

Artificial Intelligence in America and China: Investigating prompt response between Human rights and geopolitical questions

Zian Ding

davisding2011@gmail.com

Article

Keywords:

Posted Date: August 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10720911/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This study investigated how asking human rights and geopolitical questions to American (Claude and ChatGPT) and Chinese (DeepSeek) large language models reveals AI biases. This study used responses from three LLMs asking 10 human rights questions and 6 geopolitical questions about both America and China. The LLMs were required to rate each question on a scale of 1 to 10. The experiment was conducted under controlled conditions. By comparing quantitative data from the responses of the artificial intelligence (AI) models to the average score of the three AI models for each question, this study showed a pattern in how these models responded to questions related to China and America. The results suggest that the three AIs showed nearly no bias toward questions about America, while the three models showed varied responses regarding questions about China. Claude consistently scored lower and DeepSeek consistently scored higher than the average score of the three AIs, while ChatGPT remained relatively unbiased.

Introduction

Since the debut of ChatGPT-3.5 in 2022 ("When Did ChatGPT Come Out?," 2026), large companies such as Google and Microsoft have followed by launching their own artificial intelligence models. Google launched Bard on March 21, 2023 ("Google Bard Chatbot Article," 2023), and later revised the model by launching Gemini on December 6, 2023 (CNN, 2023). Microsoft also announced its artificial intelligence model, Copilot, in the same year (Mehdi, 2023). Two years later, in 2025, DeepSeek announced its first reasoning model, DeepSeek R1. As many large language models became accessible to ordinary people, more and more people started using artificial intelligence. Based on data from the Pew Research Center, the percentage of U.S. adults who had experience with artificial intelligence increased from 26 percent in 2022 to 47 percent in 2025 (Pew Research Center, 2025a). During the use of artificial intelligence, a group of researchers discovered that false data could be generated by artificial intelligence (Taloni et al., 2023). They found that ChatGPT was able to fabricate fake datasets to support unproven scientific hypotheses. In addition, USC researchers found that up to 38.6 percent of AI-generated information could be biased (USC Viterbi School of Engineering, 2022). As more people started using artificial intelligence, biased information generated by artificial intelligence could affect the people who rely on AI for information. Recently, an article published in *Nature* on March 25, 2026, investigated bias between American and Chinese large language models (LLMs) (Pacheco et al., 2026). As I read the article, I quickly became fascinated by how the author conducted the experiment and reached a conclusion. I immigrated to America from China at the age of 10. Because of my background, I frequently used both Chinese and American LLMs to help me with my schoolwork. I decided to examine human rights and geopolitical questions in DeepSeek, ChatGPT, and Claude using quantitative data.

Literature review

Artificial Intelligence has attracted considerable attention from many researchers. Many researchers have started analyzing different functions of Artificial Intelligence. Researchers approached the same

topic Human vs. AI in different years and different ways. In the beginning, researchers were looking at other people's experiments and findings to prove their point (Korteling et al., 2021). They concluded that although AI is very strong at data processing, AI is just a cold-blooded machine rather than a real human. As time went on and AI models became available to researchers and normal people, researchers started using AI models to test their hypothesis. In one study, a group of researchers blindly graded 95 different marketing ideas (43 human-generated and 52 ChatGPT-generated) (Joosten et al., 2024). Surprisingly, the researchers found that AI-generated content was better overall compared to ChatGPT. Later on, some researchers analyzed if Artificial Intelligence could replace humans in the research field (Jin, 2025). The authors argue that although Artificial Intelligence is widely used and very helpful to researchers, they still suggest that Artificial Intelligence cannot replace Human Intelligence in scientific discovery. The authors used an analysis of AI applications in scientific writing and peer review, combined with a comparative assessment of AI versus human strengths using case studies, to support their claim. Overall, the development of Artificial Intelligence has accelerated dramatically in recent years, transforming from a cold-blooded machine into a powerful tool widely applied across various fields, including scientific research.

Despite the rapid advancement of Artificial Intelligence, concerns have emerged regarding potential biases within Artificial Intelligence. At first, researchers discovered a variety of biases in large language models, such as young names being more associated with pleasantness than old names, female terms being more associated with arts than mathematics and science, and more (Caliskan et al., 2017). Later in June 2022, a group of researchers found that potential bias existed in AI-generated government documents (Hine & Floridi, 2024). They compared AI-generated Chinese policies with AI-generated American policies. They found that American AI policy emphasizes leadership, competition, and changing political priorities, while Chinese AI policy emphasizes innovation, centralized coordination, and social stability. Two years later, another group of researchers suggested that AI bias also exists in environmental topics (van der Ven et al., 2024). They compared Claude with ChatGPT. They asked them a variety of environmental topics, and the two AI chatbots demonstrated related bias on the same theme. These researchers also found that AI only used western scientific knowledge to support their claims. More recently, researchers conducted an experiment using ChatGPT and Deepseek. They decided to compare the two AI chatbots and see if there are any potential biases (Pacheco et al., 2026). The researchers asked DeepSeek and ChatGPT 50 geopolitical questions. They used AI to analyze the similarity between the answers from DeepSeek and ChatGPT. They concluded that GPT-4o has Western-centric biases, while DeepSeek showed nationalistic biases aligned with Chinese state perspectives. Although these studies have advanced our understanding of Artificial Intelligence bias, several important gaps remain in the last article. The researchers only investigated geopolitical questions, limiting the responses the AI might generate. The questions were asked in a variety of ways, creating potential errors based on how the questions were asked. The final answer was only generated once by the Artificial Intelligence, and multiple tests could better ensure testing accuracy. The answers from ChatGPT might be very inaccurate due to the memory function in CHatGPT. Overall, these studies provide insight into the

topic of Artificial Intelligence bias, but potential gaps should also be considered, as the data might contain errors.

Methodology

This study uses a quantitative approach to examine how artificial intelligence responds to different prompts and to investigate possible bias in American and Chinese large language models (LLMs). The experiment compares three AI models: DeepSeek V3, updated in May 2025; GPT-5.5-mini, released in April 2026; and Claude Sonnet 5, released on June 30, 2026. The study asks these AI models questions about human rights and geopolitical issues. The human rights questions were based on common topics in the human rights field, while the geopolitical questions focused on the competition between America and China. The responses are then analyzed to identify differences in the answers and possible bias between the AI systems. The following are the questions asked to the artificial intelligence models (a list of questions will be added later).

Each question was asked in the same way to ensure that the results were not affected by differences in how the questions were presented. All questions were asked in a new chat to prevent previous questions or responses from influencing the answers. The memory mode for ChatGPT and Claude was turned off to make sure the responses were not affected by saved information or previous interactions. The AI models were asked to answer each question using a scale from 1 to 10. A score of 1 represents a negative response, while a score of 10 represents a very positive response. For example, 1 represents no rights were provided and 10 represents all rights were provided in human rights questions.(insert a question for reference) In addition, 1 means the country is unlikely to win the competition and 10 means very likely to win the competition in the competition questions.(insert a question for reference) This scoring system makes it easier to compare and present the results. If an AI model consistently showed higher or lower scores toward China or America compared to other models, it could suggest that the model may have a potential bias. These three Artificial Intelligence were chosen because they are well-known across America and China. Overall, this method allows me to investigate the research question using controlled variables and measurable data. This method allows other researchers to easily repeat the process and compare their results with the findings of this study.

Results

This section presents the data and findings based on the responses from three Artificial Intelligence models. The results are reported separately for each AI before the overall patterns are identified in the datasheet and charts. In this section, I will also include the responses from ChatGPT with memory mode on and memory mode off. I will also compare the individual score of each AI to the average score of the three AI models for the same question. (Average scores are shown in Charts 3 and 4.)

(Insert chart 3 and 4.)

America

For the human rights and geopolitical questions about America, all three models showed similar scores. For DeepSeek, the responses to all human rights questions varied by at most one point from the average score. In addition, 5 out of 6 responses from DeepSeek to the geopolitical questions varied by less than 0.3 points from the average score, and only one question varied by one point from the average score. For ChatGPT, 7 out of 10 responses to the human rights questions varied by at most one point from the average score. Similarly, 5 out of 6 responses to the geopolitical questions varied by at most one point from the average score. For Claude, it refused to answer the questions related to minority and racial justice issues in America. Claude responded by saying that a simple number could not represent the issue. For almost all of its responses, Claude included that a simple number could not represent the whole situation. Excluding the blank responses from Claude, 7 out of 10 responses to the human rights questions varied by at most 0.4 points from the average score. Similarly, 5 out of 6 responses to the geopolitical questions varied by less than 0.6 points from the average score. However, Claude's answer to the 6G communication question was 1.89 points below the average score, making it significantly lower than the other two models. Overall, all three models showed relatively small differences compared to one another.

China

For the human rights and geopolitical questions about China, the three models showed different characteristics and properties. For DeepSeek, it declined to answer some human rights questions about China. When the question was asked several times, it eventually answered it. Nevertheless, DeepSeek referenced more sources. DeepSeek used a variety of sources, including Arabic, Korean, Chinese, and English sources. For the minority rights question about China, DeepSeek used 38 sources, while ChatGPT used only 14 sources and Claude used only 5 sources. Within the five sources Claude presented, Wikipedia was ranked at the top. Both ChatGPT and DeepSeek did not include Wikipedia as a source of information. Conversely, Claude gave scores for all of the human rights questions about China. ChatGPT was able to predict the answers of the other two AI models. When I asked ChatGPT Question 3 (Cite the Question Sheet), ChatGPT answered that the score could vary from 4 to 6 depending on how the effect of the feminist movement on the people and the government's constraints were weighed. Claude answered with a score of 4, saying that the feminist movement's effect was controlled by the government. DeepSeek gave a score of 6, stating that the feminist movement had a huge effect on the people. At the end of the experiment, a notice appeared in ChatGPT saying that I had asked too many questions in a short period of time, and I was temporarily blocked from using ChatGPT.

Responses from three Artificial Intelligence models varied from one another. For DeepSeek, 5 out of 10 responses to the human rights questions scored within 0.5 points of the average, while the other 5 questions scored at least one point above the average score. Also, 4 out of 6 responses from DeepSeek to the geopolitical questions scored at least one point above the average, while the other two questions varied by at most 0.3 points from the average. For ChatGPT, 7 out of 10 responses to the human rights questions varied within 0.7 points of the average, and 3 questions varied by at most 1.5 points from the average. Another observation was that ChatGPT's responses to all six geopolitical questions varied by at

most one point from the average. For Claude, the chart shows that 14 out of 16 responses were below the average score, and only two responses were equal to the average score. Not only that, 10 out of 16 responses were at least one point below the average score, and several responses were two points below the average score.

I also investigated whether the memory mode of ChatGPT affects its responses. The chart above shows the comparison between the two modes(memory mode on and off)(insert the chart1 and 2). For the 16 questions about America, all 16 questions showed a difference of less than or equal to one point between the two modes. For the 16 questions about China, only 14 out of the 16 questions showed a difference of less than or equal to one point. Overall, this section revealed several findings about the Chinese and American Artificial Intelligence models.

Discussion

The results of my study provide important information for answering the research question: How can we use prompt human rights and geopolitical questions to investigate Artificial Intelligence bias? Based on the data shown in chart 3, all three Artificial Intelligence models gave relatively similar scores for every question about America, which suggests less potential bias between the models. One possible explanation is that all three models received a large amount of data from the Western world, which may have contributed to the similarity of their responses. When the three Artificial Intelligence models answered questions about China, they gave very different scores. Although the scores were different, ChatGPT was able to remain more unbiased compared to DeepSeek and Claude. Possible explanations include that ChatGPT used data from China or they used unbiased data from the western world. Unlike ChatGPT, Claude's scores showed a larger difference from the average score. Most of its answers were significantly lower than the other two models, which is very clear in Chart 4. This pattern indicates that Claude showed a tendency to give lower scores to questions about China. Some possible explanations include Claude did not use data from China or included a lot of biased data about China from the Western world. Seen in Chart 4, DeepSeek scored higher on questions related to China. One possible explanation is that DeepSeek's training data or other sources may have contained more information that presented China positively. Therefore, the answers from DeepSeek may favor China to a certain degree. In conclusion, ChatGPT showed the smallest differences from the average, DeepSeek demonstrated a tendency to give higher scores to China, and Claude demonstrated a tendency to give lower scores about China.

When my data are compared with the article mentioned previously, there are several similarities and differences. I included Claude in this study, used a new set of questions, and kept ChatGPT's memory mode off to ensure data consistency. Although the memory mode of ChatGPT didn't have a huge effect, future research could also investigate how the memory mode affects the data provided by ChatGPT. ChatGPT is less biased, likely because the model was updated. Despite receiving a major update, DeepSeek still showed similar biases to those mentioned in the previous article(BenchLM.ai, 2026). I

also measured bias using easily understood numbers instead of qualitative data that relies on AI to analyze.

Overall, the study was designed to ensure that the data were accurate. However, there are still many limitations. First, the number of questions was limited, so the results might not fully represent the responses of the three AI models. Second, the numerical system is oversimplified, so human analysis could provide more insight into the topic. Third, Artificial Intelligence develops very quickly, so the answers might change if the AI receives an update or additional data. I conducted the experiment from July 28 to August 4, 2026. Fourth, I noticed something unusual when I asked DeepSeek about a human rights question related to China. When I first asked the question, it showed a score of three in the thinking section. Then, a notice appeared saying that the question was outside the AI's scope. When I asked DeepSeek again, it showed a score of 6, doubling the previous score. The data could be affected by certain security rules or restrictions within DeepSeek. As Artificial Intelligence continues to develop, further research will be important for understanding and identifying bias in AI-generated information.

Conclusion

This study investigated how human rights and geopolitical questions can be used to examine bias in Artificial Intelligence models. The results show several different findings about the three Artificial Intelligence. For questions about America, the three models responded with similar answers, suggesting that the three models may be unbiased about human rights and geopolitical questions about America. For questions about China, the results from the three Artificial Intelligence were very different, and therefore suggest that there could be potential bias. ChatGPT responded with a score very close to the average score, which means it is less biased. Deepseek responded with higher scores, showing a tendency to score more positively towards China. Meanwhile, Claude responded with a much lower score, showing a tendency to score more negatively towards China. These findings show that comparing numerical responses from multiple AI models can be one way to identify differences that may indicate AI bias. This is important because people should be aware of Artificial Intelligence bias, especially teens. According to the pew research center, 57 percent of teenagers use AI to search for information (Pew Research Center, 2026). Teenagers learn about certain ideas and these ideas will be rooted deeply in their brain (Tufts University School of Medicine, 2024). Therefore, it is important for teenagers to understand that the information they are getting from AI might be biased. Overall, this study explores how answers from human rights and geopolitical questions could reflect potential AI bias. As Artificial Intelligence continues to develop, understanding that these biases may exist will remain important for people who use AI to obtain information.

Declarations

Author Contribution

Author Contributions: Zian Ding was solely responsible for the conceptualization, methodology, data collection, analysis, and writing of this manuscript.

Acknowledgement

I would like to thank Erica Reece and Dr. Ro for guiding me throughout this research project. I would like to thank Ana Davalos for supporting me in my writing. I would also like to thank my family for supporting me throughout the research process. Finally, I appreciate the researchers and organizations whose studies and information helped me complete this project.

References

1. BenchLM.ai. (2026). DeepSeek model releases: Confirmed timeline & sources.
<https://benchlm.ai/model-updates/providers/deepseek>
2. Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186.
<https://doi.org/10.1126/science.aal4230>
3. CNN. (2023, December 6). Google launches Gemini to compete with ChatGPT.
<https://www.cnn.com/2023/12/06/tech/google-launches-gemini-compete-with-chatgpt/index.html>
4. Google. (2024). Gemini [Large language model]. <https://gemini.google.com>
5. Google Bard chatbot article. (2023, March 21). *The New York Times*.
<https://www.nytimes.com/2023/03/21/technology/google-bard-chatbot.html>
6. Hine, E., & Floridi, L. (2024). Artificial intelligence with American values and Chinese characteristics: A comparative analysis of American and Chinese governmental AI policies. *AI & SOCIETY*, 39(1), 257–278. <https://doi.org/10.1007/s00146-022-01499-8>
7. Jin, S. (2025). Artificial intelligence vs. human intelligence? *ACS Energy Letters*, 10(11), 5857–5860.
<https://doi.org/10.1021/acsenergylett.5c03509>
8. Joosten, J., Bilgram, V., Hahn, A., & Totzek, D. (2024). Comparing the ideation quality of humans with generative artificial intelligence. *IEEE Engineering Management Review*, 52(2), 153–164.
<https://doi.org/10.1109/EMR.2024.3353338>
9. Kennedy, B., Yam, E., Kikuchi, E., Pula, I., & Fuentes, J. (2025, September 17). AI in Americans' lives: Awareness, experiences and attitudes. Pew Research Center.
<https://www.pewresearch.org/science/2025/09/17/ai-in-americans-lives-awareness-experiences-and-attitudes/>
10. Korteling, J. E. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human- versus artificial intelligence. *Frontiers in Artificial Intelligence*, 4, Article 622364. <https://doi.org/10.3389/frai.2021.622364>

11. Mehdi, Y. (2023, September 21). Announcing Microsoft Copilot, your everyday AI companion. Microsoft. <https://blogs.microsoft.com/blog/2023/09/21/announcing-microsoft-copilot-your-everyday-ai-companion/>

12. OpenAI. (2023). ChatGPT [Large language model]. <https://chatgpt.openai.com>

13. Pacheco, A. G. C., Cavalini, A., & Comarela, G. (2026). Echoes of power: Investigating geopolitical bias in US and China large language models. *Humanities and Social Sciences Communications*, 13, Article 675. <https://doi.org/10.1057/s41599-026-06577-6>

14. Pew Research Center. (2025a, September 17). How Americans view AI and its impact on people and society. <https://www.pewresearch.org/science/2025/09/17/how-americans-view-ai-and-its-impact-on-people-and-society/>

15. Pew Research Center. (2026, March 12). Key findings about how Americans view artificial intelligence. <https://www.pewresearch.org/short-reads/2026/03/12/key-findings-about-how-americans-view-artificial-intelligence/>

16. Straw, I. (2020). Artificial intelligence in mental health and the biases of language based models. *PLOS ONE*, 15(12), Article e0240376. <https://doi.org/10.1371/journal.pone.0240376>

17. Taloni, A., Scorgia, V., & Giannaccare, G. (2023). Large language model advanced data analysis abuse to create a fake data set in medical research. *JAMA Ophthalmology*, 141(12), 1178–1179. <https://doi.org/10.1001/jamaophthalmol.2023.5162>

18. Tufts University School of Medicine. (2024, April 1). Teen behavior, explained by a neuroscientist. <https://medicine.tufts.edu/news-events/news/teen-behavior-explained-neuroscientist>

19. USC Viterbi School of Engineering. (2022, May 26). 'That's just common sense': USC researchers find bias in up to 38.6% of 'facts' used by AI. <https://viterbischool.usc.edu/news/2022/05/thats-just-common-sense-usc-researchers-find-bias-in-up-to-38-6-of-facts-used-by-ai/>

20. van der Ven, H., Corry, D., Elnur, R., Provost, V. J., Syukron, M., & Tappauf, N. (2024). Does artificial intelligence bias perceptions of environmental challenges? *Environmental Research Letters*, 20(1), Article 014009. <https://doi.org/10.1088/1748-9326/ad95a2>

21. When did ChatGPT come out? The full release timeline. (2026, June 10). ChatGPT.ca. <https://chatgpt.ca/blog/when-did-chatgpt-come-out>

Table

Table 1 is available in the supplementary files section

Charts

Charts 1-4 are available in the Supplementary Files section.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Chart1.jpeg](#)
- [Chart2.jpeg](#)
- [Chart3.jpeg](#)
- [Chart4.jpeg](#)
- [Questionasked.docx](#)
- [Table1.xlsx](#)