On the Relation between Opinion Change and Information Consumption on Reddit

Flavio Petruzzellis,[1] **Francesco Bonchi,**[2] **Gianmarco De Francisci Morales,**[2] **Corrado Monti**[2]

[1]Università degli Studi di Padova, Padova, Italy
[2]CENTAI, Turin, Italy
flavio.petruzzellis@studenti.unipd.it, bonchi@centai.eu, gdfm@acm.org, corrado.monti@centai.eu

Abstract

While much attention has been devoted to the causes of opinion change, little is known about its consequences. Our study moves a first step in this direction by looking at Reddit, and in particular to the subreddit `r/ChangeMyView`, a community dedicated to debating one's own opinions on a wide array of topics. We analyze changes in online information consumption behavior that arise after a self-reported opinion change, by looking at the participation to a set of sociopolitical communities. We find that people who self-report an opinion change are significantly more likely to change their future participation in a specific subset of those communities. Specifically, there is a significant association (Pearson $r = 0.46$) between using propaganda-like language in a community and the increase in chances of leaving it. Comparable results (Pearson $r = 0.39$) hold for the opposite direction, i.e., joining these same communities. In addition, the textual content of the post associated with opinion change is indicative of which communities will be joined or left: a predictive model based only on the text of this post can pinpoint these communities with an average precision@5 of 0.20. Our results establish a link between opinion change and information consumption, and highlight how online propagandistic communities act as a first gateway to internalize a shift in one's sociopolitical opinion.

Introduction

Opinion dynamics is the field of study that deals with how people's opinions form and evolve in a social context (French 1956; Harary 1959). This branch of social psychology has received growing attention, also due to the widespread adoption of social-media platforms and their effect on opinion formation. Most of the effort has been devoted to understanding the *causes* of opinion changes, by analyzing the dynamics of the opinion formation process, simulated by means of *agent-based models* (DeGroot 1974; Friedkin and Johnsen 1990; Axelrod 1997; Deffuant et al. 2000). Only a few studies that analyze real-world social-media data have recently emerged (Xiong and Liu 2014; Tan et al. 2016; Monti, De Francisci Morales, and Bonchi 2020; Naskar et al. 2020). While empirical studies of opinion change are lacking due to the complexity of measuring such a construct (Flache et al. 2017), studying the *externalities* of opinion change is even

more challenging, for several reasons. First, the effects may be delayed in time, which complicates the data-gathering process. Second, while significant, the effects might be small and hard to capture, which requires large amounts of data to tease them out. Third, there may be several confounders to the analysis. For all these reasons, limited attention has been devoted to the consequences of opinion change.

This work is a first step towards filling this gap. We seize an opportunity provided by the subreddit `r/ChangeMyView`, an online community dedicated to debating one's opinions on a wide array of topics. On it, a user posts an opinion they hold, and other people in the community argue against it in the comments by providing alternative viewpoints. Most importantly, the original poster explicitly marks which comments succeeded in changing their view by awarding them a *delta* (Δ). These features makes `r/ChangeMyView` a natural testbed for testing hypotheses about opinion dynamics. The main question we tackle in this paper is the following: *how does opinion change relate to behavioral change?* In particular, the behavioral change we focus on is in *information consumption*, measured as participation to a set of sociopolitical subreddits. Such participation "profile" is important as it represents one's information diet, and is a reliable proxy for, e.g., political affiliation (De Francisci Morales, Monti, and Starnini 2021) or health choices (Kumar et al. 2021).

To answer this question we join data from `r/ChangeMyView` (`CMV`) with user activity on a set of politically-charged subreddits. First, we design a model to predict future changes in a user's participation to subreddits based on that user's submission on `CMV`. This model answers the more specific question: is there a relationship between the change of opinion and change of behavior in a user? By using only information about the topic of that single submission, the model achieves a precision similar to what can be reached by using information about the past history of participation in communities, which is more intuitively related to the future activity of a user (Massachs et al. 2020). We then focus our analysis on characterizing those subreddits which are more likely to see a change in participation after an opinion change. In this case, we ask wether there are specific communities which are more likely to be subject to behavior change. For a handful of subreddits, we find a statistically significant effect on the odds of a participation change after an opinion change. In other words, some communities are

more likely to be abandoned or joined after a user reports a change in their opinion. For instance, the subreddit of Donald Trump's supporters `r/The_Donald` is roughly 20% more likely to be abandoned after a user reports an opinion change. Finally, we characterize these communities by looking at their language. We quantitatively show that there is a correlation between the usage of language typical of propaganda and satire and how much the participation in the subreddit is sensitive to opinion change.

Our results suggest that opinion change is indeed consequential and has a measurable effect on behavioral patterns within the online platform. It is known that online behaviors are predictive of other offline behaviors (Yamamoto, Kushin, and Dalisay 2015; Mejova et al. 2022). Therefore, understanding opinion change online is a data-rich task with important implications, such as generating valuable proxies for other social constructs (e.g., political participation). The present study thus represents an important validation for the field of opinion dynamics.

Related work

Reddit's `r/ChangeMyView` has consistently been used as a data source for empirical studies about online debate and persuasion dynamics in recent years. Studies have focused on the structure of online debates, which are inherently different from offline ones, since they are often asynchronous and anonymous. Öcal, Xiao, and Park (2021) analyse a relatively small sample of manually annotated discussions on `CMV` to study reasoning processes in online debates. They identify the most frequent patterns used by users to provide evidence to support their arguments in online discussions. They find that personal experience is often provided as supporting evidence for arguments, even when people claim to be experts in a field, and especially when they are discussing domestic politics. The corpus-based study by Musi (2018) uses `CMV` discussions to enrich the theoretical standpoint that identifies concessions as one of the most effective strategies to change people's mind in a discussion. The findings of the study suggest that concessions in online debates mainly serve two purposes: to prevent counterarguments and to introduce them.

A different perspective is taken by studies about the effectiveness of peoples argumentative strategies. Tan et al. (2016) analyze a corpus of discussions on `CMV` to study the dynamics of persuasive discourse online. Their findings highlight the importance of the formal features of discussions: independently of the topic, the language used to present one's arguments and the dynamics of interaction between the users involved in the debate prove to be important features in the task of predicting the successful change of perspective of the author. Priniski and Horne (2018) focus on online debate specifically regarding socio-moral issues, and test the hypothesis that it is inherently harder to change one's mind about such topics. Their findings suggest that providing factual evidence supporting one's arguments is useful to change people's mind. However, this effect disappears when debating socio-moral topics, despite the fact that factual evidence is provided more often in debates about these topics. Finally, Monti et al. (2022) analyze the social dimensions of language on `CMV`, and find that messages expressing the pragmatics social dimensions of knowledge, similarity, and trust, are most likely to lead to opinion change.

To the best of our knowledge, this is the first study that employs data from `r/ChangeMyView` discussions to provide empirical evidence about the *consequences* of online debates on users online behavior.

Data

Reddit is a discussion forum organized in topical communities called subreddits. All users must have a pseudonymous account in order to participate. Users can post submissions in these subreddits, and comment on other submissions and comments, thus creating a tree structure for the overall discussion. In addition, users can also upvote a submission to show approval, appreciation, or agreement (and their opposites with a downvote). The score of a submission is the number of positive votes minus the number of negative votes it has received. Differently from other social media such as Facebook or Twitter, Reddit's homepage is organized around subreddits, and not on user-to-user relationships (Cinelli et al. 2021). As such, the subreddits chosen by a user represent the main source of the information they consume on the website.

The subreddit `r/ChangeMyView` is a community on Reddit dedicated to challenging one's own opinions, with over 1.3M subscribers at the time of writing. Specifically, posters offer to debate their stance on a particular topic, and commenters respond by trying to change the poster's opinion. When the poster acknowledges that a comment has succeeded in changing their opinion—even partially—they leave a comment with the codified symbol 'Δ'. The presence of such labelling makes this forum an ideal place to study opinion change. We collect public data from this subreddit from the Pushift collection (Baumgartner et al. 2020), for a 7-year time period starting from 2013, which results in $39\,845$ posts by $26\,074$ authors with $2\,422\,379$ comments on them.

We are interested in understanding whether this declared change in opinion is related to a change in information consumption behavior in the future. Previous work (Massachs et al. 2020; Balsamo, Bajardi, and Panisson 2019; Balsamo et al. 2023) has showed that participation in subreddits is a valuable proxy for a wide range of characteristics of the user. For this reason, we take the subreddits they participate in—their communities on Reddit— as a meaningful measure of the behavior of the user. In particular, participating in a subreddit affects a user's information diet: by subscribing to a new subreddit, a user sees its posts on their Reddit homepage. To reduce the spread of effect that might occur due to different topics of discussions, we limit our scope to *sociopolitical* subreddits. We follow the definition of sociopolitical given by Moy and Gastil (2006), thus including topics related to political figures, parties, institutions, and cultural, social, or national issues. We use the classification of sociopolitical subreddits by Monti et al. (2022): they label the 2000 most popular subreddits in 2019 as either sociopolitical or not (by looking at their description and a sample of their posts), and obtain a set of 51 sociopolitical subreddits (see Appendix).

To keep the topic of opinion change events we analyze consistent with the sociopolitical theme, we restrict our analysis to posts on `CMV` with sociopolitical content. We develop

a supervised post classifier based on a sample of posts from the subreddits we labeled as sociopolitical. Specifically, we build our dataset by taking a random sample of 50 posts per month for each sociopolitical subreddit, and a sample of equal size from the others we analyzed and labeled as non-sociopolitical. The model is a logistic regression with L1 regularization, and the features are simple n-gram counts ($n \in [1, 3]$). The classifier obtains an average F1 score of 89.5% on a held-out test set. We also build a set of 200 posts from `CMV` that we manually label as sociopolitical or not: on this validation set, the classifier obtains an F1 score of 82%. We then proceed to use this classifier on all `CMV` posts. Out of the 65k `CMV` posts with textual content, we identify 46k as sociopolitical.

Then, for each `CMV` user who authored a sociopolitical post, we gather the set of subreddits they participate in. We define participation as writing at least one comment in that subreddit in a given time period, which is a strict proxy for having joined the community as it excludes 'lurkers'. To answer our research questions, we separate participation *before* and *after* a user's post on `r/ChangeMyView`. Furthermore, we limit the time frame in which we observe behavioral change by considering only comments written at most one year before or after the date of the user's post on `CMV`. We do so in order to consider in our analyses only events of online behavior change that are relatively close in time to the reported opinion change, and that can thus be considered related to it more reliably (Balsamo et al. 2023). We indicate with $b_{u,s}$ and $a_{u,s}$ the number of comments authored by user u on subreddit s before and after (respectively) their post on `r/ChangeMyView`, limited to the time frame described above. Using these values, we define:

- L_u as the set of subreddits that user u *left* after their post on `r/ChangeMyView`; that is, $L_u = \{s \mid b_{u,s} > 0 \land a_{u,s} = 0\}$.
- J_u as the set of subreddits that user u *joined* after their post on `r/ChangeMyView`; that is, $J_u = \{s \mid b_{u,s} = 0 \land a_{u,s} > 0\}$.
- C_u as the set of subreddits that *changed* for user u after their post on `r/ChangeMyView`, that is $C_u = L_u \cup J_u$.
- S_u as the set of subreddits that user u *stayed on* after their post on `r/ChangeMyView`; that is, $S_u = \{s \mid b_{u,s} > 0 \land a_{u,s} > 0\}$.

We gather this data for all the `r/ChangeMyView` posts and their authors. By using this data we define an activity profile associated to a post as a vector Z that indicates, for each sociopolitical subreddit, whether the author of the post *joined*, *left*, or *stayed* in that subreddit. Finally, we further filter our dataset by considering only one post per user, in order to consider the user (rather than the post) as our statistical unit. If a user has assigned more than one delta on `CMV`, we consider the closest one to the midpoint in the dataset's timespan, so to have the best possible conditions to build the user profile after having applied the one-year filter on the time of comments. This decision does not alter the dataset significantly, as 84.1% of authors have assigned only a single delta. The final dataset contains $\approx$ 29k users and Table 1 reports some of its statistics.

Table 1: Dataset statistics for the subsets of `CMV` posts with (Δ) and without ($\neg\Delta$) a delta: number of users $|U|$, number of posts $|P|$, average ($\mathbb{E}$) and standard deviation (σ) of number of non-zero elements in user profiles $|Z|$.

User set	$|U|$	$|P|$	$\mathbb{E}(|Z|)$	$\sigma(|Z|)$
CMV Δ	13 871	18 843	4331	4434
CMV $\neg\Delta$	15 026	21 005	4713	4687

Table 2: Comparison between activity on sociopolitical (SP) and other subreddits for different groups of users: participants in `r/ChangeMyView` (CMV), participants in `r/AskReddit` (AR), users who self-report opinion change (Δ), and users who do not ($\neg\Delta$). Values represent the average over considered users.

User set	Number of subreddits	Comments		Score	
		SP	Others	SP	Others
CMV	107.3	418.8	2460.0	8.9	7.4
AR	45.6	94.5	663.4	10.9	8.1
CMV Δ	109.8	402.2	2562.4	9.6	7.8
CMV $\neg\Delta$	104.4	438.5	2341.9	8.2	7.0

Textual corpus

We use two sources of textual information in our analyses: the text of the submissions of users on `CMV` and a sample of the textual contents of the sociopolitical subreddits we consider. We apply the same pipeline of text preprocessing to both sources: we strip URLs, include only words that are written with latin alphabet and are at least 2 characters long, then lemmatize by using the `Python` library `nltk` (Bird, Klein, and Loper 2009). We use this preprocessed textual data to build a tf-idf representation of both posts and subreddits. We use the `Python` library `scikit-learn` (Pedregosa et al. 2011), filter stopwords, include unigrams and bigrams, and limit the number of features to lemmas having a minimum document frequency of 10, which results in a vocabulary of 85 078 lemmas.

Characterizing `r/ChangeMyView`

To better understand this dataset, as a preliminary analysis we compare the behavior of `r/ChangeMyView` participants to participants in a generalist subreddit, `r/AskReddit`, the most popular board on the site. In this subreddit, users ask and answer questions on any topic, without any intent to challenge their opinion. These characteristics make it is a suitable control group. We compare these two groups in terms of activity, breadth of interests, and community feedback, both on the selected sociopolitical subreddits and on the rest.

Table 2 reports this comparison (computed on a random sample of 10% of `r/AskReddit` users for efficiency). `CMV` users are more active on the platform, both in terms of number of comments and breadth of interests, with more than double the number of participated subreddits. This difference is to be expected as `r/AskReddit` is a landing point for

new users, while `CMV` users are more engaged with the platform. The fraction of activity on sociopolitical subreddits is slightly higher for `CMV` (14.5%) than for the others (12.5%). This result corroborates our choice of sociopolitical subreddits, and suggests that our set of `CMV` users is indeed more interested in sociopolitical topics. Instead, the community feedback, as measured by the average score of users' comments, is similar, albeit slightly lower for `CMV`. Finally, the distribution of comments across the different sociopolitical subreddits follows approximately the same ranking in the two groups (Spearman $r = 0.97$, not shown). Therefore, the selected group of users does not appear to be biased towards a specific topic w.r.t. the general Reddit population.

As a final preliminary analysis, we investigate the difference in general behavior between the group of `CMV` users who assigned a Δ and the ones who did not. The two groups do not significantly differ in terms of average activity on Reddit, both on sociopolitical and general subreddits. Also the number of subreddits and community feedback are comparable. However, as we show next, we are able to identify a relationship between the event of opinion change, codified by the assignment of Δ, and a change in the future activity of an individual in terms of information consumption.

Results

We use the collected data to answer the following research questions:

RQ1: Is there a relationship between the change of opinion and change of behavior in a user?

RQ2: Are there specific communities which are more likely to be subject to behavior change?

RQ3: If so, how can we characterize these communities more affected by opinion change?

To answer these questions, we train a classifier aimed at predicting the future behavior of the user by using information about its opinion change as input, and show that the two are indeed related. Then, we identify a subset of communities that are significantly affected by opinion change. We do so by comparing the chances of behavioral changes before and after the event of opinion change. Finally, we characterize this subset of communities, and find that they employ a particular language akin to propaganda and satire.

As a starting point, we compare the rate of change in user profiles between `r/ChangeMyView` users and a propensity-score matched one of similar size coming from `r/AskReddit`. The propensity score takes into account the global user activity, the specific user activity on the 51 sociopolitical subreddits, and the timespan of the user on Reddit. Overall, users in the `r/ChangeMyView` group are more likely to experience behavioral change in the future (OR = 1.511) than users posting on `r/AskReddit`. Therefore, a fundamental component of opinion change and the subsequent behavioral change is engagement and open-mindedness by the user.

Relationship between behavioral and opinion change (RQ1)

Our first goal is to understand whether the topic of the discussion on which an individual changes their minds is related to the communities that they will switch in the future. We also wish to compare the relevance of the topic of discussion against other similarly relevant pieces of information, such as the past activity of the user. To investigate these aspects, we frame the task as a multi-class classification one. The task consists in associating the subreddits for which the participation status will change in the future (i.e., pass from activity to inactivity, or vice versa) to each author of a submission in the Δ group. The possible subreddits where change can happen are 51—i.e., the set of sociopolitical subreddits. To tackle this task, we use an ad-hoc, feed-forward neural network and compare two possible input features.

The first type of input contains *textual information* about the post of the user in `CMV`. It is composed of two concatenated vectors: (i) the tf-idf representation of a user's submission to `r/ChangeMyView`, and (ii) the cosine similarity between such a vector and the tf-idf representation of each of the candidate subreddits. These latter representations are obtained by vectorizing a random sample of submissions in a given subreddit as a single document.

The second type of input for the model contains information about *past activity* of the user in sociopolitical subreddits. We use the aforementioned user profiles Z: each community participated by the user. Specifically, this input contains the number of comments written by a user on each subreddit before the submission on `CMV`. Such a feature vector is known to be predictive of future user behavior (Massachs et al. 2020).

For both inputs, the output layer of the neural network is composed of 51 sigmoid units. Each unit represents the probability that a specific subreddit will be left or joined by the given user after their `r/ChangeMyView` submission. We test this neural network in four versions: one without any hidden layer, and with a hidden layer of size 10, 100, or 1000. We evaluate the classifier via a nested 10-fold cross-validation procedure, to perform both a hyperparameters search on the number of hidden units and an evaluation of its performance. For each hyperparameter fold of the inner CV we select the best architecture based on its obtained Average Precision@5 (AveP@5). The selected model is then retrained and tested on the outer CV by computing AveP@1 and AveP@5. In all cases, the AveP@k metrics are computed by using micro-averaging, i.e., all users in the test set are weighted the same.

Figure 1 and Table 3 report a comparison of the results obtained by the classifier with the two inputs. By using only the textual representation of the single `r/ChangeMyView` submission, the model obtains an AveP@5 of 0.20. In other words, the classifier can identify a subreddit that is going to be switched after the opinion change event with 5 guesses in 20% of the cases. As a comparison, our alternative model based on the whole history of a given user—i.e., all the other communities that they participated in—obtains a similar AveP@5 of 0.23.

We conclude that there is a connection between the topic of the opinion change event, as measured by the textual repre-

Table 3: Performance metrics for the behavior change classifier. Average and SD of the metrics across 10 cross-validation folds for the two types of input vectors.

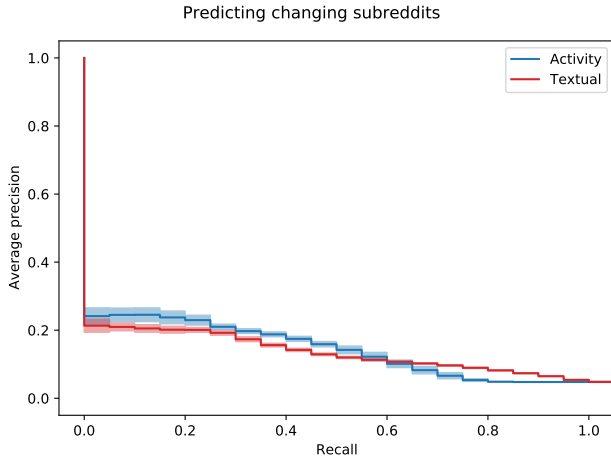| Input vector	AveP@1	AveP@5
Participation history	0.258 ± 0.032	0.229 ± 0.015
CMV submission content	0.215 ± 0.016	0.201 ± 0.009

Figure 1: Precision-recall curves of two predictive models for behavior change, the blue one uses the past activity profile of a user (Activity), and the red one uses the textual information in the post on CMV (Textual). The shaded area indicates the standard deviation of the measure across 10 cross-validation folds.

sentation of the submission, and some of the future behavioral changes of that user.

Communities affected by opinion change (RQ2)

After assessing the relationship between changes in subreddit participation and the submission on r/ChangeMyView, we investigate which changes in user behavior are more likely after the opinion change.

Let us define $p_s^{J,g}$ as the empirical probability that a user in group g will join subreddit s. Analogously, $p_s^{L,g}$ is the empirical probability that a user in group g will leave subreddit s. The two considered groups g are Δ and $\neg\Delta$. The first group represent individuals that have assigned a Δ, thus self-reporting that they experienced opinion change. The second group is the control group, i.e., individuals that did not assign a Δ in their CMV post, and hence are considered not to have experienced opinion change on that issue. Both empirical probabilities are computed from the users' profiles described in the previous section. Hence they only take into account changes in participation occurred up to one year after the submission of the user on CMV.

We estimate these probabilities by counting the occurrences of the associated events for each subreddit. Then, we consider for each probability p_s the corresponding odds $O_s = p_s/(1 - p_s)$. Finally, we compute the odds ratios be-

tween the Δ and $\neg\Delta$ groups as $OR_s^J = O_s^{J,\Delta}/O_s^{J,\neg\Delta}$ for joining a subreddit s. Therefore, OR_s^J represents the increase in the odds of joining a subreddit s after experiencing the opinion change. Similarly, OR_s^L the increase in the odds of leaving it. Figure 2 shows these odds ratios for all the considered subreddits. The figure uses log odds ratios to paint a symmetric view around zero (which represents no difference between groups, i.e., $OR = 1$).

Interestingly, the odds ratios of joining and leaving seem to be correlated, i.e., there does not seem to be a single direction for the behavior change, and it depends on the specific user. The only exception is The_Donald, for which the effect of giving a delta seems to be clearly in the direction of leaving the subreddit (and not joining it). This behavior might be explained by its quarantine and eventual ban from Reddit. In addition, we find no global effect, i.e., the Δ group does not have a higher OR of leaving or joining than the $\neg\Delta$ group.

To assess the statistical significance of the increase in odds, we perform a binomial test for each subreddit. Specifically, we test whether the occurrences of a user leaving (or joining) a subreddit are significantly greater among individuals who reported opinion change with respect to those who did not. We employ a Bonferroni correction in order to account for the large number of tests.[1] We report these results with the color of the dots and the presence of the label in Figure 2. Therein we also report information about the support of each binomial test (i.e., the number of users joining that subreddit or leaving it after opinion change), represented as the radius of the circle surrounding each dot.

While for most subreddits giving a delta has no or negative effects on behavior change, there are a handful of subreddits for which this effect is positive. From the results reported in Figure 2, we can clearly observe two distinct groups of subreddits. The first one is centered in the bottom-left quadrant, whose log odds ratios are generally non-positive and there is no significant increase. The second and more interesting group is in the top-right quadrant: these subreddits exhibit a significant increase in switching one's participation to said subreddit after a self-reported change in opinion. This group comprises 7 subreddits: gatekeeping, EnlightenedCentrism, ABoringDystopia, LateStageCapitalism, PoliticalHumor, TrumpCriticizesTrump, and The_Donald. In other words, these are the communities that are most likely to be dropped or picked up in the year following the occurrence of the opinion change.

The specific subreddits for which behavioral change follows opinion change seem to share a satyrical nature, with 'memes' as the main communication medium. For instance, PoliticalHumor describes itself as "A subreddit focused on US politics, and the ridiculousness surrounding them", while ABoringDystopia and LateStageCapitalism poke fun at modern capitalistic society, and ENLIGHTENEDCENTRISM ridicules the "hypocrisy of the centrist types". We test this hypothesis quantitatively in the following section.

---

[1] We remark that the Bonferroni method is the most conservative one, i.e., the chances of false positives are lower than other methods.
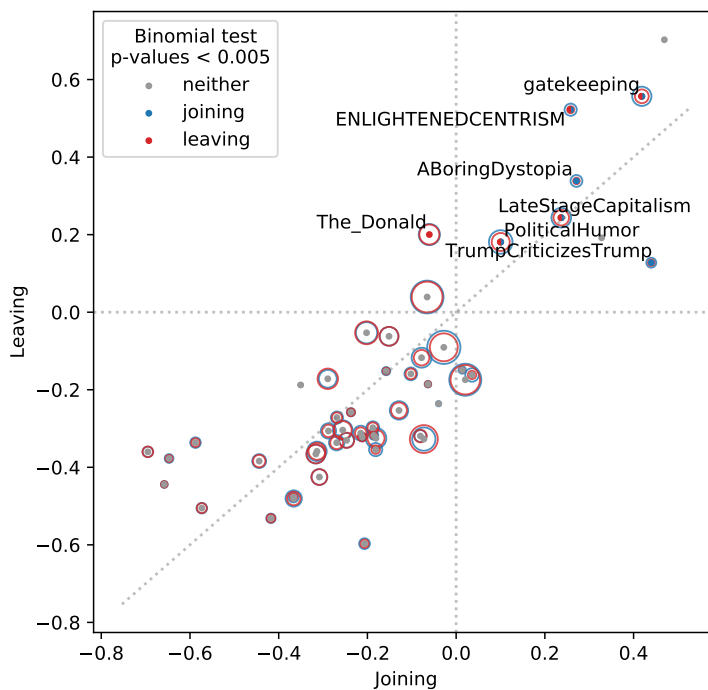
Figure 2: Log odds ratios of joining (X axis) or leaving (Y axis) each subreddit after opinion change. For each subreddit, we represent with circles area the support of their binomial test – i.e., the number of users joining that subreddit (in blue) or leaving it (in red) after opinion change. Blue dots (respectively, red dots) represent subreddits significantly ($p < 0.005$) more likely to be joined (resp., left) after opinion change. We show the name of the subreddit for all those significant in any direction. The subreddits in the top-right quadrant are those for which an opinion changed marked by a Δ is more likely to induce a behavior change.

Behavioral changes and language (RQ3)

To test our hypothesis about the nature of the subreddits for which the user participation change is significantly different in the Δ group, we look at the text of the posts within the subreddits. We leverage from the work presented by Rashkin et al. (2017) about *satire, hoaxes, and propaganda*. The authors develop a measure of 'dramatic language' to automatically classify news belonging to these categories. This measure detects the use of specific words belonging to lexical categories associated with "unreliable sources". We make use of the resources made available by the authors to measure this aspect in our data.[2]

To use these lexica in our case, we perform the following steps. We fit `scikit-learn`'s tf-idf vectorizer on Rashkin et al. (2017) original dataset, which contains both reliable and unreliable news. Here, we use the same preprocessing steps and parameterization of the vectorizer that we use on submissions and subreddits text. Then, we use such vectorizer to obtain a representation of the text extracted from each considered subreddit. Finally, we compute the dramatic language measure for each subreddit by summing up the entries that corresponds to words included in the lexica associated with unreliable sources. We normalize this measure by the number

---

[2] Lexicon available at https://hrashkin.github.io/factcheck.html

Table 4: Pearson's correlation coefficient r and explained variance R^2 between odds ratios and dramatic language measure.

Event	r	R^2
Join	0.385 **	0.149 **
Leave	0.458 ***	0.210 ***

of non-zero entries in the vector in order to take into account the different amount of text present in different subreddits.

Figure 3 shows the full scatterplot of the data (one dot for each subreddit) and the linear regression line between the odds ratio of behavior change and the dramatic language measure. The dramatic language measure is significantly correlated with the values of odds ratios presented in Figure 2, which supports our hypothesis. As shown in Table 4, Pearson's correlation coefficient between the odds ratio of each subreddit is around 0.4 (slightly lower for join, and slightly higher for leave). In other words, the Δ group of users who change opinion are more likely to leave or join subreddits that make more use of dramatic language—akin to satire, hoaxes, and propaganda.
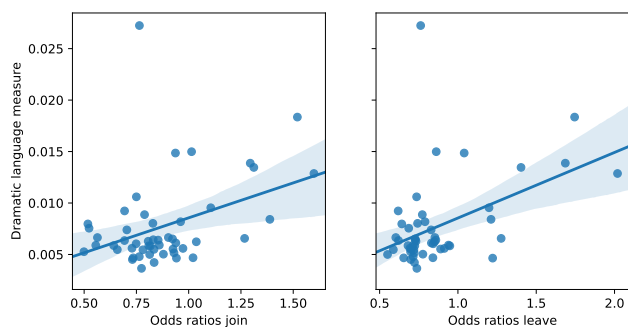
Figure 3: Regression plots which show the correlation between the dramatic language measure and the values of odds ratios OR_s^J and OR_s^L for each sociopolitical subreddit s.

Discussion

We have found that opinion change, signaled by leaving an explicit Δ mark in a discussion on `r/ChangeMyView`, is related to behavioral change, measured as change in user activity on Reddit. In particular, the text of the submission which marks the opinion change can be used as a predictor for which communities are subject to change, and works as well as the full past user activity profile—a commonly-used predictor in the literature (Massachs et al. 2020). Globally the Δ group is not any more likely to show behavioral change than the $\neg\Delta$ group. Conversely, a large effect on future information consumption behavior is due to posting on `r/ChangeMyView`. However, there are specific subreddits for which giving a Δ has an effect (i.e., those in the upper-right cluster in Figure 2). These subreddits are more vulnerable to participation change after opinion change. They are also associated to a high level of 'dramatic language', which is often found in satire, propaganda, and misinformation. Indeed, most of these communities are satirical in nature, and use 'memes' as the main medium of expression.

Starting a discussion on a sociopolitical topic on `r/ChangeMyView` demonstrates an interest in such themes by the user. In fact, we find that users who post on `r/ChangeMyView` are more likely to change their future participation to Reddit communities than those posting in a generalist community (OR = 1.511). This effect is larger than the one associated to self-declaring opinion change. As such, posting on `r/ChangeMyView` can be considered part of a politicization process (in the broadest sense). And even if the original author of the submission is not politicized a priori, at least part of the comments will be. Indeed, sociopolitical content is more common on Reddit than one would initially think. Kane and Luo (2018) show that even in subreddits that have no social or political character, it is possible to recognize traits of politicization in the discussions. Rajadesingan, Budak, and Resnick (2021) report that nearly half of all political talk takes place in subreddits that host political content less than 25% of the time. This trend is in line with an increasing politicization of typically non-political spaces that scholars have observed (Dagnes 2019). Increasing politicization has also been connected to polarization (Chinn, Hart, and

Soroka 2020): increasing interest and engagement in politics can increase polarization, while polarization also drives the politicization of non-political spaces (Garimella et al. 2018).

In this regard, satirical and humorous content (e.g., memes) have been increasingly recognized as essential tools of propaganda in modern politics (Hornback 2018). Previous research has identified satire, hoaxes, and propaganda as related concepts (Cooke 2017). For this reason, they are often clubbed together in some works, including in the one that developed the "dramatic" lexicon we employ (Rashkin et al. 2017). In this latter work, the authors distinguish these three categories (satire, hoaxes, and propaganda) in their data set, but conclude that the language bears similarities across all of them. Thus, while the dramatic lexicon is not fine-grained enough to distinguish them, it is in general typical of "unreliable (esp. hoax and propaganda)" content.

According to Laaksonen, Koivukoski, and Porttikivi (2021), political humor online has two conflicting effects. On the one hand, it favors engagement with political issues. On the other hand, satire also amplifies polarization and might repel some users. Penney (2020) also notes the relationship between political humor online and hyper-polarization.

We found empirical support for these hypotheses in our work. The usage of satirical and propagandistic language in a community makes it more likely both for it to be left in case of opinion change, but also more likely to be joined (depending on the user). The increase in odds seems to be strongly correlated between the two cases, thus bolstering the findings from Laaksonen, Koivukoski, and Porttikivi (2021) about the ambivalent nature of online political satire.

Different possible explanations can be formulated to interpret this effect. By building a shared vocabulary and sense of humor, memes and political satire help in delineating in-groups and out-groups (Elgaaied-Gambier and Mandler 2021; Buie et al. 2021). The emotional component associated to satire is also a factor that contributes to this conflicting nature. Utych (2018) has shown that when people are exposed to political communication expressed with words that have a negative emotional value, they are more likely to reject the content of the message. This effect can be exemplified by the unpleasant feeling that a meme has on a member of the out-group. Naturally, this line of reasoning can be applied in general to online propaganda. Such communities and sources of information are therefore both the first to fall in case of opinion change, as well as the main gateways.

Limitations. As with any empirical study, there are some limitations to our work, mostly due the particular dataset used. First of all, our choice of sociopolitical subreddits could be biased, and many other subreddits might be missing from the selected ones. However, we adopted a systematic process with precise criteria to select these subreddits to minimize bias. Our choice of subreddits is also consistent (by virtue of the content classifier) with the CMV posts we selected. Finally, while some subreddits might be missing, adding more subreddits does not invalidate our current analysis.

Another important limitation is that our analysis is observational, and not causal: we do not prove a causal link, but discover a relation between opinion change and a change in

information consumption. However, we do consider the arrow of time in this relation (behavioral changes follow opinion change), and the presence of a topical link (there is a connection between the text of the post and the subreddits that subsequently change). While limited, this analysis constitutes a necessary first step for any subsequent causal investigation.

Ultimately, our results are necessary limited to the platform we study. As with any empirical work, it is difficult to prove that our results generalize to different platforms, cultures, or societies. This limitation stems from the exceptionality of the CMV community, that offers a rich, codified dataset that is not directly available on other mainstream social media.

Contribution and future work. To the best of our knowledge, our study is the first step into assessing the relationship between a self-reported opinion change event, and a change in the consumption patterns of online information. As such, it lays the ground for different lines of research. Firstly, the characterization of changes in behavior could be further expanded to assess which other characteristics of a community or an information source makes it more vulnerable to opinion change. The measure of dramatic language we used offers a first, coarse-grained characterization of the communities most affected by opinion change events. Future work includes finding more precise NLP tools to better characterize the behavioral consequences of opinion change, and investigating which users that display more prominent changes in behavior. This matter has also practical consequences for fighting against misinformation (e.g., on health-related issues) where opinion change, as part of a polarization process, can make users more exposed to hoaxes (Rollo et al. 2022). Therefore, it would be helpful to verify our findings in other media and societies. Furthermore, identifying the effects of opinion change events can also be used to provide empirical bases to recognize them, contributing thus to a more empirical understanding of the processes of opinion dynamics.

Ethical considerations

As researchers working with user-generated data, our first consideration must be devoted to evaluating whether the data we used in this work was properly collected and treated. In this study, Reddit users that authored any content we used were in general aware of the public nature and accessibility of their content: the communities under study are in fact in the public domain, visible without any account or password, and have thousands of participants. We highlight that we did not recover any content willingly deleted by its author. Furthermore, no personally identifiable information was ever collected: Reddit users make use of pseudonyms, and the messages employed in this study involve only one's view on general, broad topics, making it difficult to uncover any participant's identity. Finally, all the data was used and presented only as aggregated estimates.

Further considerations must tentatively evaluate the potential broader impact of our work. Among positive outcomes, identifying the relationship between opinion change, polarization and online information consumption patterns can advance our understanding of how dis- and mis-information sources can be picked up by social media users, and help design public information campaigns. Of course, however, the same knowledge can also be used to leverage opinion change events in order to spread disinformation. Subsequently, any ethical judgement on possible applications of our work on information spreading ultimately depends on a value of merit on the information being spread. However, we believe that understanding how do we change our opinion online and how this affects our information diet helps in different ways. On the one hand, it recognizes the importance that one's views have on the external world; on the other hand, it furthers our comprehension of how we might fall victims of propaganda.

References

Axelrod, R. 1997. The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution*, 41(2): 203–226.

Balsamo, D.; Bajardi, P.; De Francisci Morales, G.; Monti, C.; and Schifanella, R. 2023. The Pursuit of Peer Support for Opioid Use Recovery on Reddit. In *International AAAI Conference on Web and Social Media*, ICWSM '23.

Balsamo, D.; Bajardi, P.; and Panisson, A. 2019. Firsthand opiates abuse on social media: monitoring geospatial patterns of interest through a digital cohort. In *The World Wide Web Conference*, 2572–2579.

Baumgartner, J.; Zannettou, S.; Keegan, B.; Squire, M.; and Blackburn, J. 2020. The pushshift reddit dataset. In *International AAAI Conference on Web and Social Media*, volume 14 of *ICWSM*, 830–839.

Bird, S.; Klein, E.; and Loper, E. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.".

Buie, H. S.; Ford, T. E.; Olah, A. R.; Argüello, C.; and Mendiburo-Seguel, A. 2021. Where's your sense of humor? Political identity moderates evaluations of disparagement humor. *Group Processes & Intergroup Relations*.

Chinn, S.; Hart, P. S.; and Soroka, S. 2020. Politicization and Polarization in Climate Change News Content, 1985-2017. *Science Communication*, 42(1): 112–129.

Cinelli, M.; De Francisci Morales, G.; Galeazzi, A.; Quattrociocchi, W.; and Starnini, M. 2021. The Echo Chamber Effect on Social Media. *Proceedings of the National Academy of Sciences*, 118(9): e2023301118.

Cooke, N. A. 2017. Posttruth, truthiness, and alternative facts: Information behavior and critical information consumption for a new age. *The library quarterly*, 87(3): 211–221.

Dagnes, A. 2019. Us vs. Them: Political Polarization and the Politicization of Everything. In *Super Mad at Everything All the Time*, 119–165. Springer.

De Francisci Morales, G.; Monti, C.; and Starnini, M. 2021. No echo in the chambers of political interactions on Reddit. *Scientific Reports*, 11(1): 1–12.

Deffuant, G.; Neau, D.; Amblard, F.; and Weisbuch, G. 2000. Mixing Beliefs among Interacting Agents. *Advances in Complex Systems*, 3: 87–98.

DeGroot, M. H. 1974. Reaching a Consensus. *JASA*, 69(345): 118–121.

Elgaaied-Gambier, L.; and Mandler, T. 2021. Me trying to talk about sustainability: Exploring the psychological and social implications of environmental threats through user-generated content. *Ecological Economics*, 187.

Flache, A.; Mäs, M.; Feliciani, T.; Chattoe-Brown, E.; Deffuant, G.; Huet, S.; and Lorenz, J. 2017. Models of Social Influence: Towards the Next Frontiers. *Journal of Artificial Societies and Social Simulation*, 20(4): 2.

French, J. R. P. 1956. A Formal Theory of Social Power. *Psychological Review*, 63(3): 181–194.

Friedkin, N. E.; and Johnsen, E. C. 1990. Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3-4): 193–206.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2018. Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship. In *The Web Conference*, WWW, 913–922.

Harary, F. 1959. A criterion for unanimity in French's theory of social power. In *Studies in social power*, 168–182.

Hornback, R. 2018. Afterword: White Nationalism, Trolling Humor as Propaganda, and the "Renaissance" of Christian Racism in the Age of Trump. In *Racism and Early Blackface Comic Traditions*, 271–293. Springer.

Kane, B.; and Luo, J. 2018. 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*, BigData, 3665–3671.

Kumar, N.; Corpus, I.; Hans, M.; Harle, N.; Yang, N.; McDonald, C.; Sakai, S. N.; Janmohamed, K. A.; Tang, W.; Schwartz, J. L.; et al. 2021. COVID-19 vaccine perceptions: An observational study on Reddit. *medRxiv*.

Laaksonen, S.-M.; Koivukoski, J.; and Porttikivi, M. 2021. Clowning around a polarized issue: Rhetorical strategies and communicative outcomes of a political parody performance by Loldiers of Odin. *New Media & Society*.

Massachs, J.; Monti, C.; De Francisci Morales, G.; and Bonchi, F. 2020. Roots of Trumpism: Homophily and Social Feedback in Donald Trump Support on Reddit. In *12th International ACM Web Science Conference*, WebSci, 49–58.

Mejova, Y.; An, J.; Morales, G. D. F.; and Kwak, H. 2022. Modeling Political Activism around Gun Debate via Social Media. *ACM Transactions on Social Computing*, 3532102.

Monti, C.; Aiello, L. M.; De Francisci Morales, G.; and Bonchi, F. 2022. The language of opinion change on social media under the lens of communicative action. *Scientific Reports*, 12(1): 17920.

Monti, C.; De Francisci Morales, G.; and Bonchi, F. 2020. Learning Opinion Dynamics from Social Traces. In *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD, 764–773.

Moy, P.; and Gastil, J. 2006. Predicting Deliberative Conversation: The Impact of Discussion Networks, Media Use, and Political Cognitions. *Political Communication*, 23(4): 443–460.

Musi, E. 2018. How Did You Change My View? A Corpus-Based Study of Concessions' Argumentative Role. *Discourse Studies*, 20(2): 270–288.

Naskar, D.; Singh, S. R.; Kumar, D.; Nandi, S.; and de la Rivaherrera, E. O. 2020. Emotion Dynamics of Public Opinions on Twitter. *ACM Transactions on Information Systems*, 38(2).

Öcal, A.; Xiao, L.; and Park, J. 2021. Reasoning in Social Media: Insights from Reddit "Change My View" Submissions. *Online Information Review*.

Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12: 2825.

Penney, J. 2020. 'It's So Hard Not to be Funny in This Situation': Memes and Humor in U.S. Youth Online Political Expression. *Television & New Media*, 21(8): 791–806.

Priniski, J. H.; and Horne, Z. 2018. Attitude change on Reddit's change my view. In *40th Annual Meeting of the Cognitive Science Society*, 2279–2284.

Rajadesingan, A.; Budak, C.; and Resnick, P. 2021. Political Discussion is Abundant in Non-political Subreddits (and Less Toxic). In *International AAAI Conference on Web and Social Media*, volume 15, 525–536.

Rashkin, H.; Choi, E.; Jang, J. Y.; Volkova, S.; and Choi, Y. 2017. Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking. In *Conference on Empirical Methods in Natural Language Processing*, 2931–2937.

Rollo, C.; De Francisci Morales, G.; Monti, C.; and Panisson, A. 2022. Communities, Gateways, and Bridges: Measuring Attention Flow in the Reddit Political Sphere. In Hopfgartner, F.; Jaidka, K.; Mayr, P.; Jose, J.; and Breitsohl, J., eds., *Social Informatics*, volume 13618 of *SocInfo*, 3–19. [best paper award]: Springer. ISBN 978-3-031-19096-4 978-3-031-19097-1.

Tan, C.; Niculae, V.; Danescu-Niculescu-Mizil, C.; and Lee, L. 2016. Winning Arguments: Interaction Dynamics and Persuasion Strategies in Good-Faith Online Discussions. *25th International Conference on World Wide Web*, 613–624.

Utych, S. M. 2018. Negative Affective Language in Politics. *American Politics Research*, 46(1): 77–102.

Xiong, F.; and Liu, Y. 2014. Opinion Formation on Social Media: An Empirical Approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(1): 013130.

Yamamoto, M.; Kushin, M. J.; and Dalisay, F. 2015. Social media and mobiles as political mobilization forces for young adults: Examining the moderating role of online political expression in political participation. *New Media & Society*, 17(6): 880–898.

Appendix

Table 5: List of considered sociopolitical subreddits. This list was obtained by manually labeling the 2000 most popular subreddits in 2019 as sociopolitical or not, following the definition given by Moy and Gastil (2006).

Sociopolitical subreddits	
ABoringDystopia	Anarchism
antiwork	AskHistorians
askphilosophy	atheism
Bad_Cop_No_Donut	Buddhism
Christianity	communism
Conservative	conspiracy
conspiracytheories	CryptoMarkets
Economics	economy
ENLIGHTENEDCENTRISM	environment
Feminism	Firearms
gatekeeping	geopolitics
history	JordanPeterson
LateStageCapitalism	Libertarian
lostgeneration	MensRights
Military	NeutralPolitics
news	OurPresident
philosophy	PoliticalDiscussion
PoliticalHumor	politics
progun	ProtectAndServe
Republican	SandersForPresident
skeptic	socialism
The_Donald	TheRedPill
TrollXChromosomes	TrueReddit
TrumpCriticizesTrump	TwoXChromosomes
ukpolitics	WitchesVsPatriarchy
worldnews	