment of historical reasoning. AI-HCTM can be further tested through Design-Based Research (DBR) across cycles of design, implementation, data collection, analysis and revision, gradually refining principles for using GenAI in ways that fit the cognitive characteristics of History.

5.4. Concluding Remarks

The central issue for History learning in the GenAI context is not simply whether AI is used, but how interaction with AI is organized so that learners continue to think for themselves. AI-HCTM proposes a structure in which GenAI supports information generation and processing, learners lead the inquiry process, historical sources provide the evidentiary reference base and conclusions must be justified with evidence. In this approach, the value of GenAI is assessed not primarily by the speed of text generation or the completeness of answers, but by its capacity to support learners in asking questions,

verifying information and accounting for their reasoning. Because AI-HCTM is currently a theoretical model, its effects on historical thinking competence require further empirical testing.

REFERENCES

- [1] Fikri, A., Mahdum, M., & Isjoni, I. (2022). The development of local history learning model based on progressivism to promote historical thinking skills. *Ta'dib*, 25(1), 105–116. <https://doi.org/10.31958/jt.v25i1.5831>
- [2] Kong, Y. (2021). The role of experiential learning on students' motivation and classroom engagement. *Frontiers in Psychology*, 12, 771272. <https://doi.org/10.3389/fpsyg.2021.771272>
- [3] Miki, Y., Kojiri, T., & Seta, K. (2015). "If Thinking" support system for training historical thinking. *Procedia Computer Science*, 60, 1542–1551. <https://doi.org/10.1016/j.procs.2015.08.263>
- [4] Ninh, T. H., Nguyen, T. B., Chu, N. Q., Phan, T. T. C., & Dang, T. T. D. (2025). Historical thinking skills from the perspective of high school students in Vietnam. *International Journal of Education and Practice*, 13(1), 312–326. <https://doi.org/10.18488/61.v13i1.4059>
- [5] Rüsen, J. (2005). *History: Narration, interpretation, orientation*. Berghahn Books. <https://doi.org/10.1515/9781782389675>
- [6] Seixas, P., & Morton, T. (2013). *The Big Six historical thinking concepts*. Nelson Education.
- [7] Miao, F., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO. <https://doi.org/10.54675/JKJB9835>
- [8] Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*. Temple University Press.
- [9] Körber, A. (2011). German history didactics: From historical consciousness to historical competencies – and beyond? In H. Bjerg, C. Lenz, & E. Thorstensen (Eds.), *Historicizing the uses of the past: Scandinavian perspectives on history culture, historical consciousness and didactics of history related to World War II* (pp. 145–164). Transaction Publishers.
- [10] Pratama, R. A., Saputra, M. A., & Hikmawaty, L. (2024). Enhancing historical consciousness in history education through integrating STEM approach and historical thinking skill. *Journal of Education and Learning (EduLearn)*, 18(1), 236–243. <https://doi.org/10.11591/edulearn.v18i1.20890>
- [11] Gibson, L. S. (2014). *Understanding ethical judgments in secondary school history classes* [Doctoral dissertation, University of British Columbia]. Library and Archives Canada.
- [12] Hung Do, Tuan Hoang, Hy Nguyen, Thuy Bui, Anh Nguyen, Quan Dao, Hoa Tran, Dung Nguyen, Phuong Nguyen, & Tuyen Bui. (2025). *AI Competency Framework for University Students*. Vietnam National University Publishing.
- [13] Ahmad, A. R. (2010). The acquisition of conceptual understanding of historical thinking in the context of multi ethnic students in Malaysia. *Historia: Jurnal Pendidikan dan Penelitian Sejarah*, 11(2), 24–37. <https://doi.org/10.17509/historia.v11i2.12326>
- [14] Đỗ, V. H. (Chủ biên), Hoàng, A. T., Nguyễn, G. H., Nguyễn, P. A., Nguyễn, B. T., Trần, Đ. H., Phạm, V. T. N., & Đào, M. Q. (2025). *Đột phá trong giảng dạy và nghiên cứu cùng trí tuệ nhân tạo*. Nhà xuất bản Đại học Quốc gia Hà Nội.
- [15] University of Social Sciences and Humanities, Vietnam National University, Hanoi (USSH). (2025, October 5). USSH announces "Artificial Intelligence Competency Framework for Students": A foundation for responsible, innovative, and empowered digital citizens.

Author bio

PHAN THI THU is a researcher in the Educational Program Development Research Division under the Center for Curriculum Development Research and Educational Assessment at the Vietnam National Institute of Educational Sciences. She received her Ph.D. in History from Hanoi National University of Education in 2021. Her research interests include modern educational approaches applied in general education, as well as issues related to world and Vietnamese history. The author can be contacted via email at thupt@gesd.edu.vn.

ORCID: 0009-0007-0001-9077.

Google Scholar: <https://scholar.google.com/citations?hl=vi&user=dNw3BPwAAAAJ>