

Hate speech on social media is fueled by users' shared moral concerns

December 16 2021



Credit: CC0 Public Domain

People whose moral beliefs and values align closely with other members of their online communities—including those on social networks Gab and Reddit—are more prone to radicalization, according to new USC

research.

Researchers at the USC Dornsife College of Letters, Arts and Sciences had theorized that a high level of consistency in similar moral concerns within online communities is linked to an increase in radical intentions—that is, readiness to participate in illegal or violent political action.

In new research published this week in *Social Psychological and Personality Science*, they found the degree of shared moral concerns or "moral convergence" within an online cluster predicts the number of hate-speech messages posted by members.

"Our research team has looked at how morality motivates people to engage in various types of behavior, from donation during a disaster to taking extreme actions, even violence, to protect their group," said study lead author Mohammad Atari, who recently defended his Ph.D. in the department of psychology at the USC Dornsife College and is now a postdoctoral fellow at Harvard University. "They feel like others are doing something morally wrong and it's their sacred duty to do something about it, even if that means posting hate speech and committing hate crimes."

Scientists first analyzed posts on an alternative social media network popular with the alt-right and right-wing extremists called Gab. The platform, which claims to champion free speech and isn't moderated for hate speech, provided the researchers with a unique opportunity to investigate the dynamics that could lead to radicalization.

They found Gab users who had a similar moral profile with their immediate group—meaning they felt similarly about core moral issues including care, fairness, loyalty, purity and authority—were more likely to disseminate hate speech and use language intended to dehumanize or

even call for violence against outgroup members.

The researchers replicated the observations in the Gab study by looking at another extremist network in the online community Reddit. They analyzed a subreddit called "Incels"—involuntary celibate men who blame women for their inability to find sexual partners—and found those who were more like-minded in their morality produced more hateful, misogynistic speech.

Working together, scientists at USC and other institutions a few years ago developed a model for detecting moralized language. It's based on a prior, deep-learning framework for a computer program that can reliably identify text evoking moral concerns associated with different types of moral values and their opposites. The values, as defined by the Moral Foundations Theory, are focused on care/harm, fairness/cheating, loyalty/betrayal, authority/subversion and purity/degradation.

Moral Foundations Theory is a social and cultural psychology theory that explains the evolutionary origins of human moral intuitions based on innate, gut feelings rather than logical reasoning.

"Morality binds us together and gives our society structure and direction for taking care of those in need, and a vision for a just and prosperous future for the group. But morality also has a dark side, in that extreme forms of it can lead to the opposite of a lot of these positive principles," said Morteza Dehghani, an associate professor of psychology and computer science. He leads USC Dornsife College's Computational Social Science Lab, where he, Atari and others investigate how morality intertwines with prejudice and hate.

Social media platforms allow extremists to find each other and, as Dehghani describes, "feed each other's visions of the world and anger towards the outgroup."

Experimental studies further revealed the role of morality

In three controlled experimental studies, the research team further demonstrated that leading people to believe that others in their hypothetical or real group shared their views on moral issues increases their radical intentions to protect the group at any cost, even by resorting to violent means. When U.S. study participants were led to believe that other Americans shared their moral views, they became more willing to "fight and die" for their country and the values it stands for.

"These findings highlight the role of moral convergence and familial-like bonds in radicalization, emphasizing the need for diversity of moral worldviews within social networks," said Atari.

But, he acknowledged, that's easier said than done. More study is needed to determine the most effective interventions for online communities to introduce different views, which may hold the key to stopping radicalization.

'Stop the Steal' had roots in online radicalization

The real-world threats posed by online radicalization were recently illustrated by the Jan. 6 storming of the U.S. Capitol. Those who were convinced the 2020 presidential election had been stolen from former President Donald Trump organized online under the hashtag #StoptheSteal on Facebook and on Gab, which served as a hub for organizing the insurrection.

These radicalization studies were already well underway before the Jan. 6 insurrection. Even so, Atari said the events of Jan. 6 further motivated the research team that had been trying to understand online

radicalization.

He added that identifying as conservative or liberal does not necessarily predict who is predisposed to radicalization. "When people are motivated by morality, regardless of their political affiliation, it clouds their judgment," Atari said.

More information: Mohammad Atari et al, Morally Homogeneous Networks and Radicalism, *Social Psychological and Personality Science* (2021). DOI: 10.1177/19485506211059329

Provided by University of Southern California

Citation: Hate speech on social media is fueled by users' shared moral concerns (2021, December 16) retrieved 2 May 2024 from <https://phys.org/news/2021-12-moral-echo-chambers-social-media.html>

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.