

Creative Thinking in Human-AI Dialogue: A Qualitative Study of History Learning with Generative AI

Shiyuan Yan

Faculty of Education, Beijing Normal University, Beijing 100875, China

E-mail: yanshiyuan01@mail.bnu.edu.cn

Abstract

The rapid development of generative artificial intelligence (generative AI) has introduced new possibilities for transforming how students interact with information, construct knowledge, and engage in inquiry-based learning. However, the educational value of generative AI depends not only on its capacity to generate content but also on how learners critically engage with and reflect upon AI-generated responses. In history education, where learning involves questioning evidence, interpreting the past, and constructing reasoned explanations, understanding how students think through interactions with generative AI has become an important research concern. Existing research has largely focused on the instructional applications and learning outcomes of generative AI, with limited empirical attention to how creative thinking is enacted through sustained human-AI dialogue during historical inquiry. This qualitative study explored how middle school students engaged with the generative AI platform Doubao during historical inquiry and how creative thinking was enacted throughout these interactions. Participants were ten eighth-grade students from a middle school in Beijing, China. Data were collected from multiple sources, including student – AI dialogue records, classroom observations and researcher field notes, and semi-structured interviews, and analysed using Braun and Clarke's thematic analysis. The analysis revealed that students' interactions with generative AI evolved from information seeking towards dialogic exploration and critical negotiation. Across these interactions, creative thinking was progressively enacted as students refined historical questions, evaluated historical evidence, considered multiple perspectives, revised historical explanations, and constructed personal historical interpretations. Rather than being used primarily for information retrieval, generative AI served as a dialogic partner through which students externalised, examined, and progressively refined their historical understanding. By conceptualising creative thinking as an observable process enacted through human-AI dialogue rather than merely a cognitive outcome, this study extends current understandings of how creative thinking can be examined in generative AI-supported history learning.

Keywords: generative AI; creative thinking; human-AI dialogue; history education; middle school student

DOI: 10.7176/JEP/17-7-05

Publication date: July 30th 2026

1. Introduction

In recent years, generative artificial intelligence (generative AI), exemplified by large language models (LLMs), has developed rapidly and become increasingly integrated into educational practice, reshaping how students access information, construct knowledge, and participate in learning activities (Alasadi & Baiz, 2023; Guo et al., 2025). By generating responses based on natural language interaction, generative AI enables learners to engage in conversational exchanges with AI systems, providing new possibilities for questioning, exploration, and knowledge construction during learning activities (Kasneci et al., 2023; Tlili et al., 2023). Through these ongoing interactions, generative AI has increasingly been discussed as a cognitive partner that can support learners' questioning, idea generation, knowledge organisation, and meaning-making processes (Banh & Strobel, 2023; Zhang et al., 2025). As these technologies become more widely adopted in schools, secondary students are increasingly using generative AI to support inquiry-based learning across different disciplines, including history education (Chen et al., 2025; Yang & Wang, 2022).

This development has particular significance for history education. Contemporary history education increasingly emphasises disciplinary inquiry, evidence-based reasoning, and historical interpretation rather than the memorisation of factual knowledge (Bertram et al., 2021). As generative AI becomes embedded in history classrooms, it creates new opportunities for students to engage with historical questions through sustained dialogue, the examination of evidence, and the exploration of multiple perspectives. Understanding how students participate in these interactions has therefore become an increasingly important issue for history education

research.

Existing research has demonstrated the educational potential of generative AI across a wide range of learning contexts. Studies have examined its role in supporting learning performance, classroom instruction, collaborative learning, and students' learning experiences, while also exploring issues such as critical thinking, originality, academic integrity, and ethical use (Chien et al., 2025; Lee et al., 2026; Shahzad et al., 2025). Within history education, researchers have increasingly investigated how generative AI supports historical inquiry, historical understanding, and classroom interaction (Bertram et al., 2021; Chen et al., 2025). Collectively, these studies have substantially advanced our understanding of the instructional applications of generative AI. However, they have focused predominantly on whether generative AI supports creative thinking rather than on how creative thinking is enacted through sustained human-AI dialogue during authentic historical inquiry. In history education, where students are expected to construct interpretations by questioning evidence, considering multiple perspectives, and refining historical explanations, understanding these interaction processes is particularly important. Examining how creative thinking is enacted during human-AI dialogue may therefore provide a richer understanding of how generative AI contributes to inquiry-based history education beyond its role in supporting learning outcomes.

To address this gap, the present study investigates how eighth-grade students engaged with the generative AI platform Doubao during historical inquiry and explores how creative thinking was enacted through sustained human-AI dialogue.

2. Literature Review

2.1 Application of Generative AI in History Education

The growing emphasis on inquiry-based history education has created new opportunities for integrating generative AI into history classrooms. Consequently, increasing attention has been paid to its application in supporting history teaching and learning (Luthfi Fahmi et al., 2025; Washburn & McCutchen, 2024).

Regarding student learning, studies have shown that generative AI can assist students in organizing source evidence, comparing historical perspectives, and generating historical interpretations (Zhang, 2025). Accordingly, recent studies increasingly describe generative AI as a resource for supporting historical inquiry and knowledge construction (Wall, 2024). In terms of classroom instruction, existing studies have explored how teachers integrate generative AI into lesson planning, classroom activity design, formative assessment, and feedback generation, suggesting its potential to support inquiry-oriented teaching practices (Patrick & Javed, 2024; Bazanis et al., 2025; Chimbunde & Moreeng, 2025).

Taken together, these studies have substantially advanced our understanding of what generative AI can support in history education. However, they have focused predominantly on the affordances of generative AI and the outcomes associated with its use. Comparatively less attention has been paid to how students engage with generative AI during sustained historical inquiry in authentic classroom contexts and how these interactions unfold over time.

2.2 Generative AI and Students' Creative Thinking

Creative thinking has long been regarded as a vital component of history learning and historical inquiry (Cooper, 2018a; Pan et al., 2023). It is not free imagination detached from historical facts, but a cognitive process of deepening historical understanding under the guidance of source evidence through posing questions, comparing different perspectives, establishing new connections, and formulating plausible historical interpretations. However, existing research has pointed out that history education has long been influenced by transmission-oriented teaching, where history is often viewed as a subject emphasising established facts and standard answers. A classroom model centred on teacher lecturing rarely provides students with learning opportunities for autonomous inquiry, multiple interpretations, and sustained reflection, which, to some extent, restricts the development of students' creative thinking (Cooper, 2018b; Kesuma et al., 2020). Consequently, researchers have begun to continuously explore new learning environments, hoping to provide students with more open history learning contexts that facilitate historical inquiry, the formation of historical interpretations, and the development of creative thinking (Pan et al., 2023).

With the gradual integration of generative AI into education, researchers have begun to examine its application in creative learning (Chen et al., 2026; Guo et al., 2025). Compared with traditional digital learning, generative AI can support learners in problem analysis, idea generation, knowledge organisation, reflective revision, and meaning construction through sustained human-AI dialogue, and has thus gradually come to be regarded as an

important cognitive resource for creative learning (Rezwana & Maher, 2023; O'Toole & Horváth, 2024). Studies have found that generative AI can simultaneously enhance performance in both the ideation and evaluation phases of brainstorming, thereby improving creative outcomes (Shaer et al., 2024). Students who used ChatGPT generated more unique and high-quality creative solutions (Urban et al., 2024). These studies indicate that generative AI can provide learners with more open knowledge resources and interactive opportunities, offering new learning support for engaging in creative thinking activities.

At the same time, some studies have raised differing views on the role of generative AI in supporting creative thinking. Scholars have pointed out that several questions remain to be addressed regarding whether generative AI can truly stimulate students' creativity (Moruzzi, 2025). For instance, learners may gradually become overly reliant on AI-generated content, leading to cognitive fixation and difficulty in forming new ideas beyond existing suggestions (Habib et al., 2024). These contrasting findings suggest that the significance of generative AI for creative learning may lie less in what it produces for students than in how students think through their interactions with it.

However, existing research on generative AI and creative thinking has primarily focused on fields such as chatbots, language learning, and STEM, with relatively limited attention to history education (Guo et al., 2025; Lee & Chung, 2024). Moreover, when creative thinking has been examined in AI-supported learning, studies have focused predominantly on whether generative AI improves or facilitates creative performance. Given the disciplinary characteristics of history education, considerably less attention has been paid to how creative thinking is enacted through sustained human-AI dialogue during authentic historical inquiry. Rather than conceptualising creative thinking solely as an individual outcome, this study examines how creative thinking is enacted through sustained human-AI dialogue during authentic historical inquiry. The research questions are as follows:

RQ1. How do middle school students engage with generative AI during historical inquiry?

RQ2. How is creative thinking enacted through students' interactions with generative AI during historical inquiry?

3. Method

3.1 Research Design

This study adopted a qualitative research paradigm to explore how middle school students engaged with generative AI during history learning and how creative thinking was enacted throughout their interactions. Given that the study sought to understand students' learning experiences and meaning-making processes in an authentic classroom context, rather than to examine causal relationships or measure learning outcomes, a qualitative approach was considered appropriate.

Purposive sampling was employed to recruit 10 eighth-grade students from a middle school in Beijing, China, as research participants. Data collection relied on three complementary sources: students' dialogue records with Doubao, classroom observations (including researcher field notes), and semi-structured interviews. Thematic analysis was subsequently employed to identify recurring themes and patterns across the data and to develop a comprehensive understanding of the ways in which creative thinking was enacted in generative AI-supported history learning.

This study strictly adhered to ethical research principles. Before the study commenced, the school, teachers, students, and parents were fully informed of the research purpose, procedures, data usage, and potential risks, and written informed consent was obtained from all participants. To protect participants' privacy, all personal information was anonymised through the use of student identification codes (S01-S10). All research data were securely encrypted, stored, and used solely for academic research purposes. Participation was entirely voluntary and unrelated to students' academic performance. Participants were informed that they could withdraw from the study at any stage without any adverse consequences.

3.2 Participants

This study involved 10 eighth-grade students from a middle school in Beijing. Participants were selected through purposive sampling to ensure diversity in their history learning experiences and prior experience using generative AI. Consistent with the exploratory nature of qualitative research, the purpose of participant selection was not to obtain a statistically representative sample, but to include students with varied learning backgrounds and AI experiences, thereby enabling an in-depth exploration of how students engaged with generative AI and how creative thinking was enacted during history learning.

To provide contextual information for interpreting the qualitative findings, participants' demographic and learning characteristics were documented, including gender, history learning interest, history academic achievement, prior experience using generative AI for learning, and previous experience using generative AI in history learning. These characteristics were collected to describe the learning context of the participants rather than to compare differences between groups.

Table 1. Demographic and Learning Characteristics of Student Participants

Characteristic	Category	Number of Participants (n)
Gender	Male	6
	Female	4
History Learning Interest	High Interest	4
	Medium Interest	4
	Low Interest	2
History Academic Achievement	High (≥ 85 points)	3
	Medium (70-84 points)	4
	Low (< 70 points)	3
Prior Experience Using Generative AI for Learning	Frequent (> 3 times/week)	4
	Occasional (1–3 times/week)	3
	Rare (Less than once per week or never)	3
Previous Experience Using Generative AI in History Learning	Yes	6
	No	4

Note. History academic achievement was determined based on students' average scores from three school history assessments completed before the study, including the previous semester's midterm examination, the previous semester's final examination, and a unit assessment conducted during the semester in which the study was undertaken.

The participant group represented diverse learning backgrounds in terms of history achievement, interest in history, and prior experience with generative AI. Some students had previously used generative AI to support history learning, whereas others had little or no such experience. This variation was intentionally sought to capture a broad range of learning experiences and engagement patterns as students engaged with generative AI during history learning. These demographic and learning characteristics were not treated as variables for comparison; instead, they provided contextual information that informed the interpretation of participants' experiences and the themes generated during analysis.

3.3 Research Tools

The three data sources introduced in Section 3.1 were selected to capture students' interactions with generative AI during history learning and to document how creative thinking was enacted throughout historical inquiry.

Doubao was selected as the generative AI platform for this study because of its widespread adoption among secondary school students in China and its capability to support continuous natural language interaction. As one of the most widely used large language model applications in Chinese educational contexts, Doubao enables students to engage in iterative dialogue, explore historical questions, interpret historical evidence, compare multiple historical perspectives, and construct historical explanations. A single AI platform was adopted throughout the study to ensure consistency in the learning environment and to minimise platform-related variation. This design allowed the study to focus on students' learning experiences and interaction processes rather than differences in AI-generated responses. Nevertheless, the findings should be interpreted within the

context of interactions with Doubao and may not be directly transferable to other generative AI platforms.

A semi-structured interview protocol was developed in accordance with the research questions. The interviews invited students to describe how they used generative AI during history learning, how they interacted with AI while exploring historical questions, and how they interpreted, evaluated, and refined historical explanations throughout the learning process. Rather than asking students to evaluate their own level of creativity, the interviews focused on students' experiences of learning with generative AI and how they understood their own inquiry processes during historical learning. Open-ended questions and follow-up prompts were used to encourage participants to elaborate on their experiences, allowing rich descriptions of students' meaning-making processes in authentic classroom contexts. The complete interview protocol is provided in Appendix A.

Classroom observations were conducted throughout the learning activities to document students' interactions with generative AI in authentic history classrooms. Consistent with the qualitative orientation of this study, the observations did not attempt to directly assess creative thinking itself. Instead, the observation protocol focused on observable manifestations of creative thinking during historical inquiry. These manifestations included students' generation of historical questions, interpretation and evaluation of historical evidence, comparison of multiple historical perspectives, construction and revision of historical explanations, and refinement of ideas through ongoing dialogue with generative AI. Particular attention was also paid to how students initiated interactions with AI, responded to AI-generated suggestions, questioned or modified AI-generated historical interpretations, and negotiated historical understanding through continuous human-AI dialogue. In addition to the structured observation protocol, the researcher maintained detailed field notes during every classroom session to document contextual information, unexpected classroom events, and students' spontaneous interactions with Doubao. These field notes served as supplementary observational data during the thematic analysis.

To enhance the credibility of the study, both the semi-structured interview protocol and the classroom observation framework were reviewed by two researchers with expertise in history education and qualitative research before data collection commenced. Based on their feedback, several revisions were made to improve the clarity, relevance, and openness of the research instruments. These revisions included refining the wording of interview questions to reduce potential leading effects and adding follow-up prompts that encouraged students to describe how their historical interpretations evolved during interactions with generative AI. The revisions enhanced the clarity and openness of the instruments and supported the collection of rich qualitative data relevant to the research questions.

3.4 Data Collection Process

Data were collected over a four-week period during regular history lessons at a middle school in Beijing, China. The data collection process consisted of three stages: preparation before the classroom activities, classroom observations during history learning, and semi-structured interviews following the completion of the learning activities.

Prior to data collection, approval was obtained from the participating school, and written informed consent was obtained from all participating students and their parents. Before the classroom activities began, students were introduced to the research procedures and the use of Doubao as the generative AI platform for history learning. To enhance the clarity and appropriateness of the research instruments, the semi-structured interview protocol and classroom observation framework were piloted with two students who were not involved in the formal study. Based on their feedback, minor revisions were made to improve the wording and sequencing of several interview questions and observation items.

During the classroom activities, students completed a series of curriculum-based history inquiry tasks by interacting with Doubao. The regular history teacher implemented all inquiry activities and managed routine classroom procedures, while the researcher remained a non-participant observer throughout the study, completing the observation protocol and maintaining detailed field notes during every session. The six inquiry activities were designed as a coherent sequence through which students progressively developed, evaluated, and refined their historical interpretations through sustained dialogue with Doubao. Rather than simply providing answers, Doubao was used to generate follow-up questions, alternative interpretations, and additional historical evidence that encouraged students to reconsider and refine their own historical explanations. A detailed description of each activity is provided in Appendix B.

A total of five classroom sessions were observed, with each session lasting approximately 60 minutes. Field notes were expanded immediately after each observation session.

Following the completion of the classroom activities, individual semi-structured interviews were conducted with all ten participating students. Each interview lasted approximately 30 to 40 minutes. Students were invited to describe and reflect upon their experiences of using generative AI during history learning, including how they interacted with the AI platform, how they interpreted and evaluated AI-generated responses, how their historical explanations evolved throughout the inquiry process, and how they understood their own learning experiences while engaging with generative AI. All interviews were audio-recorded with participants' consent and transcribed verbatim shortly after each interview.

After data collection had been completed, students' dialogue records, classroom observation records and accompanying field notes, and interview transcripts were organised and anonymised using student identification codes (S01-S10). The three data sources were then integrated and prepared for thematic analysis.

3.5 Data Analysis Methods

Thematic analysis was employed to identify recurring patterns and themes across the three data sources. Following the six-phase approach proposed by Braun and Clarke (2006), the analysis proceeded through familiarization with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the final report.

Data analysis began concurrently with data collection, with all data sources coded as they became available. Initial coding was inductive and closely grounded in the data. While the research questions and the conceptual framework informed by the literature review guided the overall analytical focus, they did not predetermine the coding categories. Codes were generated directly from close reading of the data, allowing themes to emerge from students' own descriptions of their learning experiences.

To ensure the reliability of the coding process, two members of the research team independently coded all interview transcripts and a subset of the dialogue records and observation notes. The two coders met regularly to compare their coding decisions, discuss discrepancies, and refine the coding framework. Disagreements were resolved through discussion and by returning to the original data to examine the contextual meaning of specific excerpts.

For example, when coding an excerpt from Student S03's interview—'At first I thought the AI's explanation made sense, but later I realised the historical source it cited was not accurate, so I looked it up again and revised my answer'—the two coders initially assigned different codes. Coder A categorised this as 'information verification behaviour', while Coder B categorised it as 'critical thinking'. After discussion, both coders reviewed the original transcript and agreed that the student had simultaneously engaged in verifying the AI's output and revising their own interpretation based on evidence. The excerpt was therefore coded as 'evidence-driven reflective revision', a code that was subsequently integrated into the broader sub-theme 'Revising historical explanations' during later stages of the thematic analysis (see Table 2).

Data collection continued until additional interviews and observations no longer generated substantially new insights relevant to the research questions. By the eighth interview, the core themes regarding how students interacted with generative AI and how creative thinking was enacted during history learning had stabilised, and the remaining two interviews, along with later classroom observations, served to confirm and elaborate on the existing thematic structure. Given that the sample included students with diverse levels of history achievement, interest, and prior generative AI experience (see Table 1), the attainment of saturation within this heterogeneous group suggests that the dataset provided sufficient depth and breadth to address the research questions.

4. Findings

The thematic analysis integrated data from students' dialogue records with Doubao, classroom observations and researcher field notes, and semi-structured interviews to develop an interpretive understanding of students' experiences of learning history with generative AI. Two interrelated themes were developed through the analysis. The first captures the ways in which students interacted with Doubao during historical inquiry, whereas the second illustrates how creative thinking was enacted throughout these interactions. Rather than representing separate aspects of learning, the two themes are closely interconnected and together portray students' learning as an iterative process of dialogue, reflection, and historical meaning-making. Table 2 presents an overview of the thematic structure.

Table 2. Themes

Main Themes	Sub-themes	Description
Patterns of Human-AI Interaction	Information seeking	Students primarily used Doubao to obtain historical information or clarify factual questions before continuing the inquiry.
	Dialogic exploration	Students extended conversations by responding to Doubao's follow-up questions, considering alternative interpretations, and elaborating on their initial ideas.
	Critical negotiation	Students critically evaluated AI-generated interpretations, questioned evidence, and selectively incorporated or rejected AI suggestions while constructing their own historical understanding.
Enactment of Creative Thinking during Human-AI Interaction	Historical questioning	Students refined historical questions through iterative dialogue and progressively focused their inquiry.
	Evaluating historical evidence	Students compared different pieces of historical evidence and assessed their relevance when interpreting historical events.
	Considering multiple perspectives	Students explored historical issues from different social, political, or international perspectives before reaching conclusions.
	Revising historical explanations	Students modified and refined their own historical explanations in response to dialogue with Doubao rather than simply adopting AI-generated responses.
	Constructing personal historical interpretations	Students synthesized evidence, dialogue, and reflection to construct their own evidence-based interpretations of historical issues.

4.1 Patterns of Human-AI Interaction

Through thematic analysis of students' dialogue records with Doubao, classroom observations and researcher field notes, and semi-structured interviews, three interrelated patterns of human-AI interaction were identified: information seeking, dialogic exploration, and critical negotiation. These patterns were not static; rather, students' interactions with Doubao evolved throughout the history inquiry activities. Individual students often moved between patterns depending on the inquiry task, the complexity of the historical issue, and their own prior knowledge. The three patterns therefore represent different ways in which students engaged with generative AI while constructing historical understanding, rather than discrete categories or sequential stages.

At the beginning of the inquiry activities, many students approached Doubao primarily as a convenient source of historical information. Their interactions were generally driven by the immediate need to clarify historical concepts, identify factual information, or obtain background knowledge relevant to the inquiry tasks. Classroom observations showed that these conversations were typically brief and linear, consisting of a single question

followed by a single response from the AI. Once sufficient information had been obtained, students frequently returned to the task without extending the dialogue.

S04 reflected this pattern throughout the early inquiry activities. Most of his interactions focused on obtaining factual information, such as the chronology of historical events or the background of particular policies. Reflecting on his learning experience, he explained:

“I just asked it what I needed to know. It gave me answers quickly, so I could finish the task faster.” (S04, Interview)

Although these interactions appeared relatively limited, they provided students with an accessible entry point into historical inquiry. For several participants, particularly those who were less confident about the historical content, information seeking reduced uncertainty and enabled them to establish a basic understanding before engaging in more complex historical reasoning.

As students became increasingly familiar with both the inquiry activities and Doubao’s conversational style, their interactions gradually extended beyond information retrieval. Instead of treating AI responses as final answers, many students began to ask follow-up questions, request clarification, and explore alternative historical interpretations suggested during the dialogue. The interaction itself gradually became part of the inquiry process rather than merely a means of obtaining information.

S07 described how this shift occurred during one of the inquiry activities:

“At first I only wanted an answer. But when Doubao asked me another question, I realised there were other ways to think about it. Then I wanted to keep asking.” (S07, Interview)

Rather than simply accumulating historical knowledge, students increasingly used dialogue to examine different possibilities, reconsider their initial ideas, and expand their understanding of historical events. Classroom observations suggested that this iterative interaction became progressively more common during the later inquiry activities, as students grew more comfortable sustaining conversations with the AI. In this sense, Doubao functioned not only as an information resource but also as a conversational partner that encouraged continued historical exploration.

A further pattern emerged when students actively evaluated and negotiated the ideas generated during the dialogue. Rather than accepting AI-generated explanations at face value, these students questioned the AI’s reasoning, requested additional historical evidence, compared alternative interpretations, and ultimately decided whether particular suggestions should be incorporated into their own historical explanations. Dialogue therefore became a process of negotiating historical meaning rather than simply exchanging information.

S09 illustrated this pattern particularly clearly. When Doubao proposed an interpretation that differed from her own understanding, she challenged the explanation and continued the dialogue until she felt able to justify her own position. Reflecting on this experience, she explained:

“I didn’t just accept what Doubao said. Sometimes I agreed, sometimes I didn’t. But every time I had to think about why I agreed or disagreed.” (S09, Interview)

Similarly, S01 described Doubao as “someone to discuss history with rather than someone who simply gives answers.” Throughout several inquiry activities, he repeatedly questioned AI-generated explanations, requested further justification, and revised his own ideas only after critically considering the alternatives presented during the dialogue. These findings suggest that students were not passive recipients of AI-generated information. Instead, they actively negotiated the credibility, relevance, and explanatory value of AI-generated ideas while constructing their own historical understanding.

Taken together, these findings demonstrate that students’ interactions with generative AI cannot be adequately described as simple information retrieval. Although information seeking remained an important starting point for many participants, sustained engagement with Doubao gradually encouraged more dialogic and evaluative forms of interaction. Rather than replacing students’ own thinking, human-AI dialogue increasingly became a space in which historical ideas were explored, challenged, and refined through continuous interaction.

4.2 Manifestations of Creative Thinking during Human-AI Interaction

Whereas the previous section examined how students interacted with Doubao during historical inquiry, this section addresses how creative thinking was enacted as these interactions unfolded. Across the dialogue records, classroom observations and field notes, and interview data, creative thinking was not treated as an internal cognitive ability that could be directly observed. Instead, it became evident through students’ observable learning behaviours and verbal expressions as they refined historical questions, evaluated evidence, considered alternative

perspectives, revised historical explanations, and gradually constructed their own understandings of historical events. These enactments often emerged progressively throughout dialogue rather than appearing at a single moment.

One of the earliest ways in which creative thinking was enacted was through students' continual refinement of historical questions. Rather than treating their initial questions as fixed, many students modified, narrowed, or reformulated their inquiries after engaging in dialogue with Doubao. The AI frequently responded with probing questions or pointed out issues that had not yet been considered, prompting students to reconsider what they actually wanted to investigate. This pattern was particularly evident during the first inquiry activity, in which students were asked to formulate a historical question that genuinely interested them. Classroom observations indicated that many students initially proposed broad or descriptive questions. As the dialogue progressed, however, they increasingly refined these questions by identifying ambiguities, specifying historical contexts, or focusing on particular aspects of the issue.

S03 described this process during the interview:

"At first I just wanted to know why that event happened. But Doubao kept asking what exactly I meant by 'why'. I realised there were many different reasons, so I changed my question." (S03, Interview)

Similarly, the researcher's field notes documented that several students repeatedly revised the wording of their inquiry questions before proceeding to search for historical evidence. Rather than simply generating more questions, students demonstrated a growing awareness that productive historical inquiry often begins with clarifying the problem itself.

Creative thinking was also enacted as students evaluated historical evidence and considered multiple historical perspectives through dialogue with Doubao. Rather than accepting AI-generated explanations as authoritative, many students compared different viewpoints, questioned inconsistencies, and reconsidered their own interpretations in response to alternative perspectives introduced during the dialogue. For example, during the activity requiring students to evaluate the historical significance of major events between 1949 and 1978, Doubao frequently presented criteria that differed from those initially proposed by students. Instead of immediately accepting these suggestions, many students engaged in extended dialogue, asking for further explanations or requesting additional evidence before deciding whether to modify their own judgments.

Reflecting on this experience, S06 commented:

"When Doubao gave a different explanation from mine, I didn't change my answer straight away. I wanted to know why it thought that." (S06, Interview)

Classroom observations similarly showed that students increasingly posed questions such as "Is there another way to understand this?" and "What evidence supports that interpretation?" These exchanges often led students to compare multiple historical perspectives before reaching a provisional conclusion. Importantly, students rarely described the dialogue as a process of identifying a single correct historical explanation. Instead, they regarded alternative interpretations as opportunities to reconsider their own reasoning.

Another important way in which creative thinking was enacted was through students' revision of historical explanations throughout the inquiry process. Many students described moving beyond simply collecting information from Doubao to integrating different ideas into explanations that they considered more coherent and personally meaningful. This pattern was particularly evident during the final inquiry activity, in which students were asked to summarise their understanding of China's historical development between 1949 and 1978. Rather than producing a final answer immediately, many students repeatedly revised their interpretations after discussing alternative viewpoints with Doubao. They selectively incorporated some suggestions, rejected others, and reorganised their explanations until they felt that the final interpretation genuinely reflected their own understanding.

S09 explained:

"Sometimes I agreed with Doubao, and sometimes I didn't. The important thing wasn't copying what it said. It was thinking about whether my own explanation still made sense." (S09, Interview)

Similarly, S01 reflected:

"In the end, the explanation was still mine. Doubao gave me ideas, but I decided which ones to keep." (S01, Interview)

The classroom observations supported these accounts. Students frequently deleted, reorganised, or rewrote substantial sections of their written explanations after engaging in multiple rounds of dialogue with Doubao.

These revisions were not simply corrections of factual information but represented attempts to strengthen the coherence and persuasiveness of their historical interpretations.

Taken together, these findings suggest that creative thinking during generative AI-supported history learning was enacted not through the production of unusually original ideas alone, but through students' iterative processes of questioning, evaluating, revising, and constructing increasingly coherent historical understandings in dialogue with generative AI. Rather than serving as the source of students' ideas, Doubao functioned as a conversational partner whose questions, alternative interpretations, and feedback created opportunities for students to externalise, examine, and progressively refine their own historical thinking.

5. Discussion

5.1 Key Findings

This study explored how middle school students engaged with generative AI during historical inquiry and how creative thinking was enacted throughout these interactions. Rather than positioning generative AI simply as a source of historical information, the findings suggest that students' interactions with Doubao constituted an evolving process of inquiry in which creative thinking was progressively enacted through sustained dialogue. Three key findings emerge from this study.

First, students' engagement with generative AI was dynamic rather than static. Although many students initially used Doubao primarily to obtain factual information, their interactions often became more dialogic as they gained experience with both inquiry-based learning and the AI platform. The interaction patterns identified in this study thus do not represent fixed learner types, but rather different modes of participation that students adopted at different moments in the learning process. This finding indicates that meaningful human-AI dialogue does not arise automatically through mere access to generative AI; instead, it develops progressively through sustained interaction and inquiry-oriented task design. This resonates with the view that generative AI can function as a cognitive partner supporting knowledge construction (Wall, 2024). At the same time, not all students demonstrated this progression—several participants remained primarily at the information-seeking level throughout the study. This suggests that access to generative AI alone may be insufficient to support deeper engagement, highlighting the importance of pedagogical scaffolding in helping students move beyond information retrieval toward more reflective and dialogic forms of inquiry (Chimbunde & Moreeng, 2025; Patrick & Javed, 2024).

Second, this study offers an alternative perspective on creative thinking in history learning, suggesting that it can be productively understood as an evolving inquiry process rather than solely as a measurable learning outcome. Rather than evaluating whether generative AI improved students' creativity, the analysis examined how creative thinking was enacted through students' interactions with Doubao. Across the inquiry activities, students progressively refined their historical questions, reconsidered evidence, explored alternative perspectives, revised their explanations, and constructed increasingly personal interpretations of historical events. These enactments did not appear at a single moment but emerged gradually through cycles of questioning, reflection, and revision. This finding complements existing research that has primarily assessed creativity through final products or performance outcomes (e.g., Shaer et al., 2024; Urban et al., 2024) by illuminating how creative thinking can also be understood as an ongoing process of disciplinary meaning-making. It also extends the theoretical work of Pan et al. (2023) and Cooper (2018a) on creative thinking in history education by empirically demonstrating what such a process can look like when generative AI is embedded in authentic historical inquiry.

Third, the findings underscore the importance of disciplinary context in understanding the educational role of generative AI. In history learning, students are expected not only to acquire historical knowledge but also to evaluate evidence, compare competing interpretations, and construct reasoned historical explanations (Bertram et al., 2021; Pan et al., 2023). In the context of this study, the significance of generative AI therefore resided not merely in its capacity to provide historical information, but more importantly, in its potential to serve as a dialogic tool that supported students' ongoing interpretation and reconsideration of historical understanding. This aligns with recent arguments that generative AI in history education should be viewed as more than an information retrieval tool (Wall, 2024; Zhang, 2025). Moreover, this finding suggests that the educational significance of generative AI cannot be fully understood independently of the epistemological characteristics of different disciplines. While previous research has documented the role of generative AI in supporting creative thinking across fields such as language learning and STEM (Guo et al., 2025; Lee & Chung, 2024), the present study demonstrates that in history education, the emphasis on source-based reasoning and interpretive pluralism shapes the nature of human-AI interaction in distinctive ways. Rather than functioning as a universal

instructional tool, generative AI may support learning in different ways depending on the nature of disciplinary inquiry.

Taken together, these findings call for a shift in attention from asking whether generative AI enhances creative thinking to examining how human-AI dialogue creates conditions in which creative thinking is enacted during authentic historical inquiry. From this perspective, we argue that generative AI can be productively viewed not as a substitute for students' historical reasoning, but as a dialogic partner with the potential to support their continual questioning, refinement, and reconstruction of historical understanding. At the same time, it is important to acknowledge that such benefits are not guaranteed. As Habib et al. (2024) have cautioned, learners may risk cognitive fixation or uncritical reliance on AI-generated content.

5.2 Limitations of the Study and Recommendations for Future

Several limitations of this study should be acknowledged.

First, this study was conducted in a single middle school in Beijing with ten eighth-grade students. Consistent with the purpose of qualitative inquiry, the study sought to develop an in-depth understanding of students' experiences within an authentic classroom context rather than to achieve statistical generalisation. Consequently, the interaction patterns and how creative thinking was enacted identified in this study may be shaped by the specific educational setting. Future research conducted across different schools, regions, and educational contexts would help examine the transferability of these findings.

Second, all inquiry activities were completed using a single generative AI platform, Doubao. Employing one platform ensured consistency throughout the inquiry process and enabled the analysis to focus on students' interaction experiences rather than technological differences between AI systems. However, the findings may partly reflect the conversational characteristics and response patterns of this particular platform. Future studies comparing different generative AI systems could provide a deeper understanding of how platform characteristics influence human-AI dialogue during disciplinary inquiry.

Third, data collection was conducted over a four-week period. Although this allowed the study to capture changes in students' interactions across the inquiry activities, it could not reveal how interaction patterns and how creative thinking is enacted might evolve through longer-term engagement with generative AI. Longitudinal research would therefore be valuable for understanding how students' inquiry practices develop over time as they become increasingly familiar with generative AI-supported history learning.

Finally, the qualitative design enabled an in-depth exploration of students' interaction processes but did not seek to determine how frequently particular interaction patterns occur or how creative thinking is enacted across broader student populations. Future research adopting mixed-methods designs could further examine the wider applicability of these findings and deepen our understanding of how creative thinking is enacted through students' sustained interactions with generative AI during history learning.

6. Conclusion

This study explored how middle school students engaged with generative AI during history learning and how creative thinking was enacted throughout these interactions. The findings show that students' interactions with generative AI were dynamic and evolved during the inquiry process. Rather than simply using AI to obtain historical information, students increasingly engaged in questioning, evaluating evidence, considering multiple perspectives, revising historical explanations, and constructing their own interpretations through sustained dialogue with Doubao.

The study contributes to current research by shifting attention from what generative AI can do for students to how creative thinking is enacted through human-AI dialogue in disciplinary learning. By examining students' interactions within authentic historical inquiry, this study demonstrates that the educational value of generative AI lies not merely in providing information but in supporting dialogue through which students progressively develop and refine their historical understanding.

As generative AI becomes increasingly integrated into school education, understanding students' interactions with these technologies within authentic disciplinary contexts will become increasingly important. Ultimately, the educational significance of generative AI may lie not in replacing students' thinking or generating historical knowledge on their behalf, but in creating dialogic spaces where students can question, negotiate, and continuously reconstruct their historical understanding.

Future research could further explore how human-AI dialogue unfolds across different educational contexts

and disciplinary settings, particularly how students' interaction patterns and creative thinking processes may vary with different inquiry tasks and learning environments. Longitudinal studies may also provide deeper insights into how students' ways of engaging with generative AI develop over time as they gain experience with AI-supported learning. Moreover, comparative investigations across different generative AI platforms could help clarify how technological affordances shape students' inquiry processes and historical meaning-making.

Appendix A. Semi-Structured Interview Protocol

Table A1. Main interview questions and probes

No.	Core Interview Question	Follow-up Prompts
1	Looking back on the six history inquiry activities, which activity do you remember most clearly? Could you tell me what happened from beginning to end?	• What interested you most about this activity? • What did you do first? • What happened next? • Why do you think this activity stayed in your memory?
2	Can you describe how your conversation with Doubao developed during that activity?	• What kinds of questions did you ask? • How did Doubao respond? • Did you continue asking questions, or did the conversation end quickly? • What made you decide to continue (or stop) the dialogue?
3	Was there a moment during your conversation with Doubao that surprised you, confused you, or made you stop and think? Please tell me what happened.	• What did Doubao say? • Why did that catch your attention? • What did you do afterwards? • Did you continue asking questions?
4	Can you remember a time when you did not immediately accept what Doubao suggested or explained?	• What made you hesitate? • How did you decide whether the information was reasonable? • Did you check other sources or ask additional questions? • What did you finally decide?
5	Looking back at one of the inquiry activities, how did your understanding or explanation develop from the beginning to the end?	• When did you first realize your thinking was changing? • What contributed most to that change? • Were there any ideas you decided to keep unchanged? • Why?
6	During the activities, was there a time when you began to think about a historical event or issue in a different way?	• What made you think differently? • Was it something Doubao said, or something you discovered yourself? • How did your final understanding differ from

		your initial one?
7	Thinking about the final inquiry activity, can you tell me how you arrived at your own historical interpretation of China's development between 1949 and 1978?	• Where did you begin? • Which ideas or evidence did you decide to include? • Were there any suggestions from Doubao that you accepted, modified, or rejected? • Why did you make those choices?
8	Looking back on all six activities, how would you describe the role that Doubao played during your history learning?	• Did its role change over time? • Were there situations where it was especially helpful? • Were there situations where it was less helpful? • Can you give an example?
9	Looking back on the whole experience, what do you think was the most important thing you learned from history learning with Doubao?	• Has this experience changed the way you learn history? • Would you approach a similar inquiry differently in the future? • What advice would you give another student who is engaged in history learning with Doubao? • Is there anything else about your experience that you would like to share?

Appendix B. Overview of the Generative AI-Supported History Inquiry Activities

Activity 1: Formulating a Historical Question

Students proposed an initial question about the period 1949–1978 that genuinely interested them. Doubao responded with follow-up questions that prompted students to identify ambiguities, narrow their focus, or consider neglected perspectives. Students revised their questions through multiple rounds of dialogue until they arrived at a final version. Finally, students wrote a brief reflection on how their question had evolved through this process. The final inquiry question developed in this activity became the starting point for students' subsequent historical inquiry.

Activity 2: Selecting and Defending Historical Significance

Students selected three events they considered most significant for introducing this period to an international peer unfamiliar with Chinese history. Doubao presented alternative events or suggested different criteria for assessing historical significance, encouraging students to reconsider their initial choices. Students then revised their selection—retaining one original event, replacing one—and provided an evidence-based justification for what had changed and why. The judgments formed in this activity informed students' later evaluations of historical figures and historical developments.

Activity 3: Re-evaluating a Historical Figure

Students chose a prominent historical figure from this period and wrote down their initial evaluation. Doubao presented contrasting interpretations and additional historical evidence that encouraged students to reconsider their initial view. Students continued the dialogue, questioned Doubao's suggestions and refined their evaluations through multiple rounds of discussion. Finally, students described what had changed in their evaluation and what evidence or interpretation had prompted the shift. Students' revised evaluations provided a basis for exploring broader historical relationships in the following activity.

Activity 4: Tracing Causal Connections

Students selected a significant event and mapped its influence on subsequent historical developments, supported by historical evidence. As the dialogue progressed, Doubao raised questions about indirect or long-term effects and introduced counterfactual prompts. Students iteratively revised their causal maps as new ideas emerged through dialogue, marking which connections they had identified independently and which emerged through interaction with Doubao. This activity emphasised the ongoing refinement of causal understanding, which deepened students' capacity to construct evidence-based historical interpretations.

Activity 5: Dialoguing across Time

Students selected a historical actor—such as a farmer, factory worker, or local official—and wrote down their assumptions about this person's daily life. Doubao adopted the first-person perspective of this actor, and students engaged in extended conversations, asking follow-up questions based on previous responses. After the dialogue, students reflected on how their understanding of this actor's experience had changed. This activity emphasised perspective-taking and the examination of historical experiences from multiple viewpoints through sustained dialogue, which informed the personal historical interpretation students constructed in the final activity.

Activity 6: Constructing a Personal Historical Interpretation

Students wrote a concise statement summarizing their understanding of China's development between 1949 and 1978, followed by an extended, evidence-based explanation. Doubao generated alternative interpretations and raised questions about the student's explanation, prompting further reflection. Students negotiated their final interpretation through multiple rounds of discussion, deciding which feedback to incorporate and which to set aside. Finally, students described how their interpretation had evolved from the initial to the final version. Rather than seeking a definitive historical explanation, students were encouraged to regard their interpretations as provisional, evidence-based, and open to continuous refinement through further inquiry.

References

- Alasadi, E., & Baiz, C. R. (2023). Generative AI in Education and Research: Opportunities, Concerns, and Solutions. *Journal of Chemical Education*, 100(8), 2965–2971. <https://doi.org/10.1021/acs.jchemed.3c00323>
- Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33, 63. <https://doi.org/10.1007/s12525-023-00680-1>
- Bazanis, V., Bilali, A., Karatziki, Z., Katsarou, A., & Nikolaou, G. (2025). The contribution of artificial intelligence to the learning and teaching of history in secondary education. *European Journal of Education Studies*, 12(12), 367–387. <https://doi.org/10.46827/ejes.v12i12.6423>
- Bertram, C., Weiss, Z., Zachrich, L., & Ziai, R. (2021). Artificial intelligence in history education: Linguistic content and complexity analyses of student writings in the CAHisT project (Computational assessment of historical thinking). *Computers and Education: Artificial Intelligence*, 100038. <https://doi.org/10.1016/j.jcaeai.2021.100038>
- Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chen, J., Mokmin, N. A. M., & Qi, S. (2025). Generative AI-powered arts-based learning in middle school history: Impact on achievement, motivation, and cognitive load. *The Journal of Educational Research*, 118(6), 688–700. <https://doi.org/10.1080/00220671.2025.2510395>
- Chen, P., Ma, Z., Liang, Y., & Lu, H. (2026). Impact of metacognitive scaffolding on high-school students' creativity within a GenAI learning environment. *Interactive Learning Environments*, 1–16. <https://doi.org/10.1080/10494820.2026.2677726>
- Chien, C. C., Chan, H. Y., & Hou, H. T. (2025). Learning by playing with generative AI: design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction. *Journal of Research on Technology in Education*, 57(4), 894–913. <https://doi.org/10.1080/15391523.2024.2338085>
- Chimbunde, P., & Moreeng, B. B. (2025). Reconfiguring history teachers' assessment practices in the context of ChatGPT: Charting the new terrain. *TechTrends*, 69, 607–617. <https://doi.org/10.1007/s11528-025-01062-4>

- Cooper, H. (2018a). Children, their world, their history education: The implications of the Cambridge review for primary history. *Education 3-13*, 46(6), 5–69. <https://doi.org/10.1080/03004279.2018.1483797>
- Cooper, H. (2018b). What is creativity in history? *Education 3-13*, 46(6), 636–647. <https://doi.org/10.1080/03004279.2018.1483799>
- Guo, H., Zhou, Z., & Ma, F. (2025). The influence of artificial intelligence generated content-based problem-solving on engineering students: A controlled experimental study. *Education and Information Technologies*, 30, 25969–25999. <https://doi.org/10.1007/s10639-025-13709-z>
- Habib, S., Vogel, T., Thorne, E., & Xiao, A. (2024). How does generative artificial intelligence impact student creativity? *Journal of Creativity*, 34(1), 100072. <https://doi.org/10.1016/j.vjoc.2023.100072>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & 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>
- Kesuma, A. T., Putra, H., Mailool, J., & Kristoro, H. C. A. (2020). The effects of Mansa historical board game toward students' creativity on historical subjects. *European Journal of Educational Research*, 9(4), 1689–1700. <https://doi.org/10.12973/eu-jer.9.4.1689>
- Lee, B. C., & Chung, J. (2024). An empirical investigation of the impact of ChatGPT on creativity. *Nature Human Behaviour*, 8, 1906–1914. <https://doi.org/10.1038/s41562-024-01953-1>
- Lee, S., Choe, H., Zou, D., & Jeon, J. (2026). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 34(1), 335–359. <https://doi.org/10.1080/10494820.2025.2498537>
- Luthfi Fahmi, R., Haryanto, H., Raharja, S., & Luthfi, M. I. (2025). Effect of AI-Enhanced Learning on Students' History Achievement and Engagement. *Journal of Advanced Research Design*, 138(1), 175–191. <https://doi.org/10.37934/ard.138.1.175191>
- Moruzzi, C. (2025). Artificial intelligence and creativity. *Philosophy Compass*, 20(3), e70030. <https://doi.org/10.1111/phc3.70030>
- O'Toole, K., & Horváth, E. Á. (2024). Extending human creativity with AI. *Journal of Creativity*, 34(2), 100080. <https://doi.org/10.1016/j.vjoc.2024.100080>
- Pan, A.-J., Lai, C.-F., & Kuo, H.-C. (2023). Investigating the impact of a possibility-thinking integrated project-based learning history course on high school students' creativity, learning motivation, and history knowledge. *Thinking Skills and Creativity*, 47, 101214. <https://doi.org/10.1016/j.tsc.2022.101214>
- Patrick, R., & Javed, A. (2024). AI apps that enhance history teaching. *Agora*, 59(2), 25–28. <https://search.informit.org/doi/10.3316/informit.T2024070500013000280034445>
- Rezwana, J., & Maher, M. L. (2023). Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems. *ACM Transactions on Computer-Human Interaction*, 30(5), 1–28. <https://doi.org/10.1145/3519026>
- Shaer, O., Cooper, A., Mokryn, O., Kun, A. L., & Ben Shoshan, H. (2024). AI-augmented brainwriting: Investigating the use of LLMs in group ideation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1–17). <https://doi.org/10.1145/3613904.3642414>
- Shahzad, M. F., Xu, S., & Zahid, H. (2025). Exploring the impact of generative AI-based technologies on learning performance through self-efficacy, fairness & ethics, creativity, and trust in higher education. *Education and Information Technologies*, 30, 3691–3716. <https://doi.org/10.1007/s10639-024-12949>
- 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(15). <https://doi.org/10.1186/s40561-023-00237-x>
- Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., ... Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*, 215, 105031. <https://doi.org/10.1016/j.compedu.2024.105031>
- Wall, V. (2024). “Here, ready or not!”: Teaching history in an AI world. *Agora*, 59(2), 13–16. <https://search.informit.org/doi/10.3316/informit.T2024070500012701986667387>
- Washburn, J., & McCutchen, J. M. (2024). AI meets AI: ChatGPT as a pedagogical tool to teach American Indian history. *Critical Humanities*, 2(2), 1–27. <https://doi.org/10.33470/2836-3140.1037>
- Yang, K., & Wang, H. (2022). The application of interactive humanoid robots in the history education of museums under artificial intelligence. *International Journal of Humanoid Robotics*, 20(06). <https://doi.org/10.1142/s0219843622500165>

- Zhang Li. (2025). Research on the Creation of Middle School History Teaching Situation and the Training Strategy of Students' Historical Thinking Based on Artificial Intelligence Technology. *Journal of Sociology and Education*, 1(4). <https://doi.org/10.63887/jse.2025.1.4.13>
- Zhang, Z., Zhang, J., Zhang, X., & Mai, W. (2025). A comprehensive overview of Generative AI (GAI): Technologies, applications, and challenges. *Neurocomputing*, 632, 129645. <https://doi.org/10.1016/j.neucom.2025.129645>