"I, for One, Welcome Our New Computer Overlords": Using Artificial Intelligence as a Lesson Planning Resource for Social Studies

Christopher H. Clark

University of North Dakota

Cathryn van Kessel

Texas Christian University

Due to the introduction and rapid ubiquity of artificial intelligence (AI) and AI-integrated programs that can be used by students and teachers, educational scholarship evaluating the capabilities of AI is needed. This study evaluates the abilities of three prominent AI programs —ChatGPT, Microsoft's Bing, and Google's Bard — to create high school lesson plans on the subjects of Martin Luther King, Jr., the Indian Removal Act, and climate change. The authors judge the quality of the lessons' content based on scholarship in the education field and document the process of prompting the AI to produce lessons more in line with these criteria.

In a 2011 episode of the television game show *Jeopardy!*, Ken Jennings played against IBM's Watson, an artificial intelligence (AI) designed to assist in the automation of business tasks (IBM, n.d). Jennings still holds the record for the highest average correct responses as well as the longest winning streak on that show with 74 consecutive wins. Despite his human prowess at trivia, he and another high scoring jeopardy champion, Ken Rutter, lost to Watson. Realizing his inevitable defeat, Jennings wrote underneath his Final Jeopardy answer, "I for one welcome our new computer overlords" (Hiskey, 2012). In the years since that *Jeopardy* match, AI science has continued to evolve, enabling computers to process human language more fluidly and complete a myriad of tasks.

In November 2022, a company called Open AI released ChatGPT, an AI "chatbot." ChatGPT describes itself as "a language model developed by OpenAI, which uses deep learning techniques to generate human-like text based on the input it receives. It is designed to respond to questions, complete text based on the prompt, summarize long text, and perform various other language-related tasks" (OpenAI, 2023a). Soon after, Microsoft released a ChatGPT-based AI integrated into their Bing search engine, and Google released Bard. (At the time of data collection for this study, Microsoft's AI was integrated into the Bing search engine and was not branded as a separate product, but prior to publication the name was changed to Copilot as part of Microsoft's strategy to compete with ChatGPT; Warren, 2023). In addition, a myriad of AI-based platforms tailored for specific applications have been released.

Educators might see these AIs as tools for malfeasance. Many educators, for example, worry that students will use AI to cheat, as it is capable of producing answers sophisticated enough to pass exams in both law schools and business schools (Kelly, 2023). Given that AI is increasingly prevalent in tech circles and is being integrated into everyday technological tools such as search engines, educators will need to adapt to its presence.

This paper focuses on the lesson planning abilities of three widely available and free-to-use AIs and examines the extent to which social studies educators can use AIs as a tool to create lesson plans that challenge students to think critically about social studies subjects.

Using someone else's lesson plans can be hit or miss — so much depends on the creator, whom the teacher may or may not know. Online marketplaces like Teachers Pay Teachers, as well as social media platforms like Pinterest, are widely used and can contain not only inadequate but also harmful narratives about topics and issues vital to social studies, such as the Civil Rights Movement (Rodríguez et al., 2020).

AI also offers the promise of quickly generated lessons. Early explorations of ChatGPT for teaching purposes by publications such as *Education Week* have emphasized producing basic outlines to serve as a starting point for lessons, while noting that it can reduce the time burden of some of the more menial teaching tasks, such as grading and writing letters of recommendation (Mallon, 2023; Will, 2023). Yet, with the focus on time-saving affordances of the new technology, little attention has been paid to the quality of the content in lessons produced by AI and whether AI produces lessons that repeat problematic framings. Given that large language models, the basis of the three AI we investigated, are trained using Internet content, there is a risk that AI asked to produce educational content will merely replicate potentially problematic discourses commonly found online on a variety of topics.

Our goal in this study was to contribute to emerging understandings of the potential for AI in social studies classrooms. Specifically, we created and used rubrics to assess AI's lesson planning ability when addressing topics whose coverage in schools has been subject to critique. Since many narratives commonly found in education are seen to serve the interests of dominant social or corporate groups, we chose to focus on how the content of the lessons was framed.

We engaged AI to produce lessons on two historical topics, the Civil Rights movement and Indian Removal, where both standards and texts have often reinforced problematic White or settler-colonial narratives and power structures (Sabzalian et al., 2021; Shear et. al, 2015; Woodson, 2016). We also asked ChatGPT, Bing, and Bard to create a lesson on a third topic, climate change, where student learning is often complicated by corporate-initiated discourses of climate denial (Damico & Baildon, 2022). Our investigations were framed around the following questions:

1. To what extent can ChatGPT, Bing, and Bard generate robustly critical lesson plans when prompted to on the subjects of Martin Luther King, Jr., Indian Removal, and climate change?
2. What types of prompting can help AI avoid potentially problematic discourses around these three topics?

Artificial Intelligence

While a full technical description of AI programming is beyond the scope of this paper, AI such as the three we used in this study are large language models (LLMs) trained to produce human-like responses to text prompts using large caches of data, mostly from writing found publicly on the Internet. With such large sets of data, scholars and members of the public have expressed apprehension about the content used to train the models (e.g., Bender et al., 2021), worrying that LLMs will end up reproducing the more problematic elements of Internet discourses, both in the sense that they may echo extremist or hateful positions and that they may uncritically adopt common hegemonic narratives (Gupta et al., 2024).

After the public release of ChatGPT, Microsoft soon integrated ChatGPT into its Bing search engine, combining the text response and summary features with the ability to pull information from Internet searches. Google's Bard also uses a large language model, known as PaLM-2, with different computational structures and training sets from the GPT 3-and GPT-4 used by ChatGPT and Bing, respectively. All three products, despite underlying technical differences, present a similar user experience. Users are provided a space to enter a prompt, and the AI generates a response based on that prompt. Because of the human-like quality of the responses, many users (the authors of this paper included) often rely on humanizing metaphors like "assistant" to understand AI (Gupta et al., 2024). Yet, it is important to note that all three platforms, despite the label of Artificial Intelligence, are not "thinking" in a way that resembles human cognition. Rather, their generation of text, because of the sheer volume of parameters involved in the training data, amounts to very sophisticated predictions of plausible replies to a given prompt.

Much of the published literature on AI in education comes from scholarship conducted prior to the public release of ChatGPT and other user-friendly AIs. A review of 36 studies on earlier chatbot tools (Kuhail et al., 2022) found them primarily fulfilling relatively narrow roles, such as that of didactic teacher or tutor, which differ greatly from the more generic, conversational LLMs found in recently publicized AIs. Still, many scholars have anticipated the potential impacts of more broadly functional

AI on education and considered the ethical and practical consequences of the technology.

Selwyn (2019), for example, noted that many technologies can robotically perform the role of teacher, but some interactive elements of teaching and learning are harder to replicate. Thus, the question of how much AI can influence education, according to Selwyn's argument, can be boiled down to what society envisions education to be and how much it wants AI to influence education.

Much of the impact of AI in the classroom will likely be determined on teachers' willingness to adopt AI tools. In a study of teachers in China, Wang et al. (2023) found that many factors, such as whether teachers viewed AI as a threat to their job, whether they had prior exposure to working with AI, and whether they had a vision for how to use it in the classroom, impacted teachers' likelihood of adopting the AI classroom tools. A study of South Korean language teachers' use of ChatGPT (Jeon & Lee, 2023) found that, following training on how to prompt ChatGPT, the 11 teachers in their sample reported using ChatGPT as a sort of collaborator. Teachers in the study extended their presence in the classroom with AI by using it as a student partner for activities, a teaching assistant to handle minor student inquiries, a content generator or modifier for age-appropriate materials, or a tool for evaluating student work.

Regardless of whether educators want to adopt it for educational purposes, many recognize that young people must be prepared to exist in a world with ubiquitous AI. Calls for AI literacy or Critical AI literacy (e.g., Laupichler et al., 2022; Ng et al., 2023) in both K-12 and higher education will likely grow more urgent in the coming years.

Most of the research and public discourse we found on teachers' use of AI focused on the capabilities of the programs and how teachers have been employing the technology to save time. We were unable to locate any peer-reviewed critical analyses of the content produced by LLMs within a social studies education context.

Theoretical Framework

Because AI programs can produce cogent responses to queries and requests on a number of subjects, much of the early conversations about AI focused on potential academic dishonesty and the potential declines of academic skills following AI use. Many educators, however, liken AI to other technological tools that changed how schools operate. For example, one guide for educators (Johnson et al., 2023) notes that calculators, spellcheck, autocorrect, and Google Translate are all examples of technologies that necessitated some adaptation from teachers, but are now commonplace in classrooms and on campus. In this view, AI, much like a calculator, can allow humans to offload more tedious tasks to a machine, potentially freeing up time and energy for different tasks.

We approach AI tools through a framework of technoskepticsm (Krutka et al., 2020; Krutka et al., 2022; Pleasants et al., 2023), which emphasizes that neither technology nor the societies into which it is introduced can be

considered neutral. Rather, it is important to consider the tradeoffs of technological tools (Postman, 1998/2004) and ask technoskeptical questions (Krutka et al, 2022). Such questions may include asking how a given technology shapes our actions and understandings and what that technology takes from us (e.g., personal data or opportunities to develop certain skills).

An example of a potential tradeoff in the case of AI would be if teachers, especially new or preservice teachers, use AI to generate lesson plans, they may run the risk of stunting their ability to generate questions and creative pedagogies. Further, similar to lessons from other websites like Teachers Pay Teachers (TPT), teachers may be tempted to use lessons produced by AI uncritically and without due consideration for their learning contexts or students (Rodríguez et al., 2020).

Put another way, although AI technologies can be helpful, they can have serious drawbacks that include being subject to troubling societal discourses. Technoskepticism provides a compass to not only navigate what AI is capable of (and not capable of) but also to assess how these AI technologies interact with and potentially reproduce certain narratives that can be sticky in social studies regardless of AI use. In this study, we evaluated the narratives AI produced on the topics of Martin Luther King, Jr. (MLK), Indian Removal, and Climate Change. From a technoskeptical framework, it would be unwise to rely upon AI for producing lessons without critically interrogating the results.

Evaluating AI Generated Lesson Content

Our first research question focused on AI ability to plan lessons with content that avoids problematic narratives and framings typically found around certain topics. In doing so, we drew on scholarship that suggests ways of presenting specific social studies content in ways that humanize and empower. Here, we briefly summarize the scholarship that informed the rubrics that we created to evaluate AI lessons on three topics: MLK, the Indian Removal Act, and climate change.

We chose one key research piece for each lesson topic to create the rubrics. Each of these pieces illuminates common, troubling discourses, and provides guidance for teachers regarding how to do better. These were chosen because of our perception of each piece's influence on social studies education (e.g., citations in other articles as well as conference presentations or the general prominence of the authors on those subjects) and concomitant assumed presence in teacher education. They represent highly ethical approaches to teaching these important topics.

These research pieces were also considered because they offered specific recommendations for teachers and teacher educators against which we could measure the AI's performance. Although AIs are trained via multiple sources, we wanted to see how AI held up when compared with significant counternarratives (i.e., narratives that problematize common discourses). While the scholarship we detail was based on research and theory in the field of social education, the choice of these pieces as the foundation for our evaluations of AI-produced lessons was ultimately subjective, and different sources could be applied to that task. We contend, however, that

the process of (a) establishing criteria for evaluating AI-generated lessons and (b) prompting AI to refine lessons to meet that criteria should be applicable across multiple contexts.

The Heroification of MLK

Heroification (Loewen, 2007) is the process whereby historical figures are centered in narratives and presented without any nuance. These perfect heroes are devoid of the qualities of a well-rounded human being and are consequently devoid of human interest. For our first topic, MLK, we drew on Woodson's (2016) critique of how the Civil Rights Movement is typically covered in history classes, specifically focusing on the heroification of MLK.

Woodson noted that these lessons often frame the movement as driven by larger-than-life figures like MLK to the neglect of the organizations and networks of many individuals, often with different ideas and goals than the prominent figures highlighted. Further, these lessons often view the Civil Rights Movement in isolation, without connection to current efforts at social justice. The effect of these framings is that students are taught to see social justice advocacy as something that a few heroes engage in, rather than as something every person in society is capable of.

Ongoing Indigenous Sovereignty and Nationhood

Our second topic relied on Sabzalian et al.'s (2021) commentary on standards and curricula about Indigenous peoples. They found these documents often deny sovereignty and nationhood, while creating the perception that Indigenous nations are a thing of the past rather than current, living nations (see also Shear et al., 2015). Such curricula leave students with the impression of Indigenous nations as passive historical victims and fail to make a connection with current issues impacting Indigenous peoples. There is a paucity of engagement with Indigenous knowledge and practices, and what is present in the documents critiqued by Sabzalian et al. is mostly content *about* (rather than *with* or *for*) Indigenous nations and nationhood.

Discourses of Climate Change

Teaching about climate change is often challenging because of the prevalence of industry-driven and politically driven messaging that suggests controversy over climate science or stymies conversations about solutions. Damico and Baildon (2022) provided an in-depth look at ways educators can refocus student energy on identifying climate denial discourses and promote better understanding of scientific consensus and different courses of action. These lessons, while aligning with the available science on climate change, run counter to many teachers' instincts of wanting to present both sides of issues such as climate change.

Methodology

We are unaware of any established methodologies in education research for systematically producing and critically evaluating content from AI tools. To develop our ad hoc method for interacting with AI platforms and to take advantage of the conversational nature of the AI tools, we decided to undertake something akin to a collaborative curricular design process (Lewis & Tsuchida, 1999), with several important deviations necessitated by working with AI. Instead of an in-school model, we asked AI to design a lesson that we then assessed according to the rubrics we developed from the research. An iterative process helped us to see the extent to which educators could nudge AI toward a lesson that would honor the topics and the students engaging with them. We engaged with three different AI (ChatGPT, Bing, and Bard), and compared across different AI platforms in terms of the specific lessons we queried. A summary of our process is as follows:

1. We selected the topics for the lessons in tandem with what we deemed to be the most highly relevant piece of social studies scholarship, developing a rubric for each of the topics.
2. We met over Zoom to generate the AI lessons, using ChatGPT, Bing, and Bard, independently evaluating each an initial lesson from each of the three AI with the rubric immediately after creation.
3. We then compared our assessments of each lesson, discussing our rationale for our choices, and came to a final consensus scoring for each criterion from the rubric.
4. We next chose the AI that produced the best lesson of the three and prompted it to refine that lesson further, each time assessing the revision with the rubric.
5. Then, both synchronously over Zoom and asynchronously on our own, we continued our analysis and evaluation of individual lessons, as well as how the AI performed in comparison to each other.

Our approach to the lesson-planning process with AI also loosely borrowed steps from lesson study protocols (Stigler & Hiebert, 1999). Namely, we utilized collaborative goal-setting and a cycle of evaluation and feedback. Yet, because of the nature of working with AI, some modifications of the protocol were necessary. While the two researchers were able to set broad topics and goals for the lessons, AI had no role in that part of the process. In lesson studies, all participants collaboratively plan the lesson, though in our case, AI did the bulk of the planning. Similarly, we could not engage in the crucial lesson study step of having AI teach the lesson it designed while the other members of the team observed. Lastly, while the two researchers were able to reflect on the lesson designed by AI and suggest refinements, the AI could not participate as an equal partner in this reflective process.

Rubric Development

Prior to data collection, we generated a lesson content rubric for each topic containing research-informed content criteria relevant to each topic

derived from the scholarship described above. These rubrics helped us answer both of our research questions; that is, the extent to which AI fell into problematic discourses versus meeting the criteria for the counternarratives to those discourses, as well as the extent to which AI could be prompted to better meet the criteria.

Our rubrics utilized a 1-3 rating system to evaluate each criterion. A rating of 1 was given when the content element was not present or present in such a way that did not reflect any of the desired content or framings. This rating was also applied when AI provided vague responses (e.g., "use documents that reflect multiple views on the civil rights movement"). A rating of 2 meant a developing or partial fulfillment of the criterion. This rating was given when an element was present and shows potential but did not fully meet the stated requirements or remains problematic in some way.

For example, when we asked ChatGPT to include an increased focus on systemic racism in the MLK lesson, we found that it mentioned systemic racism frequently in the text of the lesson. It was also vague about sources it would use and it was unclear if the AI understood the term or if it was just pasting the term into the lesson because we had asked it to.

The rating of 3 was used when the lesson produced by AI had the content element that we are looking for in an acceptable manner. We do not claim that lessons that receive 3s in all elements of our rubrics are perfect in terms of content, but merely that they appear to have addressed the ideas in meaningful ways.

Data Generation and Analysis

Data generation and analysis was iterative. Our conversations with the three AI (ChatGPT, Bing, and Bard) were conducted with both researchers connected over Zoom. Prior to beginning the conversations, our first step was to decide that we would continue prompting the AIs until we generated a lesson that received the maximum score on all the elements of our rubric or we reached five iterations of the lesson (whatever came first). If none of the AIs generated a lesson that scored 3s on all elements of the rubric, we picked the highest scoring lesson and used that AI to generate the remaining lessons. The reason for this maximum of five iterations is that we assume that educators, being as busy as they are, are not going to continue prompting AI for a long duration. We discussed together how many prompts seemed reasonable given our own personal experiences and dispositions, and we agreed that five prompts was the limit.

Each initial prompt was purposely phrased in a general manner (e.g., "Create a high school lesson plan about Martin Luther King, Jr."). This approach varies from much of the wisdom found in articles and videos about prompting AI, which suggests prompts include as much specificity as possible, including inputting specific standards, asking for specific types of activities, and requesting specific accommodations for student needs in the class. Because we wished to capture the "default" tendencies of the AI platforms in terms of lesson content generated and we expected that we would introduce more specificity in our follow up prompts, we opted to begin with a request that included only the topic and grade level.

We collaboratively decided how to phrase the follow-up prompts, when needed, based on the perceived strengths and weaknesses of the previous lesson generated by the AI. In general, we phrased follow up prompts to include direct requests that the next iteration of the lesson include the elements of our rubric we found lacking in the previous version. (e.g., "Modify this lesson plan to include more of a focus on systemic racism"). Each prompt was recorded on a separate rubric, and AI responses were copied into separate documents and stored in a shared Google Drive folder to allow for annotation and reanalysis, as needed.

For ChatGPT, both researchers would type the same prompt on their computers and compare results. We independently assessed each element on the rubric. Immediately upon completion of our independent assessments, we discussed each rating and, where points of difference emerged, we deliberated until a consensus scoring was reached. For Bing and Bard, we followed the same procedure but could only use one computer (that of first author Clark) due to the limited release of those AI at the time.

When prompting Bing, users are given several options of what type of conversation they wish to have. Based on our experience, the options seem to set how long the user wanted responses from Bing to be. The "More Precise" option features to-the-point answers, while the "More Creative" option elaborates more. The "More Balanced" option attempts to strike a middle ground between the two and was set as the default when we first used the chat function. To find the best option for planning lessons, we gave Bing the climate change lesson prompt using each of the options.

We found that the Precise and Balanced options produced links to lessons produced by organizations focused on climate change, but the Creative option resulted in lessons that we could evaluate using our rubrics. Thus, we used that option for generating lessons for the other two topics.

Researcher Positionality

Both researchers have white, settler colonial racial and ethnic identities, so we needed to attend to those identities, especially in relation to our first two lesson foci, MLK and the Indian Removal Act. Our identities and positionality relative to hateful and violent systems steeped in white supremacism requires constant vigilance on our part. As one example, we wrestled with the nomenclature of the "Indian Removal" Act. In the end, we kept the wording because it is the name of the legislation featured in the lesson plans, but we tried to keep asking questions during the project.

It was also important for us to draw directly from the theory and research, itself, as our basis for analysis. A "good" lesson plan was not one that we liked or felt comfortable with; rather, the criteria were derived directly from the original researchers' articles. As such, we hope their voices resound loudly in this project, and it is important to emphasize that the wisdom shared by the scholars on whose work we based our lesson rubrics cannot (and should not) be contained in a single lesson or even a lesson set. There should be a consistent emphasis on the humanity, dignity, and survivance of communities thrust to the margins throughout a social studies course (and, ideally, beyond).

Findings

With the exception of Bing's lesson on climate change, it proved difficult to obtain an AI-generated lesson plan that fully met the established criteria with a succinct prompt. In other words, AI could not escape popular, problematic discourses, even with some prompting. For the lesson on MLK, the strongest initial lesson required us to provide detailed prompts, and even then, the lesson plan could not fully meet the criteria derived from Woodson (2016). For the Indian Removal Act lesson, we assessed via Sabzalian et al. (2021), generating a lesson that acknowledged longstanding and ongoing Indigenous nationhood proved difficult, regardless of the AI we used. In contrast, using Bing we easily obtained a lesson on climate change that fully met the criteria we created via Damico and Baildon (2022). The next section describes our assessments of lessons from ChatGPT, Bing, and Bard for each of the three topics. The highest scoring lessons on each of the three topics are included in Appendices A-C.

Lessons on MLK: Unacceptable Trade-Offs

We asked AI to "create a high school lesson plan about Martin Luther King, Jr." Our evaluations of the lessons generated were informed by Woodson's (2016) article titled, "We're Just Ordinary People: Messianic Master Narratives and Black Youths' Civic Agency." Our rubric included the following:

- Criterion A: Avoids framing an individual as a "savior" or deliverer of a specific group (i.e., depicts them holistically).
- Criterion B: Highlights the networks of individuals doing a wide variety of actions that contributed to the Civil Rights Movement.
- Criterion C: Views racism as systemic (or avoids framing racism as a personal failing).
- Criterion D: Highlights diversity and/or divergent voices within the Civil Rights Movement.
- Criterion E: Connects the Civil Rights Movement to contemporary efforts at social justice.

ChatGPT produced the strongest lesson, although it took five attempts to obtain the higher score (see Appendix A).

Of all the lesson subjects, the lesson on MLK seemed to be the most difficult for the AI to capture our content goals, with even the best lesson (ChatGPT) taking the maximum number of tries without producing a lesson that scored the highest in all our criteria.

ChatGPT

Each attempt revealed a different level of connection to the criteria, although the suggested pedagogy remained fairly constant: assessing prior knowledge, primary sources, use of audiovisuals, small group work, reflection, and application to contemporary times, and the suggestions for assessment are participation and completion.

Table 1
ChatGPT's Lesson on MLK, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C	Criterion D	Criterion E
1	Create a high school lesson plan about Martin Luther King, Jr.	2	2	1	2	3
2	Modify this lesson plan to include more of a focus on systemic racism	2	1	3	1	3
3	Modify this lesson to include people who supported MLK and people within the Civil Rights Movement who disagreed with MLK	2	3	2	3	3
4	Keep the emphasis on diverse voices but modify this lesson to show how MLK worked alongside communities and organizations	3	2	1	2	3
5	Modify this lesson to include: systemic racism, people who supported MLK and people within the Civil Rights Movement who disagreed with MLK, and how MLK worked alongside communities and organizations	3	3	2	3	3

The first attempt was largely mediocre and would depend heavily on a teacher knowing what to do with the instructions ChatGPT generated. For example, the prompt, "Ask students what they know about the Civil Rights Movement and its goals," could set up a discussion about the networks of individuals doing a wide variety of actions, but not necessarily. A highlight was the connections to our contemporary times: "Ask students to write a short paragraph on how they think King's legacy has influenced American society today," and in the extension, "Have students research and analyze a current social justice issue and discuss how the strategies used in the Civil Rights Movement could be applied today."

As we attempted to get ChatGPT to modify this first lesson, there were tradeoffs. As an example, the second lesson, which we asked to focus more on systemic racism, lost any reference to folks other than MLK. Also, the lesson mentioned systemic racism but would rely on a teacher knowing what that is (i.e., ChatGPT offers no framework for defining or explaining that racism is systemic). When we asked ChatGPT to modify the lesson to include supporters as well as those who disagreed with MLK, ChatGPT put in the word "diverse," seemingly understanding what we were getting at, rather than in the second attempt where the AI just parroted the word "systemic" and tacked it onto lesson activities and topics.

Nonetheless, there was another tradeoff–that diversity of voices meant that the systemic aspect lost its explicit focus. When asked to modify the

lesson to emphasize collaboration alongside communities and organizations, the lesson lost a lot of specificity regarding racism as well as the people involved in the Civil Rights Movement.

What seemed to work best was what we affectionately called the "kitchen sink" approach. Knowing that ChatGPT would respond to the specific words we used drawn from our rubric, we asked it to "Modify this lesson to include: systemic racism, people who supported MLK and people within the Civil Rights Movement who disagreed with MLK, and how MLK worked alongside communities and organizations." This prompt brought forth the best result, but it still was not perfect, For example, discussions of systemic racism still felt forced, relying on students to find it or teachers to explain it well. Even with all of these modifications, ChatGPT could not produce a lesson that scored a 3 in all of the criteria.

Bing

Bing's lesson on MLK showed more potential than Bard's, but was still weak compared with ChatGPT's lesson. Overall, this lesson was not able to meet our criteria in meaningful ways.

Table 2
Bing's Lesson on MLK, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C	Criterion D	Criterion E
1	Create a high school lesson plan about Martin Luther King, Jr.	2	2	2	1	1

The lesson hinted at networks of individuals doing a wide variety of actions that contributed to the Civil Rights Movement: "Other leaders" are mentioned but not named. This lesson would be stronger if more people were highlighted specifically and if there were some discussion of disagreement about methods to counter injustice. Importantly, the Bing lesson talked about laws as perpetuating racism; however, a systemic analysis started and stopped with legislation. No contemporary connections were even hinted at, which is troubling given ongoing white supremacism.

Bard

Despite the prompt not asking for options, Bard made three different lesson plans — all of which were weak. The first two were surface level and vague, one focused on the "I Have a Dream Speech" and the other focused on basic information about MLK (e.g., birthdate) and a few quotations divorced from context. The third lesson fared the best on our rubric, albeit not by much, and is discussed next.

Table 3
Bard's Lesson on MLK, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C	Criterion D	Criterion E
1	Create a high school lesson plan about Martin Luther King, Jr.	1	1	1	1	2

The lesson had a vague procedure, beginning with establishing prior knowledge and transitioning to content knowledge that would rely on the teacher's own knowledge of MLK's life, work, and legacies. He is examined in isolation, and systemic racism and diverse voices within the Civil Rights Movement were completely absent. The only category that showed some promise was the last criterion regarding contemporary connections. A suggested extension activity, not in the main body of the lesson, suggested that teachers "have students research a current issue related to civil rights and write a letter to their elected officials expressing their views." This activity could be fruitful, but we would rather it be part of the main lesson, as opposed to being presented as an optional extension activity.

Lessons on the Indian Removal Act: Nationhood Ignored

We asked AI to "create a high school lesson plan about the Indian Removal Act." Our evaluations of the lessons generated were informed by Leilani Sabzalian, Sarah B. Shear, and Jimmy Snyder's (2021) article, "Standardizing Indigenous Erasure: A TribalCrit and QuantCrit Analysis of K–12 U.S. Civics and Government Standards." Our rubric included the following:

- Criterion A: Recognizes Indigenous sovereignty AND/OR Indigenous government, institutions, rights.
- Criterion B: Indicates (implicitly or explicitly) that Indigenous nationhood as the present state of affairs (not relegated to the past).
- Criterion C: Includes information on Indigenous refusals, resistances, and survivance rather than just the tragedy itself.
- Criterion D: Mentions that "removal" is an ongoing problem.
- Criterion E: References Indigenous "nations" not "groups" and uses specific terminology for relevant Indigenous nations (e.g., Chickasaw vs. "Native American tribes").

Based on these criteria, we concluded that Bing produced the strongest initial lesson and received the top score in each category after a single revision (see Appendix B). Criterion B regarding ongoing Indigenous nationhood was consistently weak across all three AI platforms on the first attempt.

ChatGPT

Table 4
ChatGPT's Lesson on the Indian Removal Act, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C	Criterion D	Criterion E
1	Create a high school lesson plan about the Indian Removal Act	1	1	2	2	3

ChatGPT's lesson scored 1 in Criteria A and B, indicating that we did not feel the lesson produced sufficiently recognized Indigenous institutions and government nor did it recognize the continued existence of Indigenous nationhood. In both cases, the lesson included vague terminology that could be interpreted as partially or fully meeting the criteria, but we judged the instructions to be too nebulous to count.

For example, in the case of recognizing Indigenous government or institutions, ChatGPT's lesson asked teachers to distribute primary source documents related to the Indian Removal Act, including "letters from Native American leaders opposing the act," which due to it not suggesting primary sources from specific leaders or impacted Indigenous nations, we judged too vague to meet the criteria. Similarly, the lesson included discussion of the "long-term consequences of the Indian Removal Act on Native American communities" but was, again, not specific enough to be clearly referring to Indigenous nations existing as present-day entities.

ChatGPT's lesson scored 2s on both Criteria C and D in our rubric. For Criteria C, about refusals and resistance, we felt that having letters of opposition from Indigenous leaders highlighted resistance to the Indian Removal Act, though, again, it did not suggest specific letters, nor did it discuss other forms of resistance. By noting the long-term consequences of the Indian Removal Act, the lesson could prompt teachers to consider Criteria D, removal as an ongoing problem. We deliberated over whether the instruction was too vague to be useful. Ultimately, we decided that, in noting consequences beyond the act itself, the lesson had barely crossed into the range of a 2 on our rubric.

For Criteria E, we assigned ChatGPT's lesson a 3 for its mention of specific Indigenous nations in places, though we also noted that in several instances the lesson defaulted to more generic phrasing of "Native American tribes."

Bing

Bing provided the strongest initial lesson on the Indian Removal Act. Overall, we were impressed with clearer language recognizing nationhood and sovereignty, both directly and indirectly. Still, the first attempt from Bing did not include any reference to the continued existence of Indigenous nations and did not address the ongoing consequences of removal. Thus, when we asked Bing for a revision, we specifically targeted

these areas of our rubric, while asking it to keep the other parts of the lesson (see Table 5 for the revised prompt).

Table 5
Bing's Lesson on the Indian Removal Act, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C	Criterion D	Criterion E
1	Create a high school lesson plan about the Indian Removal Act	3	1	3	1	3
2	Please keep this lesson plan but add something about the implications of the Indian Removal Act for contemporary Indigenous nations	3	3	3	3	3

Bing produced a revised lesson that, in our judgment, merited scores of 3 on all elements. In particular, we found several of the discussion questions suggested by Bing to be good for students to consider. Some examples included the following:

- "How do Indigenous nations exercise their sovereignty and self-determination in relation to federal and state governments?"
- "How do Indigenous nations preserve and revitalize their identity and culture in their communities?"

In general, out of the three AI, Bing presented a lesson with the potential to foster deeper understandings of Indigenous nations while teaching about the Indian Removal Act.

Bard

Table 6
Bard's Lesson on the Indian Removal Act, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C	Criterion D	Criterion E
1	Create a high school lesson plan about the Indian Removal Act	2	1	1	1	1

Google's Bard produced the weakest scoring lesson for the Indian Removal Act based on our criteria. The only criteria we rated as being even partially fulfilled was the recognition of Indigenous sovereignty and rights (Criteria A), which was achieved mostly due to discussions of "the rights of Native Americans to their own land" during a section of the lesson devoted to the ethics of the Indian Removal Act.

The remaining criteria were scored as 1 on our rubric. We did not notice meaningful mentions of Indigenous nations as existing in the present day, discussions of Indigenous resistance, considering current or long-term

challenges of removal, or references to the experiences of specific Indigenous nations.

Lessons on Climate Change: Scientific Consensus Comes Through

We asked AI to "create a high school lesson plan on climate change." Our evaluations of the lesson generated were informed by James S. Damico and Mark C. Baildon's (2022) book, *How to Confront Climate Denial: Literacy, Social Studies, and Climate Change.* Our rubric included the following:

- Criterion A: Promotes critical analysis of sources of information about climate change (e.g., funding sources, parent organizations, greenwashing etc.).
- Criterion B: Avoids climate denial discourses.
- Criterion C: Engages students in conversations about action on climate change, not conversations about the existence of the phenomena.

For the climate change prompt, Bing produced the strongest initial lesson plan (see Appendix C). Notably, the initial Bing lesson did not require any revisions based on our criteria.

ChatGPT

ChatGPT produced a 3-day lesson plan that we scored 1, 2, and 2, on Criteria A, B, and C, respectively. The lesson did not include any attempts to critically analyze sources of information on climate change (Criterion A). To its credit, the lesson noted that climate change is caused by human activities and asks students to split into groups to research the impact of various activities on climate change. In doing so, however, the lesson seems to assume students will critically analyze the information they find during their research, something which is unlikely to happen without prior preparation (McGrew et al., 2017).

Table 7
ChatGPT's Lesson on Climate Change, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C
1	Create a high school lesson plan on climate change	1	2	2

ChatGPT's lesson partially fulfilled Criteria B and C by explicitly challenging some climate denial discourses and promoting action to address climate change. For example, the lesson directly noted that human activity is the principal driver of climate change and directly discussed ways that students can take action in their communities, even prompting them to brainstorm ways they can influence local government to act. Yet,

there were also instances where climate denial thinking was still present in the lesson.

For instance, when discussing climate action, ChatGPT's lesson asked students to brainstorm actions that "individuals and communities" can take, giving "recycling, reducing plastic waste, conserving water, using public transportation or carpooling, and eating a plant-based diet" as potential responses. While individual and local action can help combat climate change, the lesson ignored the large roles of corporations and national governments in contributing to climate change through emissions and policies, with one report suggesting that 100 government and business entities involved in fossil fuels accounted for nearly 71% of emissions (Griffin, 2017). Emphasizing individual action and responsibility over corporate and government accountability, while less overt than simply denying climate change altogether, is a discourse tactic frequently employed by companies with an interest in delaying large-scale climate action policy (Damico & Baildon, 2022).

Bing

Like ChatGPT, Bing created a multiday lesson, though Bing's "lesson" could be better described as a unit that comprehensively addressed both the scientific and social/political elements of the topic. Due to its 5-day length and thoroughness, this lesson was the only one generated in the study that scored 3s in all content criteria on the first iteration. We also noted that Bing drew heavily upon lesson resources and activities from established institutions like Stanford University and the National Aeronautics and Space Administration (NASA), and provided links to these sources, allowing us to compare its lesson to existing curriculum. In this case, the connections to its source material were clear, but Bing also showed evidence of synthesizing and condensing the units available from the sources it drew upon. For example, Stanford's unit on climate change is 3 weeks long, compared to Bing's five-day unit.

Table 8
Bing's Lesson on Climate Change, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C
1	Create a high school lesson plan on climate change	1	2	2

Overall, we liked that the content of the unit blended science and social science, using both to challenge climate denial narratives. For example, one common tactic of climate denial is to highlight extreme cold as a sign that "global warming" is not happening. Yet, this unit explicitly addressed the distinction between weather (e.g., a very cold day) and climate as long-term patterns. The unit also expanded its focus beyond individual actions to reduce climate change (although the first lesson in the sequence did invite students to consider their individual carbon footprints) to talk about actions possible through policy and corporate action. Further, there is substantial discussion of the trustworthiness of information coming from various sources about climate change, which can be used to address climate denial tactics such as greenwashing.

Bard

Google Bard produced a 1-day lesson that we scored the same as the ChatGPT lesson. For Criteria A, it scored a 1. Like the ChatGPT lesson, this lesson did not engage students in critical reading of information and messaging about climate change. For Criteria B and C, we scored this lesson as a 2.

Table 9
Bard's Lesson on Climate Change, Scoring of Each Attempt

Attempt No.	Prompt Text	Criterion A	Criterion B	Criterion C
1	Create a high school lesson plan on climate change	1	2	2

We did see some potential in its discussion of policy solutions to climate change alongside individual action, as well as its extension activities that encouraged students to take political action, but we also found spaces where climate denial discourses could seep in. For example, one element of the lesson asked students to discuss both human and natural causes of climate change. A common climate denial tactic is to blame climate change on natural fluctuations in the Earth's climate over time, minimizing the human influences. If not handled carefully, a discussion of natural changes in temperature over time could leave students with the message that there is nothing that can be done about climate change.

AI-Generated Pedagogy

Although our primary purpose for this analysis was to critique the content framings of our three focal topics found in AI-generated lesson plans, we also noted interesting trends in the pedagogies employed. All the lessons generated by the AIs in our study tended to follow the pattern of activating prior knowledge during a bellringer activity, introducing content (through video, short lecture, or handout), "discussion," and a learning activity (with a group research time being the most common choice). Some explicitly noted time for reflection and wrap up, while others simply ended the lesson with the learning activity. While there is not anything necessarily wrong with this sequence of activities, it was interesting that the lessons were so formulaic, and we worried that students may quickly find the repetitive sequence dull.

The bellringer activities nearly all consisted of students writing down what they already knew about a topic or, in the case of multiday lessons, writing down what they could recall from the previous lessons in the sequence. Again, while this is not necessarily problematic, there are many different ways to activate students' prior knowledge or prime them for learning through bellringers that the AIs were not utilizing.

There were also many pedagogical moments where the AI lesson directions were vague. Bing's lesson on climate change, although scoring high on the content framings rubric, often used the word "discuss" to describe portions of the lesson where there were not many interesting

avenues for discussion. For example, at one point in the lesson, it included the instructions, "Discuss how scientists collect and analyze data on climate change using various methods and sources, such as ice cores, satellites, models, etc." In this case, the word "discuss" seems to mean "lecture" and does not incorporate any understanding of how to foster rich discussions in class (e.g., Hess, 2009).

In contrast, Bard's lesson on climate change and the Indian Removal Act, while scoring lower in the content framings, did provide focal questions or highlights to guide discussion (although, we could also critique some of the questions provided as not being particularly generative for sustained discussion).

In terms of other materials that might be helpful for teachers, the various AIs had differing strengths and weaknesses. Many lessons mentioned using video resources, but only Bing (because it is part search engine) was able to provide links to specific videos during its lessons on climate change. In terms of other suggested resources, several lessons referenced handouts or primary sources that were not provided by the AI. Occasionally, these resources were described in enough detail that a teacher could easily find them online (e.g., finding a copy of the "I have a Dream" speech), but others provided little substantial direction. One Bard lesson simply directed teachers to distribute "the handout on climate change" while providing no detail on the specific contents of the handout or even listing a handout among the needed materials for the lesson.

Bing, despite producing what we judged to be the best content in two of our three topics, made no effort to address issues of differentiation or extension. ChatGPT did provide extension activities for its lessons on MLK but not on the Indian Removal Act or climate change, and it offered no suggestions for differentiation on any of the lessons. Bard, despite scoring the lowest on our content rubrics, was the only AI to offer suggestions for differentiating the lesson on all three topics.

Discussion

Technology shapes what we do and how we do it. In the case of lesson planning, the three AIs we assessed were able to generate interesting ideas and possible source material for lessons. As helpful as that may be, our technoskeptical approach reveals that an uncritical use of AI carries with it the danger of slipping into problematic discourses. Prompting can help shift lessons toward counternarratives, but those prompts need to be specific, and it is not always possible to completely eliminate troubling aspects.

Reflecting on the different platforms in relation to the criteria developed from social studies research, Bing produced better lesson content for two of our three topics. At the time of our data gathering, ChatGPT seemed to garner the most attention, but Bing generally did a better job of tapping into existing research and resources. Bard might be good for some things (e.g., it was the only AI to reliably provide differentiation suggestions and extension activities for all topics) but performed the weakest on our content criteria.

One question we had is why climate change as a topic was so much easier for AI than MLK and the Indian Removal Act. The AI could draw from a number of open-access sources with counternarratives to debunk climate denialism and prop up research, whereas there might be fewer open-access resources for the other two topics. We theorized that the scientific consensus on climate change provided ample source material in the training data for the AIs to draw upon, whereas it is also easy for the AI to incorporate the plentiful whitewashed narratives in history and social sciences into its lesson designs.

There also has been a concerted effort to combat climate misinformation campaigns that were likely included in the AI training data or that the web-enabled AIs could discover easily online, which stands in contrast with the sneaky forms of settler colonialism and white supremacism that pervade the Internet and public discourse more generally. In short, the abundance of educational material about teaching climate change logically entails that AI can make better lessons than it would for topics with less abundant, high-quality material to draw from. Unfortunately, it is for those topics that teachers would likely need the most assistance from AI.

Ongoing Indigenous nationhood was a sticky point for all the initial lessons on all three platforms (at least on the first iteration), perhaps because it is an idea that openly challenges settler assumptions, both explicit and implicit in public discourse. Some climate change discourse does not necessarily involve changing mainstream habits and assumptions (e.g., switching from gas-powered vehicles to electric ones, but not considering bicycling), while recognizing Indigenous sovereignty clearly challenges many institutions and practices.

We were able to direct Bing toward slightly less settler-colonial framings through our follow-up prompt, but not entirely. For example, the lesson plan assumed an "us" that placed Indigenous people and nations as others outside of those in the classroom: "How does learning about the contemporary implications of the Indian removal policy help *us* respect and appreciate the diversity and resilience of Indigenous nations?" (emphasis added).

Even though we were able to get decent AI-generated lesson plans in terms of content, it was rare to obtain a strong lesson on the first try. AI can produce a less harmful lesson plan than one that simplistically parrots problematic discourses, but that process requires specific and, sometimes multiple, prompts. This indicates that strong research — in our case, from Woodson (2016), Sabzalian et al. (2021), and Damico and Baildon (2022) — will still need to be communicated in teacher education classes and professional development so that teachers can create strong, ethical lessons with or without the use of AI.

AI may help teachers create lessons quickly and generate ideas, but nothing can replace specific knowledge from educational scholars. In addition, teacher education programs should explicitly address AI as a part of teacher practice, especially in developing teacher candidates' critical eye toward AI-generated products. As with any source of curriculum or classroom materials, it is incumbent upon teachers to critically evaluate AI lessons to ensure that content is presented in a way that aligns with their classroom goals.

Even though we do have lessons on the Indian Removal Act and climate change that scored three on all elements of our rubrics, we do not recommend teachers use these lessons off the shelf without considering their students and contexts. Professional judgment about the applicability of the suggested information and activities to a given classroom is still necessary even in the age of AI. Further, although our primary focus was on content framings and not pedagogy, we noted several instances where teachers may wish to elaborate on the suggested methods themselves or prompt the AI for different choices.

If classroom teachers were to engage in a similar exercise to ours, where the AI becomes a sort of partner/sounding board in the lesson planning process, they could use the technology to generate questions, activities, and resources they might not have previously thought of. As long as it is the teacher who is directing the lesson planning and not the AI, AI can be a valuable tool for social studies educators.

Limitations

We recognize that these AI, having been publicly available for less than a year at the time of our data gathering, may not have been a finished product. In fact, prior to chatting with ChatGPT, users received a warning: "While we have safeguards in place, the system may occasionally generate incorrect or misleading information and produce offensive or biased content" (OpenAI.com/Chat). We also recognize that the rapid development of these technologies might mean that AIs develop capabilities beyond what we were able to observe in our explorations. Last, the three topics we chose represent a narrow slice of the work the AIs can produce and lesson prompts of other topics could produce different outcomes.

Another limitation stems from the lack of specific protocols to evaluate AI-generated lesson plans. Our process needed to be somewhat ad hoc because of this gap. Had we used (or developed) a different protocol, there would be different results. We hope that further research in this area will eventually illuminate different protocols and their own affordances and constraints.

Significance

In terms of this project, we found that AI can provide some lesson ideas and resources that avoid problematic content framings if used wisely. This practice requires a technoskeptical approach with a foundation provided by current social studies research to input requests that specify the sorts of content and questions the AI generates. In short, AI cannot replace meaningful scholarly work coupled with teacher education that engages with that work.

Teachers using AI to create lessons need to be specific about what they are looking for in a lesson plan, and even then, they will likely need to modify the lesson. We need scholars (e.g., Dr. Ashley Woodson for a lesson on MLK) to point us toward thoughtful content that takes a critical stance against problematic discourses. Although AI can help teachers figure out

ways of approaching a topic, nothing can replace professional judgment stemming from educational research.

That being said, when paired with professional judgment and research-informed instructions, AI was able to produce lesson plan ideas that, while not perfect, included some insightful questions and interesting activities. A thoughtful teacher would likely be able to combine the stronger elements of each of the lesson iterations, elaborate on the vague areas, and create a meaningful learning experience for their students. On the other hand, asking AI for a lesson plan on a topic without any additional guidance, as done in some early explorations of the AI's lesson planning potential (Malon, 2023; Will, 2023) and as we did in the first iterations of our instructions, seemed to result in lessons that were fairly mediocre and, depending on the topic, may end up perpetuating problematic narratives and discourses.

An over-reliance on AI for creating lesson plans means that teachers would be giving up many of the insights provided by educational scholars. Much like using sites like Teachers Pay Teachers or Pinterest (Rodríguez et al., 2020), it might be helpful to see how someone else would make a lesson, but teachers must use current research to ignore or modify aspects of those lessons if we want to do right by our students. In such a way, teacher educators might incorporate AI into a session on ways to lesson plan, but AI would be at most a tool to facilitate a thoughtful and broader process of lesson planning. Quoting Baudrillard (1993),

> Artificial intelligence is devoid of intelligence because it is devoid of artifice. True artifice is the artifice of the body in the throes of passion, the artifice of the sign in seduction, the artifice of ambivalence in gesture, the artifice of ellipsis in language, the artifice of the mask before the face, the artifice of the pithy remark that completely alters meaning. So-called intelligent machines deploy artifice only in the feeblest sense of the word..." (p. 58)

Although AI might help teachers generate ideas for lessons, both content and form, AI cannot replace the beautiful imperfections and complexities of a warm-blooded, human educator.

References

Baudrillard, J. (1993). *The transparency of evil.* (J. Benedict, Trans.) Verso.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623). Association for Computing Machinery.

Damico, J. S., & Baildon, M. C. (2022). *How to confront climate denial: Literacy, social studies, and climate change.* Teachers College Press.

Griffin, P. (2017). *The carbon majors database: CDP carbon majors report 2017.* Carbon Disclosure Project. https://cdn.cdp.net/cdp-

production/cms/reports/documents/000/002/327/original/Carbon-Majors-Report-2017.pdf?1501833772

Gupta, A., Atef, Y., Mills, A., & Bali, M. (2024). Assistant, parrot, or colonizing loudspeaker? ChatGPT metaphors for developing critical AI literacy. *Open Praxis.* Advance online publication. https://doi.org/10.55982/openpraxis.16.1.631

Hess, D. E. (2009). *Controversy in the classroom: The democratic power of discussion*. Routledge.

Hiskey, D. (2012, October 16). Where the expression "I for one welcome our new x overlords" came from. *Today I Found Out.* https://www.todayifoundout.com/index.php/2012/10/where-the-expression-i-for-one-welcome-our-new-x-overlords-came-from/

IBM (n.d.) *About.* https://www.ibm.com/watson/about

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies, 28*(12), 15873–15892. https://doi.org/10.1007/s10639-023-11834-1

Johnson, S.M., Coley, M., Bandy, J., Bradley, J., & Molevig, O. (2023). *Teaching in the age of AI*. Vanderbilt Center for Teaching. https://cft.vanderbilt.edu/teaching-in-the-age-of-ai/

Kelly, S.M. (2023, January 26). ChatGPT passes exams from law and business schools. *CNN Business.* https://www.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html

Krutka, D. G., Heath, M. K., & Mason, L. E. (2020). Technology won't save us–A call for technoskepticism in social studies. *Contemporary Issues in Technology and Teacher Education, 20*(1), 108-120. https://citejournal.org/volume-20/issue-1-20/social-studies/editorial-technology-wont-save-us-a-call-for-technoskepticism-in-social-studies

Krutka, D. G., Metzger, S. A., & Seitz, R. Z. (2022). "Technology inevitably involves trade-offs": The framing of technology in social studies standards. *Theory & Research in Social Education, 50*(2), 226-254.

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, *28*(1). https://doi.org/10.1007/s10639-022-11177-3

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence, 3*, 100101. https://doi.org/10.1016/j.caeai.2022.100101

Lewis, C. C., & Tsuchida, I. (1999). A lesson is like a swiftly flowing river: How research lessons improve Japanese education. *Improving Schools, 2*(1), 48-56. http://dx.doi.org/10.1177/136548029900200117

Loewen, J. W. (2007). *Lies my teacher told me: Everything your American history textbook got wrong*. Simon & Schuster.

Mallon, S. (2023, January 25). Teachers try this: Build a lesson plan using ChatGPT. *Education Week*. https://www.edweek.org/teaching-learning/video-teachers-try-this-build-a-lesson-plan-using-chatgpt/2023/01

Mcgrew, S, Ortega, T., Breakstone, J., Wineburg, S. (2017). The challenge that's bigger than fake news: Civic reasoning in a social media environment. *American Educator, 41*(3), 4-9, 39.

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

OpenAI. (2023a, February 5). *Explain briefly what ChatGPT is*. ChatGPT [Computer software].

Pleasants, J., Krutka, D. G., & Nichols, T. P. (2023). What relationships do we want with technology? Toward technoskepticism in schools. *Harvard Educational Review, 93*(4), 486-515

Postman, N. (2004). Five things we need to know about technological change [Speech transcript]. University of California, Davis. https://web.cs.ucdavis.edu/~rogaway/classes/188/materials/postman.pdf (Original work published 1998)

Rodríguez, N., Brown, M., & Vickery, A. (2020). Pinning for-profit? Examining elementary preservice teachers' critical analysis of online social studies resources about Black history. *Contemporary Issues in Technology and Teacher Education, 20*(3), 497-528. https://citejournal.org/volume-20/issue-3-20/social-studies/pinning-for-profit-examining-elementary-preservice-teachers-critical-analysis-of-online-social-studies-resources-about-black-history/

Sabzalian, L., Shear, S. B., & Snyder, J (2021). Standardizing Indigenous erasure. A TribalCrit and QuantCrit analysis of K-12 U.S. civics and government standards. *Theory & Research in Social Education, 49*(3), 321-359. https://www.doi.org/10.1080/00933104.2021.1922322

Selwyn, N. (2019). *Should robots replace teachers?: AI and the future of education*. Polity Press.

Shear, S. B., Knowles, R. T., Soden, G. J., & Castro, A. J. (2015). Manifesting destiny: Re/presentations of Indigenous peoples in K-12 U.S. history standards. *Theory & Research in Social Education, 43*(1), 68–101. https://doi.org/10.1080/00933104.2014.999849

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving in the classroom*. The Free Press.

Wang, X., Li, L., Tan, S. C., Yang, L., & Lei, J. (2023). Preparing for AI-enhanced education: Conceptualizing and empirically examining teachers' AI readiness. *Computers in Human Behavior*, 146, 107798. https://doi.org/10.1016/j.chb.2023.107798

Warren, T. (2023, November 15). Microsoft rebrands Bing Chat to Copilot, to better compete with ChatGPT. *The Verge*. https://www.theverge.com/2023/11/15/23960517/microsoft-copilot-bing-chat-rebranding-chatgpt-ai

Watson, T. (2023, November 15). Microsoft rebrands Bing Chat to Copilot, to better compete with ChatGPT. *The Verge*. https://www.theverge.com/2023/11/15/23960517/microsoft-copilot-bing-chat-rebranding-chatgpt-ai

Will, M. (2023, January 12). We gave ChatGPT common teaching tasks. Here's how teachers say it did. *Education Week*. *https://www.edweek.org/technology/we-gave-chatgpt-5-common-teaching-tasks-heres-how-teachers-say-it-did/2023/01*

Woodson, A. N. (2016). We're just ordinary people: Messianic master narratives and Black youths' civic agency. *Theory & Research in Social Education, 44*(2), 184-211. https://doi.org/10.1080/00933104.2016.1170645

Appendix A
ChatGPT Lesson on MLK

Lesson Plan Title: Martin Luther King Jr. and Diverse Perspectives on Systemic Racism and Collaborative Efforts in the Civil Rights Movement

Grade level: High School (9th-12th grade)

Objective: Students will learn about the life and impact of Martin Luther King Jr., explore diverse perspectives on systemic racism, and examine how King worked alongside communities and organizations, including those who supported and disagreed with him, in the Civil Rights Movement.

Materials:
- Martin Luther King Jr. biography or timeline
- Video or audio recordings of King's speeches
- Primary sources related to the Civil Rights Movement, systemic racism, and diverse perspectives (e.g. speeches, letters, interviews, photographs, documents)
- Whiteboard or chalkboard
- Writing utensils

Introduction (10 minutes):
1. Begin by asking students what they know about Martin Luther King Jr., systemic racism, and the Civil Rights Movement.
2. Use a timeline or biography to give a brief overview of King's life and accomplishments, with a focus on his work to address systemic racism and his collaborative efforts.
3. Ask students to share their thoughts on why it is important to understand diverse perspectives and work collaboratively to address systemic racism.

Main Activities (40 minutes):
1. Show a video or audio recording of one of King's speeches, such as "I Have a Dream" or "Beyond Vietnam".
2. Ask students to take notes on the main points of the speech and how King addressed systemic racism and the importance of working collaboratively to achieve social change.
3. Divide the class into small groups and assign each group a person or group who either supported or disagreed with King during the Civil Rights Movement. These might include Malcolm X, the Black Panthers, or white allies like Viola Liuzzo or James Reeb.
4. Have each group research their assigned person or group and find primary sources related to their views on systemic racism and the Civil Rights Movement.
5. Have each group share their findings with the class and discuss the impact of diverse perspectives on the Civil Rights Movement, including how King worked alongside communities and organizations to address systemic racism.

Wrap-Up (10 minutes):
1. Have students reflect on what they learned about King's collaborative efforts in the Civil Rights Movement and diverse perspectives on systemic racism.

2. Ask students to write a short paragraph on how they can work collaboratively with others to address systemic racism and create a more just and equitable society.

Assessment:
1. Participation in class discussion and group activities
2. Completion of the writing assignment at the end of the lesson

Extension:
1. Have students research and present on a current social justice issue related to systemic racism and discuss diverse perspectives on how to address it, as well as examples of successful collaborative efforts.
2. Have students create a project that demonstrates the importance of understanding diverse perspectives and working collaboratively to address systemic racism. This might include a video, podcast, or written report.

Appendix B
Bing Lesson on the Indian Removal Act

Lesson Plan: The Indian Removal Act
Objective

- Students will learn about the historical context and significance of the Indian Removal Act of 1830, which authorized the president to exchange lands west of the Mississippi River for Indian lands within existing state borders.
- Students will analyze the perspectives and experiences of different groups involved in or affected by the Indian removal policy, such as President Andrew Jackson, white settlers, state governments, and various Native American tribes.
- Students will evaluate the consequences and legacy of the Indian removal policy, especially the Trail of Tears, which was the forced relocation of thousands of Native Americans from their ancestral lands to designated territories in the west[3].
- Students will explore the contemporary implications of the Indian removal policy for indigenous nations, such as their sovereignty, identity, culture, and rights.

Materials

- Copies of the Indian Removal Act of 1830 (available online at https://www.loc.gov/law/help/statutes-at-large/21st-congress/session-1/c21s1ch148.pdf)
- Copies of primary sources related to the Indian removal policy, such as speeches, letters, petitions, treaties, maps, paintings, etc. (available online at https://guides.loc.gov/indian-removal-act/primary-sources)
- Copies of secondary sources related to the contemporary implications of the Indian removal policy, such as articles, reports, videos, etc. (available online at https://americanindian.si.edu/nk360/removal/resources)
- Graphic organizers for note-taking and summarizing (such as KWL charts, Venn diagrams, etc.)
- Writing materials or devices for reflection and assessment

Procedure

1. Activate prior knowledge and generate interest by asking students what they know or want to know about the Indian Removal Act and its impact on Native Americans. Record their responses on a KWL chart or a similar graphic organizer.
2. Provide some background information on the Indian Removal Act by reading or summarizing its main provisions and purpose. Explain that it was passed by Congress and signed by President Andrew Jackson on May 28, 1830[12], and that it was part of a broader policy of expanding white settlement and agriculture in the eastern United States at the expense of Native American lands and rights.
3. Divide students into small groups and assign each group a different stakeholder or perspective related to the Indian removal policy. For example, some possible groups are: President Andrew Jackson, white settlers in Georgia, Cherokee Nation, Choctaw Nation, Seminole Nation, etc. Provide each group with copies of relevant primary sources that reflect their assigned perspective, such as speeches, letters, petitions, treaties, maps, paintings, etc.

Ask each group to read and analyze their sources using guiding questions such as:

4. What are the main arguments or reasons for supporting or opposing the Indian removal policy?
5. What are the benefits or challenges of staying or moving to a new territory?
6. How do the sources portray or describe the relationship between Native Americans and white Americans?
7. How do the sources express or imply emotions or attitudes toward the Indian removal policy?
8. What are some similarities or differences between the sources within or across groups?
9. Have each group share their findings and perspectives with the rest of the class using a graphic organizer such as a Venn diagram or a T-chart. Encourage students to compare and contrast the different viewpoints and experiences of the various groups involved in or affected by the Indian removal policy. Facilitate a class discussion on the following questions:
10. How did the Indian removal policy reflect or challenge the ideals and values of American democracy?
11. How did the Indian removal policy affect the lives and cultures of Native Americans?
12. How did Native Americans resist or adapt to the Indian removal policy?
13. How did the Indian removal policy shape the development and expansion of the United States?
14. How do we remember or commemorate the Indian removal policy today?
15. Provide students with copies of secondary sources related to the contemporary implications of the Indian removal policy for indigenous nations, such as articles, reports, videos, etc. Ask students to read and watch these sources using guiding questions such as:
16. How did the Indian removal policy influence or determine the current status and location of indigenous nations?
17. How do indigenous nations exercise their sovereignty and self-determination in relation to federal and state governments?
18. How do indigenous nations preserve and revitalize their identity and culture in their communities?
19. How do indigenous nations address their social and economic challenges and opportunities?
20. How do indigenous nations relate to other indigenous peoples across borders and regions?
21. Have students share their insights and questions with each other using a graphic organizer such as a fishbone diagram or a mind map. Encourage students to make connections between past and present issues and experiences of indigenous nations. Facilitate a class discussion on the following questions:
22. How does learning about the contemporary implications of the Indian removal policy help us understand its historical significance?
23. How does learning about the contemporary implications of the Indian removal policy help us respect and appreciate the diversity and resilience of indigenous nations?

24. How does learning about the contemporary implications of the Indian removal policy help us become more informed and responsible citizens?
25. Assess student learning by asking them to write a reflection or an essay on one of the following prompts:
26. Imagine you are a member of one of the groups involved in or affected by the Indian removal policy. Write a letter to another group expressing your opinion on the policy and its impact on you and your people.
27. Imagine you are a historian researching the Indian removal policy. Write a summary of your findings and conclusions based on the primary sources you have examined.
28. Imagine you are a teacher planning a field trip to a historical site related to the Indian removal policy. Write a proposal explaining why you have chosen this site and what you hope your students will learn from visiting it.
29. Imagine you are an activist advocating for indigenous rights. Write a speech explaining why it is important to acknowledge and address the contemporary implications of the Indian removal policy.

References

[1]: https://www.history.com/this-day-in-history/indian-removal-act-signed-andrew-jackson

[2]: https://www.britannica.com/topic/Indian-Removal-Act

[3]: https://en.wikipedia.org/wiki/Indian_Removal_Act

[4]: https://guides.loc.gov/indian-removal-act

Source: Conversation with Bing, 5/15/2023

(1) American Indian Removal - Smithsonian Institution. https://americanindian.si.edu/nk360/removal/pdf/lesson-0-ataglance.pdf

(2) Indian Removal Act | Definition, History, Significance, & Facts. https://www.britannica.com/topic/Indian-Removal-Act

(3) Indian Removal Act - National Geographic Society. https://www.nationalgeographic.org/thisday/may28/indian-removal-act/

Appendix C
Bing Lesson on Climate Change

Objective: To help students understand the causes and effects of climate change, and the actions they can take to reduce greenhouse gas emissions and adapt to the changing climate.

Materials: Computer with internet access, projector, worksheets, poster paper, markers, scissors, glue.

Duration: Five 45-minute sessions.

Activities:

Session 1: Introduction to climate change

- Start with a bellwork exercise where students write down what they know or think about climate change. Share some responses with the class.
- Show a short video on climate change basics, such as this one: https://www.youtube.com/watch?v=EtW2rrLHs08
- Discuss the difference between weather and climate, and the main factors that influence the Earth's climate system (solar radiation, greenhouse gases, albedo, etc.).
- Introduce the concept of carbon as a key element in climate change, and explain how human activities increase carbon dioxide and other greenhouse gases in the atmosphere.
- Have students complete a worksheet on carbon sources and sinks, and calculate their own carbon footprint using an online tool, such as this one: https://www3.epa.gov/carbon-footprint-calculator/
- Review the worksheet answers and discuss ways to reduce carbon emissions at individual and societal levels.

Session 2: Climate change impacts on physical systems

- Start with a bellwork exercise where students write down some examples of how climate change affects the physical environment, such as glaciers, sea level, oceans, etc. Share some responses with the class.
- Show a short video on climate change impacts on physical systems, such as this one: https://www.youtube.com/watch?v=6UGsRcxaSAI
- Discuss how rising temperatures affect ice sheets and glaciers, and how melting ice contributes to sea level rise and changes in ocean circulation and salinity.
- Have students conduct an inquiry lab on sea level rise using ice cubes, water, food coloring, and graduated cylinders. They will measure how much water level rises when ice melts in fresh water and salt water, and compare their results with real-world data on sea level rise from NASA: https://climate.nasa.gov/vital-signs/sea-level/
- Review the lab procedure and results, and discuss the implications of sea level rise for coastal communities and ecosystems.

Session 3: Climate change impacts on biological systems

- Start with a bellwork exercise where students write down some examples of how climate change affects living organisms, such as plants, animals, humans, etc. Share some responses with the class.
- Show a short video on climate change impacts on biological systems, such as this one: https://www.youtube.com/watch?v=7wGJ4NwSb6k

- Discuss how rising temperatures affect the distribution and phenology of species, and how changes in precipitation and extreme events affect habitats and food webs.
- Have students conduct an inquiry lab on phenology using online databases of plant flowering dates, such as this one: https://www.usanpn.org/data/visualizations
- They will graph the changes in flowering dates over time for different plant species and locations, and compare their results with real-world data on temperature trends from NASA: https://climate.nasa.gov/vital-signs/global-temperature/
- Review the lab procedure and results, and discuss the implications of phenological shifts for biodiversity and ecosystem services.

Session 4: Climate change science and society

- Start with a bellwork exercise where students write down some questions or doubts they have about climate change science or its social relevance. Share some responses with the class.
- Show a short video on climate change science and society, such as this one: https://www.youtube.com/watch?v=nkMIjbDtdo0
- Discuss how scientists collect and analyze data on climate change using various methods and sources, such as ice cores, satellites, models, etc.
- Explain how scientific consensus is reached through peer review and publication of research findings, and how scientific uncertainty is communicated using language of probability and confidence.
- Address some common misconceptions or myths about climate change using evidence-based arguments, such as those from this website: https://skepticalscience.com/
- Have students complete a worksheet on evaluating sources of information on climate change using criteria of credibility, accuracy, relevance, etc.

Session 5: Climate change solutions

- Start with a bellwork exercise where students write down some examples of actions or policies that can help mitigate or adapt to climate change. Share some responses with the class.
- Show a short video on climate change solutions, such as this one: https://www.youtube.com/watch?v=0kx3qjOR9F0
- Discuss the difference between mitigation and adaptation strategies, and the role of different actors and sectors in implementing them.
- Have students work in groups to research and choose one mitigation or adaptation strategy to present to the class using a poster. They can use online sources, such as this one: https://www.drawdown.org/solutions
- They will include information on the benefits, costs, challenges, and feasibility of their chosen strategy, and provide evidence and examples to support their claims.
- Have each group present their poster to the class, and have the audience ask questions and provide feedback.
- Review the main points and lessons learned from the unit, and have students reflect on their own learning and actions regarding climate change.

Source: Conversation with Bing, 4/10/2023
(1) High School Curriculum | Climate Change Education | Stanford Doerr
.... https://earth.stanford.edu/climate-change-ed/curriculum/high
Accessed 4/10/2023
(2) Curriculum | Climate Change Education | Stanford Doerr School of ....
https://earth.stanford.edu/climate-change-ed/curriculum Accessed
4/10/2023.
(3) Climate Change Lesson Plan - University of Guelph.
https://www.uoguelph.ca/oac/system/files/Climate%20Change%20Less
on%20Plan_0.pdf Accessed 4/10/2023.
(4) Climate Change Inquiry Labs - Teacher Guide - NASA.
https://gpm.nasa.gov/education/sites/default/files/lesson_plan_files/cl
imate%20change%20inquiry/Climate%20Change%20Inquiry%20TG.pd
f Accessed 4/10