



AI Chatbots to Support Disciplinary Literacy in Social Studies for Middle School Multilingual Spanish Learners

Caio Gomes^a, Eric Junco^{*b}, Razak Dwomoh^b & Diana Bonilla^b

* Corresponding author

Email: Ericjunco22@gmail.com

a. Multilingual Teacher & AI Literacy Specialist, Clinton Rosette Middle School, Dekalb, Illinois, United States.

b. Northern Illinois University, Department of Curriculum & Instruction, Dekalb, Illinois, United States of America



10.46303/ressat.2026.10

Article Info

Received: January 11, 2026

Accepted: April 24, 2026

Published: July 5, 2026

How to cite

Gomes, C., Junco, E., Dwomoh, R., & Bonilla, D. (2026). AI Chatbots to Support Disciplinary Literacy in Social Studies for Middle School Multilingual Spanish Learners. *Research in Social Sciences and Technology*, 11(2), 66-92. <https://doi.org/10.46303/ressat.2026.10>

Copyright license

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (CC BY 4.0).

ABSTRACT

Disciplinary literacy in social studies includes multiple traditions. This study focuses specifically on historical thinking practices. A tool that can create opportunities for students' to practice disciplinary historical thinking are AI Chatbots. However, there is little research detailing how AI chatbots are used in multilingual classrooms beyond creating interactive, simulation-oriented experiences for students. In this article, we employed a MagicSchool Custom Chatbot Builder to develop a Chatbot Challenge (Desafío Chatbot) assignment assisting middle school multilingual learners (MLLs) as they practiced historical inquiry. Thus, through a Spanish Chatbot assignment, students interacted with a historical figure, researched about them, and analyzed their interactions, practicing the skills of a historian, such as critical questioning, sourcing, and contextualization, evaluating the reliability of information generated by AI, and explaining a figure's historical legacy in relation to equality, justice, and social change. Findings from this interpretive case suggest that across classrooms, the Desafío Chatbot appeared to lower barriers to participation for MLLs. Drawing on classroom artifacts, teacher notes, and observations of student engagement, this article offers an account of how disciplinary moves such as inquiry, sourcing, contextualization, corroboration, close reading, and expanding learning became visible across the assignment's stages. By engaging in planning, questioning, interaction, and reflection, students practiced historical thinking, evidence-based inquiry, chronological reasoning, and digital literacy. This article offers insight into how teachers can foster grade-appropriate engagement in historical thinking through AI chatbots.

KEYWORDS

Disciplinary literacy; multilingual learners (MLLs); AI chatbots; historical thinking; middle school; social studies.

INTRODUCTION

Despite nearly a half a century worth of scholars, teachers, and educational organizations alike emphasizing the necessity of general reading strategies, many students still struggle with the challenges of grade-level social studies texts (e.g., Bickford et al., 2020; Nokes, 2011). Since it emerged in the early 2000s, disciplinary literacy has been elevated as a means to close literacy gaps and improve student achievement. Although social studies involves multiple disciplinary traditions including history, civics, economics and geography (National Council for the Social Studies [NCSS], n.d.), each with their own distinct literacy practices (Kenna et al., 2018), this study focuses specifically on historical thinking practices as a form of disciplinary literacy within social studies. Disciplinary literacy in social studies goes beyond general reading skills and strategies to centering the skills required by historians to make sense of complex texts (Shanahan, 2014). Historians approach texts differently from general readings by close reading, sourcing, contextualization, and corroboration to build interpretations of the past (Wineburg & Reisman, 2015; Wineburg, 1991). Interpreting texts such as primary sources would require significant background knowledge, vocabulary knowledge, and problem-solving (Bain, 2006).

Thus, disciplinary literacy in social studies classrooms means students engage in the practices of historians, to do things like synthesizing and evaluating multiple sources (Nokes, 2011). In this study, disciplinary literacy is operationalized through foundational historical thinking practices such as corroboration, contextualization, and sourcing. Students often struggle to adopt these practices because disciplinary texts are linguistically complex, culturally distant, and cognitively demanding (Bain, 2006; Nokes, 2011). These cultural and linguistic demands can be intensified for multicultural students and multilingual learners (MLLs) and their engagement in the content area. MLLs can struggle with background knowledge of histories of new places and understanding of words across new contexts (Fang & Schleppegrell, 2010; Short & Fitzsimmons-Doolan, 2007).

The cultural and linguistic demands of the discipline require closer attention to responsively and equitably teach disciplinary literacy with a focus on linguistic development of MLLs to support them in reading like historians. Existing challenges to MLLs are the complexity of primary sources, especially when texts come from before the 19th century. These challenges tend to impact students' motivation and engagement when learning social studies. Without structured opportunities to apprentice into disciplinary ways of thinking, students may revert to treating history as the recall of facts rather than inquiry into evidence and interpretation (Barton & Levstik, 2004).

Social studies teachers across scholarship tend to struggle with operationalizing discipline-specific skills, often because they lack access to training (Langan & Lawrence, 2021; Trombino, 2010; Wineburg & Reisman, 2015). For instance, Soininen's (2022) research illustrates that although it was common practice for many social studies educators to employ primary sources in their instruction, they seldom encouraged students to engage in close reading and contextualize these texts due to feeling pressure to teach additional content.

Similarly, Gebhard (2010) argues that teachers are not always equipped to teach academic language development as utilized within their content areas. This leads to teaching social studies as the process of recalling facts (Barton & Levistik, 2004) rather than engaging students in inquiry-based approaches activating social studies disciplinary literacy to understand the world around them (Freire, 1970). Our research question for this study was as follows:

1. How can AI chatbots used in a multilingual setting support disciplinary literacy in a social studies classroom?
2. How do AI chatbots create opportunities for multilingual learners to demonstrate historical thinking practices?

Theoretical Framework

Historical Thinking as Disciplinary Practice in Social Studies

Students come into our classrooms with academic and cultural wealth that can be drawn from in various languages (e.g., Dwomoh et al., 2022). When engaging in content learning with MLLs, educators can center critical language awareness (CLA) and critical cultural awareness (e.g., Dwomoh et al., 2022) to support social studies and disciplinary literacy development across languages through transformative practices (Fairclough, 1999; Quan, 2021). Language can become a tool for learning, welcoming linguistic development and understanding by students and educators alike (Gebhard, 2010). This means “engag[ing] with students in ways that encourage them to read from positions that are both mainstream and resistant and to write in ways that are powerful so that they can participate in new disciplinary contexts and contribute to shaping and reshaping those contexts” (Achugar et al., 2007, p. 10).

Historical thinking practices in social studies requires students to read primary resources that can originate in a variety of languages in this multilingual discipline (Coffin, 2006). The linguistic demands include grappling with dense language and abstract phrases, using multiple sources for information, and developing arguments grounded in evidence from the texts (Wineburg 1991; Schleppegrell, 2004). One challenge is ‘reading to learn’ where students rely on information from the text to understand unfamiliar language, decode, and make interferences and connections between sources (Achugar & Carpenter, 2012). Achugar and Carpenter (2012) outline key strategies to engage in critical language awareness and multilingual literacy development such as; modeling practices, making explicit connections between language, scaffolding, and integrated reading tools while building and engaging their linguistic repertoires.

Additionally, Reisman (2012) has shown through the Reading Like a Historian curriculum that guided prompts and structured discussions help students practice sourcing and corroboration. Scholars also emphasize that writing in history is not simply about recounting events but about constructing arguments from evidence (Monte-Sano, 2010; Nokes 2017). While such tools cannot replace teacher feedback, they can generate additional practice opportunities that align with the disciplinary goal of evidence-based writing. Speaking and discussion are likewise central to historical inquiry. Mercer and Howe (2012) argue that

exploratory talk fosters reasoning and conceptual understanding across disciplines. In history classrooms, structured dialogue helps students articulate interpretations and consider alternative viewpoints.

In sum, historical and disciplinary thinking practices within social studies encourage students to read, think, write, and speak like historians. Established scholarship demonstrates that these practices are challenging but teachable with structured scaffolds (Wineburg, 1991; Reisman, 2012; Monte-Sano, 2010).

Critical AI Literacy

Driven by a 'rhetoric of inevitability' (Conrad & Kamperman, 2025), schools have begun to rapidly, and sometimes uncritically (Baer, 2025) integrate AI tools to support learning contexts from developing fundamental vocabulary, reading, and writing skills (Crompton & Burke, 2024) to disciplinary inquiry (Picasso et al., 2024; Tanksley, 2024). These new literacy contexts precipitate the need for what Long and Magerko (2020) call AI literacy, "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace" (p.598). Building upon a foundation of digital literacy (Lankshear & Nobel, 2003), scholars define critical AI literacy as a skillset that requires users to understand social and ethical implications of AI (Tacelosky et al. 2025) in addition to the skills necessary to evaluative skills to interpret and utilize informational computer technologies in social contexts (Junqueira, 2025).

Tensions involving social and ethical implications within critical AI literacy can be both disciplinary and broad. As a case in point, Tacelosky et al., (2025) explored students' experiences with generative AI in language learning and found most students viewed it as a consultation tool acknowledging the affordances of speed and efficiency generative AI provided. Conversely, a fundamental pillar of critical AI literacy troubles this same efficiency students appreciated, positing that generative AI output is neither neutral nor always reliable (Junqueira, 2025).

Because generative AI is often trained on large and uncured datasets, AI output can reflect hegemony of dominant voices and the erasure of marginalized ones (Bender et al., 2021; Tanksley, 2024). critical AI literacy calls for competencies like understanding generative AI's strengths and weaknesses, recognizing how AI makes decisions, recognizing human's role in programming AI, and critically interpreting AI output (Long & Magerko, 2020). Likening generative AI output to "stochastic parrots" which lack understanding or communicative intent, generative AI is capable of generating "text based on their internal logic or patterns, rather than the true context," a phenomenon known as hallucination (Li, 2023, p.1). As a result, some Critical AI scholars have called for a "Data Justice" approach to emphasize civic responsibility, human-centered values, and agency when helping students to examine power dynamics and biases inherent in AI systems (Benjamin, 2022; Picasso et al., 2024). Similarly, across literature, critical AI literacy calls for "critical evaluation mindsets" which can detect hallucinations and inaccuracies in generative AI outputs (Bender et al., 2021).

In review, understanding generative AI's limitations, questioning who benefits or who is harmed by AI, actively evaluating AI generated content, corroborating AI output with verifiable sources are fundamental competencies of critical AI literacy. To date, while most research involving critical AI literacy has been conducted in higher education (Tacelosky et al., 2025), critical AI literacy competencies can be transferable to secondary contexts, particularly when they are situated in disciplinary tasks. Bridging critical AI literacy with disciplinary literacy, Zhang and Botelho's (2025) study exemplified how middle school students can be trained to critically evaluate the suggestions of generative AI to refine chatbots in science and math. In order to develop skill, students used generative AI as a collaborative scaffold to design and train their chatbots within an iterative process through a platform called AI Made by You (AMBY). Generative AI provided students with suggestions at each step of the design process, and students then evaluated suggestions for scientific accuracy and compatibility with the voice of their chatbot. Through vetting appropriate vs inappropriate suggestions students both improved the quality of their chatbot and increased their ability to discard hallucinations, evaluate AI output, and deepen their disciplinary knowledge (Zhang & Botelho, 2025).

AI Chatbots can offer students low stakes environments and opportunities to make their historical reasoning and AI literacy practices visible (Tacelosky et al., 2025). When carefully designed, AI chatbots can extend this scaffolding by offering interactive practice environments where students rehearse sourcing, contextualization, corroboration, and evidence-based argumentation, not as substitutes for teacher instruction, but as supplements that make disciplinary thinking more accessible. Notably, the probability of hallucinations, bias, and lack of contextual awareness in AI output can become powerful learning opportunities. When students are taught to understand and identify biases or inaccuracies in AI output, they enact both disciplinary habits of mind such as questioning sources, verifying information through corroboration, and seeking multiple perspectives (Reisman, 2012) as well as critical AI literacy (Bender et al., 2021; Long & Magerko, 2020). When used intentionally and strategically, AI chatbots may offer educators opportunities to examine students' historical reasoning and AI literacy practices.

LITERATURE REVIEW

Artificial intelligence (AI) has increasingly entered educational spaces as an interactive tool capable of supporting historical thinking, disciplinary literacy practices, and language development. AI chatbots, in particular, are conversational agents designed to simulate human dialogue through natural language processing. Early chatbot systems, such as ELIZA (Weizenbaum, 1966), relied on rule-based scripts and keyword recognition, producing interactions that were limited in scope and depth. Contemporary chatbot systems, however, draw on machine learning and large language models, enabling more adaptive and context-sensitive responses (Adamopoulou & Moussiades, 2020).

Within educational research, chatbots have been explored for their potential to increase student engagement, provide immediate feedback, and support personalized learning experiences (Følstad & Brandtzæg, 2017; Winkler & Söllner, 2018). These affordances are particularly relevant in multilingual classrooms, where students must simultaneously navigate disciplinary concepts and language demands. Recent scholarship suggests that large language models can function as scaffolding tools by generating explanations, modeling academic language, and supporting inquiry-based tasks (Kasneji et al., 2023). At the same time, scholars caution that AI-generated content may include inaccuracies, oversimplifications, or embedded biases, underscoring the need for guided and critical use in educational contexts (Dwivedi et al., 2023).

In social studies education, disciplinary literacy emphasizes practices such as sourcing, contextualization, corroboration, and reasoning about historical causality and legacy. Conversational agents have the potential to support these practices when they are intentionally designed to foreground inquiry rather than factual recall. Pérez-Marín and Pascual-Nieto (2011), for example, argue that conversational agents can model expert reasoning and guide learners through inquiry processes. Applied to social studies, this suggests that chatbots can be positioned not as authoritative sources of historical truth, but as tools through which students interrogate perspective, context, and evidence.

Despite growing interest in AI in education, research examining the use of AI chatbots in multilingual social studies classrooms remains limited. While some research emphasizes AI's limitations (Bali, 2024) and ethical implications (Feng, 2024), existing studies about chatbots tend to position them as interactive or simulation-based tools (Washburn & McCutchen, 2024), with fewer studies attending to how multilingual learners engage critically with AI-generated historical representations. This study examines student engagement during a Spanish-language chatbot assignment in which multilingual learners interacted with a historical figure as part of a social studies inquiry task. The assignment was designed to foreground historical thinking practices—such as questioning sources, evaluating context, and considering historical legacy—while creating opportunities for students to engage critically and reflectively with AI-generated content. A detailed description of the instructional sequence and student tasks is provided in a later section of the article.

METHODS

This study employed an interpretive case study methodology grounded in Stake's (1995, 2005) constructivist approach to qualitative inquiry. Stake's (2005) case study tradition prioritizes understanding how participants make meaning within a particular context rather than generating causal explanations or generalizable findings. This orientation aligned with the purpose of the study, which sought to examine how multilingual middle school students engaged in disciplinary literacy and historical thinking while interacting with an AI chatbot embedded within a classroom inquiry task.

The case was defined as a single, bounded instructional sequence—the *Desafío Chatbot* (Chatbot Challenge) assignment—implemented within two One-Way Dual Language (OWDL) social studies classrooms at the sixth- and seventh-grade levels in one middle school. The case was bounded by time (one instructional cycle), setting (one school context), participants (multilingual middle school students), and activity (the design, programming, interaction with, and reflection on an AI-generated historical chatbot). Although the assignment was implemented in two grade levels, it functioned as a unified case because the instructional design, pedagogical goals, and analytic focus remained consistent across contexts.

Data sources consisted of classroom artifacts generated as part of regular instruction, including students' chatbot planning templates, written historical questions, chatbot interaction records, and reflective responses, as well as teacher field observations. These artifacts were treated as interpretive texts that made students' disciplinary reasoning visible. Rather than serving as measures of achievement, the artifacts provided insight into how students approached historical inquiry, evaluated AI-generated information, and navigated issues of accuracy, bias, and perspective.

Analysis followed an interpretive approach consistent with Stake's (2005) view that understanding develops through close engagement with the case and iterative refinement of emerging interpretations. Artifacts were read holistically and comparatively to identify recurring patterns in students' questioning, sourcing, contextualization, corroboration, and reflection. Attention was given to variations across grade levels, recognizing that the assignment served different instructional purposes depending on students' prior knowledge and curricular positioning. Interpretive claims were grounded in repeated observations across artifacts rather than isolated examples.

Trustworthiness was supported by examining multiple types of classroom artifacts, including students' chatbot planning materials, historical questions, chatbot interactions, and written reflections. Rather than relying on a single source of evidence, interpretations were informed by patterns that appeared across these artifacts. Interpretive claims were discussed and refined collaboratively by the authors, with attention to how the instructional design, classroom context, and students' linguistic backgrounds shaped their engagement. In keeping with Stake's (2005) approach, the analysis does not aim to produce generalizable claims about AI chatbots, but instead offers a contextualized account of how one instructional design supported disciplinary literacy and critical AI engagement in a multilingual social studies setting.

Researcher Positionality

The primary author of this study is a middle school educator who works with multilingual learners and designed and implemented the *Desafío Chatbot* instructional sequence as part of his practitioner inquiry into disciplinary literacy and AI-mediated learning. Author 1 designed and implemented the instructional sequence in two middle school social studies classrooms at the sixth and seventh grade levels, where he provided direct instruction, facilitated student

engagement with AI-mediated historical inquiry, and collected student-generated artifacts used for analysis.

Authors 2, 3, and 4 are literacy professors at a university nearby working within an initiative called project L.I.S.T (Literacy, Inquiry, Social Justice, and Technology). While collaborating with the partner school district, these professors have codesigned districtwide PD workshops and materials with Author 1. Given the overlapping interests and intersecting disciplinary areas, Author 1 invited the L.I.S.T team to coauthor this project. Author 2 is a former high school ELA teacher, and identifies as an Afro-latino scholar invested in developing capacity for social justice pedagogy and critical AI literacy within the partner district. Author 3 is a social studies methods professor at a midwestern public university. His desire for innovative inquiry-based practices in social studies classrooms drives his teaching and research for engaged citizenship. Author 4 is a former science educator who worked with multilingual and multicultural youth. As a bilingual Latina, her identity drives her work to create expansive learning spaces in the sciences and beyond where multilingual and multicultural youth feel heard, seen, and centered. Authors 1 and 2 performed the primary analysis, and all authors reflected on how their educational roles, identities, and experiences influenced the analysis.

Context

The study took place in a middle school serving approximately 570 students in grades 6–8 within a linguistically and culturally diverse community in northern Illinois. According to publicly available school report card data, approximately 74% of the school's student population qualifies for free or reduced-price lunch (ISBE, 2025). All participants in this study were enrolled in the school's One-Way Dual Language (OWDL) program, in which Spanish is the primary language of instruction across content areas. Students in this program typically enter with strong oral proficiency in Spanish and varying levels of biliteracy development in Spanish and English, as measured by district and state language assessments. The OWDL program is designed to leverage students' home language as an asset for disciplinary learning while supporting long-term academic growth.

The *Desafío Chatbot* instructional sequence was implemented in two OWDL social studies classrooms at the sixth- and seventh-grade levels, involving 41 students identified as multilingual learners (MLLs). Students demonstrated a range of English language proficiency levels, and had prior experience with inquiry-based social studies instruction, though most had limited or no prior exposure to AI-mediated learning tools. While the same instructional design was used across both classrooms, the lesson served different pedagogical purposes aligned with each grade's curricular trajectory:

- In sixth grade, the *Desafío Chatbot* activity functioned as a pre-unit instructional entry point. At the time of implementation, students were preparing for a social studies unit focused on the Code of Hammurabi and ancient civilizations. The chatbot interaction was intentionally positioned as an instructional activator, designed to introduce students to the historical figure

of Hammurabi, prompt curiosity, and surface initial questions related to governance, law, and historical context prior to formal instruction.

- In seventh grade, the *Desafío Chatbot* activity was implemented at the conclusion of a unit on the founding of the United States. Students had already studied key historical figures, events, and documents related to U.S. independence. In this context, the chatbot interaction served as a culminating inquiry experience, providing students with opportunities to synthesize prior learning, apply historical thinking skills, and reflect on historical perspective, legacy, and interpretation through simulated dialogue.

Across both grade levels, instruction was delivered in Spanish, enabling students to engage in complex historical inquiry through their primary language of instruction. This linguistic context supported students' participation in disciplinary practices such as questioning, contextualization, sourcing, corroboration, and reflection, while maintaining coherence with the OWDL program's emphasis on leveraging multilingual resources to support content-area learning.

Desafío Chatbot

The *Desafío Chatbot* assignment was designed as a structured inquiry task in which students created and interacted with an AI chatbot representing a historical figure. The activity positioned students not only as learners of historical content, but also as novice historians and emerging AI users who were required to make intentional decisions about historical representation, accuracy, and perspective. In sixth grade, students focused on Hammurabi in connection with an upcoming unit on ancient civilizations and early systems of law. In seventh grade, students selected from a set of figures previously studied during a unit on the founding of the United States, including individuals such as Thomas Jefferson and George Washington. Limiting the selection of historical figures ensured alignment with grade-level curriculum while allowing for variation in student inquiry and interpretation.

Students were explicitly taught that generating an AI chatbot that accurately represented a historical figure required intentional, precise, and critically informed programming. Rather than viewing the chatbot as an autonomous or neutral source of historical knowledge, students learned that AI-generated responses reflect the information selected, the specificity of prompt language, and the perspectives embedded in those inputs, including the potential for bias or oversimplification. To support this process, students conducted research on their historical figure by either building on previously studied classroom texts and resources (seventh grade) or consulting new instructional materials and reference sources introduced for the activity (sixth grade). Emphasis was placed on selecting accurate information, using precise language, and considering how omissions, framing choices, or vague inputs could result in biased or historically misleading responses. This framing positioned historical research and disciplinary judgment as prerequisites to AI use and reinforced the idea that AI systems reproduce, rather than independently generate, historical understanding.

Assignment Sequence

The students engaged in the *desafío chatbot* assignment in the following sequence, demonstrated with figures of the activity.

Step 1: Planning the Chatbot

Students began by completing a structured planning template in which they documented biographical information, historical context, beliefs, accomplishments, and legacy of their selected figure (see Figure 1). This planning phase functioned as both historical research and AI preparation, requiring students to synthesize information rather than copy it verbatim.

Insert Figure 1

Step 2: Developing Historical Questions

Next, students developed their own historical questions. This stage was carefully scaffolded to move students beyond factual recall toward analytical and open-ended inquiry. Each student composed at least five questions designed to address core historical thinking practices, including sourcing, contextualization, corroboration, chronology and causality, and historical legacy (see Figure 2).

Insert Figure 2

Step 3: Creating the Chatbot

Prior to student use, the MagicSchool platform was configured by the teacher, including uploading the assignment prompts, specifying learning goals, and structuring the chatbot parameters to ensure alignment with instructional intentions and historical inquiry objectives. Students then entered their completed planning information into the MagicSchool platform, which generated a chatbot simulating the selected historical figure (see Figures 3 and 4). MagicSchool provided a structured, classroom-oriented interface that allowed students to input historically grounded information without requiring technical coding. Instruction emphasized that the chatbot's responses reflected the quality and framing of the information provided by students.

Insert Figure 3 & 4

Step 4: Interacting with the Chatbot

In the fourth step, students engaged in an interactive interview with their chatbot. Students posed at least three of their planned questions and recorded the responses (see Figure 4). After each exchange, students evaluated the extent to which the answers aligned with historical evidence and considered what sources might confirm or contradict the information. This process encouraged students to adopt a critical stance toward AI-generated content similar to the stance historians take when evaluating sources.

Insert Figure 5

Step 5: Reflection

The final step involved a written reflection in Spanish. Students explained what they learned about their historical figure, what they noticed about the chatbot's responses, and what the activity revealed about the work of historians and the affordances and limitations of artificial

intelligence in historical inquiry (see Figure 5). These reflections demonstrated both content learning and metacognitive awareness of the inquiry process. By structuring the task in this way, the *Desafío Chatbot* assignment integrated disciplinary literacy in history with digital and AI literacy, enabling students to practice analyzing sources, sequencing events, and evaluating reliability while engaging in their strongest language of learning.

Insert Figure 6

RESULTS AND DISCUSSION

This case helps make visible how multilingual students demonstrated disciplinary historical thinking through the artifacts they produced while working on the *Desafío Chatbot* assignment. Rather than evaluating the activity in terms of effectiveness, improvement, or impact, this discussion focuses on how students took up the task and how historical reasoning became visible in their planning sheets, questions, and reflections. Across these artifacts, students approached history as something to be interpreted—entangled with power, morality, context, and consequence—rather than as a set of neutral facts to be recalled. One student's reflection captures this interpretive stance:

"Aprendí que él [Hammurabi] fue muy duro con sus leyes."

"I learned that he [Hammurabi] was very harsh with his laws."

Insert Figure 7

Although concise, this statement reflects an important disciplinary move: the student evaluates a historical figure's actions and considers their implications. The sections that follow interpret what such moments suggest about disciplinary literacy and historical thinking in multilingual social studies classrooms, the role of AI chatbots as mediating tools rather than authorities, instructional considerations for teaching, and the interpretive boundaries of this case.

I. Humanization and Engagement of Historical Figures

One of the most consistent patterns across student artifacts was increased engagement with historical figures as people rather than as distant textbook subjects. The *Desafío Chatbot* required students to research a historical figure and make decisions about who that person was—what they were like, what they believed in, and why they acted the way they did. In order to program the chatbot, students could not rely solely on dates or titles; they had to think about personality, beliefs, and motivations.

This shift is visible in how students described historical figures. For example, when programming their chatbot for Hammurabi, students characterized him using evaluative language rather than neutral description:

"Tienes una personalidad desalmada, valiente y determinada."

"You have a ruthless, brave, and determined personality."

"Creías en el poder, tener orden y leyes justas."

"You believed in power, having order, and just laws."

Insert Figure 8

Even though their language is simple, these statements show students moving beyond memorization toward interpretation. Importantly, students had not been explicitly taught these descriptions in advance. In order to create an accurate chatbot, they were required to research the sources provided during the activity, identify relevant information, and make decisions about how Hammurabi should be represented. Rather than listing Hammurabi only as a king or lawmaker, students considered what kind of leader he was and how his beliefs and actions shaped his society.

Similar patterns appear in seventh grade when students worked with U.S. historical figures. One student summarized George Washington's legacy not through isolated accomplishments, but through his influence on others:

"Tu legado inspiró a las personas a luchar por su patria."
"Your legacy inspired people to fight for their homeland."

Insert Figure 9

Another student framed Thomas Jefferson through his beliefs rather than biography:

"Creías en los ideales del republicanismo y la democracia."
"You believed in the ideals of republicanism and democracy."

Across grade levels, students consistently used accessible language to express ideas about leadership, belief, and influence. These descriptions suggest that the chatbot assignment supported students in thinking about historical figures as human actors with values and worldviews, rather than as flat, textbook, historical figures.

Student reflections further highlight how the conversational format increased engagement and made historical inquiry feel more personal. One student wrote:

"Me gustó hablar con Hammurabi aunque no sea él en realidad... como si de verdad estuviera hablando con él."
"I liked talking to Hammurabi even though it's not really him... as if I were actually talking to him."

Insert Figure 10

This reflection illustrates how interacting with the chatbot encouraged students to engage more deeply with the historical figure. Although students understood that the chatbot was not the real person, the experience felt interactive and meaningful. For students who often struggle with traditional historical texts, this sense of social connection appeared to support sustained attention and willingness to research. At the same time, student engagement was uneven; while some reflections demonstrated interpretive reasoning, others remained brief or primarily affective, reflecting natural variation in how students took up the task.

Importantly, the purpose of the assignment was not to treat the chatbot as an authority, but to use it as a tool that motivated students to investigate and think more carefully about the people they were studying. In this case, engagement functioned as an entry point that helped

students move beyond memorizing surface-level facts toward more thoughtful consideration of historical figures and their actions.

II. Historical Inquiry and Questioning

As students interacted with their chatbots, their questions revealed how they were beginning to think about history as something shaped by circumstances rather than isolated actions. Many students moved beyond asking for simple facts and instead asked questions that focused on why historical figures acted as they did and what was happening around them at the time.

This was especially clear in questions that centered on context. For example, following the question models provided in the assignment, several students asked Hammurabi:

“¿Qué estaba pasando en Babilonia que te inspiró a hacer el código de leyes?”

“What was happening in Babylon that inspired you to create the code of laws?”

Insert Figure 11

Others posed similar questions to U.S. historical figures:

“¿Qué estaba pasando en tu país que te motivó [a ser presidente]?”

“What was happening in your country that motivated you [to be president]?”

“Que paso en tu vida que te hizo presidente?”

“What happened in your life that made you become president?”

Insert Figure 12

These questions show students beginning to understand that laws, leadership, and political decisions do not happen in a vacuum. Notably, their questions position historical action as a response to conditions, experiences, and challenges.

Students also asked questions that acknowledged disagreement, historical tension, and criticism, such as:

“Mucha gente no está de acuerdo con su código de leyes, ¿qué les contestaría?”

“Many people do not agree with your code of laws; what would you say to them?”

Through questions like these, students recognized that historical figures were not universally accepted and that decisions were contested. Students did not attempt to resolve these tensions; instead, they identified and named them, an appropriate stance for historical inquiry at this developmental stage.

As students read the chatbot’s responses, many reflected on how those explanations helped them think differently about the past. Several students noted that learning what came before a historical decision helped them understand why it was made. For example, one student wrote:

“Esto me hizo pensar en como fue antes las leyes de Hammurabi.”

“This made me think about what it was like before Hammurabi’s laws.”

Another student reflected on changes over time, noting:

“Me hizo pensar en lo que pasaba antes y después [de la independencia de los E.E.U.U.]”

“It made me think about what happened before and after [U.S. independence].”

These reflections suggest that students were beginning to reason about historical causality and chronology. Rather than focusing only on individual events or decisions, students considered conditions leading up to historical actions and the changes that followed.

It is important to note that these responses emerged within a highly scaffolded instructional context. As multilingual learners, students were provided with explicit models, sentence starters, and example questions aligned to historical thinking skills. This level of support was intentional and appropriate, ensuring that students could engage in disciplinary reasoning without being limited by language demands. Because student responses were brief and shaped by these supports, interpretations focus on patterns across artifacts rather than depth within individual responses.

III. Emerging AI Literacy Practices

Across interactions and reflections, the chatbot was positioned unevenly. At times students interacted with the chatbot as an authoritative source, and at others, as an interpreter whose responses could be questioned or reconsidered. In many instances, students expressed uncertainty when what they had previously learned in class was inconsistent with chatbot responses, suggesting early forms of evaluative reasoning. For example, one student communicated skepticism when noticing a discrepancy stating,

“Profe, ¿no se supone que George Washington hablaba puro inglés? ¿Cómo sé que sí es él de verdad... sí hablaba español o qué?”

“Teacher, wasn’t George Washington supposed to speak only English? How do I know it’s really him... did he actually speak Spanish or what?”

In another instance, a student questioned a historical reference by saying,

“Maestro, de Hammurabi no me acuerdo haber visto eso en clase... ¿sí hizo eso de verdad o qué onda con eso?”

“Teacher, I don’t remember learning that about Hammurabi in class... did he really do that or what’s going on with that?”

Such instances illustrate emerging AI literacy practices where students regard AI-generated output as claims they need to verify. While their reasoning was not always fully developed, it reveals the potential for AI-supported activities to create opportunities for students to engage in critical evaluation alongside core historical thinking practices like sourcing, corroboration, and contextualization.

Interpretive Boundaries and Limitations

This case represents a single instructional sequence within one school context and does not claim growth, transfer, or long-term impact. Student artifacts reflect moment-in-time thinking shaped by instructional design, language supports, and the affordances and constraints of the chatbot platform. Additionally, while students demonstrated emerging disciplinary reasoning, their engagement and depth of interpretation varied. These limitations are not shortcomings of the case, but features of classroom-based inquiry with multilingual learners. Within these boundaries, the *Desafío Chatbot* made students’ historical thinking visible and provided a

structured space for engagement, questioning, and reflection. The findings offer a contextualized account of how AI-supported instructional design can create opportunities for disciplinary (historical) literacy in multilingual social studies classrooms, without positioning AI as an authority or claiming outcomes beyond the case itself.

Implications

We invite our readers to consider grade-appropriate ways their learners can engage in contextualization, corroboration, close reading, and sourcing through chatbots. The patterns observed across the two classroom implementations suggest that AI-mediated interactions can serve as a meaningful supplement to disciplinary literacy instruction by creating carefully designed, structured opportunities for inquiry when integrated into broader units of study. Several implications emerge for educators seeking to adapt similar practices in their own contexts.

First, chatbots can extend structured inquiry routines, but their usefulness in class contexts depends on the degree of scaffolding provided. For younger or less experienced learners, multimodal text sets and scaffolded reading (Waymouth & Hinchman, 2025), labeled organizers and question templates (Buehl, 2023; Reisman, 2012), and modeled exemplars help anchor their work and prevent superficial questioning (Monte-Sano et al., 2017). In classrooms where students possessed stronger background knowledge, the chatbot became a space to test, refine, and deepen existing interpretations. Teachers may therefore consider positioning chatbots differently across a unit—either as anticipatory inquiry tools or as synthesizing platforms near the end of instruction.

Second, the assignment highlights the value of deliberate instruction in sourcing and corroboration when working with AI-generated historical responses. Our theoretical framework discussed AI bias and hallucinations as powerful learning opportunities (Li, 2023), and although students in this study displayed emerging abilities to critique omissions, contradictions, or overly idealized portrayals, their reflections varied widely. This suggests that critical AI literacy should be taught alongside historical literacy practices in social studies (Tacelosky et al., 2024; Yetisensoy & Rapoport, 2023; Yetisensoy & Karaduman, 2024). Students benefit from explicit opportunities to compare, corroborate, and source AI responses with classroom texts and to view AI outputs as claims requiring verification (Lintner, 2024). Positioning AI as a claimant rather than authority may help students adopt a more agentive and constructivist role while engaging with disciplinary inquiry (Bender et al., 2021). Embedding these comparative routines in whole-class discussions could strengthen students' ability to read AI-generated content with greater skepticism and nuance.

Third, the task underscores that language choice matters for multilingual participation. Interactions conducted in Spanish provided students with immediate access to disciplinary thinking without the added barrier of navigating complex English texts. MLLs used the chatbot to rehearse academic vocabulary, practice interpretive questions, and engage with historical voices in their strongest language. Because scholarship involving dual-language and disciplinary

literacy contexts emphasize how comprehension and confidence improves when students approach learning in their primary language (Hickey & Lewis, 2015), authors recommend utilizing chatbots in ways that afford students choices in the ways they can employ language to engage in content-area inquiry.

Fourth, authors recommend considering how the continuous use of a chatbot across a unit may expand opportunities for inquiry. In the sixth-grade classroom, the decision to return to the chatbot as new questions emerged allowed students to treat it as a dynamic reference point rather than a one-time activity. This is commensurate with Yetisensoy and Rapoport's (2023) findings that technology prepares students for active citizenship, and Dwomoh's (2025) study examining how technology impacts inquiry in civics education and the interplay of *technology, inquiry, and citizenship education* in middle-grade classrooms. Dwomoh's 3's (2025) findings revealed that "these three overarching concepts [i.e., technology, inquiry, and citizenship education] are progressive and interrelated. In essence, technology enhances inquiry, and inquiry reinforces effective citizenship education" (p. 51). Moreover, the National Council for the Social Studies (NCSS) recommends inquiry as the best practice for social studies teaching and learning (2013), as it fosters student engagement, civic awareness, focused learning, and critical and creative thinking (Dwomoh, 2025). Recursive inquiry can support the development of more precise questioning and more historically grounded interpretations (Conrad et al., 2024). However, inquiry cannot be effectively implemented in middle grade classrooms without the assistance of technology (Dwomoh, 2025), such as AI chatbots.

Finally, the teacher's role remains central. Chatbots do not replace disciplinary reading, teacher modeling, or rich classroom talk. Instead, they create additional entry points for exploration when accompanied by thoughtful planning and reflection. Teachers may consider how chatbot interactions can complement primary sources, anchor investigations, or serve as low-stakes rehearsal spaces for emerging disciplinary skills. Strategic facilitation is needed to help students navigate limitations in AI-generated responses, particularly when those responses simplify conflict, obscure historical harm, or present overly harmonious narratives.

In sum, chatbots can serve as accessible tools for supporting disciplinary literacy in multilingual social studies classrooms, provided they are embedded in intentional pedagogical design. By integrating guided inquiry, critical evaluation, and multilingual scaffolds, educators can leverage these tools to expand opportunities for students to think, question, and reason like historians.

CONCLUSION

This study explores the potential AI chatbots have to support multilingual learners as they navigate language, content, and historical thinking practices within social studies classrooms. MLLs engaged with a Spanish Chatbot assignment where they researched a historical figure, interacted with a simulated version of that figure, and analyzed their interactions. Students practiced disciplinary literacy skills like sourcing, contextualization, and evaluation of AI interactions. Although all students work on skills like planning, questioning, interaction, and

reflection, to varying degrees, findings revealed how students' prior knowledge impacted their ability to evaluate AI output. Implications suggest that although LLMS can create opportunities for students to engage with disciplinary knowledge and practices, direct support is needed to help teachers model disciplinary thinking with critical AI literacy.

REFERENCES

- Achugar, M., & Carpenter, B. D. (2012). Developing disciplinary literacy in a multilingual history classroom. *Linguistics and education*, 23(3), 262-276.
<https://doi.org/https://doi.org/10.1016/j.linged.2012.05.003>
- Achugar, M., Schleppegrell, M., & Oteíza, T. (2007). Engaging teachers in language analysis: A functional linguistics approach to reflective literacy. *English Teaching: Practice and Critique*, 6.
- Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *Artificial Intelligence Applications and Innovations* (pp. 373–383). Springer. https://doi.org/10.1007/978-3-030-49186-4_31.
- Baer, A. (2025). Investigating the 'Feeling Rules' of Generative AI and Imagining Alternative Futures. *In the Library with the Lead Pipe*.
- Bain, R. B. (2006). Rounding up unusual suspects: Facing the authority hidden in the history classroom. *Teachers College Record*, 108(10), 2080–2114.
<https://doi.org/10.1111/j.1467-9620.2006.00775.x>
- Bail, C. A. (2024). Can generative AI improve social science? *Proceedings of the National Academy of Sciences*, 121(21). <https://doi.org/10.1073/pnas.2314021121>
- Barton, K. C., & Levstik, L. S. (2004). *Teaching history for the common good*. Lawrence Erlbaum.
- Baumann, J. F., & Duffy, A. M. (2001). Teacher-Researcher Methodology: Themes, Variations, and Possibilities. *The Reading Teacher*, 54(6), 608-615.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).
- Benjamin, R. (2022). *Viral justice: How we grow the world we want*. Princeton University Press.
- Bickford, J., Bickford, M., & Dwomoh, R. (2020). Middle-level students' responses to a guided inquiry of the weeping time. *History Teacher*, 53(2), 255-318.
http://societyforhistoryeducation.org/pdfs/F20_Bickford_Bickford_and_Dwomoh.pdf
- Buehl, D. (2023). *Developing readers in the academic disciplines*. Routledge.
- Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. London: Continuum.
- Conrad, K., & Kamperman, S. (2025). Building Critical AI Literacy: An Approach to Generative AI. *Thresholds in Education*, 48(2), 142-158.
- Conrad, J., Gallagher, J. L., & Chan, W. (2024). Getting critical with compelling questions: Shifts in elementary teacher candidates' curriculum planning from inquiry to critical inquiry.

Theory & Research in Social Education, 52(4), 532–574.

<https://doi.org/10.1080/00933104.2024.2340496>

Crompton, H., & Burke, D. (2024). The educational affordances and challenges of ChatGPT: State of the field. *TechTrends*, 68(2), 380-392.

Crotty, M. A. (2003). *The foundations of social research : meaning and perspective in the research process*: London ; Thousand Oaks, Calif. : Sage Publications, 2003. 1998.

Donley, K. (2024). Teaching with ChatGPT as a linguistically responsive tool for multilingual learners. *Technology in Language Teaching & Learning*, 6(3), 1719-1719.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642.
<https://doi.org/10.1016/j.ijinfomgt.2023.102642>.

Dwomoh, R. (2022). Social studies in non-western contexts: The development, appraisal, and implications of Ghana's social studies curricula. *The Councilor: A National Journal of the Social Studies*, 83(1), 2.

Dwomoh, R. K. (2025). The impact of technology on inquiry in citizenship education. In: Zajda, J., Rapoport, A. (Eds.), *Discourses of globalisation and citizenship education. Globalisation, comparative education and policy research* (vol. 46). Springer

Fairclough, N. (1999). Global Capitalism and Critical Awareness of Language. *Language awareness*, 8(2), 71-83. <https://doi.org/10.1080/09658419908667119>

Fang, Z., & Schleppegrell, M. J. (2010). Disciplinary literacies across content areas: Supporting secondary reading through functional language analysis. *Journal of adolescent & adult literacy*, 53(7), 587-597.

Feng, G. C. (2024). Best practices for responsibly using AI tools in social sciences research. *Cogent Social Sciences*, 10(1), 2420484.
<https://doi.org/10.1080/23311886.2024.2420484>

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), 38–42. <https://doi.org/10.1145/3085558>

Freire, P. (1970). The adult literacy process as cultural action for freedom. *Harvard educational review*, 40(2), 205-225.

Gebhard, M. (2010). Teacher education in changing times: A systemic functional linguistics (SFL) perspective. *Tesol quarterly*, 44(4), 797-803.

Hickey, P. J., & Lewis, T. (2015). To Win the Game, Know the Rules and Legitimize the Players: Disciplinary Literacy and Multilingual Learners. *Language and Literacy Spectrum*, 25, 18-28.

Illinois State Board of Education. (2025). *Illinois Report Card: 2024–2025 school year*.
<https://irc.isbe.net/Default.aspx>

- Junqueira, E. (2025). A systematic review of literacies and generative artificial intelligence in education: Evolving concepts, key themes, and directions for future research. *Ubiquity Proceedings*, 6(1).
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Grund, S., Kahnwald, N., Krusche, S., Larissa, P., Neuhaus, F., Rittberger, M., & Schmidt, A. (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>
- Kenna, J. L., Russell III, W. B., & Bittman, B. (2018). How secondary social studies teachers define literacy and implement literacy teaching strategies: A qualitative research study. *History Education Research Journal*, 15(2).
- Langan, E., & Lawrence, S. A. (2021). Which came first: Literacy or social studies? How primary sources can bridge the divide. *Excelsior: Leadership in Teaching and Learning*, 13(3), 182-197.
- Lankshear, C., & Knobel, M. (2003). *New literacies: Changing knowledge and classroom learning*. Open University Press.
- Li, Z. (2023). The dark side of chatgpt: Legal and ethical challenges from stochastic parrots and hallucination. *arXiv preprint arXiv:2304.14347*.
- Lin, C. C., Huang, A. Y., & Yang, S. J. (2023). A review of ai-driven conversational chatbots implementation methodologies and challenges (1999–2022). *Sustainability*, 15(5), 4012.
- Lintner T. (2024). A systematic review of AI literacy scales. *NPJ science of learning*, 9(1), 50. <https://doi.org/10.1038/s41539-024-00264-4>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI chatbots for content and language integrated learning. *Applied Sciences*, 12(7), 3239.
- Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction*, 1(1), 12–21. <https://doi.org/10.1016/j.lcsi.2012.03.001>
- Monte-Sano, C. (2010). Disciplinary literacy in history: An exploration of the historical nature of adolescents' writing. *Journal of the Learning Sciences*, 19(4), 539–568. <https://doi.org/10.1080/10508406.2010.481014>
- Monte-Sano, C., De La Paz, S., Felton, M., Piantedosi, K. W., Yee, L. S., & Carey, R. L. (2017). Learning to teach disciplinary literacy across diverse eighth-grade history classrooms within a district-university partnership. *Teacher Education Quarterly*, 44(4), 98-124.

- Moorhouse, B. L., & Kohnke, L. (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators. *System*, 122, 103290. <https://doi.org/10.1016/j.system.2024.103290>.
- National Council for the Social Studies. (n.d.). *National Council for the Social Studies*. <https://www.socialstudies.org>
- Nokes, J. D. (2017). Exploring patterns of historical thinking through eighth-grade students' argumentative writing. *Journal of writing research*, 8(3), 437-467.
- Nokes, J. D. (2011). Recognizing and addressing the barriers to adolescents' "reading like historians." *History Teacher*, 44(3), 379–404. <https://www.jstor.org/stable/41304079>
- Pérez-Marín, D., & Pascual-Nieto, I. (2011). *Conversational agents and natural language interaction: Techniques and effective practices*. IGI Global. <https://doi.org/10.4018/978-1-60960-617-6>
- Picasso, F., Atenas, J., Havemann, L., & Serbati, A. (2024). Advancing critical data and AI literacies through authentic and real-world assessment design using a data justice approach. *Open Praxis*, 16(3), 291-310.
- Quan, T. (2021). Critical approaches to Spanish language teacher education: Challenging raciolinguistic ideologies and fostering critical language awareness. *Hispania*, 104(3), 447-459.
- Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, 30(1), 86–112. <https://doi.org/10.1080/07370008.2011.634081>
- Schleppegrell, M.J. (2004). *The Language of Schooling: A Functional Linguistics Perspective* (1st ed.). Routledge. <https://doi-org.auth.lib.niu.edu/10.4324/9781410610317>
- Shanahan, C. (2014). Disciplinary comprehension. In *Handbook of research on reading comprehension* (pp. 264-284). Routledge.
- Short, D. J., & Fitzsimmons-Doolan, S. (2007). Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners.
- Soininen, S. (2022). Teaching historical thinking in practice: A study of US history teachers' views on using primary sources in AP and IB history lessons. *History Education Research Journal*, 19(1), Article 5. <https://doi.org/10.14324/HERJ.19.1.05>
- Stake, R. E. (1995). *The art of case study research*. SAGE Publications.
- Stake, R. E. (2005). Qualitative Case Studies. In N. K. Denzin, & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443–466). Sage Publications Ltd.
- Tacelosky, K., Sue Kasun, G., Shapiro, B. R., Liao, Y. C., & Harris, K. (2025). Exploring critical AI literacy in language education: A case study. *Foreign Language Annals*.
- Tanksley, T. C. (2024). "We're changing the system with this one": Black students using critical race algorithmic literacies to subvert and survive AI-mediated racism in school. *English Teaching: Practice & Critique*, 23(1), 36-56.

- Trombino, D. L. (2010). *The experiences of secondary social studies teachers with historical thinking skills*. Old Dominion University.
- Waymouth, H., & Hinchman, K. A. (2025). Fostering Equitable Disciplinary Literacy Practices. *Education Sciences*, 15(2), 225.
- Washburn, J., & McCutchen, J. M. (2024). AI meets AI: ChatGPT as a pedagogical tool to teach American Indian history. *Critical Humanities*, 2(2), 2.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Wineburg, S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83(1), 73–87. <https://doi.org/10.1037/0022-0663.83.1.73>
- Wineburg, S., & Reisman, A. (2015). Disciplinary literacy in history: A toolkit for digital citizenship. *Journal of Adolescent & Adult Literacy*, 58(8), 636-639.
- Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Academy of Management Annual Meeting Proceedings*, 2018(1), 15903. <https://doi.org/10.5465/AMBPP.2018.15903abstract>
- Yetisensoy, O., & Rapoport, A. (2023). Artificial intelligence literacy teaching in social studies education. *Journal of Pedagogical Research*, 7(3), 100-110.
- Yetisensoy, O., & Karaduman, H. (2024). The effect of AI-powered chatbots in social studies education. *Education and Information Technologies*, 29(13), 17035-17069.
- Zhang, S., & Botelho, A. F. (2025). Scaffolding AI Literacy Through Student-AI Collaboration in Chatbot Development. In *Proceedings of the 18th International Conference on Educational Data Mining* (pp. 659-662).

APPENDICES

Appendix A: Figures

Figure 1: Step 1 - Planning Their Chatbots

Paso 1: Planeando Tu Chatbot

1. Elige un personaje histórico y completa la información. Usa **fuentes confiables (textos, artículos, biografías utilizadas en clase)**. Recuerda que un buen historiador siempre verifica sus fuentes.
2. Asegúrate de que tu información sea precisa y refleje sus logros y personalidad. Si es necesario, siéntete libre de investigar en Google para obtener la información que no recuerdes.

Instrucciones para el ChatBot	
Incluye abajo una figura o dibujo de tu personaje.	Tu nombre es Tu lugar de nacimiento es... Tu fecha de nacimiento es... Tienes una personalidad valiente y determinada. Eres famoso por Un logro importante de tu vida fue... Creías en... Una fecha importante en tu vida o en tus acciones fue... Tu legado inspiró a las personas a ...

Figure 2: Step 2 - Planning Historical Questions

Paso 2: Planeando Tus Preguntas

Imagina que eres un **historiador** entrevistando a tu personaje. Cuando un/a historiador/a formula preguntas, no se queda en lo **superficial** ("¿qué hiciste?"), sino que **busca profundizar en el contexto, las causas y las consecuencias**. Para ayudarte, aquí tienes un **paso a paso** para crear tus preguntas:

1. Empieza con el **tipo de habilidad histórica** que quieres practicar.
 - **Sourcing (fuente):** Pregunta quién escribió, dijo o creó algo y con qué intención.
 - **Contextualización:** Pregunta qué estaba ocurriendo en el mundo en un momento específico de su vida y cómo eso ha influido sus decisiones.
 - **Corroboración:** Pregunta cómo se comparan diferentes puntos de vista o testimonios.
 - **Cronología y causalidad:** Pregunta qué pasó antes, qué pasó después y cómo se relacionan los eventos en el tiempo.
 - **Legado:** Pregunta qué impacto le gustaría que su vida tuviera en generaciones posteriores.

2. **Usa palabras guía** para empezar tus preguntas:
 - *Quién, qué, cuándo, dónde, por qué, cómo...*
 - Ejemplos:
 - "¿**Qué** estaba pasando en Estados Unidos en la década de 1960 que influyó en tus decisiones?"
 - "¿**Cómo** responderías a las críticas de quienes no apoyaban tu estrategia de no violencia?"
 - "¿**Qué** evento anterior a tu famoso discurso ayudó a prepararlo?"
3. **Verifica** que tu pregunta no pueda responderse sólo con un dato.
 - En vez de: "¿Cuándo naciste?" (dato simple)
 - Mejor: "¿Cómo influyó el contexto histórico del año en que naciste en tu vida más tarde?" (pregunta analítica).
4. **Escribe tus 5 preguntas abajo**, asegurándose de incluir al menos una de cada tipo de habilidad histórica.

Ejemplos:	
Sourcing (fuente):	Ejemplo: "Cuando escribiste tu discurso más famoso, ¿a quién querías convencer y por qué?"
Contextualización:	Ejemplo: "¿Qué estaba pasando en tu país en ese momento que te motivó a luchar por los derechos civiles?"
Corroboración:	Ejemplo: "Algunas personas no estaban de acuerdo con tus métodos de protesta pacífica. ¿Qué les responderías?"
Cronología y causalidad:	Ejemplo: "¿Qué evento importante ocurrió antes de tu discurso que ayudó a que la gente te escuchara con más atención?"

Figure 3: Step 3 - Creating Their Chatbots

Paso 3: Creando tu Chatbot

- Ahora crearás tu Chatbot.
- Ve a [MagicSchool a través de este link](#) y escribe tu nombre.
- En "Instrucciones para el Chatbot", escribe las instrucciones del Chatbot que desarrollaste en el Paso 1.

Instructions for Chatbot:

What does this chatbot do? How does it behave? What should it avoid doing? What tone should it respond with?

Figure 4: Step 4 - MagicSchool Interaction

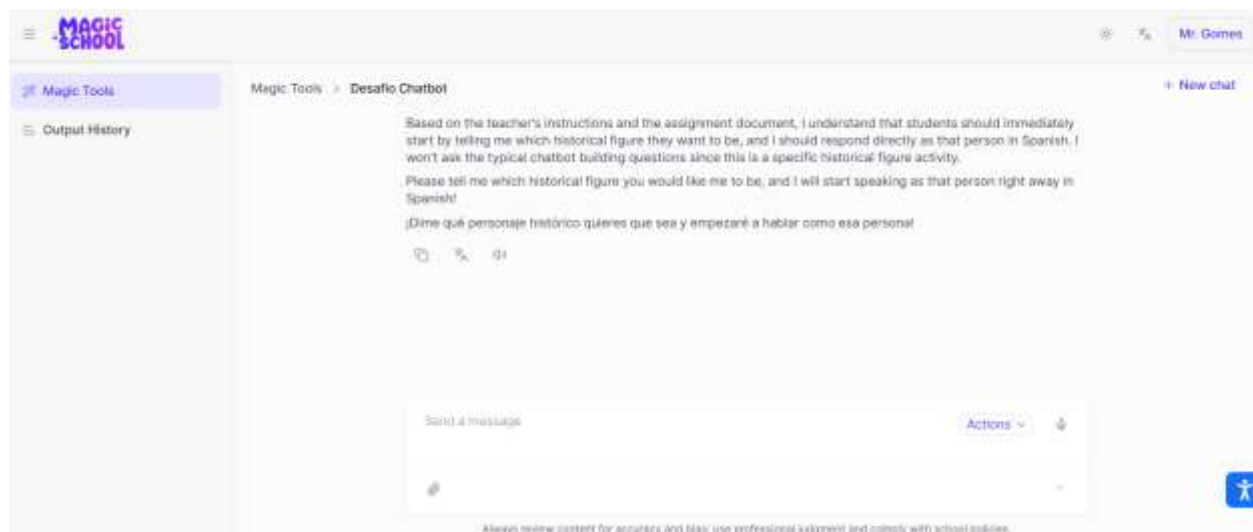


Figure 5: Step 4 - Chatbot Interaction

Paso 4: Hablando con tu Chatbot

- ¡Ahora interactuarás con tu Chatbot!
- Elige **tres preguntas** de las que desarrollaste en el Paso 2.
- Asegúrate de que tu chatbot refleje la voz, creencias y contexto histórico de tu personaje.
- Después de cada respuesta, **evalúa críticamente**:
 - ¿La respuesta **coincide** con lo que sabes de tu personaje?
 - ¿Qué **evidencia histórica** respalda o contradice la respuesta?
- Hazle estas tres preguntas a tu chatbot. Anota las respuestas que te dé el Chatbot:

Pregunta #1:

Respuesta del Chatbot:

Pregunta #2:

Respuesta del Chatbot:

Pregunta #3:

Respuesta del Chatbot:

Figure 6: Final step - Reflection

Paso 5: Reflexión

- **En mínimo tres oraciones**, reflexiona sobre lo aprendido en esta experiencia: ¿Qué descubriste sobre tu persona histórica? ¿Qué opinas de las respuestas recibidas?
- Completa la sección de Reflexión con tus aprendizajes y opiniones sobre la actividad. Sigue el ejemplo abajo:

Para mí, esta actividad fue _____, porque _____.

Al leer las respuestas de _____, aprendí que _____.

Esto me hizo pensar en _____.

Una cosa que me enseñó esta actividad sobre ser historiador/a fue _____.

Una cosa que me enseñó el chatbot sobre la IA en historia fue _____.

Figure 7: Student Work Sample #1

Al leer las respuestas de la chatbot de Hammurabi, aprendí que el fue muy duro con sus leyes.
 Esto me hizo pensar en como fue antes las leyes de Hammurabi.
 Para mi, esta actividad fue muy interesante, porque pude ver que decía Hammurabi con la chatbot.

Figure 8: Student Work Sample #2

Tu lugar de nacimiento es... Babilonia

Tu fecha de nacimiento es... 1810 BC,

Tienes una personalidad desalmada, valiente y determinada.

Eres famoso por el código de leyes.

Un logro importante de tu vida fue.. La unificación de su imperio.

Creías en... poder, tener orden y leyes justas.

Tu legado inspiró a las personas a ... desarrollar otros sistemas de leyes.

Figure 9: Student Work Sample #3

Tu lugar de nacimiento es: **Westmoreland Country, VA**

Tu fecha de nacimiento es: **February 22, 1732**

Tienes una personalidad : **Valiente y determinada.**

Eres famoso por: **Porque George Washington fue el primer presidente de los Estados Unidos.**

Un logro importante de tu vida fue: **Liderar el Ejército en la guerra de independencia.**

Creías en: **La religión.**

Una fecha importante en tu vida o en tus acciones fue... (¿qué pasó y por qué fue significativo?) **April 30, 1789 porque ese día empezó a ser el primer presidente.**

Tu legado inspiró a las personas a **luchar por su patria.**

Figure 10: Student Work Sample #4

Para mi esta actividad era divertida porque me gusto hablar con Hammurabi aunque no sea el en realidad. Al leer las respuestas de Hammurabi fueron como si de verdad estuviera hablando con el. Esto me hizo pensar en que si hablara con el me hablara igual como me hablo el robot.

*Figure 11: Student Work Sample #5***Pregunta 2 (contexto):**

¿Qué estaba pasando en Babilonia que te inspiró a hacer el código de leyes?

Pregunta 3 (corroboración):

Mucha gente no está de acuerdo con su código de leyes, ¿que les contestaría para convencerlos de que sus leyes son buenas?

*Figure 12: Student Work Sample #6***Pregunta 1 (fuente):**

¿Es verdad que te pusieron dientes de animal?

Pregunta 2 (contexto):

Que paso en tu vida que te hizo presidente?