Causal Modeling of Climate Activism on Reddit

Jacopo Lenti
Sapienza University
Rome, Italy
CENTAI
Turin, Italy
jcp.lenti@gmail.com

Luca Maria Aiello
IT University of Copenhagen
Copenhagen, Denmark
luai@itu.dk

Corrado Monti
CENTAI
Turin, Italy
corrado.monti@centai.eu

Gianmarco De Francisci Morales
CENTAI
Turin, Italy
gdfm@acm.org

Abstract

Climate activism is crucial in stimulating collective societal and behavioral change towards sustainable practices through political pressure. Although multiple factors contribute to the participation in activism, their complex relationships and the scarcity of data on their interactions have restricted most prior research to studying them in isolation, thus preventing the development of a quantitative, causal understanding of why people approach activism. In this work, we develop a comprehensive causal model of how and why Reddit users engage with activist communities driving mass climate protests (mainly the 2019 Earth Strike, Fridays for Future, and Extinction Rebellion). Our framework, based on Stochastic Variational Inference applied to Bayesian Networks, learns the causal pathways over multiple time periods. Distinct from previous studies, our approach uses large-scale and fine-grained longitudinal data (2016 to 2022) to jointly model the roles of sociodemographic makeup, experience of extreme weather events, exposure to climate-related news, and social influence through online interactions. We find that among users interested in climate change, participation in online activist communities is indeed influenced by direct interactions with activists and largely by recent exposure to media coverage of climate protests. Among people aware of climate change, left-leaning people from lower socioeconomic backgrounds are particularly represented in online activist groups. Our findings offer empirical validation for theories of media influence and critical mass, and lay the foundations to inform interventions and future studies to foster public participation in collective action.

CCS Concepts

• **Information systems → Social networks**; • **Human-centered computing → Social media**; • **Mathematics of computing →** *Probabilistic inference problems.*

Keywords

Collective Action, Reddit, Probabilistic Modeling, Climate Change

WWW '25, April 28-May 2, 2025, Sydney, NSW, Australia

ACM ISBN 979-8-4007-1274-6/25/04
https://doi.org/10.1145/3696410.3714684

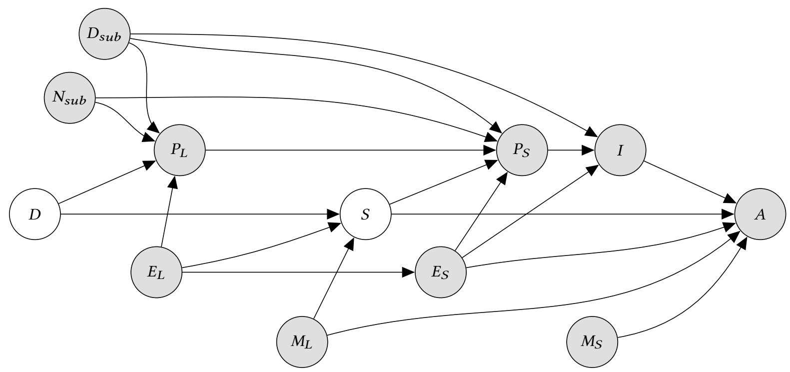
ACM Reference Format:
Jacopo Lenti, Luca Maria Aiello, Corrado Monti, and Gianmarco, De Francisci Morales. 2025. Causal Modeling of Climate Activism on Reddit. In *Proceedings of the ACM Web Conference 2025 (WWW '25), April 28-May 2, 2025, Sydney, NSW, Australia.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3696410.3714684

1 Introduction

The climate crisis is an unprecedented challenge for humanity, as the last decade has been the warmest in over 125 000 years [17]. Climate change is a collective action problem that requires mass coordination due to its global scale and to the interdependency of all agents involved [9]. In recent years, the younger generations have succeeded in mobilizing millions of people worldwide to draw attention to the climate emergency [15]. These collective events show the centrality of the climate change issue in our societies, and demonstrate that humanity can organize global collective action for a common cause.

Collective action is defined as any action that aims to improve the status, power, or influence of a group [49]. We call *climate action* the set of collective strategies intended to increase political and social pressure to quickly enact effective climate mitigation measures [17]. Since climate risks [10], responsibilities [1], and awareness [8] are spread unequally in the global population, it is natural that climate activism is also unequally spread. As such, climate action is the most pristine example of how challenges facing humanity are, first and foremost, a problem of social coordination. Therefore, investigating the mechanism of collective action and activism is fundamental not only to tackle climate change in itself, but also to understand how to address other global crises—from pandemics to misinformation.

Research in the social sciences has identified different pathways to engagement with collective action, primarily focusing on cost-benefit, collective efficacy, group-based emotions, and social identity theory [43, 44]. So far, the validity of these theories has only been assessed by survey-based studies [48], where it is hard to collect heterogeneous and balanced data. However, the abundance of human behavior data from social media opens a different perspective into these dynamics. Previous social media studies on climate action have analyzed the changes in climate debate after extreme weather events [20, 38, 42, 45], political events [11, 25] or media coverage [20]. However, these studies are either associational or focus on a single causal pathway for the phenomenon of interest. Thus, they neglect potentially important interaction effects between the key causal mechanisms involved. To gain a richer understanding of the processes sparking collective action and compare the effects of different causal mechanisms at play, a *comprehensive causal model*



Node Name	Description
D	Demographics of the user
D_{sub}	Demographics of the subreddits
N_{sub}	Number of users in the subreddits
P_L	Participation to subreddits (long term)
P_S	Participation to subreddits (short term)
S	User sympathy for activism at t_{A-W}
E_L	Reddit Engagement (long term)
E_S	Reddit Engagement (short term)
M_L	Media coverage (long term)
M_S	Media coverage (short term)
A	User activated in climate activism groups
I	Interactions with activists

Figure 1: Probabilistic Graphical Model (left) and description of nodes (right). Circles represent stochastic variables; gray nodes are observed and white ones are latent.

is needed instead. We contribute to this objective by applying a Bayesian causal approach to analyze the pathways leading people to engage with climate activism on Reddit.

We envision the path leading to collective action as a funnel, where several essential steps filter the overall population toward active mobilization. We focus on the segment of this selection process that brings users who are aware of the climate change issue into the fold of climate activist communities. This step is a fundamental filter, as it determines who—among the crowd of aware individuals that are potential targets for mobilization—is motivated enough to take a first perceptible step towards action: participating in an online forum dedicated to climate action.

We identify as our target variable the participation in one of the subreddits affiliated with the organizations that coordinated the September 2019 global climate strikes, mainly *Earth Strike*, *Extinction Rebellion*, *Fridays For Future*, and the *Sunrise Movement* (see Appendix A). The importance of Reddit in this process is demonstrated by the history of the Earth Strike movement: it was founded after a Reddit post, which led to the creation of a subreddit, and then evolved into the organization of the global Earth Stike on 27 September 2019.[1]

The goal of this work is to *quantify the relative importance of different social and information processes in causing engagement with activist communities on social media.* Based on existing literature, we identify three main potential pathways leading to this type of engagement, which we call *activation*. First, the familiarity with extreme weather events [45, 47] and the media coverage around climate change [20] affect users' involvement in the climate debate (**RQ1**). Second, the sociodemographic features affect the perception of risk due to climate change [47] (**RQ2**). Lastly, the intention to participate in collective action is affected by social influence [39, 48] (**RQ3**). We articulate our three main research questions accordingly:

RQ1: Does media coverage about climate and climate action affect activation in climate activism groups on Reddit, and over which time scale?

RQ2: Which sociodemographic characteristics affect activation?

RQ3: Do the interactions with other activist individuals affect one's activation?

To answer the RQs, we introduce a causal model based on Bayesian Networks, depicted in Figure 1, that learns multiple key determinants of engagement with activist communities from a ground truth on activation extracted from large-scale Reddit data. This is the first data-driven study that considers the interplay between different causal pathways towards collective action in the context of social media. The proposed pipeline can be flexibly repurposed for other phenomena of collective engagement from online traces.

Our results show that social influence is the main driver of engagement. Sociodemographic features influence the engagement in climate activism, as they increase the sympathy for the cause and probability of frequenting subreddits where social influence takes place. Finally, mass media coverage of climate action has a positive effect on the engagement in the short term.

2 Related Work

Social media platforms proliferate with discussions around climate change, ranging from promotion of grassroots action to expressions of climate skepticism [7, 14]. Online communities focused on the climate crisis have become a fertile ground for the development of a generational social identity [14]. Social identity is one of the main pathways for engaging in collective behavioral change [4, 41], complementing other drivers of action such as rational incentives and purely emotional triggers [4]. These strongly identity-centric communities played a significant role in organizing the global climate protests of 2019, which drawn broad political attention to the climate emergency [15].

The study of community engagement on social media has frequently been approached through classification tasks, emphasizing the power of the different models and features to predict user engagement [24, 37, 52]. Some studies rely on exact matching or propensity score matching to investigate causal links between interventions and engagement outcomes, such as joining discussion groups [36, 39]. These studies primarily focus on endogenous processes within social media, and consider treating explanatory variables as uniform sets, without attempting to model the complex causal pathways connecting them. Several studies quantify single causal effects on social media [26, 39, 46] but only a few use comprehensive causal models formalized with Bayesian Networks [2].

[1] https://www.reddit.com/r/chomsky/comments/d7i2g0/how_a_post_on_rchomsky_helped_start_the_biggest

Conversely, research on public engagement with climate issues has largely concentrated on external climate-related events and their correlation with social media responses [11, 19, 20, 25, 38, 42, 45]. The relationship between sociodemographic factors and climate activism has been explored mostly through questionnaires [18, 47, 51] and connected the personal traits of climate activists to existing psychological theories of collective action [33, 40].

Recent literature on Probabilistic Generative Agent-Based Models (PGABMS) proposed the use of Bayesian Networks to leverage observed longitudinal data for learning the causal mechanisms to be encoded in the ABMs [22], including for example the backfire effect [23, 31]. Similarly, probabilistic generative frameworks have been used to model community engagement to study dynamics such as social influence or the echo chamber effect [30, 32].

3 Model

We ground our operationalization of involvement with activist communities—the *activation* outcome we aim to build a causal model for—in the social psychology theory of collective action by Van Stekelenburg and Klandermans [48]. According to the theory, before participating in collective action, an individual needs to be 1) a sympathizer for the cause, 2) a target of mobilization attempt, 3) motivated to participate, and finally 4) a participant. We focus on the dynamics leading to step 2: how people who are *aware* of climate change become *active participants* (or *activated* for brevity) in online groups of climate activists where upcoming mobilization events are discussed, and can concretely participate to them if properly motivated [48]. This operationalization is supported by the view that activist groups are "opinion-based groups" [6, 27], whose shared identity is based on common values and beliefs towards societal issues. These opinion-based groups are primarily formed to convert broad beliefs into collective action, and indeed, having a social identity aligned with such a group is strongly associated with activism [41]. Next, we define the features that constitute our model, the causal relationships between them, and how we estimate the model's parameters.

3.1 Features

Activation. We consider our initial population of *aware* individuals to be Reddit users who mentioned the words *"climate change"* or *"global warming"* in any message between January 2016 and January 2022, which results in 1.5M users.[2] The use of these locutions indicates awareness of the climate issue by the time of posting. We define the *observation period* for each of these users to be between one year before and two years after their first use of such locutions. Note that this group of Reddit users is the whole population for which we wish to estimate the causal effects.

We select the groups of climate activists as the subreddits directly connected to political groups or movements that took part in the organization of global climate strikes. In particular, we take as reference the "Global Week For Future" event, which took place from 20 to 27 September 2019, and saw the participation of about six million people worldwide.[3] We find six groups that fulfill this criterion, which we describe in Appendix A. The three main ones are *Extinction Rebellion*, *Earth Strike*, and *Fridays For Future*, which account for more than 92% of the activated users we observe. We say that an *activation* happens when a user in our population posts their first message in one of these six communities. Standing on Social Identity Theory, we identify this moment as the first observable instance of their identification with a group characterized by a specific collective identity [4]. We consider a user *activated* if they do post such a message during their observation period (after we observe their awareness), and we represent this outcome as a binary variable A, equal to 1 for activated users and 0 for the others. In total, we obtain 4.1k activated users. To estimate our causal model, we take an equally-sized sample of non-activated users. To ensure basic comparability of the two groups, we verify that the difference between the Reddit engagement, measured as average number of comments per week, is not significant.

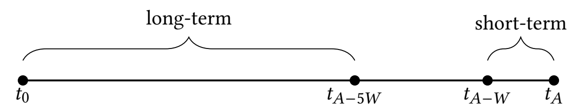


Figure 2: Temporal dynamics of the model. The final time t_A is the activation time, or the end of the observation period for non-activated users. The first three points corresponds to one year, five weeks, and one week before t_A.

Temporal dynamics. To discern the temporal scale of the causal determinants of activation, let us define t_A as the time we observe an activation, or the end of the observation period for non-activated users. We consider two distinct time periods. The *short-term* period contains the last week before t_A, i.e., the interval $[t_{A-W}, t_A]$. The *long-term* one, instead, starts one year before t_A (indicated with t_0) and ends five weeks before t_A, i.e., $[t_0, t_{A-5W}]$. Figure 2 illustrates this temporal split. We leave a four-week gap to separate more clearly the effects between these two periods. However, removing the gap by extending the long-term period to t_{A-W} does not change our results (see Appendix B).

Sympathy. To distinguish how factors in the long- and short-term periods affect activism, we introduce a measure of how likely an individual is to join an activist group at t_{A-W}, the beginning of the short-term period. We call this quantity the *sympathy for the cause* (S), defined as the first step towards mobilization in social psychology Van Stekelenburg and Klandermans [48]. As sympathy cannot be observed directly, we use a continuous latent variable whose estimation strategy is explained at the end of this section.

We can thus divide the factors we wish to study into long-term ones, whose effect on A is mostly mediated by S, and short-term ones, that affect A after t_{A-W}. In particular, we consider:

- the user's sociodemographics characteristics D, which precede all other factors;
- the participation of a user to other subreddits during the long-term (P_L) or short-term (P_S) period;
- the media coverage of climate-related news, divided into long-term (M_L) and short-term (M_S);
- interactions with activated users, that we indicate with I;
- the overall engagement with the Reddit platform (an essential confounder for our target variable A), denoted by E_L and E_S.

[2] We remove all the bots by leveraging a list of known Reddit automated accounts.
[3] https://www.theguardian.com/environment/2019/sep/27/climate-crisis-6-million-people-join-latest-wave-of-worldwide-protests

The rest of this section explains the operationalization and the assumptions behind each of these factors, and how they are linked. Figure 1 depicts the resulting causal graph.

Sociodemographic scores. We characterize users on four sociodemographic axes based on the methodology developed by Waller and Anderson [50]: affluence (poor to rich), partisanship (left to right), gender (male to female), and age (young to old). For each axis, they assign a score to each of the 10k most popular subreddits in their data set. To integrate this methodology into our probabilistic model, we indicate such subreddit scores as D_{sub}, consisting of a $K \times 4$ matrix, where K is the number of subreddits we consider. Then, we represent the sociodemographics of a given user as a 4-dimensional vector D that we estimate as a latent variable, and then we standardize as z-scores. Thus, the 4 values of D for a user represent their likelihood of belonging to a specific combination of demographic groups: poor-rich, left-right, male-female, and young-old.

Subreddit participation. To estimate the user sociodemographics latent variable D, we use user participation in the set of K subreddits that we consider, in the long-term period. In other words, we assume that the observed subreddit participation is a noisy realization of a user's real sociodemographics. To determine this set, we first restrict our attention to the 1000 most popular subreddits in our dataset. We intersect it with those considered by Waller and Anderson [50], thus resulting in $K = 826$ subreddits with sociodemographic scores. For each user, we define two observed variables P_L and P_S, representing their participation in these subreddits in the long and short term, respectively. For a given user, each of those two variables is a binary vector with dimension K, with values set to 1 for the subreddits where the user posted a message during the specified period.

Subreddits attributes. As a factor affecting the participation of users in subreddits, we consider not only the sociodemographic scores D_{sub}, but also the subreddit's overall popularity. We quantify popularity as the z-score of the log-transformed number of users in our dataset who participate in the subreddit by posting or commenting. We denote this variable with N_{sub}.

Media coverage. We wish to investigate the role of media coverage in the mobilization of activists. Furthermore, we wish to account for the geographic location of Reddit users and their proximity to the news, when possible. To do so, we use the Global Database of Events, Language, and Tone (GDELT),[4] an open database of geolocated news articles. We consider three themes that are directly related to climate change: "climate", "climate action", and "natural disasters". The first theme includes general news about the climate; the second focuses on climate activism (e.g., protests or campaigns); the third involves climate hazards, such as floods or heat waves, that are often a consequence of climate change [34]. Since Reddit is predominantly English-speaking and US-centered,[5] in order to consider a homogeneous media landscape, we only consider news located in the United States or Canada. For each administrative area (i.e., US state or Canadian province), we measure the weekly media attention on a theme as the fraction of news published in that area on that theme. This way, we obtain a vector of three variables for any period and area.

[4]http://www.gdeltproject.org
[5]More than half of its traffic comes from the US or Canada: https://www.similarweb.com/website/reddit.com

User georeferencing. We wish to associate, whenever possible, Reddit users in our population to their home location. To do so, we rely on a method developed by Balsamo et al. [3], which leverages user participation in location-based subreddits. To identify such subreddits, we first gather all the subreddits whose name matches with an existing location and clean them manually to filter out spurious matches [29], thus obtaining a list of 311 location-based subreddits in the US and Canada. Then, we assign a user to an area if all the location-based subreddits they wrote into within our observation period belong to that area. This way, we can assign a location in the US and Canada to 25.7% of our population. The proportion of georeferenced users is similar between activated users and our sample of non-activated users (24.9% vs. 26.6%). Furthermore, the number of users assigned to each area is proportional to the population of that area, thus suggesting strong reliability and homogeneous distribution (Appendix C). Now, we can associate the location-based media coverage with each user. For all the users without location, we use the median media coverage per week across all locations, as a standard imputation strategy for missing values. Therefore, we obtain for each user two variables, M_L and M_S, respectively for the long-term and the short-term period. Each variable is a vector of three elements, corresponding to the three news themes ("climate", "climate action", and "natural disasters"). M_L and M_S are standardized to their z-scores.

Interactions. One of our objectives is to study how direct interactions with activists affect the chances of activation. A direct interaction (u, v) occurs when a user u leaves a comment below a comment or submission by v. We collect all direct interactions between activated and non-activated users in both directions ((u, v) and (v, u)). We define a binary variable I to express whether a given user had at least one interaction with an activist user in the short-term period. Since we expect capturing this kind of causal relationship over long periods to be fickle, we restrict ourselves to considering only the effect of short-term interactions.

Reddit engagement. Finally, as an essential confounder for the participation in activist subreddits, we consider the overall Reddit engagement of a user. More active users are necessarily more likely to interact with other users and visit new subreddits. We measure this engagement as the logarithm of the average number of comments per week in each period for a given user. Thus we denote E_L their engagement in the long-term period, E_S short-term one. The variables are standardized to their z-scores.

3.2 Structure

Next, we describe the structure of the causal model, explaining each causal relationship between variables. Figure 1 depicts the Directed Acyclic Graph (DAG) associated with the causal model. We strive to include in the model all those relations that we cannot exclude through reasoning, in order to control the confounders and close possible backdoor paths [35]. Then, by fitting the model, we can assess which relationships are not substantiated by data.

Independent variables. Data of media coverage, M_L and M_S, comes from external sources, and we consider them as independent of user activity on Reddit. For this reason, these variables do not have any parents, and are directly observable. The observed data about subreddits—namely, their demographic scores D_{sub} and their overall popularity N_{sub}—are also unaffected by user variables. Regarding the user, since the sociodemographics are intrinsic and predate all other factors, D does not have any parents.

We consider E_L, the observed engagement with Reddit in the long-term period, to be independent of all other factors. Similarly, we consider its short-term period analogue, E_S, to be affected only by its long-term predecessor ($E_L \to E_S$). As such, we model it as a Gaussian variable dependent on E_L, whose parameters (coefficient, bias, and variance) are to be estimated

$$E_S \mid E_L \sim \mathcal{N}(\beta_{E1}E_L + \beta_{E0}, \theta_E). \quad (1)$$

Long-term participation. The demographics of a user D influence the subreddits they participate in, and thus $D \to P_L$. This effect interacts with the sociodemographic scores of the subreddit D_{sub}; for instance, left-leaning users are more likely to participate in subreddits categorized as left-leaning. P_L is also influenced by N_{sub} because, in general, users are more likely to join popular subreddits, so we include $N_{sub} \to P_L$. Moreover, more engaged users are more prone to write in any subreddit, and we model this effect with $E_L \to P_L$. We model these causal relations with a logistic function: the participation in a subreddit is a Bernoulli random variable whose probability is given by

$$P(P_L \mid D_{sub}, N_{sub}, D, E_L) = \sigma(\beta_{P1}D_{sub} \cdot D + \beta_{P2}N_{sub} + \beta_{P3}E_L + \beta_{P0}), \quad (2)$$

where $\sigma(x) = \frac{1}{1+e^{-x}}$ and $\beta_\bullet$ are coefficients to be estimated.

Sympathy. As mentioned, S denotes the sympathy for the cause—how likely is an individual to join an activist group—at t_{A-W}. As such, it is influenced by the demographics D, the Reddit engagement E_L, and the media coverage M_L, that all precede it temporally. In fact, given that a person's sociodemographics shape their experience of the world, our model assumes that the influence of sociodemographic D on A is mediated by their internal belief S. Thus, we represent S as a continuous variable, drawn from a normal distribution defined by

$$S \mid D, E_L, M_L \sim \mathcal{N}(\beta_{S1}D + \beta_{S2}E_L + \beta_{S3}M_L, \theta_S). \quad (3)$$

Short-term participation. Finally, we include the effect of community retention [52] that models the fact that individuals are more likely to write on a subreddit if they contributed to them in the past ($P_L \to P_S$). We hypothesize the same effect of popularity and demographics that we consider in the long-term to be possible in the short term via $N_{sub} \to P_S$ and $D_{sub} \to P_S$. Analogously, it is reasonable that subreddit participation is affected by the overall Reddit engagement in this period ($E_S \to P_S$). Finally, since P_S is measured in the short-term period $[t_{A-W}, t_A]$, we hypothesize that it could be affected by the sympathy for the cause S, which reflects a measure at t_{A-W}. To avoid estimating the effect of sympathy on each of the $K = 826$ subreddits we consider, which would be prone to overfitting and hard to interpret, we model this relation by aggregating subreddits over the four sociodemographic axes. Thus, we obtain a term $\beta_{p1}S \cdot D_{sub}$, where β_{p1} is a four-element vector that expresses the propensity of a user with high sympathy for the activism cause to participate in subreddits with a given demographic profile

$$P(P_S \mid D_{sub}, S, P_L, N_{sub}, E_S) = \\ \sigma(\beta_{p1}S \cdot D_{sub} + \beta_{p2}P_L + \beta_{p3}N_{sub} + \beta_{p4}E_S + \beta_{p0}). \quad (4)$$

Interactions. Users who are more engaged on Reddit have a higher probability of interacting (by chance) since interactions are induced by comments. Hence, we add the relation $E_S \to I$. The probability of interacting with some activated user depends on the specific subreddit frequented, which is captured by the causal relationship $P_S \to I$. In particular, we model this effect via the sociodemographic scores of the subreddits, assuming that the probability of interaction depends on the sociodemographic scores of the subreddits rather than the specific identity of the subreddit. This design choice enables a more robust model: since the number of subreddits K is quite large, estimating so many parameters would be prone to overfitting. Therefore, we introduce the effects $D_{sub} \to I$. Combining these factors, we obtain

$$P(I \mid D_{sub}, P_S, E_S) = \sigma(\beta_{I1}P_S \cdot D_{sub} + \beta_{I2}E_S + \beta_{I0}) \quad (5)$$

Activations. The literature on opinion dynamics, social influence, and community engagement tells us that interacting with people can have a direct impact on the probability of activation [12, 13, 39], and we capture this social influence via $I \to A$. Next, we express the role of S by introducing the relation $S \to A$. Since S represents how likely a user is to join an activist group at t_{A-W}, this relation connects S to the actual activation at t_A. The sociodemographics D precede S, temporally and causally, so their effect on A is mediated by S, with no direct effect on A. Then, to assess the role of media coverage on activation we introduce $M_S \to A$, expressing how the media in the week before the activation affects A. The media coverage in the long-term period, M_L, can affect the sympathy (at t_{A-W}, the end of that period) but could also have a delayed, direct effect on A ($M_L \to A$) [45]. Finally, the effect $E_S \to A$ acts as a confounder, since the activation corresponds to visiting a new subreddit, and users who are more engaged with Reddit in this period may activate by chance.

$$P(A \mid I, M_S, M_L, S, E_S) = \\ \sigma(\beta_{A1}S + \beta_{A2}I + \beta_{A3}M_S + \beta_{A4}M_L + \beta_{A5}E_S + \beta_{A0}). \quad (6)$$

3.3 Parameter estimation

To estimate the model's parameters and latent variables, we fit the model to our Reddit data using Stochastic Variational Inference [16]. During the hyperparameter tuning phase, we select: (*i*) the Adam optimizer with a learning rate equal to 0.05 to maximize the evidence lower bound (ELBO), and (*ii*) a Normal distribution with zero mean and variance 0.1 as the variational distribution. To discard results reached in suboptimal local minima of the parameter space, we run the optimization with 10 random restarts and choose the restart with the best fit to the observed activations. We measure the fit via accuracy, namely the average fraction of correct predictions of 100 posterior predictive samples of A. Recall that our goal is to assess causal relations: we are not interested in reaching the highest accuracy, but only in obtaining the most reliable model that respects the previously defined causal assumptions.

While the hyperparameter selection is fully data-driven, the choice of the prior distribution encodes our existing knowledge of the generative process in the context of Reddit. In particular, we focus on the latent sympathy variable S. The variance $var(S)$ of the prior of S regulates the level of stochasticity of the user's sympathy—namely the extent to which sympathy is shaped by individual factors lying outside our model. At one extreme, $var(S)$ is zero, which makes S a deterministic function of its parents. This assumption is unrealistic since users adjust their beliefs outside of Reddit and it is impossible to capture all the variance of their sympathy. At the other extreme, $var(S) \to +\infty$ indicates that the opinion of the users is determined exclusively by factors outside of the model. Thus, we present the results for different values of $var(S)$ within these two extremes, setting it to 0.01, 1, and 100, to account for low, medium, and high levels of individual variation. Results do not significantly change for different values of $var(S)$.

All the priors of our model are (multivariate) Normal distributions. We set positive mean for the priors of $N_{sub} \to P_S$ and $N_{sub} \to P_L$, under the assumption that users have a higher propensity to write in popular subreddits. The other priors are centered in 0 and have covariance equal to the identity matrix.

4 Results

We now answer our research questions, using the coefficients estimated by our model. In our Bayesian framework, in fact, each parameter provides new knowledge on the investigated process, both about the latent variables in the system and the causal paths within the model. As introduced in Section 3.3, we analyze the model under three different assumptions, ranging from a lower level of individual variation ($var(S) = 0.01$) to a medium ($var(S) = 1$) and a higher level ($var(S) = 100$). The role of $var(S)$ is regularization: higher values trivially allows for a higher accuracy, as the model gets increasingly more free to align each individual S with activation (see Appendix D). For this reason, a comparison of the accuracy across these three models is meaningless, as an increase of $var(S)$ is sufficient to increase the accuracy. Our study of the accuracy is limited to in-sample analysis, since an essential part of our model is the estimate of individual parameter for each user (e.g., their sociodemographic features). Therefore, a train-test split would be unfeasible. Ultimately, however, the goal of our model is to assess the significance of causal relations to answer our RQs, not to obtain the highest accuracy in a prediction task.

Figure 3 presents the relevant estimated parameters, grouped by the variable they affect. We observe general robustness among the different choices of $var(S)$, as the coefficients describe similar causal effects, with minor differences in magnitude. Next, we analyze these parameters to answer our research questions.[6]

4.1 Media coverage (RQ1)

For the media coverage variables, we have multiple pathways towards activation. One is $M_L \to S \to A$. Figure 3b shows that the coefficient of sympathy to activation (i.e., β_{A1} in Equation (6)) is significantly positive for all models, as expected. Thus, we consider the relationship $M_L \to S$, expressed by the coefficients β_{S3} in Equation (3), depicted in Figure 3a. These coefficients represent how long-term media coverage affects sympathy at the end of that period. The other pathways are the direct ones $M_S \to A$ and $M_L \to A$ (β_{A3} and β_{A4} in Equation (6), shown in Figure 3b). Since the media coverage variables are composed of three distinct themes, we discuss each one separately.

The broader "Climate" theme does not present significant effects, except in $M_L \to A$ for one model instance, thus suggesting a positive, yet mild effect of climate media coverage on activism.

Instead, "Climate Action"—that is, the amount of media attention devoted to the climate activists actions—shows significantly positive effects in the same pathway towards activation for all considered models (Figure 3b). This result suggests that media coverage of climate mobilization increases the likelihood of individuals to activate in the week following the news. However, we also observe a negative coefficient in the long-term variables for "Climate Action", for two out of the three instances of the model in $M_S \to A$ (Figure 3b), and for all instances in $M_L \to S$ (Figure 3a).

We hypothesize that these negative coefficients are due to the bottom-up dynamics of the first wave of protests, that was successful despite not being covered by media during its build-up stage on Reddit. In fact, two major spikes of activations are present in our dataset, the first occuring in November 2018 when `r/EarthStrike` was created, and the second in September-October 2019, around the Global Climate Strike. The first event is endogenous to Reddit and followed a long period of low media coverage on climate action (the third week of November 2018 is at the 3$^{\text{rd}}$ percentile of climate action media coverage). Conversely, the second event occurred in times of high media coverage, and this increased media attention might have influenced the engagement in climate activism. However, this peak of activations lasted a few weeks, and the periods with the highest long-term media coverage of climate action are not associated with an increased number of activations.

Regarding the media theme of "Natural Disasters", Figure 3a shows a significant postive effect in the pathway $M_L \to S$ for one of the three instances of the model. In the same model instance, the coefficient for $M_S \to A$ is slightly negative, suggesting that awareness of climate emergencies might contribute more to the process of opinion formation on climate topics than to trigger activations directly.

In summary, media impact activations in three ways. Firstly, the organization of climate activism movements on Reddit was a grassroots initiative, spawned independent of media coverage. Secondly, coverage about climate action has a positive effect on the individuals activation, with this effect decaying after about a week. Lastly, some media coverage of extreme weather events may have a positive long-term impact on sympathy towards activism.

4.2 Demographics (RQ2)

From Figure 3a, we gather that having specific sociodemographics (i.e., 'left', 'poor', and 'old' users) increases the sympathy S. Given the positive effect of S on activation (Figure 3b) these sociodemographic features impact activation, mediated by sympathy.

Figure 3c shows that the sympathy S has an impact on P_S, as it increases the probability of participating in 'old' and 'left' subreddits. The participated subreddits also impact the probability of interacting with other activists (Figure 3d) and the interactions have a positive effect on A (Figure 3b). Thus, the sociodemographic features affect the activation, mediated by S, P_S and I.

Finally, perhaps obviously, the user sociodemographics D impact subreddit participation P_L, as being 'left' increases the probability of participating in 'left' subreddits, and analogously for all the

[6]The code to analyze our data and reproduce the results is open and available at https://github.com/jaclenti/climact/tree/main.

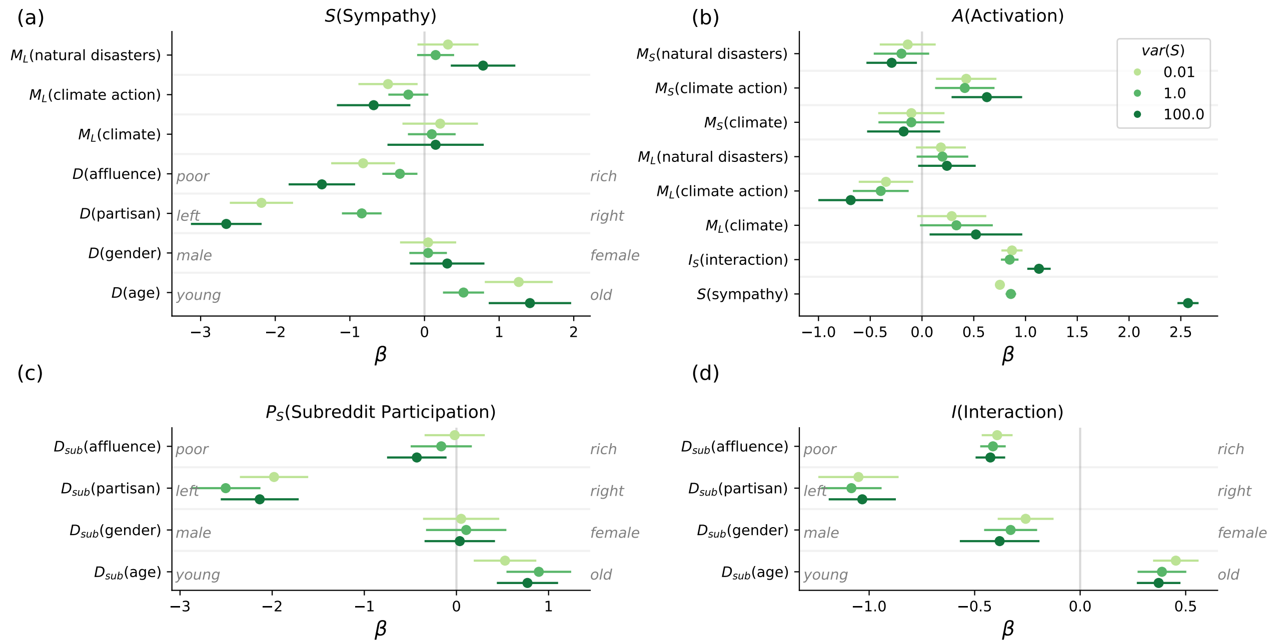


Figure 3: Mean values of the estimated parameters, with 95% credible intervals. (a) The coefficients are the causal effects of the parents of S, Equation (3). These contribute to the weighted average of the mean of the normal extraction of S. (b) The coefficients are the log-odds in Equation (6). (c) The coefficients are $\log \beta_{p1}$ in Equation (4). They represent the interaction term between D_{sub} and S. When the coefficient is positive, an increase of S increases the odds of writing in a subreddit with a positive score in that sociodemographic category. (d) The coefficients are $\log \beta_{I1}$ in Equation (5). They represent the interaction term between P_S and D_{sub}. When the coefficient is positive, an increase of $P_S \cdot D_{sub}$ increases the odds of interacting with an activist.

sociodemographic dimensions. Moreover, we estimate a positive effect of retention, as P_L affects P_S. As seen before, P_S has an effect on I, that impacts A, so we conclude that the effect of D on A is also mediated by P_L, P_S, and I.

Note that, in the context of this study, 'old' is relative to the demographics of the platform (and similarly for the other dimensions). In particular, considering the demographic of Reddit, an age label towards 'old' is not in contrast with the Youth Climate Strikes, as 64% of Reddit users are 18 to 29 years old [5]. A qualitative analysis of the sociodemographic scores indicates that the subreddits with lower age scores are typically about video games and entertainment, while politically and culturally engaged communities are more typical of older ages. In particular, some popular subreddits with high correlation between P_L and A, such as `r/environment` and `r/collapse`, have high scores for age. Moreover, the 15 subreddits with the highest correlation between P_L and A with a sociodemographic score have a 'left' partisan score, including some with an anarchist inclination such as `r/COMPLETEANARCHY`, `r/collapse`, and `r/Anarchism` or socialist (`r/ChapoTrapHouse`).

4.3 Interactions (RQ3)

Figure 3b shows that the interactions positively impact activation. This positive effect of I on A can be interpreted as already net of any homophily effect due to the common sympathy, since the effect of P_S on I is controlled by S. That is, given two users with the same level of sympathy who participate in similar subreddits, if one interacts with an activist, the odds of activation increase by approximately 2.7 times. This strong positive effect can be understood in the context of the history of the Earth Strike on Reddit. In fact, the users first joinin `r/EarthStrike` migrated from other subreddits where they decided to coordinate towards global action.

4.4 Summary of findings

Thanks to the comprehensive causal model proposed in this work, we can finally draw comparative conclusions about the determinants of climate activism online.

For RQ1, the only consistent effect is by short-term bursts media coverage of climate action ($M_S \rightarrow A$) that decays in the long-term ($M_L \rightarrow A$). Conversely, the path $M_L \rightarrow S \rightarrow A$ is not consistently significant. For RQ2, the age, partisanship, and affluence of users affect A via $D \rightarrow S \rightarrow A$. In particular, belonging to a lower economic stratum, having left-wing political preferences, and being on the older side of Reddit's population, all increase the probability of activation. Similarly, in the path $D \rightarrow S \rightarrow P_S \rightarrow I \rightarrow A$, an increase in sympathy increases the likelihood of participating in subreddits with a 'poor', 'left', 'male', and 'old' user base. These characteristics of the subreddits increase the probability of interacting with some activists, thus raising the probability of activation. All four sociodemographic features have a significant impact on A through the path $D \rightarrow P_L \rightarrow P_S \rightarrow I \rightarrow A$, as the participation in 'poor', 'left', 'male', and 'old' communities is necessarily affected by the sociodemographic features of the users. Finally, focusing on RQ3, interactions have a strong, positive, direct effect on the activation.

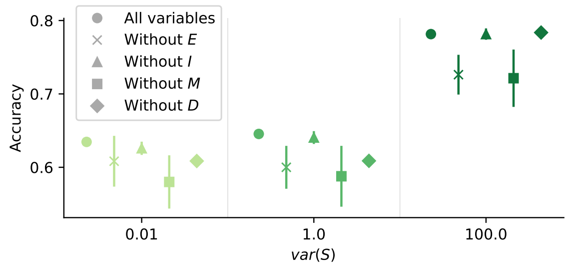


Figure 4: For each value of $var(S)$, we compare the distributions of accuracy of 100 samples of the predictive posterior for A. The error bars represent one standard deviation of the accuracy. For each $var(S)$, the left-hand dot refers to the experiment using all the variables of the model. The other dots refer to the experiment run after the removal of a variable, both long- and short-term.

That is, interacting with an activist outside of activist communities considerably increases the odds of becoming activated.

Compared to the effects of media coverage, both sympathy for the cause and direct interactions with activists have stronger effects, with coefficients approximately twice as large. Sympathy, in turn, is strongly affected by the sociodemographics of the user, especially their political partisanship, affluence, and age. The computation of the total path coefficients confirms all the results highlighted in this section (see Appendix E).

4.5 Ablation

Last, we perform an ablation analysis to check the effect of neglecting some parts of the model in our results. We run the model removing one group of variables at a time: (*i*) E_L and E_S, (*ii*) I, (*iii*) M_L and M_S and (*iv*) D. Figure 4 shows that the removal of the media coverage has the highest impact on the accuracy across the three models. When removing E, we also note that the effects of D_{sub}(affluence) and D_{sub}(gender) switch signs. This indicates that not controlling for user engagement might generate misleading effects on the interactions. In particular, while the full model captures that writing in 'male' and 'poor' subreddits increases the probability of interaction, the ablated model captures that writing in 'female' and 'rich' subreddits increases it.

This effect is indeed confounded by the positive correlation of 0.38 and 0.42 between the engagement on Reddit and the participation in 'female' and 'rich' subreddits, respectively (Appendix F), and highlights the importance of developing a comprehensive causal model. All the other effects maintain the same signs.

We also compared our results with the ones obtained by removing the gap between the short- and long-term windows, and obtained a correlation of 0.997 between the estimated coefficients in the two settings (Appendix B). This robustness check shows that the gap period between long- and short-term past is not necessary.

5 Discussion

We developed a rich and comprehensive causal model to study the interplay between different determinants of climate activism on Reddit. This work represents a first attempt to apply a multi-causal model to social media data. Our results show that the simultaneous inclusion of several causal pathways into the same model provides new insights into the underlying social process. For instance, the interactions play a central role in the engagement in climate activism communities. However, part of this effect is explained by the confounder of the frequented subreddits, which is controlled by the sympathy for the climate cause. Similarly, we show that the sociodemographic features influence the activation. However, the ablation analysis reveals that if we omit to control for the engagement on Reddit, some effects change drastically and flip signs, due to the correlation between the engagement and specific sociodemographic features. Finally, we illustrate that the exclusion of media coverage from the model significantly decreases its accuracy.

The contribution of this work is two-fold. First, we establish an operative pipeline to causally test hypotheses in computational social science using large-scale empirical data (Reddit and GDELT). The pipeline targets the existing gap between behavioral models in social science, which are usually validated via survey or lab studies, and social media analysis, which typically draws causal conclusions only by focusing on the temporal dimension.

Second, we provide new insights into the phenomenon of climate activism online. We unveil a temporally complex effect of media coverage of climate activism. On the one hand, the effect of media attention on climate action rapidly decays, with a strong short-term effect and a negative effect in the long term. On the other hand, media coverage around climate and natural disasters is persistent over a long period. These results hint at a trend effect for climate strikes, and a memory effect for people experiencing extreme events [45]. Users participating in left-leaning and lower-income communities have a higher propensity to engage in climate activism, mediated both by their sympathy towards the cause and by the subreddits they frequent, which expose them to interactions with other activists. We acknowledge that these sociodemographic features are an oversimplification of the users' real attributes, as they depend on the scores previously assigned to subreddits. In the future, it would be interesting to study how different estimators of these factors affect the results. Last, social influence has the strongest impact on user activation.

We underline that the model aims to uncover *how* or the *why* people engage in activism, and it is not intended to *predict* user engagement. For these reason, we do not approach the task through machine-learning, which could improve the accuracy of our estimates by leveraging spurious (i.e., non-causal) correlations in the data. For instance, we did not measure the out-of-sample accuracy, because the latent sociodemographic features of the users are estimated during model training, and we cannot predict the activation of a user without injecting them into the model.

We recognize that active participation in a climate activist community is not necessarily linked to joining a climate strike. However, our framework is based on well-grounded sociological theories that apply to the online context too [47]. From the moment a user becomes active in such an activist community, they will become possible targets of mobilization by the group. The remaining steps of the process happen outside of the social media sphere, becoming harder to study [28] and highlighting the need for broader methodologies to understand the full spectrum of social coordination, from digital engagement to physical mobilization and collective action.

References

[1] 2022. Demographics of Emissions. *Nature Climate Change* 12, 3 (March 2022), 209–209. https://doi.org/10.1038/s41558-022-01325-5 (Cited on 1)

[2] Rajeev Airani and Kiran Karande. 2022. How social media effects shape sentiments along the twitter journey? A Bayesian network approach. *Journal of Business Research* 142 (2022), 988–997. (Cited on 2)

[3] Duilio Balsamo, Paolo Bajardi, and André Panisson. 2019. Firsthand opiates abuse on social media: monitoring geospatial patterns of interest through a digital cohort. In *The World Wide Web Conference*. 2572–2579. (Cited on 4)

[4] Sebastian Bamberg, Jonas Rees, and Sebastian Seebauer. 2015. Collective climate action: Determinants of participation intention in community-based proenvironmental initiatives. *Journal of Environmental Psychology* 43 (2015), 155–165. (Cited on 2, 3)

[5] Michael Barthel, Galen Stocking, Jess Holcomb, and Amy Mitchell. 2016. Reddit news users more likely to be male, young and digital in their news preferences. *Pew Research Center* (2016). (Cited on 7)

[6] Ana-Maria Bliuc, Craig McGarty, Katherine Reynolds, and Daniela Muntele. 2007. Opinion-based group membership as a predictor of commitment to political action. *European journal of social psychology* 37, 1 (2007), 19–32. (Cited on 3)

[7] Kaiping Chen, Amanda L Molder, Zening Duan, Shelley Boulianne, Christopher Eckart, Prince Mallari, and Diyi Yang. 2023. How climate movement actors and news media frame climate change and strike: Evidence from analyzing twitter and news media discourse from 2018 to 2021. *The International Journal of Press/Politics* 28, 2 (2023), 384–413. (Cited on 2)

[8] Kimberly C Doell, Boryana Todorova, Madalina Vlasceanu, JB Bak-Coleman, Jay J Van Bavel, and Ekaterina Pronizius. 2024. The International Climate Psychology Collaboration: Climate change-related data collected from 63 countries. (2024). (Cited on 1)

[9] Ottmar Edenhofer. 2015. *Climate change 2014: mitigation of climate change*. Vol. 3. Cambridge University Press. https://archive.ipcc.ch/pdf/assessment-report/ar5/wg3/drafts/fgd/ipcc_wg3_ar5_summary-for-policymakers_may-version.pdf (Cited on 1)

[10] Johannes Emmerling, Pietro Andreoni, Ioannis Charalampidis, Shouro Dasgupta, Francis Dennig, Toon Vandyck, Simon Feindt, Dimitris Fragkiadakis, Panagiotis Fragkos, Shinichiro Fujimori, et al. 2024. A Multi-Model Assessment of Inequality and Climate Change. (2024). (Cited on 1)

[11] Max Falkenberg, Alessandro Galeazzi, Maddalena Torricelli, Niccolò Di Marco, Francesca Larosa, Madalina Sas, Amin Mekacher, Warren Pearce, Fabiana Zollo, Walter Quattrociocchi, et al. 2022. Growing polarization around climate change on social media. *Nature Climate Change* 12, 12 (2022), 1114–1121. (Cited on 1, 3)

[12] Noah E Friedkin. 1998. *A structural theory of social influence*. Cambridge University Press. (Cited on 5)

[13] Noah E Friedkin and Eugene C Johnsen. 1990. Social influence and opinions. *Journal of Mathematical Sociology* 15, 3-4 (1990), 193–206. (Cited on 5)

[14] Franziska Gaupp and Sibel Eker. 2024. Climate Activism, Social Media and Behavioural Change: A Literature Review. (2024). (Cited on 2)

[15] Heejin Han and Sang Wuk Ahn. 2020. Youth mobilization to stop global climate change: Narratives and impact. *Sustainability* 12, 10 (2020), 4127. (Cited on 1, 2)

[16] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. 2013. Stochastic variational inference. *Journal of Machine Learning Research* (2013). (Cited on 5)

[17] IPCC. 2022. *Summary for Policymakers*. Cambridge University Press, Cambridge, UK, In Press. https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii In Press. (Cited on 1)

[18] Md Mofakkarul Islam, Andrew Barnes, and Luiza Toma. 2013. An investigation into climate change scepticism among farmers. *Journal of Environmental Psychology* 34 (2013), 137–150. (Cited on 3)

[19] Tunazzina Islam, Ruqi Zhang, and Dan Goldwasser. 2023. Analysis of Climate Campaigns on Social Media using Bayesian Model Averaging. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 15–25. (Cited on 3)

[20] Andrei P Kirilenko, Tatiana Molodtsova, and Svetlana O Stepchenkova. 2015. People as sensors: Mass media and local temperature influence climate change discussion on Twitter. *Global Environmental Change* 30 (2015), 92–100. (Cited on 1, 2, 3)

[21] Jouni Kuha and John H Goldthorpe. 2010. Path analysis for discrete variables: the role of education in social mobility. *Journal of the Royal Statistical Society Series A: Statistics in Society* 173, 2 (2010), 351–369. (Cited on 11)

[22] Jacopo Lenti, Corrado Monti, and Gianmarco De Francisci Morales. 2024. Likelihood-Based Methods Improve Parameter Estimation in Opinion Dynamics Models. *Proceedings of the 17th ACM International Conference on Web Search and Data Mining* (2024). (Cited on 3)

[23] Jacopo Lenti, Fabrizio Silvestri, and Gianmarco De Francisci Morales. 2024. Variational Inference of Parameters in Opinion Dynamics Models. *arXiv preprint arXiv:2403.05358* (2024). (Cited on 3)

[24] Joan Massachs, Corrado Monti, Gianmarco De Francisci Morales, and Francesco Bonchi. 2020. Roots of trumpism: Homophily and social feedback in donald trump support on reddit. In *Proceedings of the 12th ACM conference on Web Science*. 49–58. (Cited on 2)

[25] AV Mavrodieva, OK Rachman, VB Harahap, and R Shaw. 2019. Role of social media as a soft power tool in raising public awareness and engagement in addressing climate change. Climate, 7 (10), 122. (Cited on 1, 3)

[26] Stefan D McCabe, Diogo Ferrari, Jon Green, David MJ Lazer, and Kevin M Esterling. 2024. Post-January 6th deplatforming reduced the reach of misinformation on Twitter. *Nature* 630, 8015 (2024), 132–140. (Cited on 2)

[27] Craig McGarty, Ana-Maria Bliuc, Emma F Thomas, and Renata Bongiorno. 2009. Collective action as the material expression of opinion-based group membership. *Journal of Social Issues* 65, 4 (2009), 839–857. (Cited on 3)

[28] Yelena Mejova, Jisun An, Gianmarco De Francisci Morales, and Haewoon Kwak. 2022. Modeling Political Activism around Gun Debate via Social Media. *ACM Transactions on Social Computing* (April 2022), 3532102. https://doi.org/10.1145/3532102 (Cited on 8)

[29] Yelena Mejova and Nicolas Kourtellis. 2021. Youtubing at home: Media sharing behavior change as proxy for mobility around covid-19 lockdowns. In *Proceedings of the 13th ACM Web Science Conference 2021*. 272–281. (Cited on 4)

[30] Marco Minici, Federico Cinus, Corrado Monti, Francesco Bonchi, and Giuseppe Manco. 2022. Cascade-based echo chamber detection. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 1511–1520. (Cited on 3)

[31] Corrado Monti, Gianmarco De Francisci Morales, and Francesco Bonchi. 2020. Learning opinion dynamics from social traces. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 764–773. (Cited on 3)

[32] Corrado Monti, Giuseppe Manco, Cigdem Aslay, and Francesco Bonchi. 2021. Learning ideological embeddings from information cascades. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 1325–1334. (Cited on 3)

[33] Sally Neas, Ann Ward, and Benjamin Bowman. 2022. Young people's climate activism: A review of the literature. *Frontiers in Political Science* 4 (2022), 940876. (Cited on 3)

[34] Rebecca Newman and Ilan Noy. 2023. The global costs of extreme weather that are attributable to climate change. *Nature Communications* 14, 1 (2023), 6103. (Cited on 4)

[35] Judea Pearl. 2009. *Causality*. Cambridge university press. (Cited on 4)

[36] Shruti Phadke, Mattia Samory, and Tanushree Mitra. 2021. What makes people join conspiracy communities? role of social factors in conspiracy engagement. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–30. (Cited on 2)

[37] Hemant Purohit, Yiye Ruan, Amruta Joshi, Srinivasan Parthasarathy, and Amit Sheth. 2011. Understanding user-community engagement by multi-faceted features: A case study on twitter. In *WWW 2011 Workshop on Social Media Engagement (SoME)*. (Cited on 2)

[38] Nicholas Roxburgh, Dabo Guan, Kong Joo Shin, William Rand, Shunsuke Managi, Robin Lovelace, and Jing Meng. 2019. Characterising climate change discourse on social media during extreme weather events. *Global environmental change* 54 (2019), 50–60. (Cited on 1, 3)

[39] Giuseppe Russo, Manoel Horta Ribeiro, and Robert West. 2023. Stranger Danger! Cross-Community Interactions with Fringe Users Increase the Growth of Fringe Communities on Reddit. *arXiv preprint arXiv:2310.12186* (2023). (Cited on 2, 5)

[40] Anandita Sabherwal, Matthew T Ballew, Sander van Der Linden, Abel Gustafson, Matthew H Goldberg, Edward W Maibach, John E Kotcher, Janet K Swim, Seth A Rosenthal, and Anthony Leiserowitz. 2021. The Greta Thunberg Effect: Familiarity with Greta Thunberg predicts intentions to engage in climate activism in the United States. *Journal of applied social psychology* 51, 4 (2021), 321–333. (Cited on 3)

[41] Maxie Schulte, Sebastian Bamberg, Jonas Rees, and Philipp Rollin. 2020. Social identity as a key concept for connecting transformative societal change with individual environmental activism. *Journal of Environmental Psychology* 72 (2020), 101525. (Cited on 2, 3)

[42] Matthew R Sisco, Valentina Bosetti, and Elke U Weber. 2017. When do extreme weather events generate attention to climate change? *Climatic change* 143 (2017), 227–241. (Cited on 1, 3)

[43] H Tajfel. 1978. Social categorization, social identity and social comparison. *Differentiation between Social Groups: Studies in the Social Psychology of Intergroup Relations* (1978). (Cited on 1)

[44] Henri Tajfel. 1979. An integrative theory of intergroup conflict. *Intergroup relations: Essential readings* (1979). (Cited on 1)

[45] Maddalena Torricelli, Max Falkenberg, Alessandro Galeazzi, Fabiana Zollo, Walter Quattrociocchi, and Andrea Baronchelli. 2023. How does extreme weather impact the climate change discourse? Insights from the Twitter discussion on hurricanes. *Plos Climate* 2, 11 (2023), e0000277. (Cited on 1, 2, 3, 5, 8)

[46] Lauri Valkonen, Jouni Helske, and Juha Karvanen. 2023. Estimating the causal effect of timing on the reach of social media posts. *Statistical Methods & Applications* 32, 2 (2023), 493–507. (Cited on 2)

[47] Sander Van der Linden. 2015. The social-psychological determinants of climate change risk perceptions: Towards a comprehensive model. *Journal of environmental psychology* 41 (2015), 112–124. (Cited on 2, 3, 8)
[48] Jacquelien Van Stekelenburg and Bert Klandermans. 2013. The social psychology of protest. *Current Sociology* 61, 5-6 (2013), 886–905. (Cited on 1, 2, 3)
[49] Martijn Van Zomeren and Aarti Iyer. 2009. Introduction to the social and psychological dynamics of collective action. , 645–660 pages. (Cited on 1)
[50] Isaac Waller and Ashton Anderson. 2021. Quantifying social organization and political polarization in online platforms. *Nature* 600, 7888 (2021), 264–268. (Cited on 4)
[51] Jaesun Wang and Seoyong Kim. 2018. Analysis of the impact of values and perception on climate change skepticism and its implication for public policy. *Climate* 6, 4 (2018), 99. (Cited on 3)
[52] Justine Zhang, William Hamilton, Cristian Danescu-Niculescu-Mizil, Dan Jurafsky, and Jure Leskovec. 2017. Community identity and user engagement in a multi-community landscape. In *Proceedings of the international AAAI conference on web and social media*, Vol. 11. 377–386. (Cited on 2, 5)

Acknowledgments

LMA acknowledges the support from the Carlsberg Foundation through the COCOONS project (CF21-0432).

A Subreddits

To capture the climate activism movements, we consider all the organizations that endorsed Earth Strike. We select the organizations with a subreddit with the same name, and we focus only on the subreddits that explicitly mention climate action in their bio (or in their official Webpage if the bio is missing). This is our final list of subreddits, and the relative bio (as of 2024/07/10):

- `r/FridaysForFuture`: "A sub dedicated to the international movement of students who skip class on Fridays to demand action to prevent climate change."
- `r/SunriseMovement`: "Sunrise is a movement to stop climate change and create millions of good jobs in the process."
- `r/ExtinctionRebellion`: "This is our darkest hour. Humanity finds itself embroiled in an event unprecedented in its history, one which, unless immediately addressed, will catapult us further into the destruction of all we hold dear: our nations, its peoples, our ecosystems, and the future of generations to come. The science is clear: we are in the midst of the sixth mass extinction event on planet Earth and we will face catastrophe if we do not act swiftly and robustly[...] The Extinction Rebellion is a lawfully abiding direct action movement challenging inaction over dangerous climate change."
- `r/EarthStrike`: "Earth Strike is a grassroots labor-environmental movement focused on organizing a GLOBAL GENERAL STRIKE TO SAVE THE PLANET! [...] This subreddit is for discussion about Earth Strike, a general strike movement for climate action."
- `r/ClimateRealityProject`: "Our mission is to catalyze a global solution to the climate crisis by making urgent action a necessity across every level of society."
- `r/350ppm`: "A subreddit for the organizing of the Climate Action Movement and various Campaigns."
- `r/sierraclub`: [from the webpage] "Sierra Club is working tirelessly to protect wildlife and wild places, ensure clean air and water for all, and fight the devastating effects of climate change. Sierra Club is the most historic grassroots environmental organization in the country. For more than 130 years, we've gathered millions of activists and volunteers to fight for the places, people, and planet we all love."

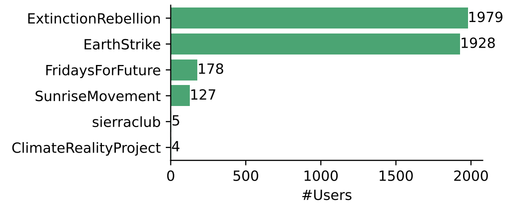


Figure A.1: Number of activated users per subreddit in our final user selection.

We excluded subreddits endorsing Earth Strike that are not focused on climate action but rather on environmental or social action in general: `r/Amnesty`, `r/AmnestyInternational`, `r/greenpeace`, `r/LushCosmetics`, `r/198Methods`, `r/MarchForScience`, `r/OxfamAmerica`, and `r/WWF`,

B Robustness

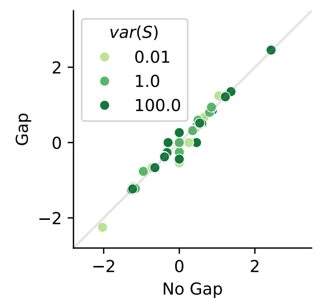


Figure B.2: Dots represent parameters or latent variables of the model, excluding D and S. We compare the estimates of the mean of these parameters, with and without the gap between long- and short-term periods. The high correlation (0.997) confirms that the choice of the endpoint of the long-term period between t_{A-W} and t_{A-5W} does not affect the results.

C Georeference

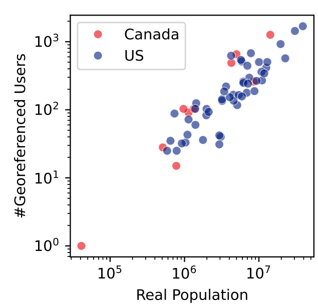


Figure C.3: Population of US States (blue) or Canadian provinces (red) vs. the number of georeferences per administrative area obtained from our sample of 17k users.

D Sympathy

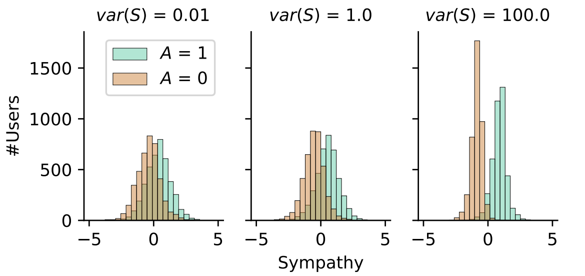


Figure D.4: The plots show the distributions of S at the varying of $var(S)$. The separation of the distributions increases with $var(S)$, which reflects the increased alignment with A.

E Total Path Coefficients

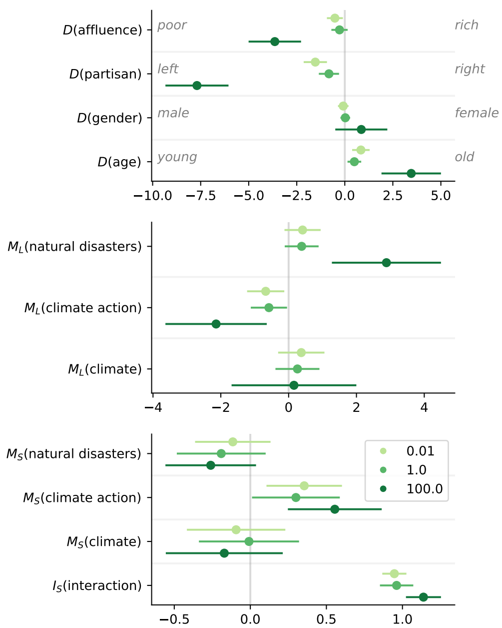


Figure E.5: The errorbar show the average total path coefficients, and the corresponding credible intervals, of each variable on A. The results are based on the linear approximation of the causal model. Notice that the comparison of the effect is possible within each plot, but not between the plots.

To account for both the direct and indirect effects of the variables on A, we study the total path coefficients. While the use of total path coefficients is established in linear models, their computation in non-linear models remains an open research question [21]. Therefore, we approximate the problem by assuming linearity for all effects, similar to previous studies [21]. To address the research questions (RQs), we focus on the following effects on A:

- D affects A through $D \rightarrow S \rightarrow A$, $D \rightarrow S \rightarrow P_S \rightarrow I \rightarrow A$ and $D \rightarrow P_L \rightarrow P_S \rightarrow I \rightarrow A$.
- M_L affects A via $M_L \rightarrow A$ and $M_L \rightarrow S \rightarrow A$.
- I directly affects A, $I \rightarrow A$.
- M_S directly affects A, $M_S \rightarrow A$.

Since we apply an assumption of linearity to a non-linear model, the path coefficients should be interpreted carefully, focusing on their sign rather than their magnitude. Furthermore, we cannot compare the magnitudes of the effects of variables entering the DAG at different depths.

We calculate indirect effects by multiplying the individual coefficients along each path. However, we observe a disproportionate influence of D_{sub} on these coefficients. Specifically, an increase in $D(\text{age})$ raises the probability of participating in all subreddits with positive $D_{sub}(\text{age})$, both in long-term and short-term past. Each of these subreddits independently influences I, and consequently A. Under the linear approximation, this results in the summation of $K = 826$ effects corresponding to the K subreddits. In the original non-linear model, these effects are moderated by the non-linear function $\sigma(\cdot)$. To address this, we average the effects of the subreddits for all paths involving D_{sub}. This approach allows us to present the effects of the variables in terms of their impact on the probability of joining an average subreddit. The path coefficients are shown in Figure E.5.

F Sociodemographic Features

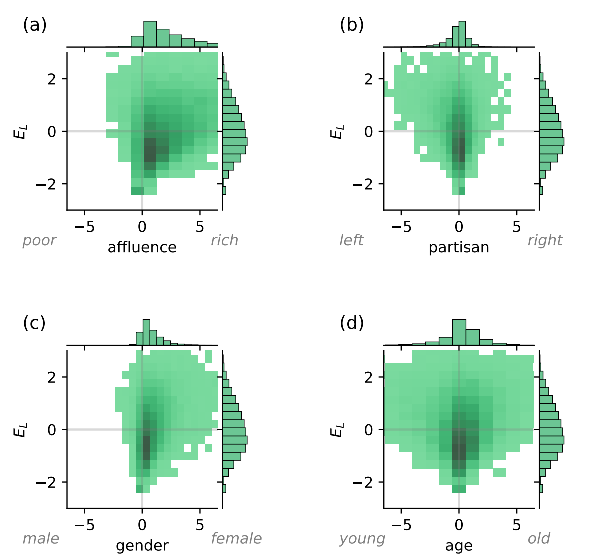


Figure F.6: Joint distributions of the 4 sociodemographic scores and the z-scores of the Reddit engagement. The sociodemographic features are proxies computed as the dot product between the vector of subreddit users participate in long-term and the sociodemographic scores of the subreddits. These values are different from D, which is a latent variable estimated by the model. A positive correlation is noticeable in (a) and (c), highlighting that users who participate more in more 'rich' and 'female' subreddits tend to comment more. (a), (b) and (c) are also asymmetric, showing that our dataset is biased toward 'rich', 'left', and 'female' users.