Information, Communication & Society

ISSN: 1369-118X (Print) 1468-4462 (Online) Journal homepage: www.tandfonline.com/journals/rics20

Collective victimhood narratives in far-right communities on Telegram

Constantine Boussalis, Callum Craig & Aaron F. Rudkin

To cite this article: Constantine Boussalis, Callum Craig & Aaron F. Rudkin (2025) Collective victimhood narratives in far-right communities on Telegram, Information, Communication & Society, 28:8, 1407-1424, DOI: 10.1080/1369118X.2024.2391815

To link to this article: https://doi.org/10.1080/1369118X.2024.2391815

INFORMATION, COMMUNICATION & SOCIETY
2025, VOL. 28, NO. 8, 1407–1424
https://doi.org/10.1080/1369118X.2024.2391815

Collective victimhood narratives in far-right communities on Telegram

Constantine Boussalis, Callum Craig and Aaron F. Rudkin

Department of Political Science, Trinity College Dublin, Dublin, Ireland

ABSTRACT

Feelings of collective victimhood have been demonstrated to have a strong effect on ingroup bias, outgroup hostility and support for violence. The use of narratives stirring these feelings in far-right communications is especially concerning given their inclusion in the manifestos of several mass killers across Europe and North America. However, scholars still have little knowledge on the reach of such narratives as well as the extent to which major salient events increase attention to collective victimhood messaging among far-right followers. To address these gaps, we analyze the use of collective victimhood narratives on the popular secure instant messaging service, Telegram, which has exploded in popularity in response to mainstream platforms' attempts to moderate extremist speech. We develop a supervised machine learning algorithm to predict the presence of these discourses in text from over 18.5 million messages that were extracted from 1,870 far-right Telegram channels. We then use these data to test what impact the George Floyd protests and the storming of the US Capitol had on the frequency of collective narrative discussions on far-right Telegram. Our findings suggest that both events coincided with a significant increase in the use of victimhood narratives, thus providing insight into the radicalization process of far-right communities online.

ARTICLE HISTORY
Received 1 December 2023
Accepted 3 July 2024

KEYWORDS
Far-right; social media; Telegram; extremism

1. Introduction

Collective victimhood involves the self-perception of a group as having been harmed by another group or groups (Vollhardt, 2012, p. 137). This is an important aspect of far-right communication as their narratives depend on conspiratorial claims that privileged social groups (e.g., white, American males) are being victimized by an array of actors ranging from immigrants and ethnic minorities to cabals of satanic elites. In this study, we analyze how these narratives are used on Telegram, one of the most popular and least moderated online spaces used by the far-right. This secure instant messaging service

CONTACT Callum Craig craigca@tcd.ie Department of Political Science, 3 College Green, Trinity College Dublin, Dublin 2 D02 XH97, Ireland

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1369118X.2024.2391815.

has proven to be a popular choice among extremists and those users who were banned from mainstream social media platforms, enjoying significant growth following attempts by Twitter and other platforms to crack down on hate speech and the spread of conspiracy theories (Urman & Katz, 2020). Understanding how these spaces utilize victimhood narratives is important due to their role in radicalization as well as their ability to filter into mainstream discussions. For example, the right-wing former Fox News television host turned social media influencer Tucker Carlson has recently begun to make implicit use of the 'Great Replacement' victimhood narrative popularized by the online far-right, and this was then defended by the hard-right of the Republican Party. These narrative claims that white people are being replaced and overpowered due to ethnic minorities having more children and mass migration into the state. This same narrative was included in manifestos of several far-right terrorists radicalized in online spaces including Anders Breivik, Brenton Tarrant and Payton Gendron who murdered 10 people in a supermarket in a predominantly black area in Buffalo New York.

Specifically, this paper assesses how far-right Telegram users discuss collective victimhood and how these changes in response to the salient events of the January 6th 2021 insurrection and the George Floyd protests of Summer 2020. Our study builds upon findings from experimental studies which suggest that if an individual is reminded of the harm committed by their own ingroup or that an outgroup is suffering, this can cause a reflexive rise in the individual's sense of ingroup collective victimhood. Our contribution to this literature centers on the development of a machine learning classifier used to detect collective victimhood claims in a novel dataset of English speaking, primarily American, far-right Telegram channels. We then use these classifications to evaluate via a regression discontinuity in time (RDiT) design if messages are more likely to contain victimhood narratives following the aforementioned major events. The results demonstrate that both the George Floyd protests and the January 6th insurrection led to a statistically significant increase in victimhood claims within these far-right communities on Telegram.

1.1. Collective victimhood

Collective victimhood is grounded within Social Identity Theory, which argues that when individuals categorize themselves as belonging to a group, they attach their own self-image to how this group is perceived by insiders and outsiders (Tajfel et al., 1979, p. 40). Individuals will generally 'strive to maintain or enhance their self-esteem' and as a result will seek to maintain or enhance perceptions of their ingroups (Tajfel et al., 1979, pp. 40–41). This can do through differentiating the ingroup from other groups in an attempt to make the group identity seem more positive in comparison to other 'inferior' groups. This allows the group to create a more positive self-image, increasing intragroup solidarity but also potentially politicizing the intergroup relations.

Simon and Klandermans (2001) argue that there are three key stages in the politicization of identity. First, a shared sense of grievance emerges and this is then attributed to an adversary as the source of this grievance which then leads to a call for societal change or action (Simon & Klandermans, 2001, p. 324). Collective victimhood literature focuses on this process, assessing how the attribution of harm to an outgroup creates or reinforces intergroup hostility. There is an important distinction between *victimhood* and *victimization*, in that one is perceived and the other is based in reality (Vollhardt, 2012, p. 137). Despite this distinction, the perception of having been victimized regardless of the legitimacy of this claim can greatly affect attitudes and behaviors. This is especially important for the type of conspiratorial victimhood narratives that dominate far-right discourses as they portray dominant ingroups such as white men as being victimized by conspiring or savage outgroups. A sense of collective victimhood creates intragroup solidarity whilst increasing hostility to those who are seen to be carrying out the victimization (Noor et al., 2017, p. 125). This has been termed a 'siege mentality' in literature studying conflict scenarios, where groups who have been targeted and feel under attack advocate for more militant and extreme responses (Bar-Tal & Antebi, 1992a,b). However, this process can also operate regardless of actual experiences of being attacked or in conflict with another group. Portraying marginalized outgroups as violent criminals or invaders are core aspects of far-right politics. Groups and individuals refer to Jihadi attacks and misinterpreted or entirely false racial crime statistics to justify violence against or oppression of outgroups to protect the ingroup (Bloch et al., 2020; Marcks & Pawelz, 2022). Portrayal of refugees as potential criminals, terrorists or 'invaders' are some of the core rhetorical devices used to mobilize against policies that offer aid to some of the most marginalized groups in societies (Greussing & Boomgaarden, 2017; Pope, 2017).

The literature on emotional and cognitive effects of collective victimhood is broad and goes beyond these siege mentality studies. Collective victimhood has been found to alter how individuals within groups assess trustworthiness, being more forgiving and trusting of ingroup members even following betrayals than they are towards those from outside of the group (Rotella et al., 2013). Schori-Eyal et al. (2017) find that those who view ingroup victimization as salient and perpetual were more likely to assume negative intentions from outgroup members in ambiguous situations. Sustained attitudes of distrust and assumptions of hostility against outgroups will create and sustain intergroup hostility as feelings of victimhood continue to be perpetuated. As previously discussed, these feelings of victimhood can also increase support for harsher policies and violence against the outgroup (Bar-Tal & Antebi, 1992b; Marcks & Pawelz, 2022).

Groups can also reflexively use victimhood narratives, to in effect compete for victim status (Young & Sullivan, 2016). Work building on findings from social dominance

Table 1. Far-right U.S. communities on Telegram with most influential channels.

No.	Community label	Prominent channels
1	Trump/MAGA	Donald J. Trump, Project Veritas, Lin Wood, OfficialMcAfee, James O'Keefe, The Trumpist
2	Qanon (Trafficking)	John F. Kennedy Jr., Take The Oath, Adrenochrome and Human Trafficking, Q Patriots
3	Proud Boys / Alt-Right	Roosh Valizadeh, Gab.com, NWO IS BOLSHEVISM, Proud Boys, ZeroHedge, Bellum Acta,
4	COVID denial	The Great Re-Opening, David Avocado Wolfe, Robert W Mallone MD, COVID VACCINE VICTIMS, Freedom ideas no covid vaccines
6	Qanon (Great Awakening)	We Are The Faithful, Great Awakening Channel, US Patriots, QAnon Central, WWG1WGA Here, Awaken We Are
7	Crime / Far-Right Journalism	Andy Ngo, West Coast News and Discussion, Western Heritage, Lord Vinnie Sullivan, WokeIntelDrops

orientation have found that socially dominant groups can view calls for equality or justice as attacks on their group and respond by emphasizing their own victimhood and suffering (Danbold et al., 2022; Kahalon et al., 2019). This process is important for understanding many contemporary victimhood narratives especially from the right (Table 1).

Sullivan (2012) finds in two studies that individuals respond to claims of their ingroup victimizing another by making stronger ingroup claims of discrimination and victimhood. This shows how victimhood narratives of powerful groups can become reflexive attempts to maintain group (and therefore self-) esteem whilst ignoring very real and quantifiable differences in validity of victimization claims. Frustration can also arise that the outgroup is gaining supposed moral benefits from their victimhood status whilst the ingroup's claim is ignored, as observed by work assessing Antisemitism in post-austerity Greece (Antoniou et al., 2020).

1.2. Far-right on Telegram

The predominance of victimhood narratives among the far-right has been noted by many scholars (Korolczuk, 2020; Mudde, 2007; Oaten, 2014). This has led to a small body of literature analyzing their use in specific parties, movements or by individual leaders (Chouliaraki & Banet-Weiser, 2021; Marcks & Pawelz, 2022; Sengul, 2021). Urman and Katz (2020) investigate the extraordinary growth of far-right actors on Telegram and ask to what extent Telegram has served as an alternative social media platform in the wake of mass bans of far-right users on mainstream social media sites. The results of a network analysis of more than 53,000 Telegram channels shows how the far-right on Telegram is mainly structured along national and ideological lines. Numerous far-right subcultures are represented including those influenced by 4chan, neo-Nazis, QAnon, the alt-right as well as Trump/MAGA related channels. As Guhl and Davey (2020) demonstrate, this is not relegated only to the online far-right but also established organizations linked to hate speech and extremist violence such as The Proud Boys and numerous militias. As Urman and Katz (2020) show, the impact of a mass banning on the growth of the far-right on Telegram can be significant. The banning of Trump from Facebook and Twitter along with the wave of bans related to January 6 seem to have turbocharged the Telegram far-right. This is supported by (Bryanov et al., 2021) who find that censorship, electoral fraud, and other Trump related topics were dominant.

The growth of Telegram raises questions about the effectiveness of 'deplatforming' as a response to extremist content and users. Deplatforming involves the banning of users by moderators of a platform for reasons such as hate speech, inciting violence, or otherwise expressing views deemed unacceptable. Whilst it is usually effective at removing problematic users from the specific platform, there are concerns that it simply drives users into more extreme spaces where they can no longer be tracked effectively (Chandrasekharan et al., 2017; Jhaver et al., 2021; Rogers, 2020). Telegram allows for the sharing of text, links, and embedded files (audio, images, and video) either in a social or broadcast media format. In our sample, most channels follow the broadcast format (i.e., the channel owner sharing content with all who joined the channel) though a very small number allow followers to also post on the channel. In many ways these are similar to the core functionality of Twitter/X, but Telegram is able to provide this service despite much more limited resources while remaining decentralized with a high level of security and user anonymity. In comparison, platforms like Gab or BitChute attempt to imitate complex and expensive competitors such as Facebook and YouTube incurring the high costs of video-hosting, live-streaming and securing user data in a centralized social media environment. Essentially, Telegram is a simple platform with a simple purpose that seems fully able to fulfill the essential functions of Twitter in allowing users to share text, images and links to followers securely while remaining solvent (Rogers, 2020). The main limitation of Telegram is that it has a smaller user base than mainstream platforms, but results from Urman and Katz (2020) show a remarkable ability for deplatformed far-right actors to maintain a similar number of followers post-migration. Not only is there a large enough user base to allow users to maintain relevancy through Telegram, it is also consistently attracting high profile names either due to the individual getting banned or joining in protest of others getting banned. This includes the hard-right of the Republican party such as Marjorie Taylor Green and Donald Trump Jr, as well as central influencers of the far-right such as Alex Jones, Milo Yiannopoulos, and Richard Spencer. The ability for celebrities to draw in new users and build communities around themselves may help to make Telegram an effective response to deplatforming strategies (Åkerlund, 2020). If users and communities successfully migrate to these dark corners where they can speak freely to an audience of similar size, we also need to understand if and how their posted content varies, including their use of victimhood narratives.

However, there are limitations in the existing literature: to our knowledge, no large-scale content analyses of the use of victimhood narratives by the far-right have been conducted. Existing quantitative research primarily relies on experiments, whereas we leverage observational data to assess how real events influence the use of victimhood narratives. On the other hand, qualitative research focuses on speeches or other materials produced by a small number of highly influential figures such as political elites. We contribute to the field by utilizing computational methods to detect and analyze victimhood narratives from a large corpus of Telegram messages. This allows us to assess the online communication of far-right posters including both influential elite and 'regular' posters since January 2020.

In this paper, therefore, we ask the following questions:

(1) How prevalent are instances of collective victimhood narratives among far-right users on Telegram?
(2) To what extent do salient political events cause far-right Telegram users to utilize more victimhood narratives?

While we have weak *a priori* expectations for the first question, we do expect that major events such as the murder of George Floyd and the attacks on the U.S. Capitol would to lead to more discussion of collective victimhood among users within far-right Telegram communities. Little work has been carried out on the types of events that may trigger feelings of victimhood, and none in the context of the far-right. Experimental work has indicated even very minor events or the framing of events as targeting of the ingroup are capable of activating these feelings even when ingroup and outgroup identifications are fairly weak or arbitrary (Brewer & Pierce, 2005). In the case of January 6, building from social identity and collective victimhood theories we assume that victimhood will be activated as the ingroup's own actions will be varyingly justified, denied or mitigated whilst the subsequent response of the outgroup will be framed as unjustified, draconian, and unfair (Bilali & Vollhardt, 2019). Similar dynamics are also likely drivers of collective victimhood in the case of the George Floyd murder and subsequent Black Lives Matter protests. Competitive victimhood and related theories such as racial resentment and white backlash also suggest that dominant ethnic groups will perceive status threat by even peaceful or non-confrontational moves to redress prior injustice which is also a likely extreme driver of feelings of victimhood in the context of country-wide demands for racial justice (Abramowitz & McCoy, 2019; Wilkins & Kaiser, 2014). These theoretical assumptions coupled with the high saliency of the two events provide justification to focus our analysis around these two events. We begin the study by detailing the collection and filtering of our Telegram corpus; the selection of training data; and the pipeline used to build, train, apply, and validate our classifier. We then discuss the modeling strategy used to test our hypotheses. This is followed by a presentation of descriptive statistics and the main results from our estimated statistical models. The paper ends with a discussion of the implications of our results.

2. Data & methods

2.1. Data collection

We wrote a series of programs to automatically interact with and extract messages posted on over 1,870 Telegram channels that have been identified as far-right (or far-right affiliated) in Europe and North America. The data collection strategy follows a snowball sampling approach, whereby approximately 50 'seed' accounts were harvested from public sources. As Telegram has no public directory facility, it is impossible to enumerate all relevant channels; moreover, Telegram has no public 'follow' system. Instead, we identified additional channels when their messages were shared by already-scraped channels. When admitting a new channel to the corpus, all messages in the channel were recorded and analyzed for out-bound links to other Telegram channels. The next channel admitted was the channel that had been linked to most broadly and most often by existing channels. By *homophily*, these channels were likely to have substantively similar content to the seed channels.

This approach described above is effective at quickly traversing a large, connected network, but it has limitations in that it is both under-inclusive (because channels whose messages are not shared will never be identified) and over-inclusive (some channels included will be general-interest channels rather than far-right channels).

Harvesting continued until the qualitative relevance of admitted channels to the project declined. Using metadata of channels such as the username and the channel description as well as qualitatively checking a random sample of messages we manually removed channels that were obviously irrelevant. These irrelevant channels included Iranian propaganda sources, jihadi content, and cryptocurrency/NFT investment ad-vice channels. Despite using only American channels in our seed list we incidentally collected channels from the UK, France, Germany, Italy, Austria, Ireland, United States, Canada, Australia, and South Africa in a variety of languages. Making use of the CLD2 R package, we check the predominant language used in each channel and excluded channels that did not use English in 80% or more of the messages (Riesa & Giuliani, 2013). It is important to note that many channel owners choose to remain anonymous, making it difficult to pinpoint their exact origins. Many, however, can be easily recognized through their channel names and the content they share. As a result, we can only categorically distinguish the geographic origins of channels based on the content and language of the channel, with some channels in our sample remaining ambiguous with respect to geographic origin. Our full corpus, therefore, consists of approximately 18.5 million messages grouped into 2.90 million documents from 1,870 channels (once irrelevant and inaccessible channels are removed).

2.2. Network analysis

Under the assumption that ideologically aligned channels will refer to each other more frequently we perform a network analysis to cluster channels. We begin by scanning every message posted to each channel in our corpus and isolating those that link to other Telegram channels in our corpus. This allows us to create a weighted directed network graph of relations between these channels. By homophily, the same assumption underpinning our initial sampling strategy, channels that are more densely connected together form intellectual subcommunities. We use the hierarchical agglomeration method described in Clauset et al. (2004) to detect clusters of channels based upon the frequency by which they link to each other. Note that this approach is entirely independent of the text analysis presented elsewhere in the paper: only links between channels enter this analysis.

We find six semantically coherent clusters (communities) of channels. Each of these clusters represents a certain sub-group or faction within the broader far-right ecosystem.[1]

2.3. Classification of victimhood narratives

The large size of our Telegram message corpus provides an ideal opportunity to employ scalable, computer-assisted methods to accurately and reliably classify the presence of collective victimhood narratives. The field of machine learning offers a range of statistical methods for achieving this objective. In this paper, we use supervised learning methods (see Grimmer & Stewart, 2013, p. 9) which rely on discriminative models to predict categories or 'classes' which in our case, are whether a message contains a collective victimhood narrative or not.

2.3.1. Development of training dataset

The first task in the machine learning pipeline is to develop a training dataset that will serve as input for the set of models we train. The objective here is to provide annotated examples of messages that contain victimhood narratives and ones that do not to a range of different models whose performance is evaluated and from which the best performing algorithm is selected for the remainder of the study. To construct this training set, two members of the research team manually annotated a selection of $n = 2,\ 959$ messages. Details of the sampling strategy, coding procedure and inter-coder reliability are provided in Appendix A.

2.3.2. Machine learning classification

A priori, there is very little theoretical reason to believe a particular method is best suited to classify large text corpora. Our goal is achieve a high level of accuracy without over-fitting. Moreover, our interest is in using the presence of a victimhood narrative as an *input* to model heterogeneity between groups and across time; thus, beyond face validity we are unconcerned with specifically which features predict a victimhood narrative. In light of these considerations, we treat this as a 'horse race' between machine learning techniques. We fit Random Forests, Logistic LASSO, Logistic Elastic Net, and neural networks (using the RoBERTa feature encoder) to assess goodness of fit (Breiman, 2001; Devlin et al., 2018; Liu et al., 2019; Tibshirani, 1996; Zou & Hastie, 2005). Full descriptions of parameterizations and outcomes of this approach are discussed in Appendix C. We find that the neural network approach results is the highest F1 (class-balanced accuracy) at 0.701 and (AUC area under the curve) at 0.868. Having selected the most promising model, we fit the remaining $n = 2,\ 905,\ 200$ messages and output predicted probabilities of being a victimhood narrative.

2.4. Statistical modeling approach

To test our hypotheses that salient events (the murder of George Floyd and January 6th Capitol Insurrection) led to increased discussion of collective victimhood narratives on far-right Telegram channels, we rely on a regression discontinuity in time (RDiT) design (see Hausman & Rapson, 2017). Regression discontinuity designs focus on observed differences in an outcome variable among units that are situated approximately below and above some given threshold value of an independent variable (also commonly referred to as the 'forcing variable') that determines the assignment to treatment. The intuition here is that we can assume that for units that are very close to this threshold, it was randomness (or as good as randomness) that led to their being 'assigned' a treatment (or not). Generally speaking, if certain assumptions about the as-if random nature of the process hold, then the assignment mechanism of the threshold is not confounded and therefore comparisons between the observations that are situated around the threshold mimic the sort of comparisons we would be able to make in a controlled experiment.

In our particular case, the 'forcing variable' is *time* and we do not have a cross-sectional dataset. Both these peculiarities of our dataset complicate matters somewhat. That is, we are forced to opt for a special case of regression discontinuity which exploits variation in the time dimension without being able to rely on a 'non-treated' comparison series. One shortcoming of RDiT is that the date of a threshold event may not be randomly assigned, but rather the consequence of some predetermined process. In this situation, it may be difficult to defend the 'local randomization' argument of the classical regression discontinuity design (Hausman & Rapson, 2017). More specifically, a threat to our inference is that posters (or postings) 'sort' around the discontinuity threshold: meaning that postings just before the threshold and just after are not the same, but for the event studied. Nevertheless, our inference can be defended on several grounds. First, the events studied – and particularly the murder of George Floyd and subsequent protests – occurred as-if randomly with respect to the knowledge of the posters. Second, posters are continuously posting before and after the events; it is very unlikely that any poster felt an obligation to defer posting until a major event happens (thus, they are unlikely to 'sort in time'). We expect only that the events in question will change the decision calculus about the types of rhetoric posters choose to post, which is the effect being studied. Finally, the events discussed are incredibly major news stories, so it is unlikely that a 'bundled treatment' occurs (where other causes take place contemporaneously with the events studied). Another potential limitation of the RDFiT approach is that unobserved variables as well as time-series characteristics (e.g., serially correlated errors) may lead to biased inferences. Unfortunately, our dataset does not include meaningful unit-specific covariates for which we can account for. However, we do account for serial dependence by using Telegram channel level clustered standard errors in all statistical models below. We also control for the categorical network cluster of each document, which are described above.

To sum up, we conduct a regression discontinuity in time analysis, treating our events of interest as exogenous shocks to the far-right discourse. In each case, we dynamically select optimum bandwidths for the regression discontinuity, and verify the choice of bandwidth is not sensitive by running alternate, manually selected bandwidths (Calonico et al., 2017).

3. Results

3.1. Descriptive statistics

Of the 2, 823, 601 documents that were processed by the model, 271, 887 documents (9.6%) had a victimhood prediction greater than 0.5. This suggests that discussions of victimhood remain an important part of the overall far-right discourse – yet not a dominant one. The frequency of these discussions also varies by sub-community. Figure 1 displays the proportion of documents that our model predicted to contain a collective victimhood narrative, across six far-right sub-communities. Here we can see that over

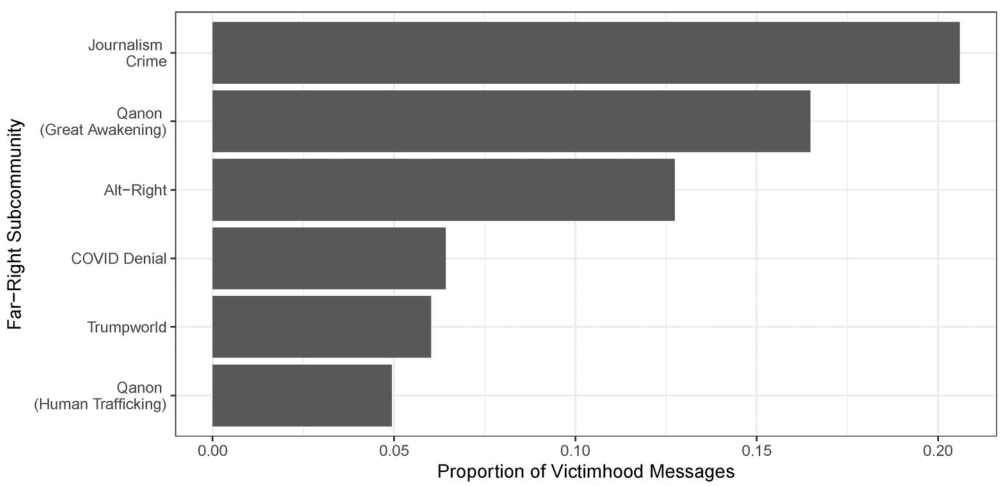

Figure 1. Proportion of documents that contain a collective victimhood narrative by far-right sub-community.

20% of the documents produced by channels that fall within what we interpret as a right-wing journalism and crime reporting cluster focus on narratives surrounding collective victimhood. Given that high profile channels within this cluster such as Andy Ngo are largely dedicated to reporting on violence committed by left-wing radicals and fear-mongering about perceived far-left conspiracies, this finding makes intuitive sense (Strickland, 2020). We also observe quite large proportions for non-human trafficking related Qanon channels as well as channels within an alt-right cluster likely driven by the extreme and conspiratorial narratives prevalent in these spaces. Although all clusters have quite a large proportion of victimhood narratives, the Covid-19 denial and Trump clusters may veer closer to more mainstream right-wing and conservative ideologies than the other clusters and thus use less extreme rhetoric. The relatively low level of victimhood narratives in the Qanon human trafficking cluster is surprising given the subject matter and ideology represented. However, we suspect that this finding is because messages in these channels largely fail to attribute the harm to an outgroup within the message even if the context of the channel might imply that they're referring to trafficking supposedly endorsed or carried out by certain outgroups (i.e., Democrats, leftists and immigrants). This means that our classifier may struggle to pick up these narratives as it is entirely dependent on what is included within the text rather than using this external context.

We also assess the entry of channels and general activity throughout our dataset overtime. We see a sustained largely linear rise in monthly post counts throughout the period. In contrast however, we see a sudden sharp increase in new channels coinciding with the 6 January 6 insurrection and the banning of Trump and numerous other far-right influencers and political figures due to their encouragement of the attacks and prior election denial. This raises some empirical challenges for our RDiT analysis as we seek to assess the effect of the failed insurrection on a far-right ecosystem that also experiences a sudden transformation in audience and content producers due to this shock. We discuss these challenges further in the discussion section (Figure 2).

We next focus our attention to the temporal variation of the salience of collective victimhood over our sample period. The top pane of Figure 3 displays the weekly share of

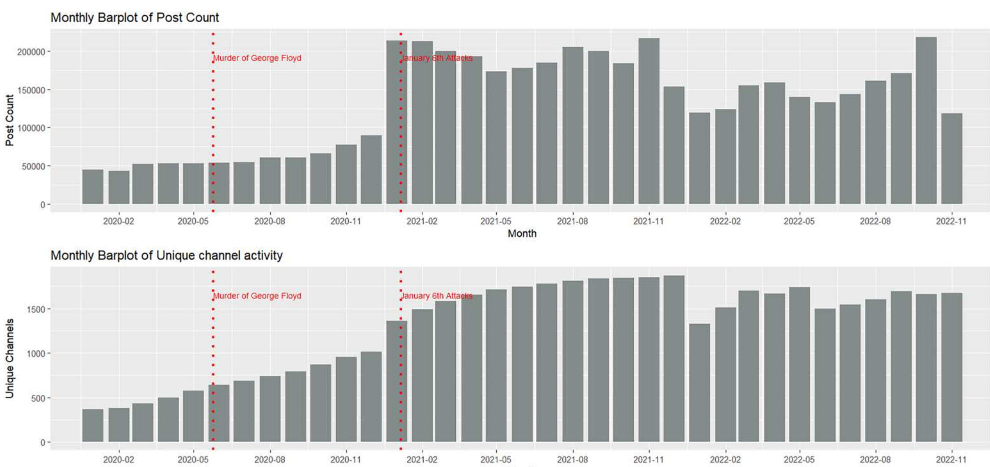

Figure 2. Messages and unique channels per month.

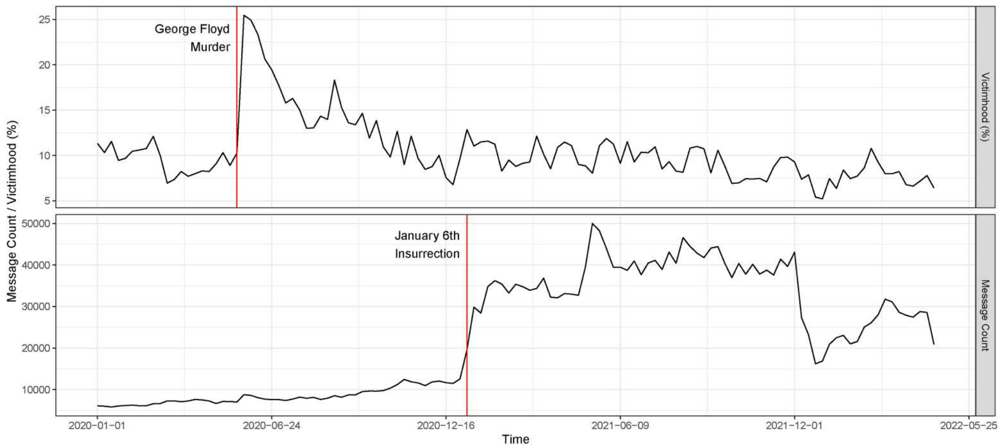

Figure 3. *Top pane:* Weekly share of documents that contain a collective victimhood narrative. *Bottom pane:* Number of documents produced by our sample of far-right channels over time.

documents in all far-right channels that feature collective victimhood from the weeks of 1 January 2020 until 25 April 2022. What stands out immediately is the large spike in victimhood claims just following the murder of George Floyd on 25 May 2020, going from an average of around 10% to over 25%. The attention to victimhood remained relatively high during the summer and autumn weeks, albeit in a steady decline back to the pre-murder average levels. We should also note how there also seems to be a pronounced (yet not as large) spike in victimhood discussions following the 6 January Capitol Insurrection. The bottom pane of Figure 3 reveals an interesting story regarding this event however: we observe a very large increase in the volume of messages produced among far-right channels in the aftermath of the Insurrection, with sustained high levels throughout the remaining period of observation. Notably, the plot reveals how the increase in activity begins to occur just before the Insurrection, but also how attention continues to spike following the attacks. What is also important to note is that while there is an increase in activity following the murder of George Floyd, the January 6th aftermath dwarfs this period with respect to increased messaging overall. This is likely due to the subsequent crackdowns on mainstream social media platforms against pro-insurrection voices causing their migration to Telegram. This migration flow was further intensified by the temporary shutdown of Parler and its removal from app stores due to its use by the Capitol attackers.

3.2. Regression discontinuity in time results

The murder of George Floyd on 25 April 2020 sparked weeks of intense protest across major American cities. We treat this as an exogenous shock to far-right discourse. Our results suggest an 8.6% increase (in raw terms) in the propensity for messages written after the shock to be classed as victimhood claims. The optimum bandwidth selected by the model is 23.8 days (effectively comparing documents from April 2020 to documents from May 2020; after clustering adjustments, an effective sample size of $n = 52,413$). We manually run bandwidths ranging from 10 days through 62 days and find

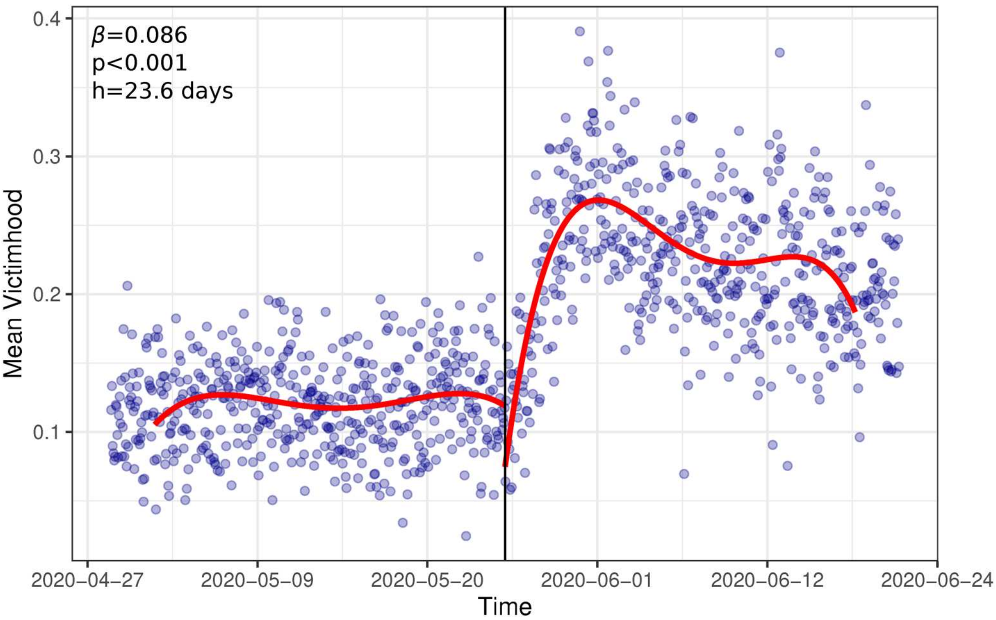

Figure 4. Plot showing regression discontinuity in time result around the murder of George Floyd, 25 May 2020.

consistently positive and generally significant results. Figure 4 visually depicts the results, while Table 2 presents the estimated treatment effect.

Table 2 depicts results from a regression discontinuity in time analysis of the effect of George Floyd's murder on victimhood discourse in our sample. N_{eff} denotes the 'effective' sample size used to estimate the effect, after adjustments for standard error clusters. Standard errors are clustered by channel ID. Bold text reflects the optimum model-selected bandwidth, 23.8 days.

Our next analysis focuses on the attacks on the US Capitol by supporters of then-President Donald Trump. We apply a similar modeling strategy around the week of January 6th, 2021. Specifically, we set the threshold to the hour of the day when Trump delivered his speech at the 'Save America' rally on the Ellipse in Washington, D.C. Our results show that there was a 3.7% increase in the propensity of collective victimhood messaging following the Capitol attacks. The optimum bandwidth for this analysis is 44.3 days. This model is therefore relying on an effective comparison of documents posted between November 2020 and February 2021, which left this model with an effective sample size of 278,094 documents. As with the previous analysis, we run models with bandwidths ranging between 10 and 62 days and find consistently positive and statistically significant

Table 2. Regression discontinuity in time: murder of George Floyd.

h (bandwidth)	*Neff*	*β*	se	95% C.I.	*p*-value
23.8	**52,413**	**0.086**	**(0.020)**	**(0.046, 0.126)**	**< 0.001**
10	22,248	0.014	(0.019)	(−0.023, 0.051)	0.451
14	31,354	0.043	(0.020)	(0.003, 0.083)	0.031
21	46,506	0.078	(0.020)	(0.038, 0.118)	< 0.001
28	60,866	0.095	(0.021)	(0.054, 0.135)	< 0.001
62	133,763	0.115	(0.021)	(0.073, 0.157)	< 0.001

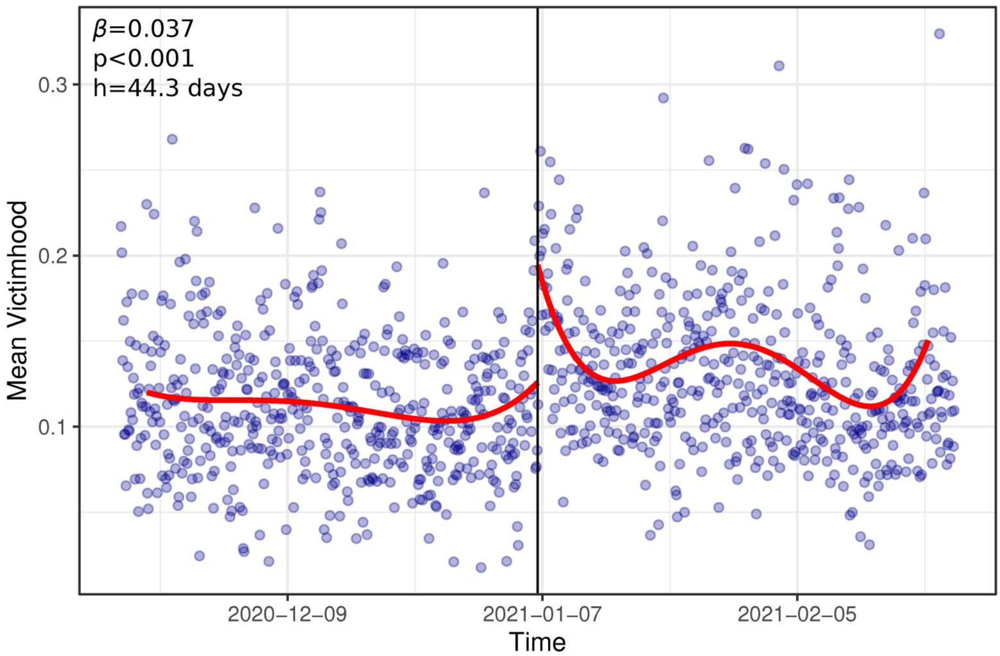

Figure 5. Plot showing regression discontinuity in time result around the Capitol Insurrection, January 6th, 2021.

effects for all bandwidth specifications. We illustrate the results for the optimal bandwidth in Figure 5 and also provide a summary of full results in Table 3.

Table 3 depicts results from a regression discontinuity in time analysis of the effect of the Capitol Insurrection on victimhood discourse in our sample. N_{eff} denotes the 'effective' sample size used to estimate the effect, after adjustments for standard error clusters. Standard errors are clustered by channel ID. Bold text reflects the optimum model-selected bandwidth, 44.3 days.

4. Discussion

Our results indicate that both events of interest drove a rise in victimhood claims on far-right Telegram. We observe statistically significant positive effect of these events on victimhood mentions across all model specifications. There are two competing explanations for why we might observe this effect. First, the result could be demonstrating the salience and emotive effect of the event on the individuals operating the channels. Second, it could also be driven by a sudden change in the audience through processes such as the banning

Table 3. Regression discontinuity in time: capitol insurrection.

h (bandwidth)	*Neff*	*β*	se	95% C.I.	*p*-value
44.3	**273,094**	**0.037**	**(0.010)**	**(0.017, 0.056)**	**< 0.001**
10	49,873	0.062	(0.012)	(0.038, 0.086)	< 0.001
14	75,598	0.047	(0.012)	(0.024, 0.070)	< 0.001
21	115,324	0.035	(0.011)	(0.013, 0.057)	0.002
28	162,951	0.032	(0.011)	(0.011, 0.053)	0.003
62	388,766	0.040	(0.009)	(0.021, 0.058)	< 0.001

of Donald Trump and other figures from mainstream social media platforms leading to a migration of followers to alternative platforms, thus creating a larger and different far-right Telegram as shown in Figure 2. The largest influx of new channels occurs in January 2021 and the months after this. This means that it is difficult to disentangle the effects of new channels likely driven by bans on other platforms from the event's effect on the already existing audience. Additionally, as channels interact with each other and share content, changes in the overall far-right Telegram ecosystem also influence channels that existed prior. Unfortunately, we cannot examine the effects of new users within channels, as all but a small number of channels prevent followers from replying and instead act as broadcasting platforms for the channel owners to share content.

Despite these limitations and challenges, we observe that the number of active channels remains relatively stable in the months surrounding the murder of George Floyd, during which we see a significant spike in victimhood narratives. This stability allows us to reasonably assume that the entirety of the January 6 spike cannot be explained solely by a change in the audience. It therefore appears that these events activated feelings of victimhood among the far-right on Telegram.

While we can observe that the events had an effect independent of audience change, drawing comparisons between the two events remains a challenge due to the aforementioned audience shifts. However, this does not mean that the two events are qualitatively the same. The reactions to George Floyd's murder are likely driven mostly by emotional responses to the mobilization and political action of their ideological or supposed racial enemies, whereas the reaction to 6 January pertains to a failed uprising by relevant ingroups and the subsequent response by institutions and society. Although we see different results empirically for the two events (i.e., the killing of George Floyd prompts a greater but less sustained rise in victimhood narratives than 6 January), we cannot effectively measure how much of this difference is due to the type of event versus the changes in audiences and the channels that are active.

5. Conclusion

This study analyzes how a large set of far-right Telegram communities utilized victimhood narratives over the period January 2020 to April 2022. We tested the extent to which two key events (the murder of George Floyd and the storming of the US Capitol) resulted in an increase in the use of victimhood narratives. To achieve this, we developed a machine-learning classifier trained on a novel training set of collective victimhood narratives and utilized a regression discontinuity in time analysis. Our findings suggest that, as expected, both of these major events coincided with a marked increase in the proportion of victimhood narratives, and that this effect is more pronounced in the case of the George Floyd murder and subsequent protests.

With respect to limitations, we can highlight three main sources in the current study. First, the training set might not be classifying all 'types' of collective victimhood narratives that are present in the wider far-right Telegram ecosystem. That is, the model may have not been introduced to alternative victimhood narratives that were not typically discussed in the channels selected for the training set (see Appendix A). Further, the source of error may also be due to our choice of retrieving example documents from the selected 27 channels for the training dataset based on their assigned topic from a topic model.

Alternatively, we may indeed have all 'types' of narratives in our training set, but the model simply finds it more difficult to classify some subset.

Second, on the inference side of the study, the use of a sharp discontinuity design for our data might not be optimal in the sense that both the events in our study featured discrete shocks (i.e., a murder and a televised insurrection of the Capitol), however the effects clearly had long-term characteristics as well due to continuing media coverage, protests and investigations. Future research should address these concerns by implementing a fuzzy RDFiT design.

Lastly, although we can be reasonably confident that some of the observed effect is driven by activation of collective victimhood, we cannot determine how much of this is driven by a changing audience and ecosystem especially in the case of the January 6 attacks. Nonetheless, this study offers a number of important contributions to the field of political communication and the literature on far-right politics. We conduct the first large scale quantitative analysis of far-right collective victimhood narratives with the largest far-right Telegram dataset to date. In particular, we are the first to our knowledge to have applied supervised machine learning methods to classify such narratives within the online far-right ecosystem. With these predictions in hand, we are also the first to test whether highly salient exogenous shocks exert influence on the spread of this dangerous discourse within the online far-right space. Our results suggest a marked upsurge in victimhood narratives following the murder of George Floyd. While some of this may be driven by incidents of violent unrest, it also indicates the use of competitive victimhood claims from more socially privileged groups even in response to peaceful protests against racialized police brutality in keeping with expectations from the competitive victimhood literature. Although we are limited in our ability to compare the impact of the two events as discussed previously, we also note a smaller but more sustained rise in victimhood narratives activated by the 6 January attacks. Despite those seen as the ingroups by the far-right are the aggressor in this event, the rise in feelings of victimhood among the ingroup is not unexpected. This is because the offences committed by the ingroup will either been seen as justified or minimized by the ingroup whilst 'attacks' against the ingroup such as arrests, bans, or public condemnations will be more psychologically affective even if comparatively much less extreme than trying to overturn an election by force.

Future research could examine the complex interactions between extreme narratives, content producers and shifting audiences especially as Elon Musk's Twitter/X suddenly and unexpectedly reversed many of the previously effective moderation tools. Other researchers could also perhaps control for some of these ecosystem effects by tracing individual users across platforms if they could overcome the significant challenges posed by collecting historical data from users who were driven to other platforms by having their accounts deleted or banned.

Note

1. Cluster 5 was excluded as non-relevant. The cluster contained flat-earth, moon landing, Tartarian empire, and other new-age conspiracy theories.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Irish Research Council: [Grant Number]; NORFACE: [Grant Number].

Notes on contributors

Dr Constantine Boussalis holds a Ph.D. in Political Science (World and Comparative Politics) and a M.A. in International Studies from Claremont Graduate University as well as a B.A. in Political Science from California State University, Los Angeles. His research interests include environmental politics, political communication, and computational social science. Before joining the Department of Political Science at Trinity College Dublin in 2014, Constantine served as Empirical Research Fellow at Harvard Law School from 2011 [email: craigca@tcd.ie].

Aaron Rudkin is a Political Science PhD candidate with UCLA, Research Fellow at Trinity College Dublin, and incoming Research Fellow at MIT. He has worked on projects including: ExId, Nationscape, DeclareDesign, Voteview/NOMINATE.

Callum Craig is a Political Science PhD candidate at Trinity College Dublin. His current research delves into diverse topics, including communication strategies of US police unions, extremist discussions on Telegram, and the use of collective victimhood by US politicians and in media reports. Collaborating with colleagues from Exeter University, University of Copenhagen and Ulster University.

References

Abramowitz, A., & McCoy, J. (2019). United States: Racial resentment, negative partisanship, and polarization in Trump's America. *The Annals of the American Academy of Political and Social Science*, *681*(1), 137–156. https://doi.org/10.1177/0002716218811309

Akerlund, M. (2020). The importance of influential users in (re) producing Swedish far-right discourse on Twitter. *European Journal of Communication*, *35*(6), 613–628. https://doi.org/10.1177/0267323120940909

Antoniou, G., Dinas, E., & Kosmidis, S. (2020). Collective victimhood and social prejudice: A post-holocaust theory of anti-semitism. *Political Psychology*, *41*(5), 861–886. https://doi.org/10.1111/pops.12654

Bar-Tal, D., & Antebi, D. (1992). Beliefs about negative intentions of the world: A study of the Israeli siege mentality. *Political Psychology*, 633–645. https://doi.org/10.2307/3791494

Bar-Tal, D., & Antebi, D. (1992b). Siege mentality in Israel. *International Journal of Intercultural Relations*, *16*(3), 251–275.

Bilali, R., & Vollhardt, J. R. (2019). Victim and perpetrator groups' divergent perspectives on collective violence: Implications for intergroup relations. *Political Psychology*, *40*(S1), 75–108.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.1111/pops.12570

Bloch, K. R., Taylor, T., & Martinez, K. (2020). Playing the race card: White injury, white victimhood and the paradox of colour-blind ideology in anti-immigrant discourse. *Ethnic and Racial Studies*, *43*(7), 1130–1148. https://doi.org/10.1080/01419870.2019.1648844

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brewer, M. B., & Pierce, K. P. (2005). Social identity complexity and outgroup tolerance. *Personality and Social Psychology Bulletin*, *31*(3), 428–437. https://doi.org/10.1177/0146167204271710

Bryanov, K., Vasina, D., Pankova, Y., & Pakholkov, V. (2021). The other side of deplatforming: Right-wing telegram in the wake of Trump's twitter ouster. In *International conference on digital transformation and global society* (pp. 417–428). Springer.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for regression-discontinuity designs. *The Stata Journal: Promoting Communications on Statistics and Stata*, *17*(2), 372–404. https://doi.org/10.1177/1536867X1701700208

Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., & Gilbert, E. (2017). You can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 1–22. https://doi.org/10.1145/3134666

Chouliaraki, L., & Banet-Weiser, S. (2021). Introduction to special issue: The logic of victimhood.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6).

Danbold, F., Onyeador, I. N., & Unzueta, M. M. (2022). Dominant groups support digressive victimhood claims to counter accusations of discrimination. *Journal of Experimental Social Psychology*, *98*, 104233. https://doi.org/10.1016/j.jesp.2021.104233

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Greussing, E., & Boomgaarden, H. G. (2017). Shifting the refugee narrative? An automated frame analysis of Europe's 2015 refugee crisis. *Journal of Ethnic and Migration Studies*, *43*(11), 1749–1774. https://doi.org/10.1080/1369183X.2017.1282813

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Guhl, J., & Davey, J. (2020). A safe space to hate: White supremacist mobilisation on telegram. *Institute for Strategic Dialogue*, *26*.

Hausman, C., & Rapson, D. S. (2017). *Regression discontinuity in time: Considerations for empirical applications. Technical report*. National Bureau of Economic Research.

Jhaver, S., Boylston, C., Yang, D., & Bruckman, A. (2021). Evaluating the effectiveness of deplatforming as a moderation strategy on twitter. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–30. https://doi.org/10.1145/3479525

Kahalon, R., Shnabel, N., Halabi, S., & SimanTov-Nachlieli, I. (2019). Power matters: The role of power and morality needs in competitive victimhood among advantaged and disadvantaged groups. *British Journal of Social Psychology*, *58*(2), 452–472. https://doi.org/10.1111/bjso.12276

Korolczuk, E. (2020). The fight against 'gender' and 'LGBT ideology': New developments in Poland. *European Journal of Politics and Gender*, *3*(1), 165–167. https://doi.org/10.1332/251510819X15744244471843

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized Bert pretraining approach. arXiv preprint arXiv:1907.11692.

Marcks, H., & Pawelz, J. (2022). From myths of victimhood to fantasies of violence: How far-right narratives of imperilment work. *Terrorism and Political Violence*, *34*(7), 1415–1432.

Mudde, C. (2007). *Populist radical right parties in Europe*. Cambridge University Press.

Noor, M., Vollhardt, J. R., Mari, S., & Nadler, A. (2017). The social psychology of collective victimhood. *European Journal of Social Psychology*, *47*(2), 121–134. https://doi.org/10.1002/ejsp.2300

Oaten, A. (2014). The cult of the victim: An analysis of the collective identity of the English defence league. *Patterns of Prejudice*, *48*(4), 331–349. https://doi.org/10.1080/0031322X.2014.950454

Pope, P. J. (2017). Constructing the refugee as villain: An analysis of Syrian refugee policy narratives used to justify a state of exception. *World Affairs*, *180*(3), 53–71. https://doi.org/10.1177/0043820018757542

Riesa, J., & Giuliani, I. (2013). Cld2: Compact language detector 2.

Rogers, R. (2020). Deplatforming: Following extreme internet celebrities to telegram and alternative social media. *European Journal of Communication*, *35*(3), 213–229. https://doi.org/10.1177/0267323120922066

Rotella, K. N., Richeson, J. A., Chiao, J. Y., & Bean, M. G. (2013). Blinding trust: The effect of perceived group victimhood on intergroup trust. *Personality and Social Psychology Bulletin*, *39*(1), 115–127. https://doi.org/10.1177/0146167212466114

Schori-Eyal, N., Klar, Y., & Ben-Ami, Y. (2017). Perpetual ingroup victimhood as a distorted lens: Effects on attribution and categorization. *European Journal of Social Psychology*, *47*(2), 180–194. https://doi.org/10.1002/ejsp.2250

Sengul, K. (2021). 'It's ok to be white': The discursive construction of victimhood, 'anti-white racism' and calculated ambivalence in Australia. *Critical Discourse Studies*, 1–17.

Simon, B., & Klandermans, B. (2001). Politicized collective identity: A social psychological analysis. *American Psychologist*, *56*(4), 319. https://doi.org/10.1037/0003-066X.56.4.319

Strickland, P. (2020). Antifa and America's revamped Red Scare — aljazeera.com https://www.aljazeera.com/features/2020/9/29/antifa-hysteria-americas-revamped-red-scare [Accessed 17-05-2024].

Sullivan, D. (2012). Competitive victimhood as a response to accusations of ingroup harm doing. *Journal of Personality and Social Psychology*, *102*(4), 778–795. https://doi.org/10.1037/a0026573

Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational Identity: A Reader*, *56*, 65.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Urman, A., & Katz, S. (2020). What they do in the shadows: Examining the far-right networks on telegram. *Information, Communication & Society*, 1–20.

Vollhardt, J. R. (2012). *Collective victimization*. Oxford University Press.

Wilkins, C. L., & Kaiser, C. R. (2014). Racial progress as threat to the status hierarchy: Implications for perceptions of anti-white bias. *Psychological Science*, *25*(2), 439–446. https://doi.org/10.1177/0956797613508412

Young, I. F., & Sullivan, D. (2016). Competitive victimhood: A review of the theoretical and empirical literature. *Current Opinion in Psychology*, *11*, 30–34. https://doi.org/10.1016/j.copsyc.2016.04.004

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320.