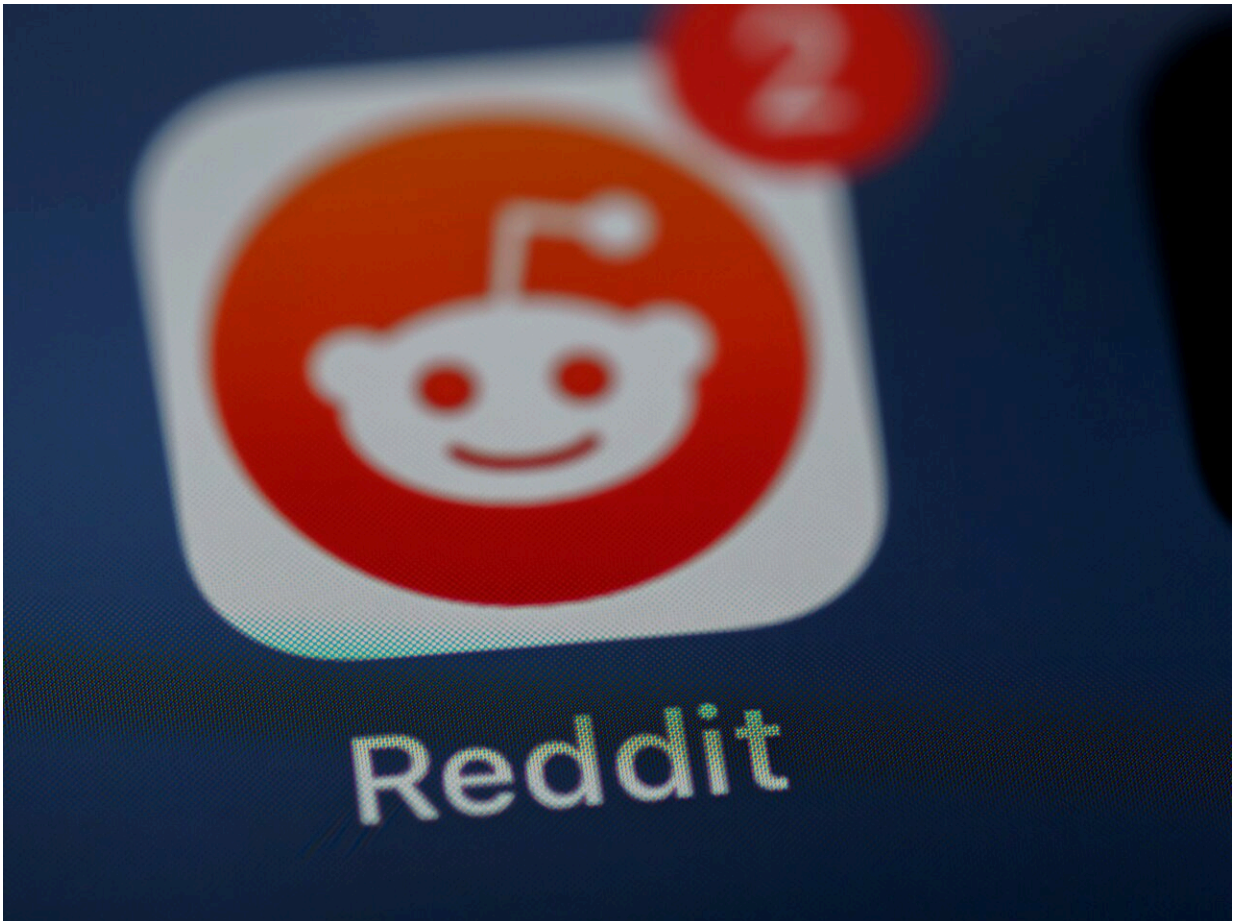![PHYS.ORG]

# New research finds that social group values are most readable in the way we write

December 12 2022



Credit: Unsplash/CC0 Public Domain

Analyzing the style of language used by social groups could offer insight

into their values and principles that goes beyond what they publicly say about themselves.

New research has found that the way in which groups communicate online reveals how the group views itself—and by looking at how group members write, it is possible to track how a group's values change over time.

The work, conducted by psychologists at the Universities of Exeter, Bath and Lancaster, could offer social scientists a new method of studying group behavior and social dynamics. It might also help authorities to address harmful behaviors such as the spreading of conspiracy theories or to better predict when groups are at risk of radicalization or criminal behavior.

"The study of group dynamics often involves asking people to tell us about their group—for example, what does being an England football supporter mean to you?" says Dr. Alicia Cork from the University of Bath. "The drawbacks of this, in simple terms, is that these questions can be quite ambiguous to answer, and it is tricky to articulate what particular group memberships may mean to a person. However, when you concentrate on the style of communication used by group members, you can access the overarching principles that guide the behavior of those members."

The Ph.D. research was conducted in three distinct stages. In the first, the academics tested whether similar groups communicated in similar ways. Based on five existing group types—vocational, political, religious/ethnic, relational, or stigmatized—they selected three communities on the online forum Reddit for each type.

For example, for the relational type, they chose communities specific to fatherhood and motherhood, but also a more general one focused on

relationships; for the political type, they selected a [political party](#) (Conservative), a social movement (Feminist) and a shared ideology (Libertarian); and for the stigmatized type, they selected a [transgender community](#), one focused on [alcohol abuse](#) and one discussing homelessness.

After gathering one year's worth of Reddit comments from the 15 online communities, the researchers ran the words through specialist software, which computed the writing style of each group. This took into account elements such as formality, or how emotive/reflective the language was. They then compared the groups' writing styles to understand how similar or dissimilar they were.

Plotting their relative positions on a graph, they found that groups clustered closely with their counterparts, suggesting that similar groups do in fact communicate in similar ways. The exception to this was the communities related to stigma, which showed greater variation in the way they communicated.

In the second phase, the team sought to more closely map the way a group communicates with the explicit values that they hold. Analyzing the posts against ten well-established values such as benevolence, security, conformity, power and achievement, they were able to build a picture of how the groups most likely described themselves. They then compared this to their relative clustering position.

"The results indicated that our vocational identities such as entrepreneurs and white-collar salespeople expressed high levels of achievement but fairly little benevolence," said Dr. Miriam Koschate-Reis, Associate Professor of Computational Social Psychology, from the Institute for Data Science and AI (IDSAI) at Exeter.

"Groups that expressed benevolence were consistent with the relational

types—mothers, fathers, and those focused on relationships. Interestingly, we also found that while the political and religious/ethnic groups tended to espouse conformity, the latter tended to express greater benevolence in their language."

The final stage of the project investigated whether it was possible to use this linguistic technique to chart the evolution and development of a group's collective self-understanding over time. Focusing upon the transgender identity, they analyzed Reddit data going back to 2011, and found that it had migrated away from fellow stigmatized groups (communities discussing homelessness or alcohol abuse), towards those defined more by the theme of collective action, such as those in the political and religious/ethnic groups.

"The changes we observe in the linguistic style of this transgender group appears to correspond with the gradual politicization and de-stigmatization of that group," said Prof Mark Levine, of the Department of Psychology at Lancaster. "And this is consistent with what we know of the development of LGBTQ+ rights over that period."

The researchers say that this linguistic method could enable better identification of emerging political movements, such as those that start out as mutual support groups but then become radicalized. It also enables them to learn about groups that may wish to disguise their true nature such as criminal organizations hiding in plain sight on online forums.

"Understanding what groups stand for is integral to a diverse array of social processes, ranging from understanding political conflicts to promoting public health behaviors," adds Prof Richard Everson, a Professor of Machine Learning at the University of Exeter. "The results from these three studies provide compelling evidence that suggests that linguistic style can be used to understand the collective self-understanding of social groups."

"Collective self-understanding: A linguistic style analysis of naturally occurring text data" has been published in the journal *Behavior Research Methods*.

Provided by University of Exeter