Social Networks xxx (xxxx) xxx

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

The rise of #climateaction in the time of the FridaysForFuture movement: A semantic network analysis

Caterina Suitner [a], Leonardo Badia [b], Damiano Clementel [c,1], Laura Iacovissi [d,1], Matteo Migliorini [e,1], Bruno Gabriel Salvador Casara [a], Domenico Solimini [f], Magdalena Formanowicz [g], Tomaso Erseghe [b,*]

[a] Department of Developmental Psychology and Socialization, University of Padova, Italy
[b] Department of Information Engineering, University of Padova, Italy
[c] Human Inspired Technology Research Centre, University of Padova, Italy
[d] Bosch Industry on Campus Lab, University of Tübingen, Tübingen, Germany
[e] INFN, Istituto Nazionale di Fisica Nucleare, Padova, Italy
[f] Department of Mathematics "Tullio Levi-Civita", University of Padova, Italy
[g] Center for Research on Social Relations, SWPS, University of Social Sciences and Humanities, Warsaw, Poland

ARTICLE INFO

Keywords:
Climate change
Semantic networks
Collective action
Network analysis
Community detection
Social identity
Empowerment
Future orientation
Hashtag activism

ABSTRACT

We investigate the psycho-linguistic features of the online discourse over climate change, focusing on its modifications throughout the years 2017–2019 as a result of collective actions emerging and spreading worldwide. We seek to understand the emerging connection between digital activism and the psychological processes related to its social drives. To this end, a semantic network is derived from the social platform Twitter, and its evolution is traced over time, tracking textual proxies of social identity and empowerment. Original proposals are made to identify communities and highlight the most important semantic contents of the corpus from a network perspective. These evaluations on semantic communities of related concepts further detail the shift in the rhetoric of collective actions. Finally, we explore projection of the ingroup to the future in the online discourse about climate change, which can point to developments of pro-environmental campaigns.

1. Introduction

A recent survey on climate change perception in the US (Leiserowitz et al., 2020) informs that there is a steady increase of people concerned with the climate issue, and that the number of climate deniers is in decline. Accordingly, social discourses about climate change have been generating significant engagement over the last few years, especially over online platforms. In 2019, a spike of global interest was reached through the actions that many credited to 16-year old Swedish activist Greta Thunberg. As a result of the "School strikes for climate" and "Fridays for future" initiatives, many voiced dissent against the world governments and their passive behavior toward the problem of anthropogenic climate change, claiming that, to revise the civic agenda about the environmental policies, a shift from a descriptive approach of climate change toward an action oriented discourse is needed.

To understand the evolution of collective content created in public discourses about climate change, and more specifically in online climate action, we use Twitter as a source, and network analysis as a methodological framework. As a matter of fact, topics involving anthropogenic global warming are actually discussed more on social platforms than on traditional media (Boykoff, 2011). Twitter is here considered as an arena (and window) of the online protest around climate change, providing an accessible ground, for both actors and researchers, to structure relations among different individuals, issues, and events over time. By focusing on this social medium, we acknowledge the role of the Internet and online communities as nurturing, fueling, and reflecting the protests (Vasi and Suh, 2013). In Twitter, even individuals without a previous celebrity status can become spokespersons of social instances such as environmentalism (Abidin et al., 2020). This is the case of Greta Thunberg, whose Twitter account is followed by almost 5 million users

* Correspondence to: Dipartimento of Ingegneria dell'Informazione, via G Gradenigo 6/b, 35131, Padova, Italy.
E-mail addresses: caterina.suitner@unipd.it (C. Suitner), leonardo.badia@unipd.it (L. Badia), damiano.clementel@studenti.unipd.it (D. Clementel), laura.iacovissi@uni-tuebingen.de (L. Iacovissi), matteo.migliorini@pd.infn.it (M. Migliorini), brunogabriel.salvadorcasara@unipd.it (B.G. Salvador Casara), domenico.solimini@studenti.unipd.it (D. Solimini), mformanowicz@swps.edu.pl (M. Formanowicz), tomaso.erseghe@unipd.it (T. Erseghe).
[1] Formerly with the Department of Mathematics "Tullio Levi-Civita", University of Padova, Italy.

https://doi.org/10.1016/j.socnet.2022.06.003

as of March 2021. She uses tweets as one of the main communication channels to spread awareness about climate change and call to action.

Since the content on Twitter is arranged through argumentation tags (i.e., hashtags), it is possible to build rhetorical networks towards which the on-line protest gravitates (Bastos and Zago, 2013). Here, in the context of climate change, we aim at characterizing the entire semantic network of related hashtags and more general online content (Breslin and Decker, 2007) over the course of three years. For the sake of avoiding biases, we start from neutral hashtags, popular throughout the entire observation span, and we infer the semantic similarity in the network of tweets through topic-sensitive ranking (Haveliwala, 2003). This approach allows identifying semantic communities of the conversations on climate change and, specifically, differentiating communities that focus on climate action. This further leads us to the development of a methodological proposal, an original contribution of the present paper called Personalized PageRank Projection (PPRP), which allows assigning tweets to semantic communities (i.e., topics).

We analyze the language used to discuss about climate issue under the socio-psychological lens of the collective action theoretical framework (Van Zomeren and Iyer, 2009). According to this theory, collective action is defined as any action addressing a goal that surpasses individual interest (Van Zomeren et al., 2008). When people find the issue at hand as unfair or morally wrong, they may undertake protests, but only if specific psychological instances are met. The most important predictors of engagement is whether they *identify* with the topic or consider it important. A second crucial predictor is whether they have a sense of *efficacy*, i.e., a belief that their actions can contribute to a broader change. Importantly, both predictors were found to be positively related to the environmental engagement (Bamberg et al., 2015). As we quantify the textual proxies of the aforementioned concepts (i.e., identity and efficacy), we are able to spot certain trends that are stronger for the precise context of climate action semantic community. The application of this theoretical framework to the collective social discourse moves the socio-psychological discussion on collective action a step forward, by acknowledging the social construction of a common meaning. We therefore move beyond the individual level analysis which has characterized the majority of the extant literature, which has mainly focused on attitudes, beliefs, and actions of single individuals (McAdam, 2017), while neglecting the "collective" part of collective action. Moreover, we here tackle the dynamism of this meaning making process by comparing of different time frames, which reveals whether language reflects socio-psychological motives and processes that change depending on the semantic community (Borge-Holthoefer and Arenas, 2010).

The rest of the paper is organized as follows. Section 2 reviews the state-of-the-art, while Section 3 describes the technical procedures used for the analysis. The main results are presented in Section 4 and discussed in Section 5, which concludes the paper.

2. State of the art

2.1. Collective action: Why and when people protest?

The socio-psychological literature on collective action explores what makes single individuals move toward social change. Any collective action mobilizes people to reach a shared goal, be it a strike, a march, an online petition or a vote, with the aim of improving the condition of a group (Van Zomeren and Iyer, 2009). Often personal and group interests are congruent, however engaging in a collective action requires strong identification with the topic and a belief that one is capable in achieving one's goal, to overcome the effort involved in the action and prioritize the group utility over the individual costs. Therefore, we focus on two central psychological processes that the literature is consistently showing as predictors of protest engaging, namely *affiliation* (or identity) and *empowerment*.

The social identity theory (Tajfel, 1974) is a seminal framework to explain the reasons behind group dynamics. According to this theoretical perspective, the inclination to behave in terms of group *affiliation* relates to the extent to which the social identity is relevant and important to the person. The value and emotional significance of this belonging influences what people do and think in the context of social relations, and explains the profound reasons that prompt people to act for the interest of the group (van Zomeren et al., 2018). A solid corpus of evidence highlights that group identification is associated with both intentions and behaviors related to collective action. For example, social identification was linked to intention to participate in collective actions for the elderly (as in study 1 by Simon et al., 1998), LGBT minorities (study 2 by Simon et al., 1998; Stürmer and Simon, 2004), and women rights (Kelly and Breinlinger, 1995). Ellemers et al. (1999) suggested that high identifiers are more concerned about and committed to group goals and interests than lower identifiers, who are more committed to their individual goals and interests. Importantly, affiliation is also associated to actual behaviors (Foster, 1999). Kawakami and Dion (1993) showed that people are more likely to engage in positive collective action, such as asking the group for help, than negative individual action, e.g., leaving the group, when their group membership is salient.

The second central drive for engaging in collective action is *empowerment* (Drury and Reicher, 1999). Empowerment refers to the sense that the goal can be achieved, and is also labeled as effectiveness (Hornsey et al., 2006), efficacy (Van Zomeren et al., 2008), or agency (Jasper, 2004). Scholars conceptualized empowerment at different scopes, i.e., personal, group, or goal levels. At the personal level, empowerment corresponds to the individual perception of being able to contribute the cause (e.g., Tagkaloglou and Kasser, 2018). At the group level, empowerment refers to the idea that enough people can be mobilized to achieve the goals (Berman and Wittig, 2004; Stürmer and Simon, 2004) or that a group can collectively reach a social change (Bandura, 2000). Thus, empowerment refers to the belief that the collective goal can be achieved (Lee Fox and Schofield, 1989; Tyler and McGraw, 1983). What all these definitions share is that a lack of perceived effectiveness prevents people from engaging in collective action. The idea is that there is no point in protesting if there is no hope for future change (Abramson and Aldrich, 1982; Verba and Nie, 1972). Only when people feel that goals are achievable, they consider joining a collective action such as a union meeting (Flood, 1993), or support a petition for bilateral disarmament (Lee Fox and Schofield, 1989).

Affiliation and empowerment predict both online and offline actions, such as signing an online or pen-and-paper petition (Brunsting and Postmes, 2002). Importantly, they are also relevant from the perspective of the environmental action (Fritsche et al., 2018). In terms of affiliation, several studies confirmed the relation between social identification and environmentalism (Dono et al., 2010; Fielding et al., 2008; Brügger et al., 2011). Moreover, Rees and Bamberg (2014) showed that the intention to participate in a local pro-environment action was stronger among those who reported a high sense of community related to their neighborhood. Schmitt et al. (2019) argued that the engagement in environmental activism was predicted by the extent respondents self-identified as politicized environmentalist. Similarly, empowerment can give strong motivations for climate change activism (Fritsche et al., 2018). The salience of the threat associated to climate change enhances collective efficacy (Hornsey et al., 2015), which in turn promotes collective action (van Zomeren et al., 2010) such as taking part in a neighborhood-based climate protection initiative (Rees and Bamberg, 2014).

While affiliation and empowerment are motives common to both environmental (Fritsche et al., 2018) and general collective action tendencies (van Zomeren et al., 2010), orientation to the future may be critical for the collective action related to the environment (McAdam, 2017), since the consequences of climate change need a long-term appraisal to concern people (Sarigöllü, 2009). The literature on the time frame issue in this context provides support for an association between pro-environmental attitudes and future orientation; for a meta-analysis see Milfont et al. (2012). For example, Corral-Verdugo et al. (2006) specifically assessed Zimbardo's time perspective inventory (Boyd and Zimbardo, 2005) and showed that future orientation, namely the tendency to foresee future events, was positively associated to how participants reported water preserving behaviors such as conserving water while washing dishes. Along the same lines, considering future consequences was associated to higher pro-environment attitudes, such as support to public transports (Joireman et al., 2001, 2004) and sustainable consuming (Lindsay and Strathman, 1997).

However, most studies on collective action grounded on the theory of social identity measure relevant indices at the *individual* level, and only observe the participation of single units. Thus, the main assessment concerns the personal contribution of group members to the collective action, whereas the *group* dynamic of meaning, creation, and evolution over time is somehow neglected. The adoption of the analytical instruments of network science represents a progress toward filling this gap.

2.2. Social networks: Beyond the sum of the people

Socially driven actions assume a central role in collective change and the achievement of common goals. This perspective overcomes both the individual and the group as a collection of individuals and rather approaches social processes as a network whose entirety is grasped only if regarded as an entity that is more than the sum of its parts (Robins and Pattison, 2005). From this standpoint, collective action is embedded as a complex substrate of social processes (Lewin, 2016). In fact, the rhetoric of an online discourse is a global phenomenon, whose evolution is promoted by countless micro-interactions at the individual level, which are better represented with a holistic characterization (Kirby et al., 2014). Congruently, the push toward collective action also happens through multiple gradual nudges that are hard to identify through individual actions (Stürmer and Simon, 2004).

Network science offers a conceptual and methodological framework to analyze such interdependencies. It is a cross-disciplinary field of study investigating complex systems exhibiting a fundamental characteristic of networked interconnection (Barabási et al., 2016; Newman, 2001), and its application to the construction of shared reality tackles the co-creational perspective of activism within social media (Lewis et al., 2010). In this sense, network science permits investigating holistic motives starting from network dependencies, e.g., from semantic interrelations that appear in the Twitter social media (Hellsten and Leydesdorff, 2020), and can be readily applied to the socio-psychological threads previously mentioned as motives for the collective engagement within climate change. Indeed, social media have several important features that make them one of the best contexts for studying the attitudes of people and the rhetoric of messages around climate change.

2.3. Social media as reality mirrors

Since the second decade of the 21st century, online media have been more popular than traditional mass media like television and newspapers, and their use is increasing over time (Newman et al., 2017). Social media allow a many-to-many communication exchange, whereas traditional media are only one-to-many. Thus, millions of people with different backgrounds can express themselves in social media by sharing their opinions, forming online groups, and organizing both offline and online collective actions.

Online social networks create the opportunity for large scale collective action with its content triggering cascade effects of social influence. Importantly, the persuasion processes exerted within social media crosses the borders of the virtual reality affecting actual behaviors. For example, an experimental study involving 61 million online users investigated the impact of the voting behavior of one's Facebook friends (Bond et al., 2012). Specifically, they are more likely to vote when they see that their Facebook friends already voted, and they are also more inclined to share their own behaviors, turning into collective actors themselves and influencing other users.

Even though digital activism faces specific hurdles, related to anonymity, lack of accountability, heterogeneity, uncertainty and emotional detachment (Jagers et al., 2019), the ease, velocity, and vastity of information sharing facilitates the creation and mobilization of huge social communities fostering the discussion and conceptualization of the social issue (Keller, 2012; Pudrovska and Ferree, 2004). The so-called hashtag activism is the ground for awareness rising and public debating on several causes and targets, including protest to defend the rights of racial minorities (`#BlackLivesMatter`, `#Ferguson`), to promote gender equality (`#DressLikeAWoman`, `#HeForShe`), to fight hate speech (`#StopFundingHate`) or economic inequalities (González-Bailón and Wang, 2016). Given the number of stakeholders involved in the social media arena (e.g., journalists, politicians, activists, business actors), communicating with the same hashtags, the investigation of the semantic networks around such hashtags allows for the understanding of the rhetoric involved in the general discourse, capturing the interplay of this multitude of actors and perspectives.

Among social media, we identify Twitter as particularly suited to analyze online rhetoric systems, since it one of the biggest platforms for micro-blogging and, to some extent, also combines instant messaging, social networking, and status communication (Ross et al., 2011). Daily interactions on Twitter can be viewed as the signal of a distributed network of human sensors where the value is a product of its interconnected structure (Boyd and Crawford, 2012). Also, the messages shared on Twitter have a condensed structure which is well suited for text analysis methods because tweets have to express brief but complete and meaningful concepts, which is often highlighted by the use of keywords (Bastos and Zago, 2013; Kirilenko and Stepchenkova, 2014). The popularity of Twitter among a variety of social actors, from individuals to organizations, allows to capture a discourse which is co-created by ordinary citizens, politicians, journalist, activists, and experts. Moreover, messages on Twitter are publicly available, and Twitter even allows for the web-scraping of the messages through its API; thus, we can collect data in aggregate anonymized form without affecting the privacy of the users.

A feature of Twitter which is particularly relevant for the present study is the use of hashtags for semantic and channel tagging and meta-communication. Hashtags can be used to organize on-line content and offer a very brief opinion toward a topic; tweets often use hashtags in both ways by putting multiple hashtags in a single message. Previous literature showed that hashtags on Twitter are also employed to brand advocacy movements and archive messages for the movements (Saxton et al., 2015; Bruns and Burgess, 2011).

The nature of hashtags enables creating a semantic network in which two hashtags are connected when they are part of the same tweet (Hellsten and Leydesdorff, 2020). The structure of this hashtag network reflects the dependency among messages, and with community detection techniques we may describe the discussion revolving around specific topics. Moreover, by leveraging the presence of hashtags in a tweet, it is possible to formalize and test a variety of societal and psycho-linguistic phenomena within the network of messages connected by hashtags (i.e., common topics). For example, Xiong et al. (2019) analyzed a network of tweets related to the feminist initiatives to examine the co-creation process of meanings in the `#MeToo` movement, identifying the core themes of this collective action, among which there are rhetoric on obstacles to gender equality, encouragement to act, and promotion of specific events. Another contribution about social semantics is presented by Gallagher et al. (2018) who compared `#BlackLivesMatter` and `#AllLivesMatter` networks, offering quantitative proof of content injection, i.e., structural mimicry of the latter network from the first aimed at hijacking the issues brought in the social conversation. In general, the application of semantic network analysis to understand online collective action is a growing and promising field of investigation.

Despite all these contributions, we note that the present paper fills a gap in several ways. First of all, the psychological processes have not yet received enough attention in this kind of investigations since it has primarily studied the responses of single individuals and averaging or summing them. We here move a step forward by testing the social processes behind collective action with an approach that focuses on the ability of meaning creation by the community, i.e., on the core shared outcomes. Second, thanks to the PPRP method we developed, we are able to look into the two-fold interaction of the semantic networks and the online community, assessing how the latter shapes the former, which in turn may enable a prediction over the future evolution of activism. Finally, the topic of climate change, regardless of its considerable on-line presence, was not thoroughly investigated through these approaches in the past.

3. Methods

We now describe the technical procedures used for the analysis. Data collected from social networks might suffer from methodological biases depending on how they are processed. This is especially true for topics, such as climate change, that are linked to political issues or are strongly related to specific time events. Therefore, in Section 3.1 we explain our methodology to extrapolate robust data from our corpus of tweets, also satisfying the important property of "event-neutrality" to allow time comparisons over different time frames.

Then, in Section 3.2, we explain how we analyzed these tweets to capture the underlying psychological processes. In addition to the concepts of affiliation, group-identity, and empowerment, we also analyze temporal perspective. Sections 3.3 and 3.4 detail the construction of the networks and the community detection procedures. It is worth mentioning that we did a separate analysis for the three subcorpuses relating to different years, as well as the whole corpus, so as to enable time comparisons.

Finally, association of tweets to communities is made through the PPRP method, which is detailed in Section 3.5 and later discussed in the results section to show its effectiveness.

3.1. Data collection

To analyze the online semantic network about climate change, we considered messages posted on the social media site Twitter, also referred to as "tweets". They can be downloaded through the free web APIs[2] that enable textual searching by specifying various parameters, such as language, time range of posting, or the presence of specific hashtags.

We limited our scope to tweets in the English language and we identified three analogous time intervals for the years from 2017 to 2019, namely:

- March 1st, 2017 to April 19th, 2017
- March 1st, 2018 to April 19th, 2018
- March 1st, 2019 to April 19th, 2019

We chose the same period within a year so as to limit the influence of seasonal events. Each of these time intervals lasts 50 days, which is the longest span that can be retrieved by sampling a batch of 100 tweets per day, summing up to 5000 tweets per year. Daily batches were uniformly sampled over each of the 24 h (in UTC time), to limit the biases of time versus location. With this intervals choice, we also meant to capture the semantics of climate change discourses around two main events, namely the U.S. withdrawal from Paris Agreement in June 2017, and the first Strike for Climate on the 15th of March 2018 (see Fig. 1).

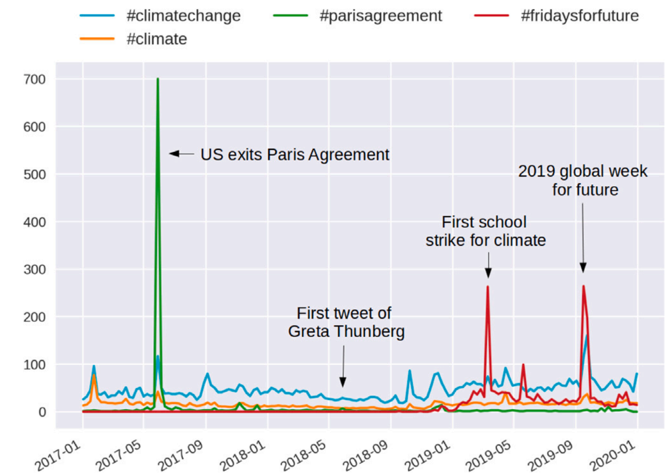

Fig. 1. Historical twitter trends for some hashtags related to climate action where values represent 1/10,000th of 1% of tweets.

Having selected the time intervals of interest, we approached collection of tweets relevant to our study. To that end, we aimed first at identifying a set of hashtags that were generally related to the climate issue, covering a wide range of perspectives on that matter, and representing the topic in a comprehensive and unbiased way. From now on, we are referring to this class of hashtags as "neutral". The strategy of casting a wide net allows us to avoid sampling biased towards specific events or topics. Building the network upon hashtags that were not related to specific collective actions, events, celebrities, or organizations was fundamental to obtain a balanced comparison among the years and an unbiased measure of the importance of semantic communities related to the climate change discourse. For instance, if the network was created from hashtags related to events of a precise year such as `#parisagreement` or specific topics such as `#recycling`, the centrality of the hashtags would be overestimated for the year in question. Similarly, a network generated from hashtags related to a specific movement (e.g., `#fridaysforfuture`), would overestimate the centrality of hashtags more related to it (as shown in Fig. 1).

In order to compile such a list of neutral hashtags, we adopted the following procedure. First, we carried out a search over the three time intervals with the sole hashtag `#climatechange` to identify the most relevant hashtags connected to the climate issue in 2017, 2018, and 2019, separately. A shortlist was built by joining together the 20 most frequent hashtags of each year, and discarding the ones related to a specific event as well as the ones related to a specific aspect of climate change. Event-neutrality and importance of hashtags in the shortlist were verified by a Historical Twitter Trends search,[3] with the aim of selecting only hashtags whose frequency was stable along the chosen intervals (event-neutrality). Moreover, we only selected hashtags that were highly ranked (importance). The resulting selection identified the top ranked neutral hashtags `#climatechange`, `#climate`, `#sdgs`, `#sustainability`, `#environment`, `#globalwarming` as appropriate for the search. Importantly, we included both terms climate change and global warming that tend to be associated with different stances on the climate attitudes spectrum (Shi et al., 2020).

In conclusion, the data collection for the study involved collecting tweets that were based on the selected neutral hashtags for the chosen time intervals. We considered these hashtags a good starting point for building the network because of their stability during 2017, 2018, and 2019, a general absence of peaks during the chosen intervals, and their overall relevance toward the topic of climate change. Their historical Twitter trend is depicted in Fig. 2, from which it is evident that the most relevant hashtag is by far `#climatechange`.

[2] http://developer.twitter.com/en/docs

[3] http://www.trendsmap.com/historical

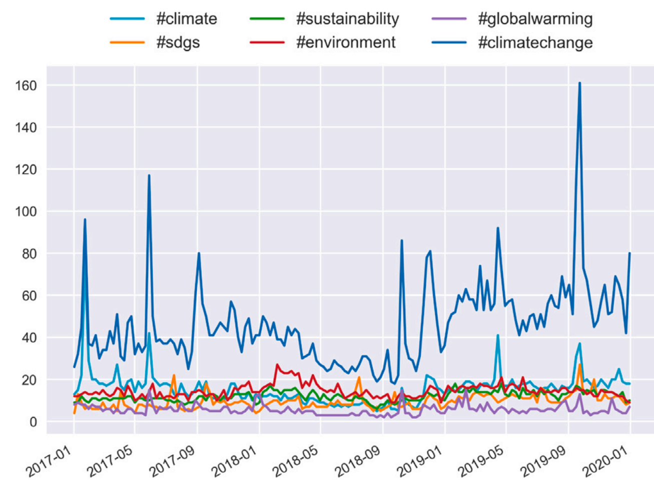

Fig. 2. Historical twitter trends for the selected neutral hashtags.

3.2. Social identity and empowerment metrics

We analyzed all the collected tweets using the Linguistic Inquire and Word Count 2015 (LIWC, Pennebaker et al., 2015), a well-established tool for detecting linguistic proxies of psychological processes in text samples (Hawkins et al., 2017).

LIWC enables a dictionary-based quantitative content analysis in which every message receives a score on several word categories based on the number of words belonging to the specific category adjusted for the total number of words within the message. Coherently with our research questions and hypotheses, the focus of our analysis was on the following main concepts:

(a) Affiliation. The LIWC score for the category *affiliation* (e.g., ally, friend, social) was used for measuring the ingroup community orientation within the text. This proved to be a reliable index of implicit motives for affiliation (Schultheiss, 2013).

(b) Group-identity salience. The frequency of personal pronouns can be used to assess the salience of group membership. In particular, the first person plural pronouns (i.e., *we*) mark the sense of belonging (Zhang, 2010). Michinov et al. (2004) experimentally manipulated group-identity salience and showed a resulting increased use of first-person plural pronouns. Indeed, this type of word has been already analyzed as a marker of social identity of the online action `#occupywallstreet` (Smith et al., 2015).

(c) Empowerment. We computed the empowerment scores aggregating with a mean the LIWC scores for the categories *power*, *achieve*, *reward*, *insight* and *cause*. Each of those component dictionaries comprises a collection of words that were carefully constructed to represent the content of a category in a way that resembles the way people use language. For example, the category of power includes words such as (command, privilege, request), achieve (advance, attain, progress), reward (benefit, goal prize), insight (perspective, realize, question), and cause (affect, because, change). Previous studies (e.g., Decter-Frain and Frimer, 2016; Pietraszkiewicz et al., 2019) reported that these categories are good proxies of agency, which "refers to a person's striving to be independent, to control one's environment, and to assert, protect and expand one's self" (Abele et al., 2008). Agency is related to intelligence, skill, creativity, achievement, power, mastery, and assertiveness, whereas the lack of agency refers to being weak, submissive, incompetent, and likely to fail (Fiske et al., 2007). Thus, agency can be assimilated with empowerment.

(d) Temporal perspective. We measured the orientation of tweets to the past or future using the specific LIWC categories of *past* (e.g., ago, did) and *future focus* (e.g., will, soon).

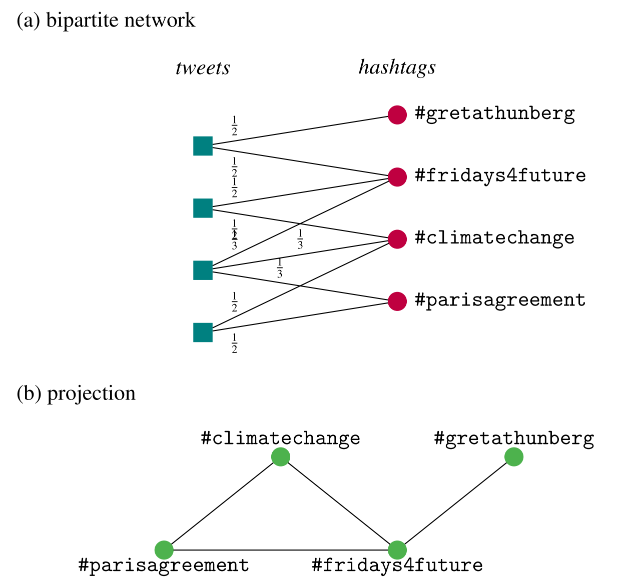

Fig. 3. Illustration of (a) the bipartite network of tweets and hashtags, and of (b) the projection onto a network of hashtags; note that the links in the bipartite network (a) are weighted in such a way that each tweet node sees links with equal weights summing up to one (i.e., normalized); also, the projection network (b) activates a link only in case the hashtag nodes have at least one tweet node in common.

3.3. Climate change network construction

We infer the essence of the climate change social network by exploiting the following rationale: hashtags not only briefly render the semantic contents of the tweets explicitly reported by users, they are also metadata tags that cross-reference such content between tweets, so that their inter-dependencies constitute implicit holistic information on how meaning is socially organized on what constitutes a topic. The semantically structured representation of tweets interdependencies, i.e., the presence of co-occurring hashtags, is captured by a bipartite graph linking each tweet to those hashtags that appear in the tweet. This is the basis for building direct hashtags connections by means of a projection, i.e., by connecting two hashtags if they appear in the same tweet (Hellsten and Leydesdorff, 2020), as we detail in later Section 3.4.

With this idea in mind, we built four different bipartite network, namely:

(a) **yearly networks** - tweets belonging to the **same year** are connected by the hashtags they have in common; three bipartite networks $\mathcal{B}_{2017}$, $\mathcal{B}_{2018}$, and $\mathcal{B}_{2019}$, corresponding to the years 2017, 2018, and 2019, respectively, where the connections are active only among the tweets that belong to one specific year and the hashtags that appear in those tweets; these networks are used for evaluations on a year-by-year basis, e.g., for the study of a temporal evolution;

(b) **whole network** - tweets belonging to **any year** are connected by the hashtags they have in common; a bigger bipartite network $\mathcal{B}_{\text{all}}$ collecting the tweets and hashtags from all the years (2017, 2018, and 2019); this network serves as a benchmark for extracting average values, e.g., for the identification of communities that can then be temporally studied through the networks in a).

Technical details. Hashtags extraction is obtained by using Python's part of speech (POS) non deterministic tagger (Gimpel et al., 2011; Owoputi et al., 2013). To avoid the presence of super-hubs that hide the social structure and complicate the analysis, we discarded the hashtags used in the tweet search. This practically limits the number of effectively used tweets to $N_{2017} = 3459$, $N_{2018} = 4031$, and $N_{2019} = 3931$.

We also applied a weighting to the links of the bipartite graph representation, see Fig. 3(a), so as to better retain the original information as well as to guarantee a correct network projection (Zhou et al., 2007). Specifically, the links departing from each tweet toward the hashtags are equally weighted and the sum of their weights is normalized to one, to identify the tweet as the central entity in our study (Fan et al., 2007). In the whole network $\mathcal{B}_{\text{all}}$, an additional weight inversely proportional to the number of effectively used tweets per year was employed to equalize temporal effects.

3.4. Obtaining the social structure via community detection

In the online discussion over climate change, the identification of the social structure is derived by investigating the interdependecies between tweets revealed by the presence of common hashtags. Topics are identified as communities (i.e., clusters) of hashtags that are strongly correlated through the tweets. As a matter of fact, the particular structure of semantic networks implies the presence of locally dense clusters (Steyvers and Tenenbaum, 2005), i.e., communities or topics in our framework, that might reveal important features of the network itself. The identification of communities, i.e., groups of hashtags more strongly connected among themselves than with the rest of the network, is performed on a network of hashtags where links capture the correlations between hashtags through their presence in the tweets, so as to be able to exploit robust community detection techniques available from the literature (Fortunato, 2010). This rationale is uniquely applied to the whole network $\mathcal{B}_{\text{all}}$, to reveal consistent patterns, and/or indirectly monitor modifications, over the years, which could not be appropriately captured if a yearly perspective was used, i.e., by exploiting the yearly networks $\mathcal{B}_{201\text{x}}$.

Technical details. The projection of the bipartite network $\mathcal{B}_{\text{all}}$ (collecting tweets and hashtags) into an all-hashtags network $\mathcal{P}_{\text{all}}$ is performed by following the classical approach of Newman (2001), see also (Zhou et al., 2007), that consists on a product between the adjacency matrix representing the network with itself. The resulting network connects hashtags that appear in the same tweet and associates to the connection a weight proportional to the hashtag-tweet link as well as to the number of common occurrences. Note that the projection $\mathcal{P}_{\text{all}}$ activates a link only between those hashtags that appear together in a tweet at least once, as depicted in the example of Fig. 3(b).

The method selected for the identification of hashtags communities, from the projected network $\mathcal{P}_{\text{all}}$, is Louvain modularity (Blondel et al., 2008; Lancichinetti and Fortunato, 2009). The implementation used follows its extension described in (Lambiotte et al., 2008). Compared to more sophisticated but less usable solutions, the Louvain approach was chosen for its generality, reliability, scalability, and robustness (Fortunato, 2010). We fine-tuned the Louvain algorithm on our $\mathcal{P}_{\text{all}}$ data, finding the best value for the resolution and threshold parameters. The resolution allows the algorithm to retrieve arbitrary large communities, instead of having them increase proportionally to the size of the analyzed network. The threshold discriminates which communities are to be kept based on the number of nodes they contain. All nodes included in communities with a size below the threshold are considered noisy outputs and therefore merged in a unique “mixed topics” group. Since the algorithm is not deterministic, we used different seeds for each couple of values considered. We evaluated the algorithm performance by looking at the stability of the obtained communities in terms of size and content, i.e., hashtags.

As this is pertinent to what follows (e.g., see later Fig. 4), we incidentally remark that the relevance of communities inside each year is quantified through a PageRank approach (Page et al., 1999; Gleich, 2015), i.e., we identify the importance of a community as the sum of the PageRank centrality scores of the hashtags belonging to that specific community. The projection matrix used for this purpose is the yearly projection $\mathcal{P}_{201x}$ of the corresponding bipartite graph $\mathcal{B}_{201x}$, since in this case the yearly perspective (as opposed to the whole perspective) is the relevant one.

3.5. Assigning tweets to communities: the PPRP approach

Tweets can themselves be assigned to those topics (i.e., communities) that better highlight their content, allowing for an in-depth analysis of topics through linguistic markers of collective mobilization that are available from the tweets contents (affiliation, identity, empowerment, etc.). The more successful this assignment of tweets to communities is, the more distinguishable the socio-psychological implications that can be associated to each community (i.e., topic) will be, and the most relevant the overall analysis. To this aim, tweets are assigned to communities by using an original approach named Personalized PageRank Projection (PPRP).

Operatively, PPRP identifies for each tweet a similarity (or closeness) score with respect to each community of hashtags. The tweet will then be assigned to the community it is most similar to. Similarity is roughly based on the idea that the tweet belongs to the community that includes the highest fraction of hashtags appearing in the tweet, this being equivalent to the (local) WOMP approach of Stram et al. (2017)[4]. More wisely, PPRP generalizes the idea (to make it global), and iteratively applies the same rationale in a PageRank fashion, by exchanging the information across the bipartite network through a number of WOMP iterations, from hashtags to tweets and back to hashtags, repeatedly. In this way, PPRP is applying the rationale at a network level, and is for example able to more finely associate tweets in all those occasions where community fractions are comparable (e.g., when half of the hashtags belong to a community and the other half to another community, in which case WOMP is not able to take a fair decision). In a sense, PPRP provides to WOMP the same generalization that PageRank provides to degree centrality. The performance gap between WOMP and PPRP will be discussed in later Section 4.4. We incidentally remark, as this is meaningful later on in the selection of tweets of Table 1, that PPRP similarity captures the level of adherence to a community, i.e., how closely related the tweet is to a specific topic, but is not able to measure the centrality or importance of the tweet inside the community, as this is related to the PageRank centrality of the hastags.

Technical details. PPRP similarity is built on the bipartite network $\mathcal{B}_{\text{all}}$ by exploiting the PageRank algorithm in a form that is suitable to measuring centrality with respect to a specific topic, as originally proposed by Haveliwala (2002, 2003). Hence, the idea is to activate the PageRank teleportation vector only on a selected community, so as to assess the similarity of nodes (tweets) with respect to that topic (community). Interestingly, this method is effective, and finds justification in a number of literature findings. It follows the rationale of Zhou et al. (2007) and Stram et al. (2017) to implement the idea of Larremore et al. (2014) that one-mode projections should be avoided, and more elaborate information is needed to properly measure dependencies between nodes. Furthermore, the idea of exploiting the similarity measure of a topic specific PageRank for clustering purposes has been proven to be effective in Avrachenkov et al. (2008), Cho and MuLee (2010), Tabrizi et al. (2013).

4. Results

We discuss the results alongside three main avenues. First of all, in Section 4.1 we analyze the outcome of our proposed community detection strategy on the hashtags. This is discussed both qualitatively, to map the online discourse about climate, and also quantitatively,

[4] We warn the reader that in Stram et al. (2017) the WOMP technique aims at identifying an efficient projection $\mathcal{B} \to \mathcal{P}$ from a bipartite into a single-population network. Nevertheless, its constituent modules apply a *projection* from one population of the bipartite network to the other (hashtags and tweets, respectively, in our specific case), and its exactly these constituent modules what we are using in the paper under the name WOMP.

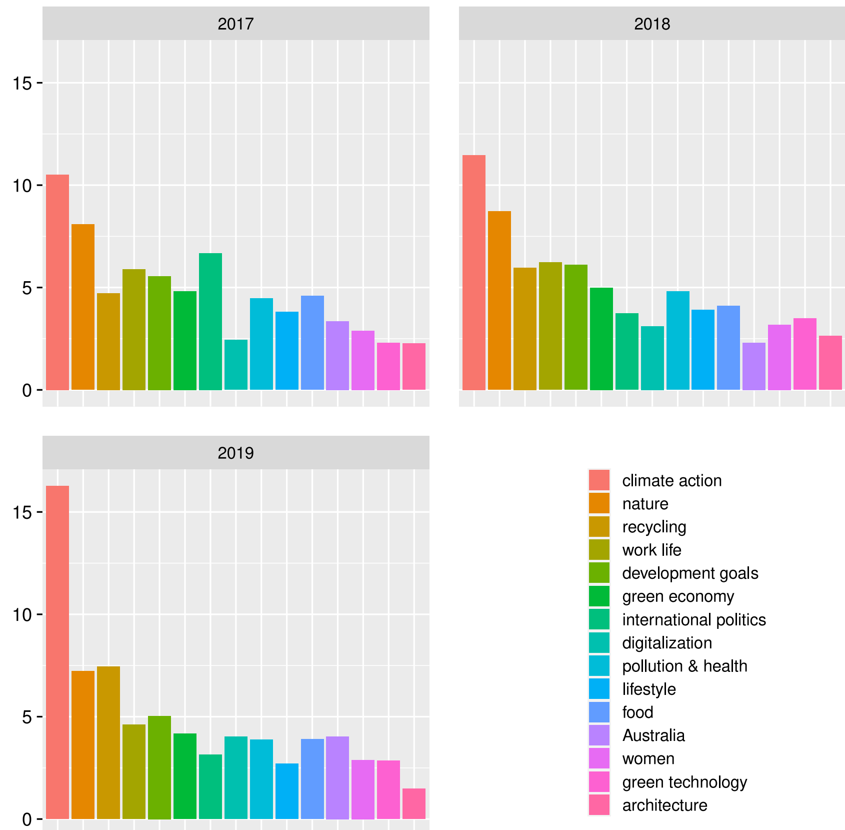

Fig. 4. Relevance of communities over the years 2017–2019 measured by PageRank centrality (0%–100% scale); see a description of communities in Table 1.

to highlight the interrelationships among the communities and also argue about their overall significance. Especially, the characterization of the most relevant communities (thus, discarding tweets about unrelated issues with just similar lexicon) enables the subsequent analysis presented in Section 4.2, revolving around linguistic markers of affiliation/empowerment, to prove that these traits are increasing over time. In Section 4.3 we bring forward the analysis of temporal comparisons, for both different datasets pertaining to different years, as well as discussing their focus towards the past or the future, which can be further extended to a community-wise investigation. Finally, in Section 4.4 we pursue the methodological goal of evaluating the effectiveness of the PPRP method in comparison with state-of-the-art proposals.

4.1. Community detection on hashtags

We performed the community detection procedure described in Section 3.4 and we report in Table 1 the 16 communities with size bigger than 200. Naming and descriptions are provided as the authors' interpretation of the semantic meaning of each community, based on the most descriptive hashtags that belong to it. One representative tweet is selected according to both the PPRP similarity value and the average PageRank centrality of its hashtags, ranked according to the product of centrality and PPRP similarity, to take into account both relevance and adherence to the community; the highest ranked tweet with meaningful/understandable text was selected. In particular, Table 1 provides the list of all communities, also including the 16th that is residual to the entire analysis as it contains a plethora of unrelated topics, not connected to climate change at all or in a very marginal fashion with hashtags of little overall importance. A pictorial representation of the community structures in 2017, 2018, 2019 is available in Fig. 5. Here, the size of a hashtag correlates with its individual relevance. Only the most relevant hashtags are shown. We also computed the relevance of these communities on a per-year basis as reported in Fig. 4.

Especially, for the sake of brevity, we concentrate here on the description of the most relevant (while still pertinent) communities, meant as the ones that obtained the highest score in Fig. 4 when considering the entire span of 3 years. In this spirit, our network analysis demonstrates that, even though the individual hashtag `#climateaction` alone heavily increased in relevance in 2019, its entire community also grew as a result, thereby suggesting a mutual interaction that is not limited to a self-standing trend of a single concept.

The most relevant communities are the ones listed first in Table 1, which reports them in increasing order of ranking. The first one, dubbed "climate action", is indeed the most relevant throughout the years, but soars in importance especially in 2019 (see Fig. 4), thereby confirming the increased visibility of these topics within the online social rhetoric. Remarkably, this trend is present albeit our data collection started from only neutral hashtags that did not involve collective action per se, thereby proving that the underlying thread of call to action can be inferred when discussing climate online.

Another community is "nature", comprising the namesake hashtag and several other items related to the beauty of the planet. Here, the focus is more contemplative rather inviting to action, even though we believe that some hashtags hint at an underlying sense of preservation of the natural environment that can be generically connected to online activism.

Third community, which we labeled as "recycling", relates to pragmatic actions to address anthropogenic climate change, albeit the focus is now distributed on concrete actions individually performed and generally detached from collective protest.

The fourth community is "work life" and it relates to the climate change action in the workplace. As such, involves mostly hashtags unrelated to environmental changes.

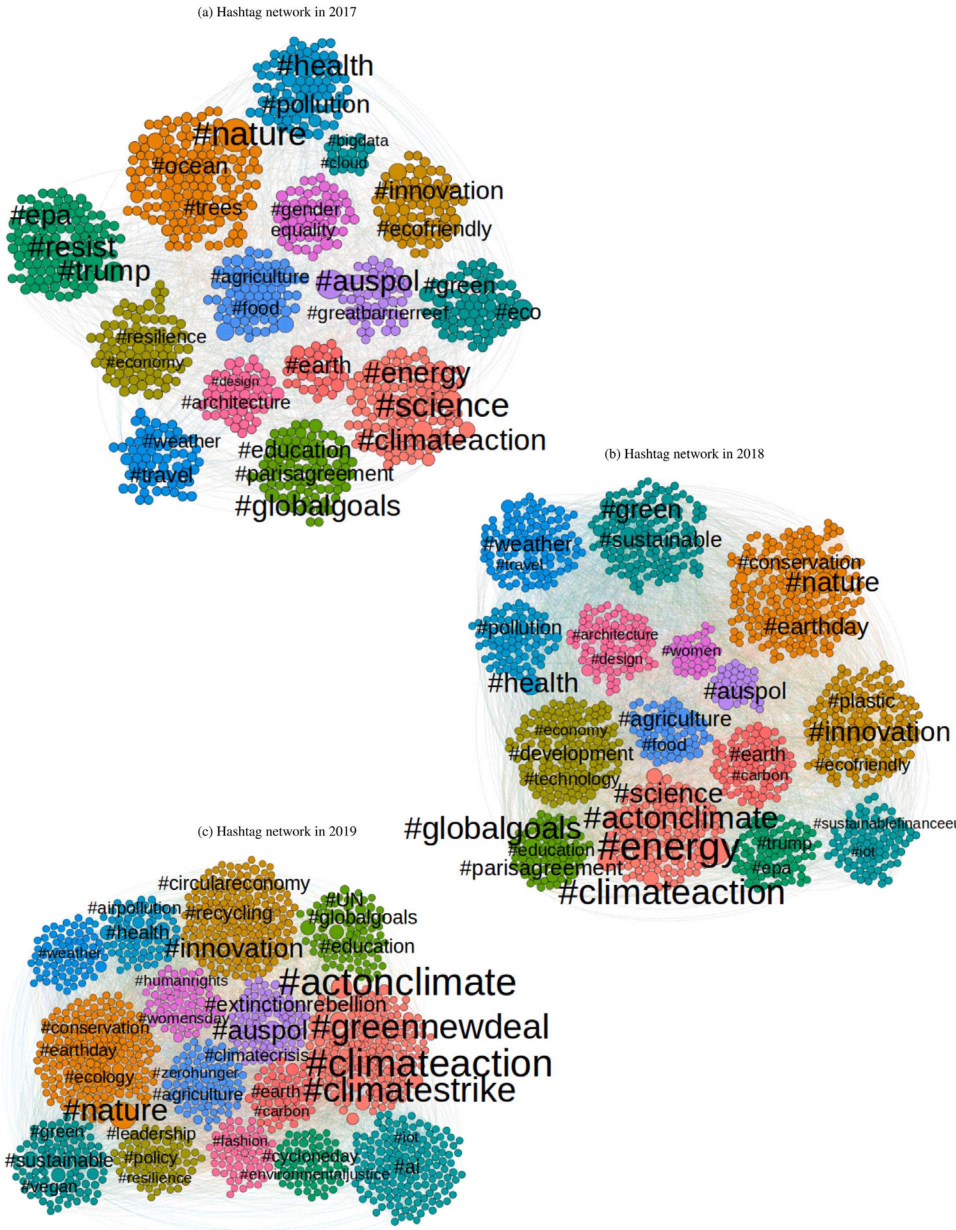

Fig. 5. Gephi graphical representation of communities (same colors as in Fig. 4) where the hashtag dimension corresponds to its PageRank centrality in the corresponding year.

"Development goals" community is named after one of the UN initiatives. It refers to global goals, the agenda of sustainable development, the international efforts on mitigation of greenhouse gas emissions, and, accordingly, to Paris Agreement.

The sixth community addresses and promotes eco-friendly products and green merchandising, and for this reason we labeled it as "green economy". It seems to be connected to branding and marketing and the environmental concerns are not strictly related to global warming.

The seventh community, "international politics", focuses on issues in the international arena, including broader references to environmental justice and political actors. This community was particularly relevant in 2017, possibly because of the US withdrawal from the Paris agreement during that year, which prompted international signals of disapproval (Payne, 2018).

The other communities scored below 5% of relevance in all the years considered (see Fig. 4), and will therefore not be discussed further.

A snapshot on the interdependencies between different communities (over the entire three years span) is available from Fig. 6, displaying communities as nodes in a *directed* graph where edges are weighted according to the PPRP similarity values which, in this context, represent the correlation between communities. In Fig. 6, the edge width is set proportionally to the corresponding PPRP value, and only the strongest links are shown. Arrows denoting the links directions are not displayed for the sake of readability, but the direction can be inferred from the link color, which indicates the node from which the link departs. Node sizes are set according to the PageRank centrality (i.e., the relevance) of each community. Distance on the graph roughly represents similarity of the communities. Interestingly, the "Climate action" call-for-action appears to be mainly correlated to the "Australia" call-for-action (as one would expect), as well as to "International politics" and "Green tech" communities, which are displayed closer then the rest. It is also of relevance to note that, in the chosen representation no output links generate from "Climate action" (because they are weak), this being a further confirmation of cohesion inside the community, i.e., of the reliability of the chosen community detection approach.

Table 1
Description of the communities.

#	Community name	Brief description	Descriptive hashtags	Descriptive tweet
1	Climate action	Calls to action related to climate change	#climateaction, #actonclimate, #energy, #science, #cdnpoli, #renewableenergy, #renewables, #greennewdeal, #climatestrike	To all school kids fighting for our planet, standing for climate change and are holding your govt to account - we tautoko you 100%! @MaramaDavidson @jamespeshaw #ClimateChange #ClimateAction #SS4C #climatestrikenz #nzclimatestrike https://t.co/EC636sMRM5
2	Nature	Photos ad videos about naturalistic environments and animals	#nature, #earthday, #conservation, #biodiversity, #oceans, #ecology, #trees, #forests, #wildlife	Mbachile #Visiterlafrique #olympus #myolympus #nature#getolympus #OlympusInspired #omd #zuiko42mm #climatechange #ecology #streetphotography #island #zuiko #landscapephotography #comores #Comoros https://t.co/nmw8dcH2eg
3	Recycling	Business solutions for the circular economy, and recycling techniques	#innovation, #circulareconomy, #plastic, #sustainabledevelopment, #recycling, #ecofriendly, #recycle	@mttw_page says: #Learn from this to achieve a better world by 2030 #Reduce #Reuse #Recycle #Upcycle #UHC #Goalkeeper #ClimateChange #ClimateAction #Climattitude #ClimateAmbition #YesWeCan #GetInvolved #SDGs #GlobalGoals #TeachSDGs #StepUpTheFight #DoOneThingToday #MushinToTheWorld https://t.co/Rv2Kgftq5w
4	Work life	Professional life and working environment aspects	#leadership, #employment, #creativity, #partnerships, #decentwork, #career	CREATIVE WORK: Respect the Dignity of All Types of Work https://t.co/8trFtZRHNf #creativity #millennials #boomers #YoungAdults #selfies #students #employment #workers #money #unemployment #satisfaction #technology #Innovation #sustainability #compensation #income #poverty https://t.co/hM5p7YzDPn
5	Developments goals	2030 Global Goals for Sustainable Development	#globalgoals, #education, #parisagreement, #un, #2030agenda, #community, #migration, #teachsdgs	Which of the #GlobalGoals are you supporting as an African Youth? #SDGs #TheAfrikanLegacy #TheAfricaWeWant #YouthForAfrica #YouthForChange https://t.co/B4ZMNzvPcS
6	Green economy	Promoting green and eco-friendly products	#green, #eco, #sugarcane, #ecofashion, #sustainablefashion, #vegetarian	With the #ClimateActionIncentive, we're ensuring Canadians can make ends meet while we fight the impacts of #ClimateChange. We are giving 90% of revenues directly back to Canadians & families, and investing the rest in our communities. #EnvironmentEconomy #cdnpoli https://t.co/BZIMramzvR
7	International politics	Political topics	#trump, #epa, #resist, #coal, #p2, #environmentaljustice, #tcot, #usa, #2a, #oil, #theresistance, #eu	Premier Dwight Ball set to release #budget expected to include plan for #carbon taxes, and #education for #Newfoundland and #Labrador https://t.co/BSfLFZCgqg #ClimateChange #cdnpoli #NLpoli #Newfoundland #Labrador #NLBudget
8	Digitalization	Methods and procedures for the digital transformation and innovations	#ai, #iot, #dataviz, #data, #bigdata, #digital, #smartcity, #digitaltransformation, #smarthome	Responsability of #IoT in finding new solutions for #energy sustainability and #environment https://t.co/4ZRSjppbMR
9	Pollution and health	Topics of air pollution and public health	#health, #pollution, #airpollution, #cities, #healthforall, #publichealth, #wellbeing, #airquality, #worldhealthday	#AirQuality #Alerts - For valuable #info watch the video at https://t.co/PIhaj9XI84 #Air #AirQualityIndex #Weather #WeatherObservation #Smog #Pollution #AirPollution #KnowYourAir #Smoke #Environment #ZeroEmissions #CleanAir #ClimateChange #JITDT #Planet #Health #Safety https://t.co/sHcUTtFjdW
10	Lifestyle	Big variety of free-time-related topics	#weather, #travel, #coffee, #worldmetday, #europe, #spring, #thursdaythoughts, #london, #sxsw, #snow, #summer, #noaa, #greenland	Master Luísa Afonso is finally using the pedal power washing machine! #sustainability #diy #mastercourse #energy #freeandreal https://t.co/Z5tWYbCYjo
11	Food	Food issues and food technologies	#agriculture, #food, #zerohunger, #foodsecurity, #regenerativeagriculture, #insect, #urbanfarming, #learn, #foodtech	#NationalAgDay is a time to celebrate the #climatechange, #food and #energy solutions that #farmers and #agriculture can provide. #AgDay19 https://t.co/Q2RjQRrWPK
12	Australia	Climate collective actions in Australia	#auspol, #extinctionrebellion, #climatecrisis, #greatbarrierreef, #stopadani, #australia, #extinction, #factsmatter, #ausvotes, #actnowforfuture, #brisbane	I have been a farmer for 57 years I know first hand the effects of #ClimateChange Help me help you with #ClimateAction. Vote 1 ICAN in the #Senate #Auspol https://t.co/5I2HnD5m71 via @thelandnews

(*continued on next page*)

Table 1 (*continued*).

#	Community name	Brief description	Descriptive hashtags	Descriptive tweet
13	Women	Gender-related topics	#genderequality, #women, #womensday, #gender, #internationalwomensday, #iwd2018, #sdg5, #unea4, #localgov, #solvedifferent, #women4climate	Here we are at SNDT women's college, Matunga, making these adolescent girls exert themselves into unfolding the naked truth of Disposable Sanitary Napkins. #menstruation #periods #menstrualcups #clothpads #women #saynotoDSN #taboos #sustainability #sustainablemenstruation https://t.co/yvvYPppJeS
14	Green technology	Technological and sustainable innovations	#earth, #carbon, #jobs, #blockchain, #emissions, #cleantech, #engineering, #startups, #ghg, #electric, #natural, #paris, #life, #mining	Don't miss the opportunity to #MakeOurPlanetGreatAgain ! Come to France for your postdoctoral research on #earth systems, #climatechange and #energy transition. Apply here https://t.co/4vJ5hUpk1G deadline : 6 April Postdoctoral research contracts of 12 to 24 months https://t.co/vlJXTADd1S
15	Architecture	Architecture topics	#architecture, #fashion, #design, #construction, #greenbuilding, #building, #webinar, #steamdrills, #5star, #innovative, #free, #interiordesign	BEST Conference Building Enclosure #Science & #Technology, April 15-18 in #Philadelphia https://t.co/ZSRNMVbYLm @NIBS_News #Pennsylvannia #Resilience #Sustainability #Green #GreenBuilding #Environment #Economy @NIBS_News #Pennsylvania #building #architecture #engineering https://t.co/aGUvv3ERE0
16	Other	Mixed topics	#agenda2030, #brexit, #news, #healthcare, #fracking, #ocean, #photography, #art, #wednesdaywisdom, #infrastructure, #climatejustice, #tourism	

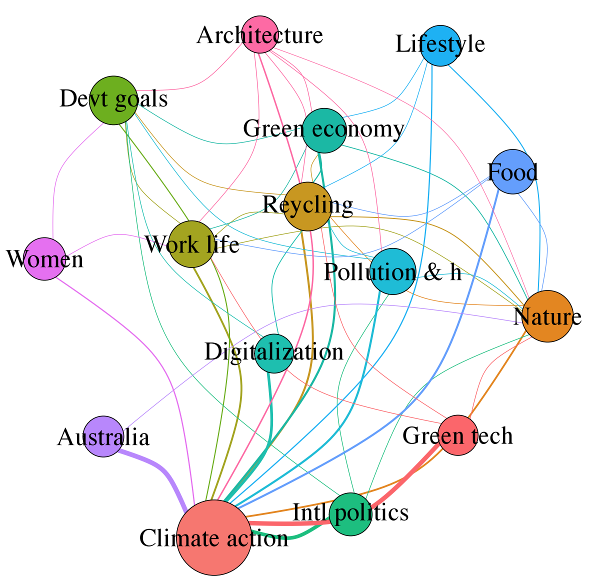

Fig. 6. Communities interdependencies: representation of communities built on a graph whose edges are weighted according to PPRP correlation values; see a description of communities in Table 1.

4.2. Socio-psychological linguistic markers

In Fig. 7 we report an overview of the linguistic markers of socio-psychological processes in our Twitter corpus over the years 2017–2019. This identifies the overall yearly trends, while in Fig. 8 the same markers are shown community-wise. Panels display the average value per tweet of *affiliation*, *we*-terms, *empowerment*, and *future focus* proxies (see their meaning in Section 3.2), respectively. The relevant highlight is that the considered proxies are linearly increasing over the years for the overall sample (Fig. 7), and that the only individual community consistently capturing this trend is that of climate action (Fig. 8). Although some other communities locally display high values, it seems to appear due to a specific focus of that community. For example high values of empowerment language in work life community may be due to the link of this community to leadership or productivity,

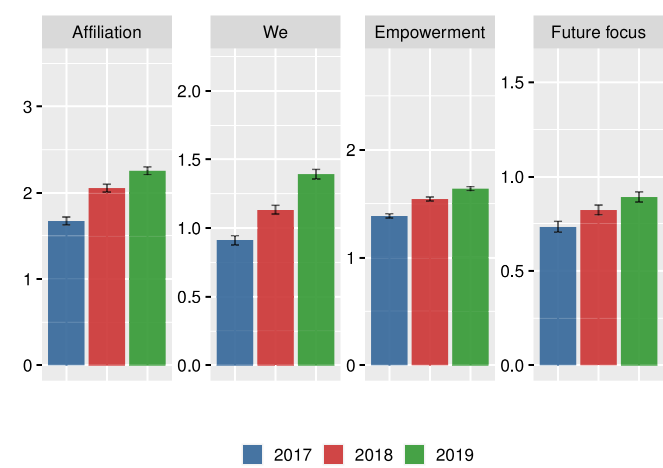

Fig. 7. Average value per tweet of (a) affiliation, (b) we, (c) empowerment, and (d) future focus measures for each of the years 2017–2019, where measures appear on a 0 – 100% scale, with histograms representing average values and error bars being constructed using 1 standard deviation from the mean.

to which empowerment is also a constituent topic. Importantly, for these communities we do not observe any systematic effect.

Given the primary relevance of climate action community, we focus on this community in the subsequent analyses and emphasize it graphically in Fig. 8. Specifically we test the effect of time comparing the means of the three year samples with an analysis of variance (ANOVA, see Tabachnick and Fidell, 2019). Table 2 indicates tests of linear trends along effect sizes indicated by Cohen's-d values (Tabachnick and Fidell, 2019), with positive (vs. negative) values signaling an increasing (vs. decreasing) pattern, and highlighting in orange or red the presence of a statistically reliable difference between the considered years. As visible from Table 2, climate action is the only community that exhibits a systematic rise in all the transitions (with the only exception of the future focus proxy in 2018 → 2019), which confirms our hypothesis. The above suggests that the online discourse about climate change is characterized by a steady increase over the three years of our corpus.

Additionally, we tested the positive linear trend of climate action using the software JMP (Sall et al., 2017). We applied two full factorial

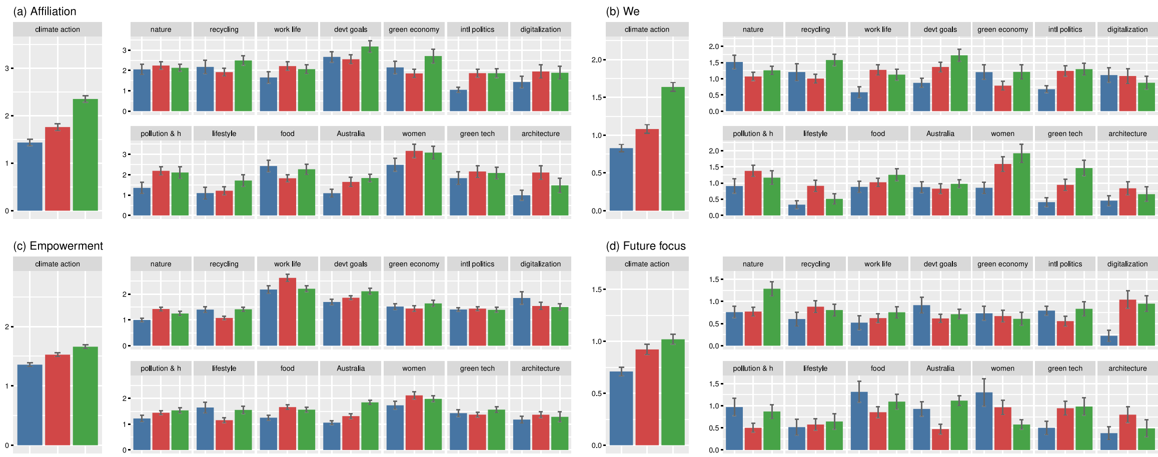

Fig. 8. Community average values of (a) affiliation, (b) we, (c) empowerment, and (d) future focus measures for each of the years 2017–2019; uses the notation of Fig. 7.

Table 2
Pairwise comparisons (Cohen's-d values, with positive and negative signs respectively indicating an increase and a decrease of the average value) per each community, and among the years, for the measures of Fig. 8: asterisks identify statistically significant differences with FDR adjusted p-value < 0.05 (red asterisk *) or with p-value < 0.01 (a double orange asterisk **); the community named "all" collects all the communities, i.e., all the nodes in the network.

	2017→2018				2018→2019				2017→2019			
	Affiliation	We	Empower	Future	Affiliation	We	Empower	Future	Affiliation	We	Empower	Future
all	0.1057**	0.1039**	0.0752*	0.0043	0.0448	0.0593	0.0609	0.0386	0.1505**	0.1631**	0.1361**	0.0429
climate action	0.1059*	0.1111*	0.1306**	0.1132*	0.1909**	0.2434**	0.1026*	0.0519	0.2967**	0.3545**	0.2332**	0.1651**
nature	0.0582	−0.1973	0.3208**	0.0325	−0.0332	0.0849	−0.1293	0.2483	0.0251	−0.1125	0.1916	0.2807*
recycling	−0.0826	−0.0879	−0.248	0.1481	0.1849	0.25*	0.2575*	−0.0403	0.1024	0.162	0.0094	0.1078
work life	0.1798	0.3021*	0.3462**	0.0543	−0.0482	−0.0639	−0.3229**	0.0702	0.1316	0.2382	0.0233	0.1245
development goals	−0.0375	0.2145	0.1259	−0.1594	0.202	0.1575	0.1921	0.051	0.1645	0.3719**	0.318**	−0.1085
green economy	−0.0956	−0.1823	−0.0597	−0.0333	0.2753	0.1852	0.1531	−0.0332	0.1798	0.0029	0.0934	−0.0665
international politics	0.2638*	0.2476*	0.0252	−0.1262	−0.001	0.0233	−0.0354	0.1477	0.2628*	0.2709*	−0.0102	0.0214
digitalization	0.1672	−0.0125	−0.237	0.4337	−0.0203	−0.0908	−0.0314	−0.0482	0.1469	−0.1034	−0.2684	0.3854
pollution and health	0.2702	0.2047	0.1695	−0.2486	−0.0262	−0.0912	0.0713	0.1959	0.244	0.1134	0.2408	−0.0528
lifestyle	0.0372	0.2538	−0.3764*	0.0311	0.1615	−0.1756	0.3052	0.0366	0.1987	0.0782	−0.0712	0.0677
food	−0.1938	0.0623	0.314**	−0.2444	0.143	0.1005	−0.0759	0.1271	−0.0508	0.1628	0.2382	−0.1173
Australia	0.1786	−0.0211	0.1929	−0.2422	0.0619	0.0654	0.4088**	0.3415*	0.2405	0.0443	0.6016**	0.0993
women	0.221	0.321	0.2952	−0.1787	−0.027	0.1449	−0.1026	−0.2071	0.1939	0.4658**	0.1926	−0.3859
green technology	0.1126	0.235	−0.0443	0.1055	−0.024	0.2269	0.1489	0.0207	0.0887	0.4619*	0.1047	0.1262
architecture	0.3601	0.1666	0.1438	0.2217	−0.2032	−0.0791	−0.0625	−0.1642	0.1569	0.0875	0.0813	0.0575
misc	0.1458*	0.0444	0.1043	0.1617*	−0.1201*	−0.0327	0.0953	−0.18**	0.0258	0.0116	0.1996**	−0.0183

Table 3
Full factorial generalized linear model outcomes: $F(d, e)$ is the Fisher value with d the degrees of freedom and e the levels of error due to the sample size; p denotes the associated statistical probability p-value; η_p^2 is the partial eta squared effect size.

(a) Fisher values for affiliation					
Variables	F	d	e	p	η_p^2
Year	10.53	2	8237	$< .0001$	.003
Community	14.25	5	8237	$< .001$	.008
Year & community	2.08	10	8237	.023	.003
(b) Fisher values for empowerment					
Variables	F	d	e	p	η_p^2
Year	6.44	2	8237	.001	.002
Community	21.93	5	8237	$< .0001$	.013
Year & community	2.95	10	8237	.001	.004

generalized linear models, using *affiliation* or *empowerment* as outcome variables. The used predictors are the variable year and the six most relevant communities, i.e., climate action, nature, recycling, development goals, green economy, and international politics. They were selected by relevance, in which case only communities scoring higher than 5% in PageRank centrality were kept (see Fig. 4), and pertinence, which excluded the spurious work life community from the shortlist.

The results are reported in Table 3, which shows that the use of both affiliation and empowerment related words changed over time (see the main effect of the *year* variable), and were used to different extents in the six considered communities (main effect of the *community* variable). The interaction between *year* and *community* further confirms that the increase across the years is uneven across the communities, which is coherent with and justifies the findings of Table 2.

4.3. Temporal perspective on climate rhetoric

Additionally, we investigated the involvement of a temporal perspective in the climate change rhetoric, and the role of identity language in shaping that discourse. We analyzed the use of words referring to the in-group (such as "we", "our", "ours") versus words referring to others (such as "they", "their", "theirs") in association to words marking the future or the past temporal frame in Fig. 9, which shows a linear regression with confidence intervals of we/they versus past/future focus markers. We observe that there is a general association between pronouns (we, they) and time (past, future) markers: the more tweets include words about time, the more also include plural pronouns. This is revealed by the presence of positive slopes in all the diagrams of Fig. 9, and may suggest an association between temporal focus and intergroup discourses where "we" and "they" are put in contrast with each other.

A few interesting aspects can be appreciated in Fig. 9 by differentiating between past and future focuses, namely:

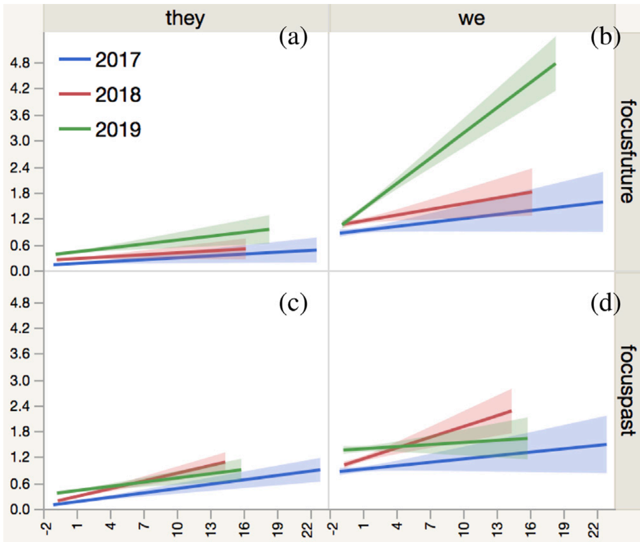

Fig. 9. Linear regressions with confidence intervals over the three considered years for we/they versus past/future focus markers.

- **past focus**, Fig. 9(c)–(d): the association seems to be more pronounced in 2018 (see the increasing 2018 red regression);
- **future focus**, Fig. 9(a)–(b): the association seems to increase over time, and is more pronounced in 2019 (see the 2019 green regression); incidentally, it is very clear that the use of future words in association with the words referring to the in-group (we) was particularly marked in 2019 (see Fig. 9(b)), which is also the year most characterized by words related to the in-group (see Fig. 7).

In Fig. 10, the general regression line of future focus (Fig. 9(b), green regression) is split into individual communities. The linguistic framing of the in-group into the future is common to all the six most relevant communities considered in Section 4.2, with the only exception of green economy community, which, however, is focused mainly on advertising products that flag the green label, and therefore is less related to the social identity and future concerns. Importantly, envisaging the in-group into the future is particularly striking in the community of climate action in which future words are the most used and clearly associated with the pronoun "we". The biggest increase in the identity related language (we) in association to time orientation (focus-future) occurred in 2019. The comparison between 2017 and 2018 (see Fig. 9) confirms that this specific use of language is not a natural evolution over time but that it is specific to the time frame in which Greta Thunberg established her influential role.

These results were further statistically confirmed through two full factorial mixed linear models to take into account that the two types of pronouns can appear simultaneously in each tweet. The two models were applied on past and future focuses, respectively, with the pronouns used as dependent variables, and the type of pronoun (we, they), year, community, and past or future temporal frames included as factors, with the type of pronoun being nested within the tweet, which was added as random factor. The models revealed that:[5]

- **past focus**: the use of both types of pronouns (we, they) was positively associated with ratio of words signaling the past, $F(1, 14442) = 56.83$, slope $b = .04$, $p < .01$; there was an increase over the years of the use of pronouns, $F(1, 14442) = 98.97$, $p < .01$, which was further characterized by the type of pronoun, $F(1, 14442) = 8.75$, $p < .01$, and by past words, $F(1, 14442) = 6.13$, $p < .01$, so that a focus on the past is positively associated with both type of pronouns; the use of words referring the past did not interacted with the type of pronoun;
- **future focus**: all the effects of type of pronoun and year already described for the past focus apply to the future focus; in addition, the role of future words had a main effect, $F(1, 14442) = 91.97$, slope $b = .05$, $p < .01$, and interacted with year, $F(1, 14442) = 28.82$, $p < .01$; a three-way interaction is appreciated between year, future and pronouns, $F(1, 14442) = 16.74$, $p < .01$.

[5] The notation used in the following was defined in Table 3; in addition, b identifies the slope of the linear model fitting.

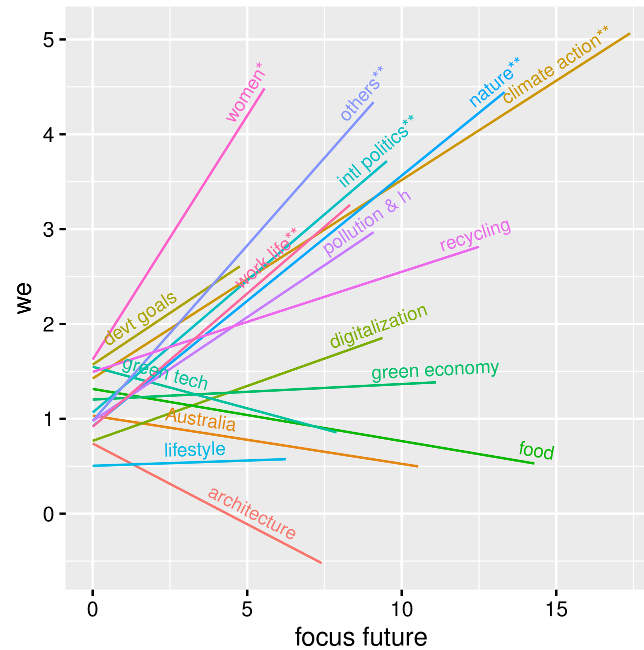

Fig. 10. Linear regression of first person plural pronouns (we) as a function of future-framed wording (focus future) by community: an asterisk denotes a $p < 0.05$ significance of the slope coefficient, two asterisks a $p < 0.01$ significance.

To sum up, our results show, consistently across all the communities, that the stronger the focus on the future, the more frequent the use of pronouns related the ingroup.

4.4. A remark on the effectiveness of the PPRP approach

We conclude this section by establishing the reliability of the proposed PPRP approach. By qualitative investigating the semantic content of tweets, we found that PPRP provides a more meaningful association with communities, especially in being able to avoid the miscellaneous community, as, for example, the tweet "Shrinking civic space, future of #governance, fragility & #sustainingpeace, climate change & #energytransition" is associated with the "Climate action" community by PPRP and to the "Misc" community by WOMP. This effect can be well motivated by the fact that PPRP is able to gain insight and robustness over multiple interactions, thus better conveying the information flow through the network.

We further challenge our approach with a more objective and quantitative test in Table 4 where PPRP is contrasted against the reference WOMP approach using a Model comparison analysis, and in particular exploiting the Bayes information criterion (BIC; Schwarz et al. 1978), a widely used criterion for determining the quality of model-based clusters (Zhao et al. 2008). Model comparison requires first to formalize a specific statistical model (e.g., a linear model fitting the data), and then apply an information criterion to compare the models in term

Table 4
$BF_{WOMP} - BF_{PPRP}$ measures; positive values indicate that PPRP is the method to prefer.

	Affiliation	We	Empower	Future
2017	11.1	110.9	206.4	−145.9
2018	45.8	2.1	280.7	26.4
2019	66.1	164.6	46.7	−8.4

of statistical evidence (i.e., supported by the data). More generally, model comparison is used in order to select the most probable statistical model among a set of candidates, given the observed data (McElreath, 2020). Specifically, BIC provides the mean error of a model's ability to predict new data and, for this reason, lower BIC implies a better fit. In the present context, BIC is separately applied to the socio-psychological target measures (affiliation, we, empower, future) on the three different years, and for each of these cases it calculates how well the community assignment predicts the target measures by assuming a linear model with categorical predictors given by the 16 communities. Table 4 reports the difference in the BIC scores between WOMP and PPRP assignments. As a lower BIC score implies that the model is to be preferred, then positive values in Table 4 indicate that PPRP is the method to prefer. Table 4 shows that the PPRP produces lower BIC scores in almost all the comparisons, meaning that the communities extracted with PPRP are generally more representative of the data.

5. Discussion and conclusions

The analysis of the increased relevance of the climate action semantic community and its traits of affiliation and empowerment brings some general conclusions. First of all, underneath the global increase in popularity of climate action discourses simply due to trending topics, our data show a substantial increase in interest related to climate action. We therefore claim a quantifiable impact from online activism in increasing the popularity of the topic, which is highly intertwined with practical and offline protests.

Also, our investigations performed at the community level seem to imply that the socio-psychological involvement of affiliation and empowerment motives, and the ingroup mobilization toward the future, are mainly related to the specific semantic group of climate action keywords, where we observe reliable increases for each pairwise comparison, and only a marginal influence by other unrelated aspects. This further corroborates the interpretation that the discussion around climate activism is evolving over time and encompasses shared social meanings that match the drives that have been identified as individual motives for collective protests. This is in line with our initial view of online discourse as a global phenomenon, whose evolution is promoted by countless micro-interactions at the individual level, in a likely recursive process where individual and collective levels interact to build a shared meaning structure. Whether this shared online collective meaning turns into offline social mobilization is not directly tested in our study. However, indirect evidence can motivate speculative reasoning to guide future empirical endeavours. In fact, the increased awareness about the anthropogenic climate change associated to Greta's activity (Sabherwal et al., 2021) and the global initiatives to protest against are reflected on the social discourse. Saberwhal and colleagues (2021) reported that participants with more familiarity with Greta Thunberg have stronger collective active intentions. This effect was mediated by the degree of collective efficacy beliefs (the belief that your group can achieve social change), a concept resembling a combination of the affiliation and empowerment categories used in our study.

Our research provides an illustration of several theoretical predictions regarding collective action, and represents a step forward in this literature for at least three reasons.

First, the sense of efficacy and identity is here for the first time assessed at the level of common meaning, rather than at the individual level. The network approach offers the possibility to investigate collective phenomena with a holistic perspective, embracing social processes as collective constructions more complex and richer than the sum of each individual contribution.

Second, we offer important insights about the change over time of the discourse on climate change, addressing the development of social movements. Distinguishing which topics are specifically central in the collective discourse is of key importance to understand the evolution of action dynamic. In fact, through the comparison of three years, we can appreciate that the increase in centrality of the climate action hashtag is sustained by an entire rhetoric community whose dominance has a growing trend over the years and whose linguistic characterization clearly signals the social processes at stage.

This points to the third and possibly most important novel aspect of the present work. Thanks to the linguistic investigation of arguments in the public discourse and their evolving over time, we can predict engagement in climate action discourse, and this illustration is far from being trivial. We are not aware of any work both in the area of climate as well as more broadly understood collective action that examined these linguistic predictors of engagement in collective discourses, fine tuning the linguistic specificity of the collective action semantic community. We therefore contribute to the identification and characterization of the cultural common ground behind mobilizing structures at their emergence stage (McAdam et al., 1999). In this way, our work offers an important diagnostic methodology for detecting exchanges about collective action in online communities and suggest communication strategies to promote actual engagement in popular mobilization to advance collective interests. So far, such a linguistic framing has received only limited attention (Morton et al., 2011). Our results may prompt future investigations empirically testing the effectiveness of the identified linguistic proxies in messages calling for climate action, while taking specific semantic clusters into account. In other words, we may inform how to communicate about climate change in ways that inspire people to take action, and how this can be tailored to specific online groups.

Even though our analysis does not detail the individual traits of online contributors on these topics, future developments may explore possible motivations and perspectives to look at the results so as to allow for socio-psychological interpretations. In particular, the increase of affiliation or empowerment terms may be gauged from different perspectives of age, gender, and/or political affiliation.

For example, climate change is generally considered as a central topic within a progressive political ideology (Cruz, 2017), which is in turn characterized by a focus on social issues and a more frequent use of affiliation words (Fetterman et al., 2015). Thus, one may argue that the increased use of affiliation terms correlates with the association between climate action and a left-wing political agenda. On the other hand, empowerment terminology, which is also shown to have an increasing trend, is often associated with conservative positions (Salmela and von Scheve, 2017). Thus, it may be worth investigating these connections further to see how much of climate change rhetoric regarding the use of affiliation and empowerment words is influenced by the positioning on the political compass, and possibly arguing against traditional pre-conceived political classifications (Ife, 2018).

More recent studies also suggest that the environmental support in relation to the political spectrum might also correlate to how individuals perceive it as juxtaposed with economic growth and individual development (Harring and Sohlberg, 2017). Future studies might apply the methods proposed here to inspect the interplay between semantic communities addressing economic and climate issues, and its evolution over time.

Another future scenario pertains to the understanding of the specific meaning creation among population strata defined by specific socio-demographic features. Collective actions are fostered when a group perceives its lower status as illegitimate and unstable (Kawakami and

Dion, 1995). For our scenario, this can be the case of younger individuals. Indeed, it was shown (de Moor et al., 2020) that they are more concerned with climate action, while being also more prone to engage in online communications.

Similar considerations can also be advanced in relation to gender. The specific increase of affiliation in the collective action discourse that developed in 2019 may be related to a stronger online rhetoric from female social media users, not necessarily limited to climate action. Women are more likely to use affiliation linguistic expressions in social media compared to men (Park et al., 2016). Yet, we speculate that the increase of empowerment words, which is not typical of female users, supports the idea that collective action motivations are the likely underpinnings of these linguistic cues, rather than gender roles.

Finally, another important contribution of the present study is to provide evidence for the intersectional role of social identity and orientation toward the future as central to the growth of the online debate on climate change. We showed an association between future orientation and ingroup mobilization. This association is increasing over time, suggesting a general trend in climate discourse to envisage the ingroup in the future. Further studies can experimentally investigate the consequences of this future oriented frame of the ingroup and test whether this feature is specific to pro-environment call to action, or whether it is critical also for other calls.

Indeed, the Social Identity Model of Pro-Environmental Action (SIMPEA) proposed by Fritsche et al. (2018) and the more general Social Identity Model of Collective Action (SIMCA) proposed by Van Zomeren et al. (2008) have a common denominator of collective identity and efficacy beliefs. In this way, it is likely that our approach applied to different, non-environmental, movements would yield similar results, that is that collective identity and efficacy beliefs would characterize tweets pertaining to other collective action movements. However, it is important to note that we have here analyzed an ascending movement, future research may explore descending trends or linguistic responses to specific events or collective efforts not specifically defined by action, but rather by inaction (see for example a strike or a boycott campaign). The identification of linguistic markers that characterize the evolution of a collective discourse over time has important implications and applications at both diagnostic and prognostic levels. A temporal analysis may track groups more likely engaged in activism of any kind, not necessarily limited to environmentalism or social justice causes, and check the progression of their rhetoric and possibly actions in the future. Accordingly, the predictive potential of our analysis, comparing a span of three years, should be taken cautiously; further studies can focus on longer time frames to build proper historical series so as to follow up on our findings in the years to come.

Funding institutions

University of Padova & University of Social Sciences and Humanities, Warsaw; Magdalena Formanowiczs contribution was funded by the OPUS 19 grant of the Polish National Science Center 2020/37/B/HS6/02587.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abele, A.E., Uchronski, M., Suitner, C., Wojciszke, B., 2008. Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. Eur. J. Soc. Psychol. 38 (7), 1202–1217.

Abidin, C., Brockington, D., Goodman, M.K., Mostafanezhad, M., Ann Richey, L., 2020. The tropes of celebrity environmentalism. Ann. Rev. Environ. Resour. 45.

Abramson, P.R., Aldrich, J.H., 1982. The decline of electoral participation in America. Am. Political Sci. Rev. 76 (3), 502–521.

Avrachenkov, K., Dobrynin, V., Nemirovsky, D., Pham, S.K., Smirnova, E., 2008. Pagerank based clustering of hypertext document collections. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, pp. 873–874.

Bamberg, S., Rees, J., Seebauer, S., 2015. Collective climate action: Determinants of participation intention in community-based pro-environmental initiatives. J. Environ. Psychol. 43, 155–165.

Bandura, A., 2000. Exercise of human agency through collective efficacy. Curr. Dir. Psychol Sci. 9 (3), 75–78.

Barabási, A.-L., et al., 2016. Network Science. Cambridge University Press.

Bastos, M.T., Zago, G., 2013. Tweeting news articles: Readership and news sections in Europe and the Americas. Sage Open 3 (3), 2158244013502496.

Berman, S.L., Wittig, M.A., 2004. An intergroup theories approach to direct political action among African Americans. Group Process. Intergroup Relat. 7 (1), 19–34.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 2008 (10), P10008.

Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D., Marlow, C., Settle, J.E., Fowler, J.H., 2012. A 61-million-person experiment in social influence and political mobilization. Nature 489 (7415), 295–298.

Borge-Holthoefer, J., Arenas, A., 2010. Semantic networks: Structure and dynamics. Entropy 12 (5), 1264–1302.

Boyd, D., Crawford, K., 2012. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Inf. Commun. Soc. 15 (5), 662–679.

Boyd, J.N., Zimbardo, P.G., 2005. Time Perspective, Health, and Risk Taking. Lawrence Erlbaum Associates Publishers.

Boykoff, M.T., 2011. Who Speaks for the Climate?: Making Sense of Media Reporting on Climate Change. Cambridge University Press.

Breslin, J., Decker, S., 2007. The future of social networks on the internet: The need for semantics. IEEE Internet Comput. 11 (6), 86–90.

Brügger, A., Kaiser, F.G., Roczen, N., 2011. One for all? Eur. Psychol..

Bruns, A., Burgess, J.E., 2011. The use of Twitter hashtags in the formation of ad hoc publics. In: Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011.

Brunsting, S., Postmes, T., 2002. Social movement participation in the digital age: Predicting offline and online collective action. Small Group Res. 33 (5), 525–554.

Cho, M., MuLee, K., 2010. Authority-shift clustering: Hierarchical clustering by authority seeking on graphs. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, pp. 3193–3200.

Corral-Verdugo, V., Fraijo-Sing, B., Pinheiro, J.Q., 2006. Sustainable behavior and time perspective: Present, past, and future orientations and their relationship with water conservation behavior. Interamerican J. Psychol. 40 (2), 139–147.

Cruz, S.M., 2017. The relationships of political ideology and party affiliation with environmental concern: A meta-analysis. J. Environ. Psychol. 53, 81–91.

de Moor, J., Uba, K., Wahlström, M., Wennerhag, M., De Vydt, M., 2020. Protest for a future II: Composition, mobilization and motives of the participants in Fridays For Future climate protests on 20-27 September, 2019, in 19 cities around the world.

Decter-Frain, A., Frimer, J.A., 2016. Impressive words: Linguistic predictors of public approval of the US congress. Front. Psychol. 7, 240.

Dono, J., Webb, J., Richardson, B., 2010. The relationship between environmental activism, pro-environmental behaviour and social identity. J. Environ. Psychol. 30 (2), 178–186.

Drury, J., Reicher, S., 1999. The intergroup dynamics of collective empowerment: Substantiating the social identity model of crowd behavior. Group Process. Intergroup Relat. 2 (4), 381–402.

Ellemers, N., Kortekaas, P., Ouwerkerk, J.W., 1999. Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. Eur. J. Soc. Psychol. 29 (2–3), 371–389.

Fan, Y., Li, M., Zhang, P., Wu, J., Di, Z., 2007. The effect of weight on community structure of networks. Physica A 378 (2), 583–590.

Fetterman, A.K., Boyd, R.L., Robinson, M.D., 2015. Power versus affiliation in political ideology: Robust linguistic evidence for distinct motivation-related signatures. Pers. Soc. Psychol. Bull. 41 (9), 1195–1206.

Fielding, K.S., McDonald, R., Louis, W.R., 2008. Theory of planned behaviour, identity and intentions to engage in environmental activism. J. Environ. Psychol. 28 (4), 318–326.

Fiske, S.T., Cuddy, A.J., Glick, P., 2007. Universal dimensions of social cognition: Warmth and competence. Trends Cogn. Sci. 11 (2), 77–83.

Flood, P., 1993. An expectancy value analysis of the willingness to attend union meetings. J. Occup. Organ. Psychol. 66 (3), 213–223.

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486 (3–5), 75–174.

Foster, M.D., 1999. Acting out against gender discrimination: The effects of different social identities. Sex Roles 40 (3–4), 167–186.

Fritsche, I., Barth, M., Jugert, P., Masson, T., Reese, G., 2018. A social identity model of pro-environmental action (SIMPEA). Psychol. Rev. 125 (2), 245.

Gallagher, R.J., Reagan, A.J., Danforth, C.M., Dodds, P.S., 2018. Divergent discourse between protests and counter-protests:# BlackLivesMatter and# AllLivesMatter. PLoS One 13 (4).

Gimpel, K., Schneider, N., O'Connor, B., Das, D., Mills, D., Eisenstein, J., Heilman, M., Yogatama, D., Flanigan, J., Smith, N.A., 2011. Part-of-speech tagging for twitter: Annotation, features, and experiments. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Vol. 2. Association for Computational Linguistics, pp. 42–47.

Gleich, D.F., 2015. PageRank beyond the web. SIAM Rev. 57 (3), 321–363.

González-Bailón, S., Wang, N., 2016. Networked discontent: The anatomy of protest campaigns in social media. Social Networks 44, 95–104.

Harring, N., Sohlberg, J., 2017. The varying effects of left–right ideology on support for the environment: Evidence from a Swedish survey experiment. Environ. Politics 26 (2), 278–300.

Haveliwala, T.H., 2002. Topic-sensitive pagerank. In: Proceedings of the 11th International Conference on World Wide Web. ACM, pp. 517–526.

Haveliwala, T.H., 2003. Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search. IEEE Trans. Knowl. Data Eng. 15 (4), 784–796.

Hawkins, I., Raymond, C., Boyd, R.L., 2017. Such stuff as dreams are made on: Dream language, LIWC norms, and personality correlates. Dreaming 27 (2), 102.

Hellsten, I., Leydesdorff, L., 2020. Automated analysis of actor–topic networks on twitter: New approaches to the analysis of socio-semantic networks. J. Assoc. Inf. Sci. Technol. 71 (1), 3–15.

Hornsey, M.J., Blackwood, L., Louis, W., Fielding, K., Mavor, K., Morton, T., O'Brien, A., Paasonen, K.-E., Smith, J., White, K.M., 2006. Why do people engage in collective action? Revisiting the role of perceived effectiveness. J. Appl. Soc. Psychol. 36 (7), 1701–1722.

Hornsey, M.J., Fielding, K.S., McStay, R., Reser, J.P., Bradley, G.L., Greenaway, K.H., 2015. Evidence for motivated control: Understanding the paradoxical link between threat and efficacy beliefs about climate change. J. Environ. Psychol. 42, 57–65.

Ife, J., 2018. Right-wing populism and social work: Contrasting ambivalences about modernity. J. Hum. Rights Soc. Work 3 (3), 121–127.

Jagers, S.C., Harring, N., Löfgren, A., Sjöstedt, M., Alpizar, F., Brülde, B., Langlet, D., Nilsson, A., Almroth, B.C., Dupont, S., et al., 2019. On the preconditions for large-scale collective action. Ambio 1–15.

Jasper, J., 2004. A strategic approach to collective action: Looking for agency in social-movement choices. Mobilization Int. Q. 9 (1), 1–16.

Joireman, J.A., Lasane, T.P., Bennett, J., Richards, D., Solaimani, S., 2001. Integrating social value orientation and the consideration of future consequences within the extended norm activation model of proenvironmental behaviour. Br. J. Soc. Psychol. 40 (1), 133–155.

Joireman, J.A., Van Lange, P.A., Van Vugt, M., 2004. Who cares about the environmental impact of cars? Those with an eye toward the future. Environ. Behav. 36 (2), 187–206.

Kawakami, K., Dion, K.L., 1993. The impact of salient self-identities on relative deprivation and action intentions. Eur. J. Soc. Psychol. 23 (5), 525–540.

Kawakami, K., Dion, K.L., 1995. Social identity and affect as determinants of collective action: Toward an integration of relative deprivation and social identity theories. Theory Psychol. 5 (4), 551–577.

Keller, J.M., 2012. Virtual feminisms: Girls' blogging communities, feminist activism, and participatory politics. Inf. Commun. Soc. 15 (3), 429–447.

Kelly, C., Breinlinger, S., 1995. Identity and injustice: Exploring women's participation in collective action. J. Commun. Appl. Soc. Psychol. 5 (1), 41–57.

Kirby, S., Griffiths, T., Smith, K., 2014. Iterated learning and the evolution of language. Curr. Opin. Neurobiol. 28, 108–114.

Kirilenko, A.P., Stepchenkova, S.O., 2014. Public microblogging on climate change: One year of Twitter worldwide. Global Environ. Change 26, 171–182.

Lambiotte, R., Delvenne, J.-C., Barahona, M., 2008. Laplacian dynamics and multiscale modular structure in networks. arXiv preprint arXiv:0812.1770.

Lancichinetti, A., Fortunato, S., 2009. Community detection algorithms: a comparative analysis. Phys. Rev. E 80 (5), 056117.

Larremore, D.B., Clauset, A., Jacobs, A.Z., 2014. Efficiently inferring community structure in bipartite networks. Phys. Rev. E 90 (1), 012805.

Lee Fox, D., Schofield, J.W., 1989. Issue salience, perceived efficacy and perceived risk: A study of the origins of anti-nuclear war activity. J. Appl. Soc. Psychol. 19 (10), 805–827.

Leiserowitz, A., Maibach, E., Rosenthal, S., Kotcher, J., Bergquist, P., Ballew, M.T., Goldberg, M., Gustafson, A., 2020. Climate change in the American mind: November 2019. PsyArXiv.

Lewin, K., 2016. Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. Hum. Relat..

Lewis, S., Pea, R., Rosen, J., 2010. Beyond participation to co-creation of meaning: mobile social media in generative learning communities. Soci. Sci. Inf. 49 (3), 351–369.

Lindsay, J.J., Strathman, A., 1997. Predictors of recycling behavior: an application of a modified health belief model 1. J. Appl. Soc. Psychol. 27 (20), 1799–1823.

McAdam, D., 2017. Social movement theory and the prospects for climate change activism in the United States. Annu. Rev. Political Sci. 20, 189–208.

McAdam, D., et al., 1999. Political Process and the Development of Black Insurgency, 1930-1970. University of Chicago Press.

Michinov, N., Michinov, E., Toczek-Capelle, M.-C., 2004. Social identity, group processes, and performance in synchronous computer-mediated communication. Group Dyn. Theory Res. Pract. 8 (1), 27.

Milfont, T.L., Wilson, J., Diniz, P., 2012. Time perspective and environmental engagement: A meta-analysis. Int. J. Psychol. 47 (5), 325–334.

Morton, T.A., Rabinovich, A., Marshall, D., Bretschneider, P., 2011. The future that may (or may not) come: How framing changes responses to uncertainty in climate change communications. Global Environ. Change 21 (1), 103–109.

Newman, M.E., 2001. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Phys. Rev. E 64 (1), 016132.

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D., Kleis Nielsen, R., 2017. Reuters Digital News Report 2017. Vol. 202017. Reuters Institute for the Study of Journalism, University of Oxford, Link: https://Reutersinstitute.Politics.Ox.Ac.Uk/Sites/Default/FileS/Digital%20News%20Report.

Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N., Smith, N.A., 2013. Improved part-of-speech tagging for online conversational text with word clusters. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 380–390.

Page, L., Brin, S., Motwani, R., Winograd, T., 1999. The PageRank Citation Ranking: bringing Order to the Web. Technical Report, Stanford InfoLab.

Park, G., Yaden, D.B., Schwartz, H.A., Kern, M.L., Eichstaedt, J.C., Kosinski, M., Stillwell, D., Ungar, L.H., Seligman, M.E., 2016. Women are warmer but no less assertive than men: Gender and language on Facebook. PLoS One 11 (5).

Payne, R., 2018. The global politics of climate change. p. 5.

Pennebaker, J.W., Boyd, R.L., Jordan, K., Blackburn, K., 2015. The Development and Psychometric Properties of LIWC2015. Technical Report.

Pietraszkiewicz, A., Formanowicz, M., Gustafsson Sendén, M., Boyd, R.L., Sikström, S., Sczesny, S., 2019. The big two dictionaries: Capturing agency and communion in natural language. Eur. J. Soc. Psychol. 49 (5), 871–887.

Pudrovska, T., Ferree, M.M., 2004. Global activism in virtual space: the European women's lobby in the network of transnational women's NGOs on the web. Soc. Politics Int. Stud. Gender State Soc. 11 (1), 117–143.

Rees, J.H., Bamberg, S., 2014. Climate protection needs societal change: Determinants of intention to participate in collective climate action. Eur. J. Soc. Psychol. 44 (5), 466–473.

Robins, G., Pattison, P., 2005. Interdependencies and social processes: Dependence graphs and generalized dependence structures. Models Methods Soc. Netw. Anal. 28.

Ross, C., Terras, M., Warwick, C., Welsh, A., 2011. Enabled backchannel: Conference Twitter use by digital humanists. J. Doc..

Sabherwal, A., Ballew, M.T., van Der Linden, S., Gustafson, A., Goldberg, M.H., Maibach, E.W., Kotcher, J.E., Swim, J.K., Rosenthal, S.A., Leiserowitz, A., 2021. The greta thunberg effect: Familiarity with greta thunberg predicts intentions to engage in climate activism in the United States. J. Appl. Soc. Psychol..

Sall, J., Stephens, M.L., Lehman, A., Loring, S., 2017. JMP Start Statistics: A Guide to Statistics and Data Analysis using JMP. Sas Institute.

Salmela, M., von Scheve, C., 2017. Emotional roots of right-wing political populism. Soc. Sci. Inf. 56 (4), 567–595.

Sarigöllü, E., 2009. A cross-country exploration of environmental attitudes. Environ. Behav. 41 (3), 365–386.

Saxton, G.D., Niyirora, J., Guo, C., Waters, R., 2015. # AdvocatingForChange: THe strategic use of hashtags in social media advocacy. Adv. Soc. Work 16 (1), 154–169.

Schmitt, M.T., Mackay, C.M., Droogendyk, L.M., Payne, D., 2019. What predicts environmental activism? The roles of identification with nature and politicized environmental identity. J. Environ. Psychol. 61, 20–29.

Schultheiss, O.C., 2013. Are implicit motives revealed in mere words? Testing the marker-word hypothesis with computer-based text analysis. Front. Psychol. 4, 748.

Shi, W., Fu, H., Wang, P., Chen, C., Xiong, J., 2020. Climatechange vs. Globalwarming: Characterizing two competing climate discourses on Twitter with semantic network and temporal analyses. Int. J. Environ. Res. Public Health 17 (3), 1062.

Simon, B., Loewy, M., Stürmer, S., Weber, U., Freytag, P., Habig, C., Kampmeier, C., Spahlinger, P., 1998. Collective identification and social movement participation. J. Personal. Soc. Psychol. 74 (3), 646.

Smith, L.G., Gavin, J., Sharp, E., 2015. Social identity formation during the emergence of the occupy movement. Eur. J. Soc. Psychol. 45 (7), 818–832.

Steyvers, M., Tenenbaum, J.B., 2005. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. Cogn. Sci. 29 (1), 41–78.

Stram, R., Reuss, P., Althoff, K.-D., 2017. Weighted one mode projection of a bipartite graph as a local similarity measure. In: International Conference on Case-Based Reasoning. Springer, pp. 375–389.

Stürmer, S., Simon, B., 2004. Collective action: Towards a dual-pathway model. Eur. Rev. Soc. Psychol. 15 (1), 59–99.

Tabachnick, B., Fidell, L., 2019. Using Multivariate Statistics, seventh ed. Pearson.

Tabrizi, S.A., Shakery, A., Asadpour, M., Abbasi, M., Tavallaie, M.A., 2013. Personalized pagerank clustering: A graph clustering algorithm based on random walks. Physica A 392 (22), 5772–5785.

Tagkaloglou, S., Kasser, T., 2018. Increasing collaborative, pro-environmental activism: The roles of motivational interviewing, self-determined motivation, and self-efficacy. J. Environ. Psychol. 58, 86–92.

Tajfel, H., 1974. Social identity and intergroup behaviour. Information (International Social Science Council) 13 (2), 65–93.

Tyler, T.R., McGraw, K.M., 1983. The threat of nuclear war: Risk interpretation and behavioral response. J. Soc. Issues 39 (1), 25–40.

Van Zomeren, M., Iyer, A., 2009. Introduction to the social and psychological dynamics of collective action. J. Soc. Issues 65 (4), 645–660.

van Zomeren, M., Kutlaca, M., Turner-Zwinkels, F., 2018. Integrating who we are with what we(will not) stand for: A further extension of the social identity model of collective action. Eur. Rev. Soc. Psychol. 29 (1), 122–160.

Van Zomeren, M., Postmes, T., Spears, R., 2008. Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. Psychol. Bull. 134 (4), 504.

van Zomeren, M., Spears, R., Leach, C.W., 2010. Experimental evidence for a dual pathway model analysis of coping with the climate crisis. J. Environ. Psychol. 30 (4), 339–346.

Vasi, I.B., Suh, C.S., 2013. Protest in the internet age: Public attention, social media, and the spread of 'occupy'protests in the United States. In: Politics and Protest Workshop. Vol. 13.

Verba, S., Nie, N.H., 1972. Participation in America: Social Equality and Political Democracy. Harper & Row, New York.

Xiong, Y., Cho, M., Boatwright, B., 2019. Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the# MeToo movement. Publ. Relat. Rev. 45 (1), 10–23.

Zhang, J., 2010. Self-enhancement on a self-categorization leash: Evidence for a dual-process model of first-and third-person perceptions. Hum. Commun. Res. 36 (2), 190–215.

Zhou, T., Ren, J., Medo, M., Zhang, Y.-C., 2007. Bipartite network projection and personal recommendation. Phys. Rev. E 76 (4), 046115.