

# New research finds biases encoded in language across cultures and history

June 14 2024

---



Credit: CC0 Public Domain

In a new study published in *Social Psychological and Personality Science*, researchers share evidence that people's attitudes are deeply woven into language and culture across the globe and centuries.

The researchers looked at connections between people's attitudes and [language](#) from 55 different topics like rich vs. poor, dogs vs. cats, or love vs. money. They used four text sources: current English writing and text, English books going back 200 years, and texts in 53 languages other than English. As a measure of people's attitudes, they used data from over 100,000 Americans; first, direct self-reports; and second, an indirect measure based on people's reaction times, often referred to as implicitly-measured attitudes.

They found that the associations picked up by large AI language models like ChatGPT match more closely with the second indirect measure rather than the attitudes they explicitly state.

"With the rise of AI and large language model applications, we as consumers, leaders, researchers, or policymakers need to understand what these models are representing about the social world," says lead author Dr. Tessa Charlesworth, of Northwestern University's Kellogg School of Management. "Do they have obvious, explicit preferences? Or do they have more hidden patterns of associations more akin to implicitly-measured attitudes?"

Mitigating these subtle biases in AI will require different approaches than looking for explicit biases. "Rather than auditing the models at the end to see if they show obvious, explicit [bias](#), we will likely need to dig deeper into the patterns in the [training data](#) itself and provide alternative examples of associations," says Dr. Charlesworth.

More broadly, "The data show that implicitly-measured attitudes are revealed in and perhaps reinforced by language, which is a key vehicle of transmitting culture," notes Dr. Charlesworth. As such, "If we want to durably address and reduce implicit bias in society, we will likely need interventions that adopt a more cultural (or macro level) focus."

While emphasizing the correlational nature, the researchers aim to continue exploring sociocultural influences. "Given that we saw some variation in which of the non-English languages showed the correlation, it is important to understand what kind of social and [cultural factors](#) could help explain greater transmission between bias and language," says Dr. Charlesworth.

The study lays the groundwork for a better understanding of the subtle ways in which attitudes become entangled with the systems of language and communication, enveloping us—both today and echoing through centuries past.

**More information:** Echoes of Culture: Relationships of Implicit and Explicit Attitudes With Contemporary English, Historical English, and 53 Non-English Languages, *Social Psychological and Personality Science* (2024). [DOI: 10.1177/19485506241256400](https://doi.org/10.1177/19485506241256400)

Provided by Society for Personality and Social Psychology

Citation: New research finds biases encoded in language across cultures and history (2024, June 14) retrieved 24 June 2024 from <https://phys.org/news/2024-06-biases-encoded-language-cultures-history.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--