Are 'person' or 'people' gender-neutral concepts? New study finds male tilt in analysis of billions of words
1 April 2022

"Our findings show that even when using gender-neutral terms, we prioritize men over women," adds co-author Adina Williams, a research scientist at Meta AI and a graduate of NYU's doctoral program in linguistics.

Bias at such a foundational level—our word choices—is potentially consequential, the researchers note.

"Conceptions of 'people' form the basis of many societal decisions and policymaking," observes Andrei Cimpian, a professor in New York University's Department of Psychology and the senior author of the paper. "Because men and women are each about half of the species, prioritizing men in our collective idea of a 'person' creates inequity for women in decisions based on this idea."

The research team examined words' meanings by considering how they are used by individuals. Specifically, the team studied how we use words expressing the concept of "person" and its gender-specific counterparts, "woman" and "man."

To test whether we're likely to think of men more often than we are of women when writing about "people," the team used artificial intelligence algorithms that learn the meaning of words based on how they are used, drawing from a language repository collected by the non-profit Common Crawl in May 2017. This repository included more than 630 billion mostly English-language words appearing on nearly three billion web pages.

The researchers considered how word meaning is related to word context and use. For example, if you hear, "Each morning, Joe boiled water in the balak for tea," you might guess that "balak" means something similar to "kettle," even though "balak" is
unfamiliar, because the words alongside "balak" ("tea," "boiled," and "water") also frequently co-occur with "kettle."

In the Science Advances paper, the researchers investigated, in three studies, the meaning of "person" and related words (e.g., "people") by taking into account adjacent words—the linguistic context.

In the first study, they compared the similarity in meaning (inferred via linguistic context) between words for people (e.g., "individual") and words for men (e.g., "he" and "male") to the similarity in meaning between words for people and words for women (e.g., "she" and "female").

They found that words for people were used more similarly, and were thus more similar in their meaning, to words for men than to words for women—and by a statistically significant margin. Put another way, the collective concept "people" overlapped more with the concept "men" than with the concept "women" in the studied words.

In the second study, instead of focusing on words for people, the team examined words denoting features central to this concept—specifically, words for traits that commonly describe what people are like. They compared hundreds of trait words identified in past research as common descriptors of people (e.g., "extroverted," "analytical," and "superstitious") to the same lists of words for women and for men from the paper's initial study.

They found that the meaning of these descriptor words in the second study were, overall, more similar to the meaning of words for men than to the meaning of words for women, with a statistically significant difference between the two. That is, common words that describe what people are like (e.g., "extroverted") are also used more similarly to words for men than to words for women.

In a third study, the researchers studied the use of verbs—a reasonable area for exploration given the initial findings. Specifically, if the collective concept "people" overlaps more with the concept "men" than with the concept "women," then words that describe what people do and what is done to them (e.g., "love," "annoy") may also be more likely to be similar in their contextual meaning to words denoting men than to words denoting women.

As with the second study, which focused on common words that describe what people are like, words that describe what people do (e.g., "run") were also used more similarly to words for men than to words for women—a difference that was again statistically significant.


Provided by New York University