The Polarizing Impact of Continuous Presence on Users’ Behavior

Wienie Strathern,1 Angelina Mooseder,1 Jiurgen Pfeffer1

1School of Social Sciences and Technology, Technical University of Munich
wienke.strathern@tum.de, angelina.mooseder@tum.de, juergen.pfeffer@tum.de

Abstract
Online political communities offer spaces where people share and exchange political views, content, and ideas. However, people seeking political exchange online are increasingly confronted with an exclusionary intensification of discussion that no longer makes it possible to participate in constructive discourse. Online political discussion forums like r/The Donald and r/ChapoTrapHouse on Reddit have been banned recently due to the proliferation of hate speech and anti-social behavior of users. Homogeneous online discussion communities have been shown to play a key role in political polarization. Spending time in these communities tends to foster strong political positions associated with conflict. In this study we analyzed social media data from online political communities on Reddit and asked: how does presence in polarizing environments influence users’ behavior? With initially equal user activity, what happens if one group of users continues to be present while the other is not? Our analysis show that continuously present users become more active, use simpler vocabulary, and employ more abusive words in their text contributions. Our results have implications for automated moderation of polarizing online communities.

1 Introduction
Millions of users actively participate in online communities and social media networks offering a platform for exchanging views on any political topic. The platforms make it possible for people to network and find a community that matches their interests, convictions and political views. By offering them a space to exchange opinions, social media operators had high hopes for these platforms. However, negative online effects such as hate speech, anti-social behavior by users and offline spillover effects such as the storming of the US Capitol in January 2021, which were fomented on politically active social media forums are studied (Gallacher, Heerdink, and Hewstone 2021). The often unrestricted and anonymous environment of online discussions can become a platform for anti-social behavior, such as online abuse or harassment (Walther 2022). The drifting apart of political opinions and attitudes on platforms has been much discussed recently due to the proliferation of hate speech and anti-social behavior of users. Homogeneous online discussion communities have been shown to play a key role in political polarization. Spending time in these communities tends to foster strong political positions associated with conflict. In this study we analyzed social media data from online political communities on Reddit and asked: how does presence in polarizing environments influence users’ behavior? With initially equal user activity, what happens if one group of users continues to be present while the other is not? Our analysis show that continuously present users become more active, use simpler vocabulary, and employ more abusive words in their text contributions. Our results have implications for automated moderation of polarizing online communities.

Our results show that continuously present users become more active, use simpler vocabulary, and employ more abusive words in their text contributions. Our analysis show that continuously present users become more active, use simpler vocabulary, and employ more abusive words in their text contributions. Our results have implications for automated moderation of polarizing online communities.

Research Questions
Reddit is a community driven discussion platform for political and polarizing discourse (Kane and Luo 2019; Phadke, Samory, and Mitra 2021). In a variety of political subreddits like-minded people come together to discuss current political topics. Given its openness and speed of communication with unrestricted ways of expression, discussions can be polarizing, heated, engaging, controversial, productive, creative, diverse in their facets. Due to its speed, its unrestricted flow of information and the constantly high number of contributions, users, if present, are continuously exposed to a variety of topics. Almost half of the Reddit users are heavy Internet users1, studies show that even short time effects can give first initial insights on changing behavior (Zhou et al. 2021). The question arises as to what presence in a polarizing environment does to people linguistically? Does it change the way users behave? Does it change the way users write? Regarding theories on polarization and radicalization in political settings (Grover and Mark 2019) we assume that prolonged dwelling in these polarizing environments affects users’ behavior in terms of activity, thus interactions. Furthermore, if we assume that polarizing environments promote homogeneity, we expect this to be reflected in linguistic patterns of complexity. With respect to studies of online radicalization in polarizing environments, we assume that people show differences in behavior and that it is reflected in their textual contributions.

Methods
To answer our research question methodologically, we ask: given the same level of activity, what happens if one group of users continues to be present while the other is not? In order to assess differences in behavior we measure interactions and text contributions, apply a linguistic metric

to determine the level of lexical diversity and apply a metric to determine the level of profane word usage in users’ posts and comments.

To test our hypothesis we analyzed social media data of three famous online political communities on Reddit. We used data from three main political subreddits alongside the political spectrum to test our hypothesis: r/The_Donald, a right-wing subreddit, r/politics, a forum for all users interested in politics, and r/ChapoTrapHouse, a left-wing subreddit. The explosive nature of the topic is shown by the fact that two of the subreddits we investigated, r/ChapoTrapHouse and r/The_Donald, were banned, i.e., the discussion site was dissolved, due to hate speech, anti-social behavior, and violations against Reddits hate policy. The user communities for these subreddits generated an extensive chat history of roughly 17 million posts and comments. Since these are structurally large communities, we want to know whether our hypothesis applies to both the right and the left, as well as to the moderate spectrum.

We compared the behavior of users who were continuously active to the behavior of similar users who were non-active for at least 14 days to determine 1) whether users change in activity and language and 2) whether this change is correlated with absence/non-activity. To answer our research question, we a) defined groups and b) measured posting and commenting activity for each group, the level of lexical diversity and profane word usage in their text comments. By comparing two groups and their linguistic attributes we were able to extract data-driven insights about the users’ activity and linguistic changes in three online political communities. In addition, we took advantage of naturally occurring variation in the degree of overall activity of users—i.e., how frequently users posted over the time period of six months—to analyze whether behavioral changes were stronger in people who are generally more engaged in a subreddit.

**Contributions.** Based on our analysis, we derive the following conclusions:

1. Continuously present people increase their activity, show more contributions.
2. Continuously present people decrease their lexical diversity, use simpler vocabulary.
3. Continuously present people increase their profanity, use more offensive language.
4. Continuously present users who are in general more active show greater behavioral changes regarding activity and lexical diversity.

To summarize, we observe that users who are continuously present on Reddit change in activity, vocabulary and emotions, while users who are inactive for a certain time show the same level of activity, vocabulary and emotions after their absence as before.

---

2 **Related Work**

Our study is motivated by work discussing polarizing environments on social media (2.1) measuring activity and language (2.2) to assess psycho-linguistic behavioral attributes.

2.1 **Polarizing Online Environments**

The unrestricted way of freedom of expression in online environments can lead to hate speech and discrimination of marginalized groups, hence causing polarizing effects. A substantial body of work explores ideologically cross-cutting discussion spaces and the role of non-political spaces for political discussion. In this work the authors examine the fact that online news consumption follows a polarized pattern, observing that users’ visits to news sources aligned with their own political leaning are substantially longer than their visits to other news sources (Garimella et al., 2021). Another study focused on identifying high and low consensus news posts on Twitter and presenting a method to automatically detect them (Babaei et al., 2018). In another study cross-cutting posting was tested. This determined that people are more likely to accept a news article containing conflicting views when it is delivered by a chatbot (Zarouali et al., 2021). A focus of this work was on news-link sharing. It shows that Reddit users’ voting and re-sharing behaviors generally decrease the visibility of extremely biased and low factual content (Weld, Glenski, and Althoff, 2021). In this paper the authors show that political conversations are less toxic in non-political subreddits (Rajadesingan, Budak, and Resnick, 2021).

Several works study the effects of social media use. For example, they examine the temporary lapse from social media platforms on behavior (Allcott et al., 2020; Kovacs, Wu, and Bernstein, 2021; Brown, 2020) and the effects of how to make online spaces more civil (Wadden et al., 2021). Reddit has become a famous community discussion platform for political discourse. Reddit data is used to measure similarity in the commenting user bases of communities (Mamí, Horta Ribeiro, and West, 2021). Another study discovers language biases encoded in the vocabulary of online discourse communities on Reddit (Ferrer et al., 2021). A study was conducted on analyzing online news sharing at scale to study bias a factual news on Reddit (Weld, Glenski, and Althoff, 2021). This work examines the social makeup of online communities to understand the social organization of online platforms on Reddit (Waller and Anderson, 2021).

2.2 **Measuring activity and linguistic patterns**

User behavior plays an important role in understanding social media platform effects. In their analysis (Jhaver et al., 2019) the authors characterized the removal explanations that are provided to Redditors and linked them to measures of subsequent user behaviors. In order to better understand political engagement, this study analyzes the political interaction network on Reddit contradicting the echo chamber narrative (De Franciscis Morales, Monti, and Starnini, 2021). A further paper addresses the social effects of content ratings on Reddit (Davis and Graham, 2021). Another study observed Redditors behavior in terms of how they interact with...
We describe the dynamics of three political subreddits, the dataset studied in this work, and basic preprocessing steps to filter out low-quality data.

In its structure and function, Reddit 4 contains elements of a discussion forum, a social network, and a news service. Users submit posts, called submissions. These can be original content, links to external content, or a combination of both. Much of the content is linked to other websites. Other users can then add comments to a post. In addition to commenting on posts, users can also rate posts and comments by upvoting them (meaning they are worthy of being seen by others) or downvoting them (meaning they should not be seen). This voting controls the display of posts and comments on the website. Overall, Reddit is divided into subreddits, which are roughly equivalent to forums or topics on other online message boards. The names of these subreddits usually describe the topic discussed (e.g., r/politics). A Pew Research Center publication gives information on the demographics of Reddit users: While only 4% of U.S. adults report using Reddit, about seven in ten of those users (78%) obtain news from the site. Overall, 2% of U.S. adults obtain news on Reddit. Both Reddit users in general and those who source news from the site tend to be young, male, and more likely to describe themselves as more liberal than the general population. About seven in ten (71%) of Reddit news users are men, 59% are between 18 and 29 years old, and 47% describe themselves as liberal, while only 13% are conservative (39% describe themselves as moderate). By comparison, of all adults in the U.S., about half (49%) are men, only 22% are between the ages of 18 and 29, and about a quarter (24%) describe themselves as liberal. Reddit news users are also heavy Internet users, with 47% reporting being online almost constantly (compared to 21% of U.S. adults overall). In this study we are interested in polarizing open online discussion forums. We would like to understand interactions and linguistic patterns. The platform offers an open space, open access and freedom of speech. There is no hierarchy – we find polarizing environments and user affiliation. Furthermore, we find informal language, a huge amount of contributions, few rules, colloquial language, and anonymity which allows for provoking behavior. Thus, the dataset is appropriate for the study of online user behavior. 5

3 Data

3.1 Dataset description

Using the Pushift Dataset (Baumgartner et al. 2020), we collected posts and comments from 146 political subreddits from January 2018 to August 2019. Fig. 1 shows the number of posts and number of comments for these subreddits (Pearson’s $r = 0.75, p = 0.0$), with both axes being log-scaled. We choose the three subreddits with the highest posting activity for our analysis: The left-wing community r/ChapoTrapHouse, the right-wing community

---

4https://www.redditinc.com/
Donald, as well as r/politics, which can be described as a meeting place for politically interested users.

**Co-occurrence network.** An additional co-occurrence analysis of users in the political subreddits confirms the structural positions of these three subreddits. Fig. 2 is based on the original two-mode network data of a user being connected to a subreddit in case the user had posted in a particular subreddit. The subsequent transformation into a one-mode network of subreddits that are connected by shared users was then reduced for the visualization to show only the top five connections for every subreddit. The use of Visone’s (Baur et al. 2002) backbone visualization procedure can reveal the different political areas in Reddit’s political discussion forums. r/The_Donald can be found in an area of the network dominated by conservative subreddits and r/ChapoTrapHouse is densely connected through shared users with a set of liberal subreddits. The central position of r/politics results from the fact that this forum is populated by users from a wide range of political views.

**Preprocessing.** For each of these subreddits, we stored all comments and posts and removed all duplicated posts, as well as posts where the author information was missing or where the contribution was created by the AutoModerator instead of a real person. While posts could include a title as well as text, comments only included text. For each post, we combined the title and text into a text field and saved the author, the day of creation and the combined text. For each comment, we saved the author, the day of creation and the text. In the following analysis we did not differentiate between comments and posts but regarded them as objects from the same type (namely contribution). For the purpose of this study, we have limited the time frame to six months and used the Reddit data from January 1 to June 30 of 2018. A brief overview of the remaining data is given in Table 1.

<table>
<thead>
<tr>
<th>subreddit</th>
<th>Users</th>
<th>Comments</th>
<th>Posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>r/The_Donald</td>
<td>126,217</td>
<td>5,204,877</td>
<td>478,614</td>
</tr>
<tr>
<td>r/politics</td>
<td>368,056</td>
<td>11,081,988</td>
<td>213,484</td>
</tr>
<tr>
<td>r/ChapoTrapHouse</td>
<td>18,719</td>
<td>974,558</td>
<td>58,286</td>
</tr>
</tbody>
</table>

Table 1: Dataset details

**4 Study Design**

Our study is motivated by the assumption that continuously present users in polarizing environments change their behavior. Continuous interactions make the user more exposed to the existing community and less to other influences, thus being repeatedly exposed to the same opinions which reinforces existing behavior and enhances further activity. We define presence as active interactions in terms of posts and comments contributions, absence in turn is defined as no activity at all for at least 14 days. Based on these assumptions we compare the behavior of present users with the behavior of similar users with absence to determine 1) whether users change with respect to their activity and language and 2) whether this change is correlated with absence.

**4.1 Setup**

The setup of this study is depicted in Fig. 3. We have built two groups showing a similar activity level in the beginning of our study (see time frame A). While the group of present users is present during the whole analysis period, the group of absent users is absent for at least 14 days (see time frame B). We then compare how each group has changed, by comparing the behavior in time frame C to the behavior in time frame A. In the following the selection process for the present and absent group is described in more detail.

**4.2 Defining groups**

We define absence on a subreddit as inactivity in posting and commenting on a subreddit for at least 14 days (i.e., a minimum 14-day difference between the creation days of two consecutive comments by the same author). To build our absent group for a subreddit, we collected all cases where a user was absent from the subreddit during the six months of our data. For each user, we analyze the time frame starting with the 7 days before an absence and ending with the 7 days after the absence. This is our investigation period for the absent user. To ensure that all users were active in the 7 days before and after absence (and not absent again), we included only those users, whose first contribution in the investigation period was created less than 14 days after the preceding contribution and whose last contribution in the investigation period was created less than 14 before the following contribution. A user could have multiple entries in the absent group when being absent multiple times. For each absent user, we collected the contributions in the 7 days before absence as well as the contributions in the 7 days after absence.

We then built the set of present users. To ensure that the comparison between present and absent users would not be influenced by inherent activity differences in both groups, we decided to choose for each absent user a matching present user with a similar activity level. To furthermore ensure that time and external events have no effect on the comparison, we decided to analyze a present user in the same time frame as his/her matching absent user. The matching process worked as follows: for each absent user, we randomly selected a user who was present during the whole investigation period of the absent user and whose number of contributions in the 7 days before the absence of the treatment user had a minimal difference to the number of contributions of the absent user in this time frame. These users represented the present group. A user could have multiple entries in the present group, when he/she was chosen as partner for multiple treatment cases. For each present user, we collected the contributions in the 7 days before the absence of his/her matching absent user (time frame A in Fig. 3) as well as the contributions in the 7 days after the absence of his/her matching absent user (time frame C in Fig. 3) . Both groups for r/The_Donald consisted of 17,157 cases, for r/politics of 52,071 cases and for r/ChapoTrapHouse of 2,788 cases.

**4.3 Measuring Activity**

**Methods.** To analyze whether continuous interactions in a polarizing environment has an affect on posting and commenting activity, we did the following: for each user in each
Figure 2: Visone’s (Baur et al. 2002) backbone visualization of the co-occurrence network of users being active in the 146 political subreddits.

Figure 3: Overview of experimental setup.

In order to understand whether continuously active/present users change in their style of language, we now wanted to test for features of linguistic attributes. We test this again for the three online political communities.

To analyze whether continuous interaction in a polariz-
Figure 4: Interquartile range (central lines represent median) of activity, lexical diversity and profanity differences.

Table 2: Tests for trends in differences in activity, lexical diversity or profanity before and after a period of absence/presence in each group as well as between groups. The Z-values represent the z-standardized test statistics of the Wilcoxon signed-rank tests and the $U_{\text{min}}$-values represent the test statistics for the Mann-Whitney U tests. Arrows indicate a statistically significant increase or decrease in activity, lexical diversity or profanity, based on the Bonferroni adjusted significance level of 0.017.

Results for lexical diversity. In the middle part of Fig. 4 we see the interquartile range of lexical diversity differences for each group, with the middle line representing the median. The statistical values of all tests can be found in Table 2. On all subreddits, continuous interaction has an effect on lexical diversity, as the difference values in the absent group were
<table>
<thead>
<tr>
<th>Differences in</th>
<th>Subreddit</th>
<th>Absent users</th>
<th>Present Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>r/The_Donald</td>
<td>$r = -.02, p = .014$</td>
<td>$r = .29, p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>r/politics</td>
<td>$r = -.01, p = .043$</td>
<td>$r = .27, p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>r/ChapoTrapHouse</td>
<td>$r = .01, p = .784$</td>
<td>$r = .27, p &lt; .001$</td>
</tr>
<tr>
<td>Lexical Diversity</td>
<td>r/The_Donald</td>
<td>$r = .01, p = .334$</td>
<td>$r = -.15, p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>r/politics</td>
<td>$r = .01, p = .015$</td>
<td>$r = -.15, p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>r/ChapoTrapHouse</td>
<td>$r = .03, p = .207$</td>
<td>$r = -.12, p &lt; .001$</td>
</tr>
<tr>
<td>Profanity</td>
<td>r/The_Donald</td>
<td>$r = -.02, p = .019$</td>
<td>$r = .02, p = .013$</td>
</tr>
<tr>
<td></td>
<td>r/politics</td>
<td>$r = .00, p = .799$</td>
<td>$r = .02, p &lt; .001$</td>
</tr>
<tr>
<td></td>
<td>r/ChapoTrapHouse</td>
<td>$r = .01, p = .609$</td>
<td>$r = -.01, p = .489$</td>
</tr>
</tbody>
</table>

Table 3: Correlations between overall activity and the differences in activity, lexical diversity or profanity before and after a period of absence/presence

By the result of the analysis, the more a user is generally active over the whole time period, the more he/she increases in activity if he/she remains present and continuously active. In contrast, there is no correlation between the activity differences and the number of contributions in the absent group on r/ChapoTrapHouse and on r/politics, and a significant, but negligible negative correlation between the activity differences and the number of contributions in the absent group on r/The_Donald.

To visualize how this would affect the differences between present and absent group, we sorted all users of each group on each subreddit according to the number of contributions and built subgroups of size equal to 10% of overall group size. We then contrasted the interquartile range of present users with the interquartile range of absent users in each subgroup. The resulting image is shown in Fig. 5. The figure shows that in the group of present users on all three subreddits more active users have a higher increase in activity than less active users.

### Results for lexical diversity
The statistical values of the correlations can be found in Table 3. On all three subreddits there is a weak negative correlation between the lexical diversity differences and the number of contributions in the present group: The more a user is generally active the more he/she decreases in lexical diversity when being continuously present in a subreddit. In contrast, there is no correlation between the lexical diversity differences and the number of contributions in the absent group on r/ChapoTrapHouse and on r/politics, and a significant, but negligible positive correlation between the activity differences and the number of contributions in the absent group on r/The_Donald.

The effect of general activity on differences between present and absent user is visualized in Fig. 6. The figure shows that in the present groups of all three subreddits more active users have a higher decrease in lexical diversity than less active users.

### Results for profanity
The statistical values of the correlations can be found in Table 3. There is a significant, but negligible positive correlation between the profanity differences and the number of contributions in the present group on r/politics and r/The_Donald, as well as no correlation on r/ChapoTrapHouse. Similarly, there is no correlation between the profanity differences and the number of contributions in the absent group on all three subreddits.

### Interpretation
We can state that the effect of continuous presence on activity and lexical diversity is influenced by...
overall activity. This means that people who are generally more active show a greater increase in activity and a greater decrease in diversity after an activity period than people who are in general less active.

![Figure 5: Activity differences of users grouped by overall activity for three subreddits.](image)

5 Discussion and Conclusion.

To investigate the question of how presence in polarizing environments influences people linguistically, we analyzed social media data from three political subreddits on Reddit. We compared the behavior of continuously engaged users to the behavior of similar users who were absent for a certain time to determine 1) whether users change regarding their activity and language and 2) whether this change is influenced by absence.

For all three online political communities we found that given the same initial level of activity, users who are continuously present increase their activity and their language changes compared to users with a non-activity. We were able to determine this equally for three of the biggest online political communities on Reddit.

As a result, we can conclude that:

1. Continuously present people increase their activity, decrease their lexical diversity and increase their level of profanity.
2. The effect of continuous presence on activity and lexical diversity is stronger for people who are in general more active.

![Figure 6: Lexical diversity differences of users grouped by overall activity for three subreddits.](image)

In other words, when dwelling in polarizing environments people engage more in discussions, become simpler in terminology and more angry.

5.1 Limitations.

As we rely on the Pushift Dataset (Baumgartner et al. 2020), we cannot assess whether there were any missing data due to data collection errors. Using social media to study human behaviour is flawed by many challenges (Ruths and Pfeffer 2014). We are aware, however, that there were posts and comments in our dataset which were deleted by users or moderators before being retrieved by Pushift. This data was marked with a "[deleted]" in the author field and had therefore to be excluded from the analysis. From the data that were available, we leveraged word lists implemented in Python in order to gain insights into the general wording of users. Furthermore, we did not identify any obvious idiosyncrasies with regard to this platform which would prevent our findings from being generalized to other online political forums although a broader comparative study would help to determine this question. Our research represents an attempt at understanding what presence in polarizing environments does on user behavior. Follow up research could leverage contextualized word embeddings (Peters et al. 2018) to identify which mentions of potentially harmful words demand attention. In addition, future work could complement our linguistically-driven approach by comparing lexical choices (Wang and Culotta 2019).
5.2 Implications and Future Work.

In our analysis we compared the behavior of present users to the behavior of similar users with absence to determine whether users’ behavior changes in polarizing environments. We found that users participate more, become simpler in vocabulary and use more offensive language. Spending inordinate amounts of time in close-knit environments is seen as a central problem and our results confirm slight changes. Based on this we would like to enrich media research on user activity and on short-term intervention techniques. Media consumption leads to changes in behavior and being active online in communities is one form of media consumption. News that is shared and active social interaction in form of discussions influence consumption behavior and this in turn influences the selection of news and interaction. That said, reinforcement effects in behavior and attitudes can occur through selective exposure (Berelson, Gaudet, and Lazarsfeld 1944; Katz 2001).

The question we want to pose here is how the mass media stream of information consumption can be interrupted and whether non-consumption in turn can influence behavior. Hence, we open the discussion for possible (automatic) content moderation techniques to counteract reinforcement effects. In particular, we think of time outs as in sports or the dead cat in British parliament discussions7. However, little is known about the effectiveness of concrete interventions such as blocking single users temporarily in online political communities. The core motivation and function of short-term interventions could be to slow down activity of single group participants to balance emotions and avoid escalation. Can a time-out have a de-escalating effect? Future work could include experiments which actively induce a user’s absence to measure the effect of time-outs caused by others. Furthermore, it would be interesting also to examine the long-term effects of online political conversations. When do these discussions lead to constructive discourses and conversations not resulting in a ban from these platforms and, given that conversations often turn on a particular instance in the discourse, what are the micro-level linguistic traits that enhance long-term stability?

References


Gallacher, J. D.; Heerdink, M. W.; and Hewstone, M. 2021. Online Engagement Between Opposing Political Protest Groups via Social Media is Linked to Physical Violence of Offline Encounters. Social Media + Society 7(1).


Kane, B.; and Luo, J. 2019. Do the Communities We Choose Shape our Political Beliefs? A Study of the Politicization
of Topics in Online Social Groups. In IEEE International Conference on Big Data, 3665–3671.


