Personal Narratives in Online Political Discussions

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science
in
Computational Linguistics

by

C S Ramakrishna Tejasvi
2019114005
tejasvi.chebrolu@research.iiit.ac.in

Advisor: Prof. Ponnurangam Kumaraguru

International Institute of Information Technology Hyderabad
500 032, India

Mar 2025

International Institute of Information Technology Hyderabad
Hyderabad, India

CERTIFICATE

This is to certify that work presented in this thesis proposal titled ***Personal Narratives in Online Political Discussions*** by *C S Ramakrishna Tejasvi* has been carried out under my supervision and is not submitted elsewhere for a degree.

_____ _____

Date Advisor: Prof. Ponnurangam Kumaraguru

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन ।
मा कर्मफलहेतुर्भूर्मा ते सङ्गोऽस्त्वकर्मणि ॥

You have the right to perform your prescribed duty, but you are not entitled to the fruits of your action. Never consider yourself the cause of the results of your actions, and never be attached to inaction.

To Amma, Nanna, and Niranjan

Acknowledgments

I have been incredibly fortunate in my journey at IIIT Hyderabad, surrounded by brilliant and compassionate mentors who played a crucial role in my growth. This thesis and the accompanying papers would not have been possible without their invaluable guidance and support.

First and foremost, I extend my heartfelt gratitude to my advisor, Prof. Ponnurangam Kumaraguru, for his unwavering emotional support and patience during times of hardship. His mentorship opened doors to invaluable opportunities and technical insights, enabling me to navigate a highly niche area of research. He not only shaped me into a better researcher but also guided me in becoming a thoughtful and responsible global citizen.

I am also deeply thankful to Dr. Ashwin Rajadesingan, my mentor at UT Austin, for his dedication and guidance. He provided me with the remarkable opportunity to immerse myself in academia at UT Austin, guiding me through countless drafts and rejections with patience and care. His leadership by example taught me not only how to conduct rigorous research but also how to excel professionally.

My sincere thanks go to my parents, brother, friends, and colleagues, whose belief in me never wavered, even when I doubted myself. A special mention goes to Harsha and Rohan, my co-authors in my first academic endeavour. Their collaboration and hard work in running experiments and building a project we could all be proud of were invaluable. Similarly, I would also like to thank Utkarsh Mujumdar and Chris Hickey for their early assistance on this project.

Lastly, I would like to express my gratitude to my friends who stood by me through thick and thin — Dhruvee, Tathagato, and Murali. They spent countless hours with me in the lab, offering a listening ear and moral support, even if it was to hear me rant. I am sure there are many others who deserve acknowledgment. To those whose names I haven't mentioned, please know that your contributions were no less significant, and I am truly grateful.

I also would like to acknowledge the generous support I received from the Center for Media Engagement, The Moody College of Communication at The University of Texas at Austin, and IHub-Data at IIIT Hyderabad.

Abstract

Engaging diverse participants in meaningful online political discussions remains a crucial challenge for contemporary democratic societies. Two central barriers to vibrant political discourse online are cross-partisan communication divides, where opposing ideological groups struggle to interact constructively, and the limited participation from politically disinclined individuals (PDIs), who often perceive themselves as lacking sufficient political knowledge or fear negative social consequences. Although existing literature presents mixed findings regarding the efficacy of personal narratives in political communication, this thesis investigates the potential of personal narratives—defined as short, first-person accounts emphasizing personal experiences—to bridge these divides and increase participation.

Experimental research has shown that personal narratives can foster empathy, respect, and trust across ideological boundaries by enabling identification and emotional engagement. Conversely, some studies suggest that personal narratives in social media environments can inadvertently reinforce in-group solidarity, strengthen partisan identities, and amplify polarization due to platform-specific affordances, such as quantifiable feedback mechanisms. To resolve these discrepancies and comprehensively evaluate the role of personal narratives in real-world settings, this research conducts a large-scale computational analysis of political discussions on Reddit.

Utilizing a BERT-based text classifier to identify personal narratives, alongside computational methods to infer user ideology and participation patterns, the analysis reveals encouraging results. First, politically disinclined individuals are significantly more likely to contribute personal narratives in political communities compared to their politically active counterparts, and such narratives notably increase their continued engagement over time. Additionally, personal narratives posted by PDIs consistently receive more positive feedback from community members relative to their non-narrative contributions, suggesting their unique value in enhancing inclusive discourse.

Furthermore, when examining cross-partisan interactions, personal narratives exhibit varied impacts: notably, right-leaning individuals using personal narratives receive more favourable feedback from left-leaning communities, indicative of reduced partisan animosity and increased openness toward opposing viewpoints. However, this positive effect is asymmetrical, as similar outcomes are not consistently observed for left-leaning individuals. Additionally, while personal narratives alone do not substantially increase the probability of replies across ideological di-

vides, they do significantly mitigate hostility levels, fostering a more respectful cross-partisan environment.

This thesis argues that personal narratives not only empower politically disinclined individuals by lowering barriers to participation but also partially alleviate polarization by fostering empathy and respect in cross-partisan settings. These findings highlight personal narratives as valuable tools for enhancing democratic engagement. Future research could further explore the conditions under which personal narratives exert their strongest positive effects, as well as examine methods to systematically encourage the use of inclusive and empathetic storytelling within contentious political environments online.

Contents

List of Figures

List of Tables

List of Related Publications

[P1] Tejasvi Chebrolu, Ponnurangam Kumaraguru, and Ashwin Rajadesingan, **"Personal Narratives Empower Politically Disinclined Individuals to Engage in Political Discussions"**, in proceedings of the *17th ACM Web Science Conference 2025 (Websci '25), May 20–24, 2025, New Brunswick, NJ, USA. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3717867.3717899*, 2025.

Chapter *1*

Introduction

Political discussions between ordinary citizens are central to a democratic society. Such discussions result in a wide range of normatively desired outcomes such as increased political knowledge [1], informed opinion formation [2] and political tolerance [3]. Political discussions with people holding opposing views is also crucial for sustaining a vibrant democracy [4]. It is through such discussions that individuals gain alternate perspectives [5], build political tolerance [3], reduce partisan animosity [6] and form informed opinions [7]. Yet, most individuals do not engage in political discussions, especially online. For example, in surveys with American participants, conducted by the Pew Research Center, one in three respondents never discussed politics with their friends [8], and 70% of Americans rarely or never post about political content on social media [9]. Worryingly, Americans come out of these discussions and feel they have less in common with whose they disagree with [10].

However, not everyone refrains from engaging in political discussions. Krupnikov and Ryan [11] find that some individuals are "deeply involved" in politics and are often very vocal in political discussions. These individuals are a hyperpartisan, affectively polarized minority group for whom politics is a central part of their lives. These individuals typically engage in hostile interactions with both outparty supporters and those within the party who they perceive as not being sufficiently extreme [12]. Although a small minority, these deeply involved individuals dominate online political discussions, making them the most visible to others [11]. Journalists, who often disproportionately focus their reporting on extreme partisans, further amplify this dynamic, creating an exaggerated perception of political polarization [11, 13]. Such polarized, active individuals often create environments hostile to moderate or politically disinclined individuals (PDIs), who prefer to remain disengaged due to fears of conflict, misinformation, or a perceived lack of political knowledge [14].

How best to address this issue of the deeply involved dominating political discourse online? This insularity can create echo chambers that deepen political polarization and diminish openness to opposing viewpoints [15]. It also marginalizes less politically active individuals, further narrowing the diversity of voices in discourse and ultimately weakening democratic engagement. The most straightforward solution is to encourage others to engage with politics online [12].

The problem with this solution is that, while the majority of other users are less polarized and more moderate, they are also much less vocal about politics [11]. How, then, can we facilitate participation from this politically reclusive group? We argue that encouraging the use of personal narratives in political discussions may facilitate greater participation among those who are disinclined towards politics. Recent research suggests that apart from concerns about maintaining social ties, a lack of political knowledge or understanding was cited as a major reason for avoiding political discussions [14]. Personal narratives may address this concern by allowing individuals to rely on their own personal experiences when engaging with politics. Further, personal narratives and storytelling can foster more inclusive political communication by presenting situated knowledge and enabling disadvantaged groups who may lack access to formal argumentative skills to effectively share their perspectives in the public sphere [16]. Similarly, Polletta, in her book, points to how storytelling can be equalizing in deliberative settings where participants have different knowledge and argumentation skills, "since everyone has his or her own story" [17].

To evaluate how personal narratives can help individuals engage in political discussions, we conducted a large-scale computational analysis of public comments in political subreddits on Reddit from 2020 to 2021. We identified two groups of users that we are most interested in: users who have rarely engaged and users who have not at all engaged in discussions in political subreddits in the previous 12 months. We formally defined these two groups as politically disinclined individuals (PDIs). We compared these users with users who are most active in these political communities. Then, we fine-tuned a transformer-based classifier to identify personal narratives in comments posted in political subreddits. Across multiple analyses, we found strong evidence linking the use of personal narratives and participation in political discussions.

First, we found that PDIs were more likely to use personal narratives in their comments in political subreddits than more politically active users. Second, comments containing personal narratives were slightly more likely to receive responses from PDIs than other comments in political subreddits. Our analysis also revealed that PDIs who employed personal narratives in a given month were more likely to continue participating in political discussions in the subsequent month compared to those who did not use a personal narrative. These results suggest that personal narratives not only attract the politically disinclined but also encourage their sustained involvement in political discourse over time.

Finally, personal narratives appear to be positively viewed by the political community members themselves. Reddit's karma scoring system allows members to upvote or downvote comments. Thus, a higher score implies that the comment is positively viewed by the community. We find that personal narratives, on average, garner higher scores than other comments by PDIs. Although there appears to be a perception gap, that is, comments by more politically active users, on average, score higher than PDIs, personal narratives appear to significantly

reduce this difference. Overall, these results present an encouraging picture of how personal narratives may facilitate participation among the politically disinclined.

Personal narratives are also a promising approach in facilitating better political discussions across ideological differences. Communication scholars suggest two routes to narrative persuasion: transportation and identification. Personal narratives can "transport" readers into the story, allowing them to become absorbed in the events and experience them vicariously, which makes them more receptive to alternate views [18]. Alternatively, personal narratives may encourage readers to identify with the characters, fostering a perspective-taking experience where they can emotionally connect and empathize with the characters' experiences, which can also make them more receptive to the characters' viewpoints [19]. Applied to cross-partisan political contexts, recent largely experimental research suggests that using personal narratives can increase respect [20], build trust [21] and reduce prejudice on deeply polarized topics [22].

However, it is unclear if the personal stories told in real-world cross-partisan contexts actually facilitate better discussions across differences. These stories may not always provide the level of immersion required to be persuasive [18]. Strong partisan identities may prevent identification with the characters in stories told by opposing partisans reducing the impact of personal narratives [23]. Crucially, personal narratives are often used to construct and negotiate identities, shifting focus from individual experiences to collective ones, reinforcing group identities [24]. While this effect may mobilize others to participate and strengthen social movements [25], it may also complicate discussions in cross-partisan settings, especially at time when strong inter-group dynamics are already present [26]. Additionally, the affordances of social media platforms, such as the ability to quantify social support through metrics like *likes* and *shares*, can amplify polarization by creating a feedback loop [27], where content that elicits strong, divisive reactions receives more visibility and engagement [28]. As a result, users may be incentivized to share polarizing stories that align with their partisan group's beliefs, further entrenching ideological divides and reducing the chance of cross-partisan understanding. As [29] conclude in their analyses of the use of stories in the 2020 US elections, "stories may have political impact less by persuading than by reminding people which side they are on."

To answer this, through a large scale computational analyses, we evaluate the prevalence and impact of personal narratives in naturally occurring cross-partisan discussions on Reddit. Specifically, we evaluated both individual replies and community feedback to personal narratives in cross-partisan contexts. First, we introduce and validate an approach to identifying the political ideology of Reddit users and subreddits. Second, we used the classifier trained earlier to perform various regressions that suggest that personal narratives at least partially fulfil the positive potential indicated by experimental research. Right-leaning users receive more favourable responses from left-leaning audiences when they employ personal narratives. Interestingly, this effect does not hold for left-leaning users. Additionally, both left- and right-leaning users tend to use personal narratives more frequently with those who share their political views than with

those from opposing sides. Furthermore, using personal narratives do not seem to increase the likelihood of eliciting a response, whether positive or hostile, from those with opposing views.

Table 1.1: Examples of personal narrative comments in our dataset, capturing individual experiences and perspectives on work, climate, healthcare costs, political identity, and gender identity. Unlike fact-based statements, these comments stem from personal lived experiences and can be shared by individuals regardless of their direct engagement in political discourse.

Example comments
I'm a cdl-a truck driver in the USA. I work 60 hours over 5 days. Teamster contract ensures that Iḿ well paid but we are paid hourly. A reduction in hours would cause hardship and with the shortage of drivers, a shortage of goods
I'm in Alabama and oh my god it was so humid yesterday. I was so unproductive from how bad it was
Buf you personally... How much is that cost? Cuz I got $500 a month coming out of my pay, then deductibles, then coverage issues, not to mention dental; vision... Oh; ya know what's not afforded in that any kind of mental health coverage...
I faked being a liberal for my years at university. They never suspected a thing.
Due to circumstances not within my control, my (gender nonconforming) partner and i have been at my fathers these past few months. My partner has been forced to present as male because he knows we're in a house full of ultra conservatives ...

Chapter *2*

Literature Review

2.1 Avoiding political discussions

Over the decades, scholars have suggested different explanations for why individuals may avoid engaging in political discussions. Eliasoph, in her book *Avoiding politics*, argues that rather than an aversion to politics, lack of political talk is more because of a fragile public sphere where "intelligence, curiosity, and generosity have evaporated" [30]. Noelle-Neumann's spiral of silence theory suggests that individuals avoid expressing their opinions when they are perceived to be unpopular [31]. Others attribute avoidance to individual predispositions such as conflict avoidance [32] and the Big Five personality traits [33]. Hostility [34], and disagreement [35] have also been identified as factors in political discussion avoidance. More recently, through a series of studies, Carlson and Settle highlight the importance of political knowledge in the decision to not engage in political discussions [14]. Specifically, in one study, they find that one of the most commonly cited reasons for political discussion avoidance was the respondent's lack of accurate political information or knowledge [14].

While a large majority avoid or only rarely engage in political discussions, a small but vocal minority of individuals who are deeply involved in politics dominate political discussions [11]. These individuals are highly politically engaged and are often extreme partisans who are hostile towards those they disagree with [11]. Thus, most political discussions that we observe online are invariably between these deeply involved individuals. A public sphere with the deeply involved largely engaging in discussions is especially problematic. Interactions predominantly between deeply involved individuals are likely to result in more extreme views and more polarization [36]. Further, the visibility of deeply involved individuals engaging in hostile discussions, coupled with an increased media focus on them, often distorts perceptions of polarization and who is the median partisan [11–13]. Finally, as Berelson et al. [37] in their foundational work on opinion formation note, a mass democracy cannot function if all individuals are deeply involved. A wider distribution of individuals based on political involvement may facilitate compromise, avert extreme partisanship and provide room for consensus and stability in democratic decision-making [37]. Thus, it is crucial that political discussions include individuals who do not usually

engage with politics. We argue that personal narratives can facilitate a wider engagement among the politically disinclined.

2.2 Personal narratives in political discussions

Personal narratives are usually first-person accounts that recount individual experiences. Engaging with personal narratives requires minimal prior knowledge or training, allowing individuals from diverse backgrounds to participate meaningfully in political discourse [16]. Personal narratives also make political discussions more accessible and create a safe space for dialogue, encouraging participation even among those hesitant to express their views publicly [38]. Moreover, personal narratives help situate individual stories within the context of larger political systems, enabling individuals to see how their experiences relate to broader societal issues [39]. Political apathy, another key reason for political inactivity, can also be countered through personal narratives as they create connections across ideological differences, encourage engagement via relatable storytelling, and highlight the impact of individual experiences within broader political and social systems [40]. Finally, personal narratives help build inclusive spaces by emphasizing individual experiences as valid forms of knowledge, which is crucial for engaging PDIs who may feel excluded from traditional political discourses [41].

Personal narratives may positively impact the substance of the political discussions as well. These narratives often humanize the narrator, either by showcasing their positive qualities or revealing their vulnerabilities, creating a sense of connection with readers [42]. This connection is further strengthened by the empathy personal narratives can evoke, making readers more open to considering opposing viewpoints [43]. Personal narratives can also inspire political action by turning personal stories into powerful tools for raising awareness and building solidarity [44]. Similarly, employing these narratives in political discourse also grants users tools for persuasion in political contexts, as these narratives have the ability to influence attitudes and beliefs, unlike data-heavy arguments, as stories often appeal directly to emotions [18]. Somewhat counterintuitively, personal narratives also enhance perceptions of rationality in political discussions. When people use personal narratives to express their political views, they are often seen as more rational, garnering greater respect even from opposing partisan groups [20]. Finally, personal narratives can amplify the voices of underrepresented groups, challenge mainstream narratives, and inspire hope while advocating for change [45]. These studies suggest that personal narratives may not only encourage PDIs to participate in political discussions but may also substantively improve the quality of the discussions.

2.3 Personal narratives in cross-partisan discussions

Personal narratives are widely regarded as one of the most common and inclusive forms of communication, as they require minimal formal training or practice [46]. Social media platforms offer numerous affordances for self-presentation and encourage individuals to share their personal experiences online [47]. However, these platforms are not simply neutral spaces that host stories; instead, the stories are themselves "restricted, enhanced, and adorned by the affordances of these platforms" [48]. Research indicates that stories told on social media sites are often 'small stories' which are typically short and episodic, describing everyday, mundane personal experiences as opposed to long-form life narratives told in face-to-face interview settings [49]. Social media affordances influence not just the structure of the stories but also the storytelling practices. Research suggests that the participatory culture inherent in social media platforms enhances reflexivity in story discussions, allowing users to critically engage with narratives and reflect on their own perspectives [50]. Georgakopoulou et al. [51] argue that three platform affordances—subreddit autonomy, a point system, and user anonymity—shape the narratives shared there. Using r/The_Donald as a case study, they find that subreddit autonomy fosters a "reflexive circulation" of stories, continually produced and modified within the community. This anonymity, alongside the point-based visibility system, encourages toxic, anti-deliberative storytelling, amplified by the feedback loop of attention and visibility these features create. Considering the social and participatory nature of storytelling on social media, studies like ours, which evaluate narratives within the context of their telling, offer valuable insights that complement experimental research in this domain.

As mentioned earlier, recent experimental research suggests that personal narratives may help bridge partisan divides. Kubin et al. [20] showed that when individuals anchor their political views in personal narratives, they are perceived as being more rational, which in turn fosters more respect among their ideological opponents. Similarly, Hagmann et al. [21] found that expressing opposing viewpoint through self-disclosing and vulnerable personal narratives enhance trust and willingness to collaborate, even more so than presenting factual data. Wojcieszak et al. [52] also found that personal narratives outperform numerical evidence when receiving counter-attitudinal messages on controversial sociopolitical topics especially when primed to view the message with an empathetic perspective. Further, qualitative research also indicates that storytelling in group settings can invoke perspective-taking that can help group members understand viewpoints different from their own, allowing for more considered deliberation [53]. These studies suggest the possibility that personal narratives can break down partisan barriers that hinder quality cross-partisan engagement.

While personal stories may unite, they may mislead and reinforce these divides [54]. Prochaska et al. [55] argue that personal stories on social media played a crucial role in supporting participatory disinformation by establishing credibility to claims of alleged election fraud during the 2020 election. These stories tapped into existing "deep stories" [56], that is, large underlying

stories, based more on feelings than fact, through which individuals make sense of their everyday experiences. In this case, these personal stories, combined with a deep story of widespread fraud and conservative oppression, mobilize support and reinforce a manufactured, misleading reality. Poletta et al. [29] also point to how personal stories, when folded into existing deep stories of conservative grievances manufactured by conservative media and amplified through social media sharing, helped to construct a form of collective identity that reinforces partisan divides. This collective identity, they argue, might explain why the truthfulness of the stories did not matter to those sharing them.

Broad research across disciplines also support the argument that personal narratives are especially adept at constructing a collective identity, which may be at odds with our goal of engaging constructively with the other side. Social movement scholarship indicates that personal narratives help mobilize and sustain participation in collective action by constructing common interests and strengthening collective identity [25, 57]. Collective identity often reinforces an "us versus them" dynamic, delineating group boundaries by fostering group solidarity while emphasizing opposition to outsiders [58]. The impact of personal narratives on engendering group identity may be more pronounced in social media platforms. Social media scholars have identified multiple affordances such as quantification [59], amplification [60] and affordances for performance [61] and for discourse [61] that strengthen collective identity. Such strong group identities may be especially problematic at a time of affective polarization, where Republicans and Democrats appear to dislike each other largely on the basis of their group partisan identities [62].

Thus, the role of personal narratives in cross-partisan discussions may be significantly influenced by the context, participatory culture and the specific affordances of the platform in which they are shared. However, there has been limited research examining personal stories within naturally occurring cross-partisan discussions on online platforms where these factors may play a role. Recently, Kubin et al. [20] found that personal experiences, rather than facts, are more effective at fostering mutual respect between opposing partisans. One of their studies used social media data. Examining comments left on abortion-related Youtube videos, they found that videos emphasizing personal experience received more positive and affirmative language than those on emphasizing facts. However, these findings do not shed light on cross-partisan reactions to personal narratives. [63], analyzing discussions on an online forum about the future of Lower Manhattan after the 9/11 attacks, found that individuals use storytelling to express a more indirect form of disagreement by reformulating others' points to advance an alternate position without seeming to disagree. However, the context of these discussions were around a fairly unifying topic of rebuilding after the terrorist attacks. The authors speculate that "in a polarized setting, deliberators would probably be suspicious of the authenticity of personal stories altogether" and remain sceptical about their capabilities in such scenarios. In this the-

sis, we empirically assess the use and impact of personal stories in precisely these polarized contexts, focusing on cross-partisan political discussions on Reddit.

2.4 Identifying personal narratives

Research on identifying and extracting personal narratives from textual data has evolved over the years. Gordon and Swanson laid the groundwork by creating a standard corpus of personal narratives from blog posts and using statistical models to classify them [64]. Subsequent research explored various statistical methods for narrative identification, focusing on different linguistic characteristics. Yao and Huang examined temporal characteristics [65], while Ceran et al. investigated the density of part-of-speech tags and named entities [66]. Researchers then expanded their approaches to incorporate semantic information. Eisenberg and Finlayson utilized verb and character features [67], and Dirkson et al. leveraged psycholinguistic features to improve narrative identification [68]. Recent studies have demonstrated the potential of transformer-based models in personal narrative identification. Ganti et al. conducted a study where they documented the performance of various transformer-based models [69], while Antoniak et al. successfully fine-tuned these models to identify narratives at both document and span levels across different domains [70]. Falk and Lapesa further validated the robustness of transformer-based models in identifying narratives in argumentation settings [71]. Generative models, including large language models, have also been used to identify personal narratives [69], but they performed worse than fine-tuning a transformer-based model.

Chapter *3*

Computational Tools and Data

3.1 Data

We analyzed public Reddit comments from 2015-2021 using the Pushshift data archive [72]. We selected political communities from Rajadesingan et al.'s extensive dataset, which comprises over 600 political subreddits identified through a large-scale analysis of political discourse on Reddit [73]. From this list, we manually identified and filtered out political subreddits that were not US-based or were fictitious, resulting in 524 subreddits. In total, our dataset comprised 26,285,619 users who participated in these political communities. To ensure data quality, we excluded moderator and bot accounts. The dataset and code is available on GitHub.[1]

3.2 Computational Tools

3.2.1 Estimating user and subreddit ideology

Since users and subreddits do not explicitly indicate their political ideology on Reddit, we need to infer this from their participation in the platform. Our approach to estimating user and subreddit ideology, similar to [74,75], relies on the existence of a latent space of political ideology based on the assumption that connections between users and entities are driven by political homophily. Concretely, we expect that users more often connect to other users and communities that are ideologically similar and less often connect to entities that are ideologically dissimilar. This homophilic behaviour has been widely observed in major social media platforms such as Reddit [75, 76], Facebook [77], and Twitter [74]. On Reddit, as per the political homophily expectation, liberal users typically comment and engage in liberal political communities and likewise, conservative users [75]. Further, users' comments receive community feedback in the form of up- or down-votes by community members. Since all Reddit comments start with a score of 1, comments that score more than 1 can interpreted as being positively viewed by the

[0] This chapter is a part of the accepted paper.

[1] `https://github.com/Ashwin-R/WebScience25-Personal-Narratives`

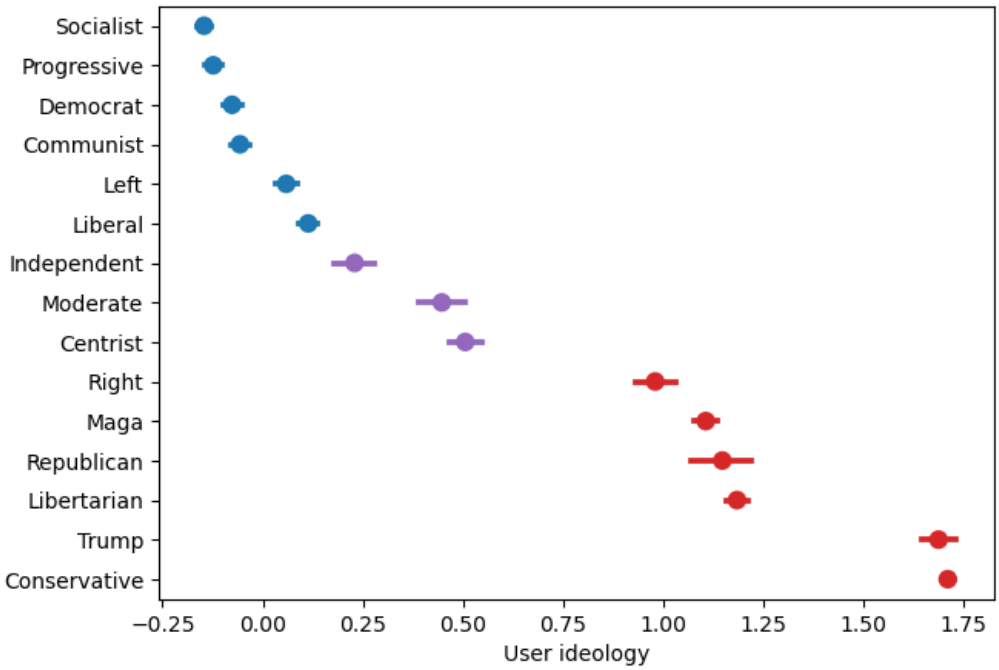
community. Thus, we expect that liberal users engage in liberal communities more often by posting more comments with positive feedback, and likewise for conservative communities.

More formally, we can represent user engagement in political communities in the form of a matrix A where $A_{i,j}$ is the number of comments with positive feedback posted by user i in political subreddit j. In each row, we represent a user in a multidimensional space defined by the number of comments with positive feedback posted in different political subreddits. Then, to identify the latent political ideology of the users, we use correspondence analysis [78]. Correspondence analysis is a dimensionality reduction technique that reduces the original data into a single axis that preserves the variance present in the data. So here, given our assumption of homophily and that we represent users based only on their participation in political subreddits, we assume that the resulting projections of the users are on a latent political ideology axis such that more liberal users lie closer to each other and further away from more conservative users who lie closer to each other. The results produced by correspondence analysis are comparable to Bayesian ideal point estimations using more computationally intensive Markov-chain Monte-Carlo methods on social media data [74]. We note that our approach is conceptually similar to [75]'s approach, which relies on a binary indicator of whether the user engaged in the subreddit as a way to represent users' behaviour. Instead, we take advantage of the additional social signals provided by the community on the users' comments to quantify only the number of comments with positive feedback.

To estimate user ideology, we again used the Reddit Pushshift dataset [72] to obtain comments on known political subreddits between 2015 and 2021 (as described earlier in the Data Collection section). We limit our ideology estimation to only users (n=1,100,475) who have made at least 10 comments with score more than 1 and participated in at least 2 of these communities. Given the size of the dataset, we follow an approach similar to [74]'s multi-stage correspondence analysis approach on a large Twitter dataset. First, we limit the correspondence analysis to users (n=123,321) who have comments with score more than 1 in at least 10 communities to identify the ideological latent space. We expect that these highly politically active users will be especially homophilic in their engagements with political subreddits, allowing for robust estimation of the ideological positions of these subreddits. Then, we project all users on to the ideological space identified in the previous step. The coordinates for these users correspond to their ideological positions. We standardize both the subreddit and user values (mean = 0, SD = 1) for easy interpretation.

Similar to [75], we validate the estimated ideology of the users based on "user flairs". On Reddit, users may assign themselves flairs as a way to identify themselves to the community. In political subreddits, certain flairs are indicative of their political ideology. For example, a flair such as "Staunch Conservative" indicates that the user is especially right-leaning, while a flair such as "Socialist" may indicate that the user is especially left-leaning. We identify commonly used terms to describe political ideology and identify users with flairs containing one

Figure 3.1: Ideology validation based on user flairs referring to their political ideology. The x-axis represents the inferred user ideology score, while the y-axis categorizes users based on their self-reported political labels. The colours indicate different ideological groupings, with left-leaning ideologies in blue, centrist ideologies in purple, and right-leaning ideologies in red. Error bars represent the variation within each group.



of those terms (n=49,636 users). We ignore flairs with multiple of these terms (for example, "Libertarian socialism"). Figure 3.1 shows the average estimated ideology of those users having a flair containing these terms. Notice that more conservative terms are located further to the right and more liberal terms are located further to the left, while centrist terms are located near the centre, suggesting that our approach produces reasonable user ideology estimates.

One particular concern with this approach is that we might estimate all users engaging in a certain subreddit as being of the same ideology. To evaluate this issue, we compare the ideological estimates of users participating in r/AskALiberal, a subreddit where users use both liberal and conservative flairs. Figure 3.2 shows the average ideological estimates of users using different flairs. Again, we find that flairs expressing more liberal views are further to the left, and flairs expressing more conservative views are further to the right.

We label users with an ideology estimate lower than 0 as liberal and users with an ideology estimate of more than 0.4 as conservative. From Figure 3.1, if we consider the users represented by the red points as right-leaning and the ones in blue as left-leaning, this threshold results in an accuracy of 81.27%. In total, we label 738,894 users as left-leaning and 275,284 as right-leaning users. We use the same threshold to identify left- and right-leaning subreddits, resulting in

Figure 3.2: Ideology validation for users in r/AskALiberal subreddit. Left-leaning identities are clustered together, with distinct separation from centrist and right-leaning identities.

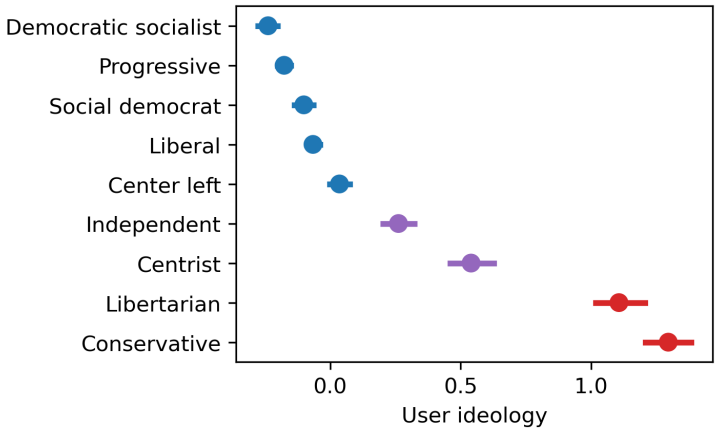
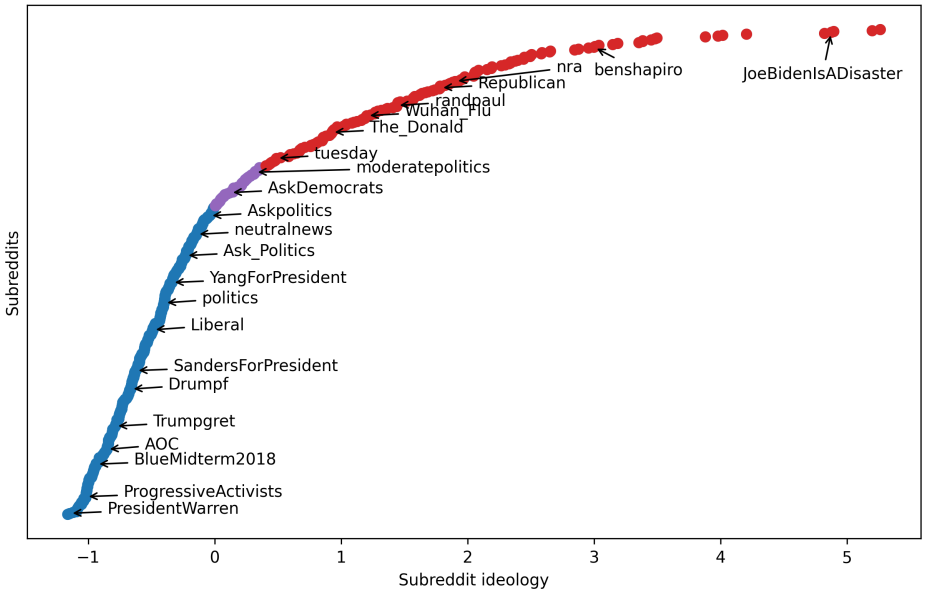


Figure 3.3: Ideology validation for subreddits. Subreddits are distributed along an ideological spectrum, with left-leaning communities clustering on one end and right-leaning ones on the other. A smoother progression is observed in the left-to-center region, while the right-leaning subreddits show more discrete jumps. Some outliers, such as particularly extreme ideological subreddits, stand further apart from the main distribution.



332 liberal and 148 conservative subreddits. In Figure 3.3, we present these subreddits and their corresponding estimated ideology values. We did not annotate all subreddits for space reasons and instead show certain subreddits across ideology values to better understand their ordering. Subreddits to the far-left generally align with more progressive democrats, while subreddits to the far-right are either aimed at derogating Democrats or are subreddits of major

right-wing personalities such as Ben Shapiro. The inter-group discussion subreddits such as r/AskDemocrats and r/moderatepolitcs are present at or near the purple colour and can be considered to be centrist or moderate.

3.2.2 Sampling copartisan and cross-partisan interactions

On Reddit, interactions between users take place in discussion threads under posts on different subreddits. We consider an interaction to be a user replying to another user. Among all interactions in all political subreddits, we randomly sample pairs such that we have 100,000 instances each of a right-leaning user replying to a left-leaning user (cross-partisan), a left-leaning user replying to a right-leaning user (cross-partisan), a left-leaning user replying to a left-leaning user (copartisan) and left-leaning user replying to a left-leaning user (copartisan). We evaluate our hypotheses and research question using this 400,000 interaction dataset.

3.2.3 Hostile cross-partisan interactions

We evaluate hostility by examining the toxicity of the cross-partisan interaction. We used the Perspective toxicity classifier [79] to identify toxic comments. The Perspective toxicity classifier defines a toxic comment as a "rude, disrespectful, or unreasonable comment that is likely to make people leave a discussion" [79]. Oh et al. [80]'s recent work auditing the Perspective toxicity classifier suggests that the classifier's "understanding of 'toxic language' is biased towards incivility than intolerance," drawing on Rossini et al. [81]'s conceptual work the two concepts. Incivility, in her work, primarily centres around the use of a rude or disrespectful tone in discourse, whereas intolerance pertains to discourse that poses a threat to democratic pluralism and core values. The Perspective toxicity classifier has been previously validated and found to produce accurate classifications on Reddit political comments [82]. If the classifier output probability is greater than 0.6, we label the comment as toxic, else we label the comment as not toxic [2]. We label a cross-partisan comment-reply discussion pair to be toxic if the reply is toxic, as we do not know who the parent comment replied to. In total, we found 16.28% of cross-partisan interactions to be toxic.

3.2.4 Identifying personal narratives

Given the size of the sampled dataset (400,000 interactions), we developed an automated approach to identifying personal narratives on Reddit. We define personal narratives minimally as simply accounts of personal experiences. We do not follow a strict Labovian [83] structural paradigm that typically view stories as temporal progressions of events about a particular place and time which have defined beginning, middle and end. Instead, we draw on narrative analysis research on social media, which suggests that stories told on these platforms are 'small stories'

[2]Varying the thresholds do not change the results substantively.

about ordinary and mundane events from everyday life which do not neatly fit into the canonical Labovian form [84]. Often, the structure of these stories lacks a clear beginning or end and may sometimes even be incomplete, inviting others to fill in [85].

We build on prior computational work on identifying personal narratives in online discussions [70, 71]. Both Falk et al. and Antoniak et al. [70, 71] finetune BERT-based classifiers on discussion data to identify personal narratives. We performed pilot testing with both classifiers. We found Falk et al. [71]'s classifier better suited for our work as Anotoniak et al. [70] adhere closely to the Labovian storytelling paradigm, which misses many of the personal narratives that we hope to identify.

We ran Falk et al. [71]'s best-performing finetuned DistilBERT [86] on cross-partisan comments and found that the classifier often identified narratives that were not personal experiences or described events that narrators saw, heard, read or listened on mass media. Thus, instead of directly using this classifier, we finetuned it further to align with our task. To obtain training data for the additional finetuning, we ran this classifier on a sample of 20,000 cross-partisan comments that contain first-person singular and plural personal pronouns such as I, we to identify *personal* stories. We also chose to only sample comments of at least 50 characters in length as we did not identify any personal narratives below that threshold in a small pilot study. From this sample, we randomly selected 1000 samples that the classifier identified as personal narratives and 1000 that it did not identify as personal narratives. This dataset was then manually labelled by two trained annotators based on our operationalization of personal narratives. Initially, two annotators labelled 1,000 samples from the first model, achieving an inter-rater reliability score of 0.79 (Cohen's κ). After reaching this agreement level, one of the annotators labelled the remaining data, resulting in a final dataset of 2,000 samples for fine-tuning. This approach enables the construction of a reasonably balanced corpus without the costly task of labelling thousands of samples, especially given the infrequent occurrence of personal narratives [64]. This approach is similar to pseudo labeling approaches devised to overcome the dependency on large labelled datasets [87].

We further fine-tuned Falk et al. [71]'s `distilBERT` classifier on this dataset of 2,000 samples, of which 842 were personal narratives. The model was trained using a batch size of 32 over 10 epochs. We employed the AdamW optimizer [88] with a learning rate of $5e-5$. A 5-fold cross-validation approach was used to evaluate the model. The evaluation metrics averaged over the 5 folds are presented in Table 3.1.

Our use for this classifier is to compare the scenarios of cross-partisan interactions using personal narratives against those that do not use personal narratives. Therefore, we evaluate the actual true positive rate—the proportion of instances in which a comment is indeed a personal narrative when classified as such—and the actual true negative rate—the proportion of instances in which a comment is not a personal narrative when the classifier identifies it as not a personal narrative. To do this, we randomly sample 10,000 unseen comments (without

any constraints on minimum length and personal pronoun use) and obtain classifier labels. Then, we sampled 100 personal narratives and 100 non-narratives from this dataset. One of the annotators manually labelled this new dataset. The annotations resulted in a true positive rate of 80% and a true negative rate of 98%. This provides confidence that our classifier can generalize well to new comments. Crucially, these results suggest that differences that we measure in the hypothesis tests ought to be seen as a conservative estimate and that the differences are likely to be higher if we had a higher true positive rate.

In the 400,000 dataset, we find 4.11% of the replies in cross-partisan interactions to contain a personal narrative. We only evaluate the reply in the interaction to ensure that we examine personal narratives in a cross-partisan context, as the parent comment could be replying to a copartisan.

Table 3.1: Mean performance metric per fold for the personal narrative classifier. The classifier achieves a macro average f-1 score of 0.82.

Metric	**Not Personal Narrative**	**Personal Narrative**	**Macro Average**	**Weighted Average**
Precision	0.8425	0.7986	0.8288	0.8302
Recall	0.8716	0.7728	0.8147	0.8325
F1 Score	0.8446	0.8032	**0.8211**	0.8219
Support	232	168	400	400
Accuracy	0.8320			

3.2.5 Identifying Politically disinclined individuals (PDIs)

To identify PDIs, we examined two groups for each month of our analysis: users who rarely engaged in political subreddits in the previous 12 months and users who never engaged in political subreddits in the previous 12 months.

We identified users who rarely engaged in political subreddits as follows - For each month in our analysis period (January-December 2021), we identified users who commented at least once in a political subreddit in that month but were in the bottom 25th percentile (lowest quartile) based on the number of comments they have made in political subreddits over the preceding 12 months. For example, for January 2021, we calculated the total comments made by each user who has commented in political subreddits from January 2020 to December 2020. Then, we identified the users in the lowest quartile as *least politically active users*. Users in the highest quartile were considered to be the *most politically active users*.

This approach, however, excludes an important group of individuals—those who were active on Reddit but not in political subreddits during the previous 12 months. We identified these users as follows - For each month, we identified users who commented at least once in a political

subreddit but did not comment in political subreddits in the preceding 12 months. To avoid selecting users who were new to Reddit, we filtered out users who did not comment in any (not just political) subreddit in the preceding 12 months. We consider these users to be *political newcomers*.

Chapter 4

Personal Narratives Empower Politically Disinclined Individuals to Engage in Political Discussions

4.1 Hypotheses

We evaluate how personal narratives may affect political engagement. As discussed earlier, individuals often feel disenfranchised in political discourse, often due to a perceived lack of sufficient political knowledge to contribute meaningfully [14]. However, personal narratives can help overcome this barrier by reframing topics into relatable personal stories, offering unique perspectives that make the subject matter less intimidating and more accessible [89]. Similarly, the approachability of personal narratives likely engages less politically active users more than other kinds of content. Therefore, we hypothesize the following:

> **H1:** PDIs are more likely to use personal narratives than the most politically active users.
>
> **H2:** Personal narratives are more likely to receive responses than other comments from PDIs.

Personal narratives may also foster a sense of community and belonging, creating bonds that encourage sustained participation [89, 90]. Based on this, we hypothesize that:

> **H3:** PDIs who use personal narratives in political communities are more likely to engage in political communities the following month than those who do not use personal narratives.

Next, we evaluate community perceptions of personal narratives in political discussions. On Reddit, comments can be upvoted or downvoted, with each upvote increasing a comment's score by one point. A higher score indicates a positive reception by the community compared to comments with lower scores. Given the PDIs limited experience and engagement with political communities compared with highly politically active users, we expect that their contributions are likely less valued in political communities. Therefore, we expect that:

H4: Comments by PDIs are likely to score lower than comments by the most politically active users.

Personal narratives often provide new information or viewpoints that are unique to the discussion. This novelty in personal narratives is often valued in user-contributed comment sections. For example, the New York Times explicitly chooses to highlight personal stories in their comments section. [1] Therefore, we hypothesize that:

H5: Personal narrative comments by politically disinclined individuals are more likely to score higher than other comments by PDIs.

Finally, if H4 is true, we evaluate if personal narratives can bridge the perception gap, that is, the difference in positive perceptions of content posted by PDIs and most politically active users. We pose the following research question:

RQ1: Can personal narratives reduce the perception gap between the content posted by PDIs and most politically active users?

4.2 Analysis and Results

4.2.1 H1: Using personal narratives

H1 states that PDIs are more likely to use personal narratives than the most politically active users. To evaluate H1, for each month in our analysis period, January - December 2021, we first sampled 10,000 comments from the least politically active users, 10,000 comments from political newcomers, and 10,000 from the most politically active users. This resulted in 360,000 comments. Next, we ran the political narratives classifier on the 360,000 comments to determine which comments were personal narratives. We then conducted a random-effects logistic regression, modelling whether a comment was classified as a personal narrative or not (dependent variable). The primary independent variable of interest was the type of user (least politically active, most politically active, political newcomer). We included random effects to account for variability across subreddits and months. To account for potential confounding factors, we included the following control variables: whether the subreddit was COVID-related or not[2] and visibility-related measures such as whether the comment was a top-level comment in the discussion thread and cube-root transformed post score[3] (as higher scores likely imply more visibility based on Reddit's ranking algorithms). For very few comments (n=8), we did

[1] `https://archive.nytimes.com/www.nytimes.com/times-insider/2014/04/17/a-comments-path-to-publication/`

[2] To account for the potential impact of the pandemic on political discussions. Given the uniqueness of the pandemic, the dynamics of political discussions may have been different.

[3] We applied a cube-root transformation to account for the skewness in karma score since cube-root is well-defined for positive, negative and zero values unlike log transformations [91].

not have the post scores in the original pushshift dataset. We imputed those post scores by calculating the median (1040) of all post scores in our dataset.

The regression analyses were performed using the *lme*4 R package [92]. The marginal means estimation and the planned contrasts to test the hypotheses were conducted using the *emmeans* R package [93].

Figure 4.1: Percentage use of personal narratives across user activity (95 % CI). PDIs (Least active users and political newcomers) are more likely to use personal narratives than the most politically active users.

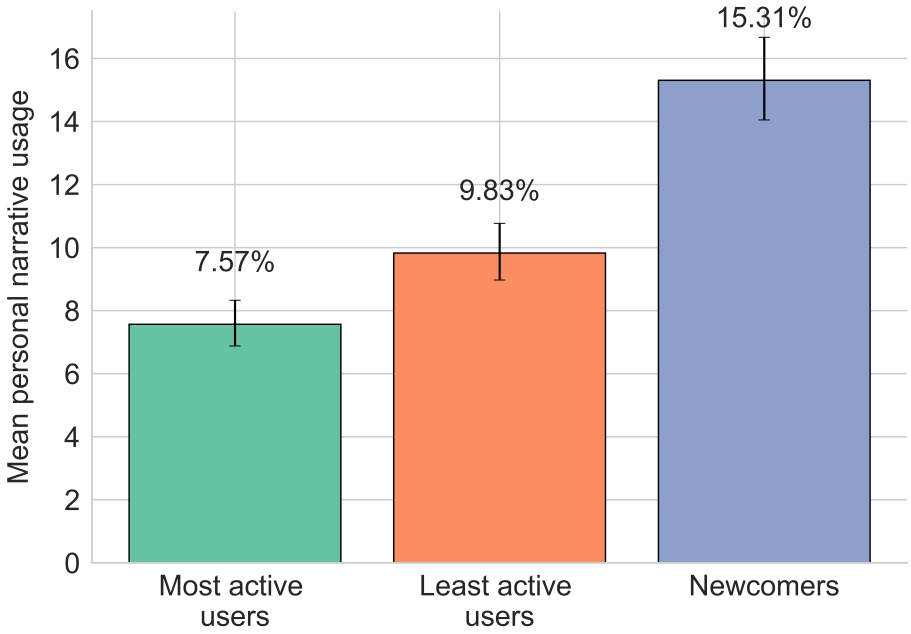


Figure 4.1 shows the estimated marginal mean probability of using a personal narrative for each type of user. On average, political newcomers were most likely to use personal narratives, with a probability of 15.31% (95% CI 14.05–16.67%), followed by the least politically active users at 9.83% (95% CI 8.97–10.77%), and finally the most politically active users at 7.57% (95% CI 6.88–8.33%). The differences between these groups were statistically significant. Specifically, the least politically active users (OR = 1.33, SE = 0.025, z-ratio = 15.464, $p < 0.001$) and political newcomers (OR = 2.20, SE = 0.038, z-ratio = 46.206, $p < 0.001$) were significantly more likely to use personal narratives compared to the most politically active users. Together, these results support H1.

Interestingly, the difference between the least politically active users and political newcomers was also significant (OR = 0.60, SE = 0.009, z-ratio = -32.988, $p < 0.001$), suggesting that political newcomers are more inclined to rely on personal narratives to engage in political

discussions than even the least politically active users. The regression table for this analysis is included in the Appendix (Table 7.1).

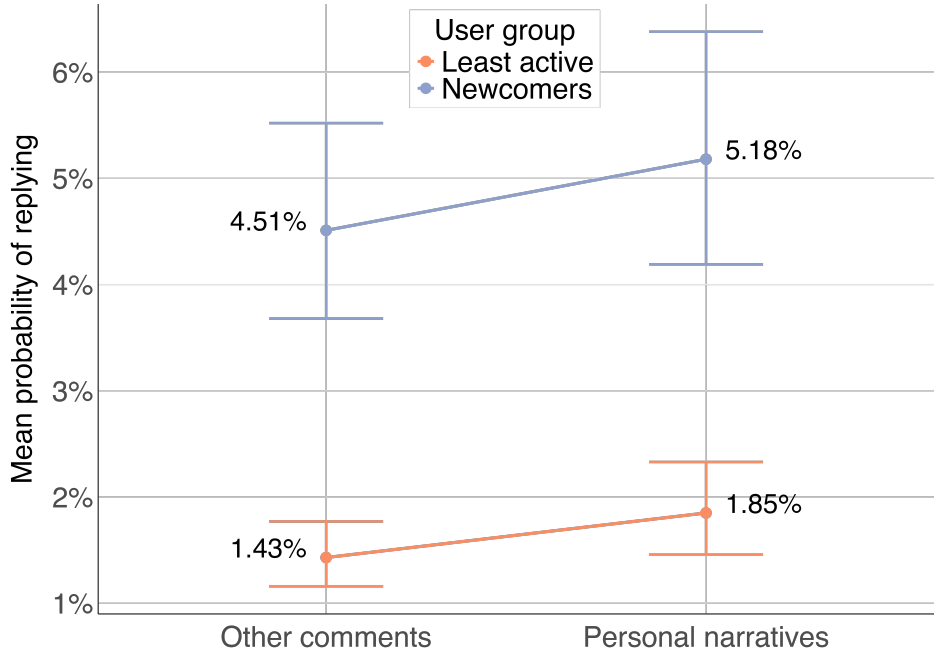
4.2.2 H2: Replying to personal narratives

H2 states that personal narratives are more likely to receive responses from PDIs than other comments. To evaluate H2, given that we are focused on replies, we constructed a dataset by randomly sampling 20,000 comments from each month in 2021, selecting only those with at least one direct reply. This resulted in a dataset of 240,000 comments and 348,279 replies. We ran the personal narrative classifier to identify personal narratives in the 240,000 comment dataset. We identified 13,259 personal narratives (5.5%) in this dataset. For each comment, we identified the number of political newcomers and least politically active users (PDIs) who replied to that comment. We then modelled the propensity of a PDI replying to a comment as a random-effects binomial regression. The number of unique PDIs who reply to a comment was modelled as the number of successes, and the total unique users who reply to the same comment was modelled as the number of Bernoulli trials in a binomial distribution.

The primary independent variables of interest were whether the comment being replied to was a personal narrative or not (personal narrative indicator) and user type who replied to the comment (political newcomer or least politically active user). In addition to the control variables used to evaluate H1, we included the cube-root transformed comment score as the visibility of the comment based on the comment score is likely correlated with replying behaviour. We also included an interaction term between the personal narrative indicator and user type to account for potential differences in the relationships between the replying behaviour of the two user groups and the use of personal narratives. Random effects were included to account for variability across subreddits and months.

Figure 4.2 shows the estimated marginal mean probability of receiving responses from PDIs for comments with and without personal narratives. On average, comments with personal narratives (M = 1.85%, 95% CI 1.46–2.33%) had a slightly higher probability of receiving a reply from the least politically active users than comments without personal narratives (M = 1.43%, 95% CI 1.16–1.77%). This difference, albeit small, was statistically significant (OR = 1.30, SE = 0.075, z-ratio = 4.508, $p<0.001$). Similarly, comments with personal narratives (M = 5.18%, 95% CI 4.19–6.38%) had a slightly higher probability of receiving a reply from political newcomers than comments without personal narratives (M = 4.51%, 95% CI 3.68–5.52%). This difference, albeit small, was also statistically significant (OR = 1.16, SE=0.040, z-ratio = 4.120, $p<0.001$). Together, these results support H2. The regression table for this analysis is included in the Appendix (Table 7.2).

Figure 4.2: Probability of receiving a reply from PDIs (95 % CI). Personal narrative comments are slightly more likely to get a response from PDIs.



4.2.3 H3: Sustaining political participation

H3 states that PDIs who use personal narratives in political communities are more likely to engage in political communities in the following month than those who do not use personal narratives. To evaluate H3, we identified if the PDIs used a personal narrative in political subreddits for each month in 2021 using the personal narrative classifier. We then checked whether these users returned to comment in political subreddits in the subsequent month. Data from December 2021 was excluded as we do not know if the users who participated in December 2021, returned in January 2022. In total, this dataset included 225,317 least politically active users and 724,096 political newcomers. For each unique user in each month, we identified whether they used a personal narrative in at least one of their comments during that month and whether they participated in a political discussion in the following month. We then conducted a random-effects logistic regression, modelling whether the user commented in a political subreddit in the subsequent month (dependent variable). The primary variable of interest was whether the user used a personal narrative in any of their comments in that particular month (personal narrative indicator) and the user type (least politically active and political newcomers). We also included an interaction term between the personal narrative indicator and user type to account for potential differences in the relationships between returning the subsequent month and the use of personal narratives by the two user groups. We included a random effect to account for variability across months.

Figure 4.3: Probability of returning next month to political discussions for PDIs (95% CI). PDIs are more likely to return to political discussions in the next month if they employ personal narratives in the current month.

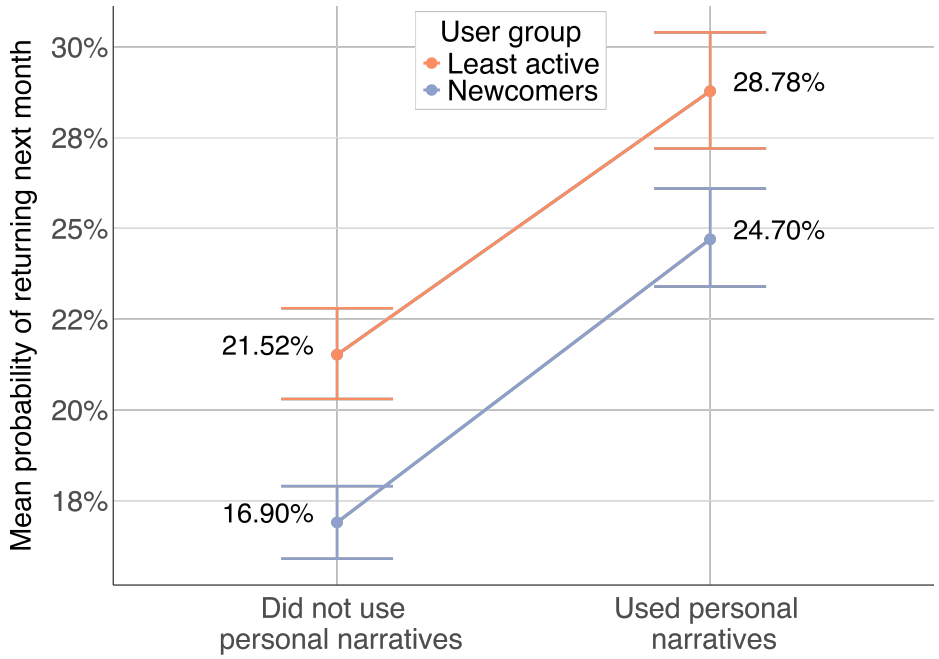


Figure 4.3 shows the average probability of the least politically active users and political newcomers returning to political subreddits the following month, based on whether they used a personal narrative in their comments in a particular month. On average, the least politically active users who used personal narratives in a given month ($M = 28.78\%$, 95% CI 27.24%–30.37%) were more likely to engage in political discussions in the subsequent month compared to those who did not use personal narratives ($M = 21.52\%$, 95% CI 20.32%–22.76%). This difference was statistically significant (OR = 1.47, SE = 0.022, z-ratio = 26.26, $p < 0.001$). Similarly, on average, political newcomers who used personal narratives in a given month ($M = 24.70\%$, 95% CI 23.36%–26.07%) were more likely to engage in political discussions in the subsequent month than those who did not use personal narratives ($M = 16.90\%$, 95% CI 15.91%–17.93%). This difference was statistically significant (OR = 1.613, SE = 0.012, z-ratio = 64.294, $p < 0.001$). Together, these results support H3. The regression table for this analysis is included in the Appendix (Table 7.3).

4.2.4 H4, H5, RQ1: Community perception of personal narratives

H4, H5 and RQ1 evaluate the perceptions of using personal narratives. H4 states that comments by PDIs are likely to score lower than comments by the most politically active users. H5 states that personal narrative comments by PDIs are more likely to score higher than other

comments by PDIs. RQ1 asks how the score of personal narrative comments made by PDIs and the most politically active users vary.

To evaluate these hypotheses and the research question, we use the same dataset used to evaluate H1. We used a linear regression model to model the cube-root transformed score of the comment as the dependent variable. We include the type of user posting the comment (political newcomers, least politically active, most politically active) and whether the comment was a personal narrative or not (personal narrative indicator) as independent variables. We also included an interaction term between user type and the personal narrative indicator to model how the relationship between different user types and the cube-root transformed comment score varies based on whether a comment is a personal narrative or not. To account for potential confounding factors, we included the same control variables we used to evaluate H1. Random effects were included to account for variability across subreddits and months.

Figure 4.4: Mean comment scores of users across different comment types (95% CI). PDIs are likely to score lower than comments by the most politically active users. Also, personal narrative comments by PDIs are likely to score higher than other comments made by the same users.

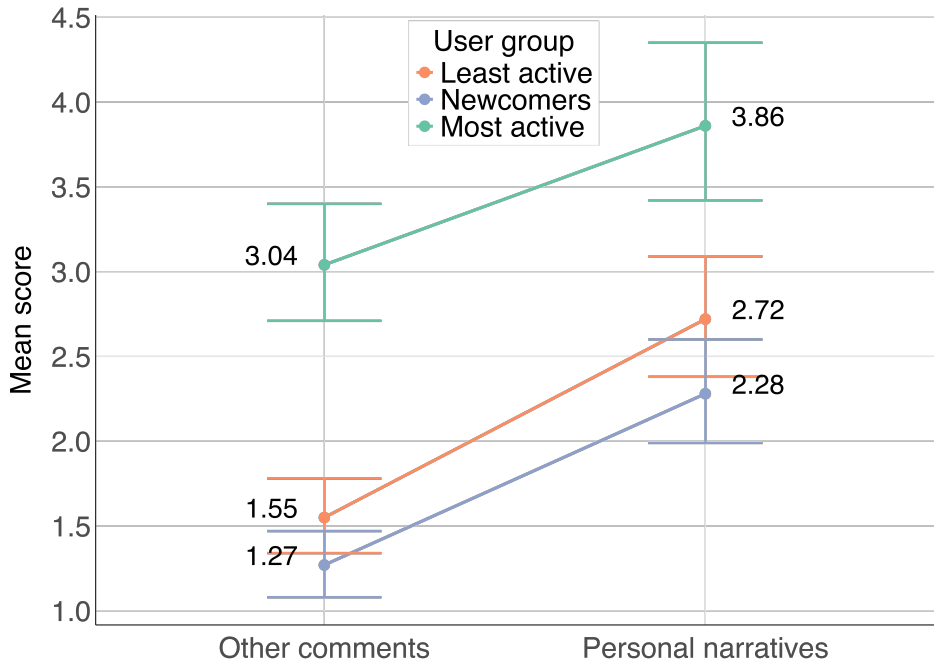


Figure 4.4 shows the estimated mean comment scores of comments across users with different political activity levels (least politically active vs. most politically active vs. political newcomers) and comment types (comments with personal narrative vs. other comments). We observe that comments made by the least politically active users and political newcomers had lower estimated mean scores than those made by the most politically active users for both comments containing

a personal narrative (least politically active users: $M = 2.72$, 95% CI 2.38–3.09, political newcomers: $M = 2.28$, 95% CI 1.99–2.60, most politically active users: $M = 3.86$, 95% CI 3.42–4.35) and those that do not (least politically active users: $M = 1.55$, 95% CI 1.34–1.78, political newcomers: $M = 1.27$, 95% CI 1.08–1.47, most politically active users: $M = 3.04$, 95% CI 2.71–3.40). These differences were statistically significant (see Appendix table 7.5 for details). Together, these results support H4. The differences between mean scores of comments containing personal narratives and those that do not, authored by the least politically active users (M = 0.24, SE = 0.0148, z-ratio = 16.26, $p < 0.001$) and political newcomers (M = 0.23, SE = 0.0119, z-ratio = 19.62, $p < 0.001$) are also statistically significant. These results support H5. The regression table for this analysis is included in the Appendix (Table 7.4).

Further, answering RQ1, we compare the difference between scores of personal narrative comments by PDIs and the most politically active users with the difference between the scores of other comments by the same group of users. We find that, on average, the difference in mean scores between the least politically active users and the most politically active users is significantly smaller when both groups use personal narratives compared to when they do not (difference in cube-root scale: 0.119, SE = 0.022, z-ratio = 5.320, $p < 0.001$). We observe similar smaller differences between political newcomers and the most politically active users as well when both groups use personal narratives (difference in cube-root scale: 0.113, SE = 0.021, z-ratio = 5.512, $p < 0.001$). These results indicate that personal narratives could help reduce the disparity in positive reception between content posted by PDIs and the most politically active users.

Chapter *5*

Personal Narratives In Cross-Partisan Political Discussions

5.1 Hypotheses

First, we evaluate the question of the prevalence of personal narratives in cross-partisan political discussions. Personal narratives, by definition, disclose some personal, sometimes intimate details of the narrator's life. Though not specific to personal narratives, qualitative research suggests that people avoid posting content about themselves in cross-partisan spaces online for fear that such information could be weaponized to attack them [94]. Instead, personal narratives are often shared with others with a common cause to express solidarity and build a collective identity [24], factors more associated with the ingroup identification than with engagement with the outgroup. Relatedly, research on online self-disclosure suggests that individuals are more likely to disclose personal experiences in supportive forums than in other discussion spaces [95]. Thus, we expect that individuals are less likely to share their own experiences with an opposing partisan than with a copartisan. Therefore, we evaluate the following hypothesis:

> H1: Individuals are less likely to reply to an opposing partisan with a personal narrative than to a copartisan.

Next, we examine how personal narratives are received by opposing partisans. On Reddit, subreddit members may upvote or downvote content that they view favourably or unfavourably, respectively. Reddit tracks these votes in the form of karma points which acts as a kind of social feedback. Thus, a reply that receives more karma points can be interpreted as being seen more favourably by the subreddit. If a reply received more karma points in a cross-partisan subreddit, that is, the political ideology of the subreddit is different from that of the user who is replying, then we can interpret higher karma points as being viewed more favourably by opposing partisans. As discussed in the previous section, in experimental settings, personal narratives appear to cut through the partisan divide and are viewed positively by opposing partisans. Therefore, using karma points as an indicator of cross-partisan favorability, we test the following hypothesis:

H2a: Replies containing personal narratives are likely to receive higher karma points than other replies in cross-partisan communities.

On the other hand, as discussed in the previous section, in real-world settings where partisans are deeply polarized, the personal narratives that individuals share may build ingroup solidarity and highlight intergroup differences, which may further deepen the faultlines. Thus, these comments may be viewed more unfavourably by the other side. Therefore, we test the alternate hypothesis:

H2b: Replies containing personal narratives are likely to receive lower karma points than other replies in cross-partisan communities.

Next, we evaluate the effect of personal narratives on their ability to engage and elicit a response from opposing partisans. There is strong evidence that personal narratives typically engage those who are exposed to them. Researchers argue that narratives uniquely engage the readers both cognitively and emotionally [96]. In health communication, [97] note that narratives' ability to engage even those who are hard-to-reach, such as those with low involvement and low knowledge. Also, stories' ability to convey disagreement in less direct and confrontational ways may help manage potential conflict and facilitate higher engagement among even conflict-averse individuals who typically opt out of talking politics [98]. Further, in studying small scale online groups, [53] observed that 'argument stories' where there was some disagreement expressed received some of the highest engagement, resulting in more responses compared to most other messages. Thus, we evaluate the following hypothesis:

H3: Cross-partisan replies containing personal narratives are more likely to prompt a response from the author of the original comment compared to other replies.

Finally, we examine the hostility of replies to personal narratives. Intuitively, if H2a is true, that is, opposing partisans are appreciative of personal narratives, then it is likely that less hostility will ensue. At the same time, if H2b is true, then it is likely that more hostility will ensue. However, evaluations by the community may differ from those of an individual user's reply. Given the relative lack of research in this domain, we pose the following research question:

RQ1: How does the original commenter's hostility level differ when responding to a cross-partisan reply that includes a personal narrative compared to one without?

5.2 Measurement variables

5.2.1 Personal narrative indicator

To answer H1, we run the personal narrative classifier on the reply comments in the interactions and use the output labels (binary indicator) as the dependent variable. For the other hypotheses and research question, this indicator is an independent variable.

5.2.2 Karma points

To answer H2a and H2b, we use the karma points accrued on the replies in the interactions as the dependent variable. For the other hypotheses and research question, we use the karma points accrued on the parent comments in the interactions and on the top-level post as control variables.

5.2.3 User interaction partisanship indicator

If the partisanship of the user authoring the parent comment matches with that of the user who is replying, then we label the interaction as left copartisan or right copartisan depending on whether the users were left or right-leaning respectively, else we label the interaction as left-right or right-left depending on if the reply author is right leaning or left leaning respectively. For all hypotheses and research question, we use this indicator as an independent variable as the indicator allows us to compare between copartisan and crosspartisan interactions.

5.2.4 Reply-community partisanship indicator

If the partisanship of the community matches with that of the user who is replying, then we label the community as copartisan, else we label it as cross-partisan. We use this indicator as a control variable for all hypotheses and research question except H2a and H2b. Instead, for H2a and H2b, we only consider interactions where this indicator is cross-partisan as we evaluate the favorability of personal narratives in cross-partisan communities.

5.2.5 Additional control variables

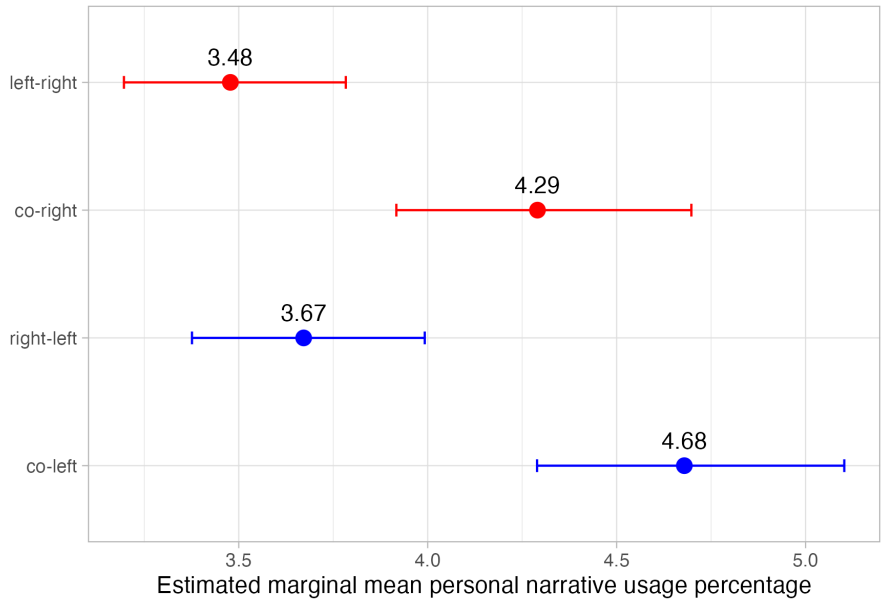
To accurately compare replies containing personal stories and those that do not, we use additional control variables. We control for the length of the comments and the level of the discussion thread (whether the discussion took place at the top-level thread or at one of the lower levels).

5.3 Results

We start with H1, which concerns the use of personal narratives in copartisan and cross-partisan settings. H1 predicted more use of personal narratives in copartisan interactions than cross-partisan interactions. The data support this prediction. We use the full 400,000 samples of interactions to evaluate this hypothesis.

We use a mixed-effects binomial regression to evaluate this prediction, modelling the probability of a reply to a comment containing a personal narrative with a random effect for the subreddit where the interaction took place and the user interaction indicator indicating the type of interaction (left-copartisan, right-left, right-copartisan, left-right). We also control for

Figure 5.1: Percentage of interactions where reply contains personal narratives (95% CI). The data suggests that co-partisan interactions (co-left and co-right) tend to have a higher percentage of personal narratives compared to cross-partisan interactions (left-right and right-left). Among these, co-left interactions exhibit the highest personal narrative usage (4.68%), while left-right interactions show the lowest (3.48%).

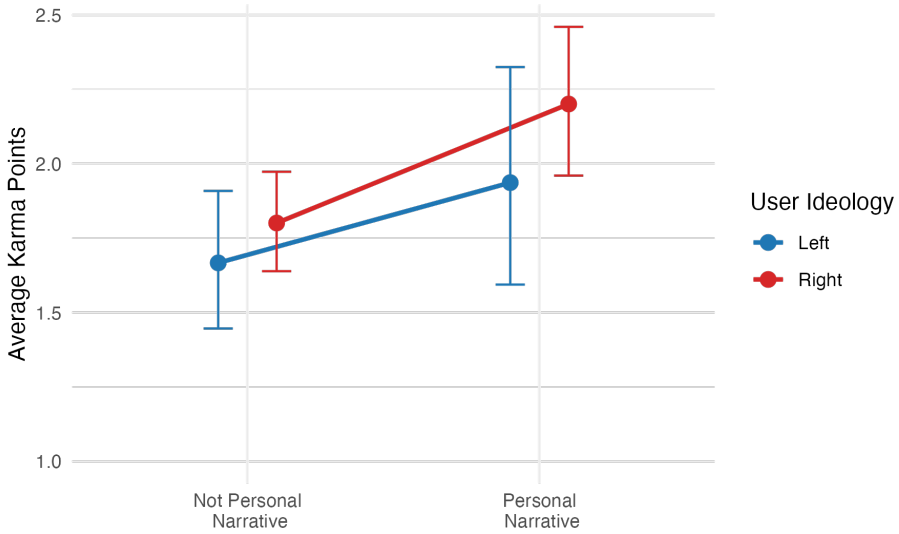


the length of the parent comment, the karma points on the parent comment and the discussion thread level.

Figure 5.1 shows the estimated marginal mean probability of using a personal narrative in each kind of interaction. We find that, on average, a left leaning user uses a personal narrative 4.68% (95%CI 4.29-5.10) of the time when replying to other left-leaning users compared to 3.67% (95%CI 3.38-3.99) of the time when replying to right-leaning users. This difference is statistically significant ($OR = 1.29$, $SE = 0.04$, $z\text{-}ratio = 8.01$, $p < 0.001$). We also find that, on average, a right leaning user uses a personal narrative 4.29% (95%CI 3.92-4.70) of the time when replying to other right-leaning users compared to 3.48% (95%CI 3.20-3.78) of the time when replying to left-leaning users. This difference is also statistically significant ($OR = 1.24$, $SE = 0.04$, $z\text{-}ratio = 6.76$, $p < 0.001$). Therefore, H1 is supported.

H2a and H2b are concerned with how personal narratives are received by opposing partisans. H2a predicted that opposing partisans view personal narratives more favorably than other comments while H2b predicted the opposite. Our data shows that H2a is supported and H2b is not supported for right-leaning users. Personal narratives by right-leaning users are viewed more favourably than their other contributions in cross-partisan communities. For left-leaning users, H2a and H2b are not supported. We observe no such statistically significant difference for personal narratives by left-leaning users in right-leaning communities. From the 400,000

Figure 5.2: Average karma score earned by opposing partisans when using and not using personal narratives (95% CI). The figure shows that both left- and right-leaning users receive higher karma scores when using personal narratives compared to when they do not. However, right-leaning users consistently receive higher karma than left-leaning users in both conditions.

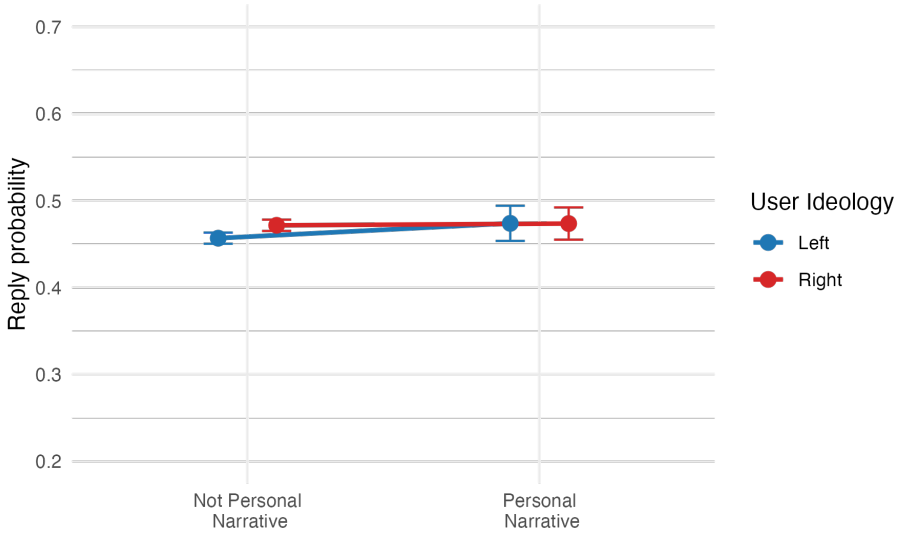


sampled interactions, we selected only cross-partisan interactions that took place in subreddits where the partisanship of the reply user and the subreddit were not the same (n=88,140).

We use a mixed-effects linear regression to evaluate this prediction, modelling the karma score of a reply to an opposing partisan with a random effect for the subreddit where the interaction took place and an interaction effect between the personal narrative indicator and the type of cross-partisan interaction (right-left or left-right). We also control for the karma score of the parent, the karma point of the top-level post and the discussion thread level. We apply a cube-root transformation to account for the skewness in the karma score of the parent and reply comments since cube-root is well-defined for positive, negative and zero values, unlike the conventional log transformations [91]. We apply $\log(x+1)$ transformation to account for skewness in the score of the top-level posts, as they can have only a value of zero or higher.

Figure 5.2 shows the interaction plot between the ideology of the user posting the reply and whether the reply contained a personal narrative. We find that, in left-leaning subreddits, a right leaning user's reply to a left-leaning user earns an average karma score of 2.20 (95%CI 1.96-2.46) when it contains a personal narrative compared to 1.80 points (95%CI 1.64-1.97) when it does not contain a personal narrative. This difference is statistically significant (estimate (in cube-root scale):-0.084, $SE = 0.017$, $z\text{-}ratio = -4.92$, $p < 0.001$). Thus, for right-leaning users, H2a is supported, and H2b is not supported. However, for left-leaning users, the difference in karma points between using and not using a personal narrative is not statistically significant

Figure 5.3: Likelihood of receiving a reply when using and not using personal narratives (95% CI). The figure shows a slight increase in the probability of receiving a reply when using personal narratives compared to not using them. However, the difference is small, and both left- and right-leaning users exhibit similar trends, with no clear ideological distinction in how personal narratives impact reply engagement.



(estimate (in cube-root scale):-0.061, $SE = 0.029$,$z\text{-}ratio = -2.050, p = 0.16$). Therefore, for left-leaning users, H2a and H2b is not supported.

H3 and RQ1 concern the reaction of the user to whom the opposing partisan directs their personal narrative. H3 evaluates if personal narrative replies are more likely to elicit a response from the user who posted the original comment. From the full dataset of 400,000 interactions, we only consider the subset of interactions that are cross-partisan to evaluate H3 (n=200,000). Our data shows that for both left- and right-leaning users, using personal narratives in cross-partisan interactions does not increase the likelihood of receiving a reply. Therefore, H3 is not supported for both left and right-leaning users.

We use a mixed-effects binomial regression to evaluate this prediction, modelling whether the original parent commenter replied or not (1 or 0) with a random effect for the subreddit where the interaction took place and an interaction effect between the personal narrative indicator and the type of cross-partisan interaction (right-left or left-right). We also control for the karma score of the parent and the reply comment, the karma score of the top-level post, the discussion thread level and whether the subreddit was copartisan or not with respect to the reply user. Again, we apply a cube-root transformation to account for the skewness in the karma score of the parent and reply comments. We apply $\log(x + 1)$ transformation to account for skewness in the score of the top-level posts.

Figure 5.4: Probability of receiving a toxic reply when using and not using personal narratives (95% CI). The figure shows that the likelihood of receiving a toxic reply remains relatively stable regardless of whether a personal narrative is used. While left-leaning users see a slight increase in toxicity when using personal narratives, right-leaning users experience a minor decrease.

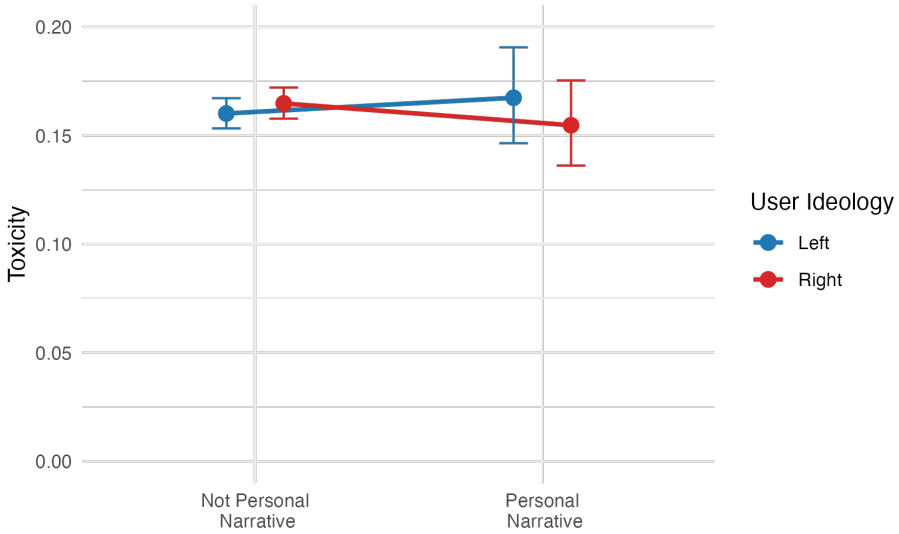


Figure 5.3 shows the interaction plot between the ideology of the user posting the reply and whether the reply contained a personal narrative. The difference in the probability of receiving a response from the parent comment author when replying with a personal narrative compared to when not replying with a personal narrative is not statistically significant for both left-leaning (OR:0.93, $SE = 0.04$,*z-ratio* $= -1.67, p = 0.34$) and right-leaning users (OR:0.99, $SE = 0.04$,*z-ratio* $= -0.22, p = 0.99$). Thus, H3 is not supported.

RQ1 evaluates how the hostility of the response differs between replies with and without personal narratives. From the full dataset of 400,000 interactions, we only consider the subset of interactions that are cross-partisan and received a response from the original commenter to evaluate RQ1 (n=92,764). Our data shows that, for both left- and right-leaning users, there is no difference in the hostility of the responses received for personal narratives compared to non-narratives.

We use a mixed-effects binomial regression to evaluate this research question, modeling whether the original parent commenter replied with a toxic comment or not (1 or 0) with a random effect for the subreddit where the interaction took place and an interaction effect between the personal narrative indicator and the type of cross-partisan interaction (right-left or left-right). We use the same controls we used to evaluate H3.

Figure 5.4 shows the interaction plot between the ideology of the user posting the reply and whether the reply contained a personal narrative. The difference in the probability of a

response from the parent comment author being toxic when replying with a personal narrative compared to when not replying with a personal narrative is not statistically significant for both left-leaning (OR: 0.95, $SE = 0.08$,*z-ratio* $= -0.66, p = 0.91$) and right-leaning users (OR:1.08, $SE = 0.08$,*z-ratio* $= 1.00, p = 0.75$).

Chapter *6*

Discussion

6.1 Politically Disinclined Individuals

The results of this study indicate that personal narratives are a promising way to encourage politically disinclined individuals to engage in political discussions. We find that PDIs are more likely to use personal narratives (H1) and are slightly more likely to respond to them (H2). Further, PDIs who use personal narratives are more likely to return to future political discussions (H3). Also, personal narratives appear to positively contribute to political discussions as well. We find that comments by PDIs containing personal narratives are more likely to be viewed positively than other comments by them (H5). Although comments by PDIs, on average, score lower points than those by more politically active users (H4), personal narratives appear to reduce this difference in scores (RQ1). Overall, these results highlight that personal narratives can positively impact both participation in and the overall quality of online political discussions. Below, we discuss the implications of our findings.

Support for H1 and H2 highlights the inclusivity and accessibility of personal narratives as a form of communication. As discussed earlier, many individuals avoid engaging in political discussions because they do not know enough about the issue [14]. These results suggest that personal narratives provide a low-barrier entry point into political discussions, reducing the reliance on prior political knowledge or expertise. Yet traditionally, political discussion spaces privilege rational arguments over personal stories and testimonials [16]. Similarly, many online political communities restrict the use of personal narratives. For example, the rules in the r/neutralpolitics community explicitly state that "if you're claiming something to be true, you need to back it up with a qualified source. There is no 'common knowledge' exception, and anecdotal evidence is not allowed." Our results suggest that such an approach to content moderation, while facilitating a fact-based discussion space, likely has an unintended effect of turning away individuals who do not typically engage in political discussions.

To survive, all communities, especially political ones, need to attract and retain newcomers in order to counter inevitable user churn and burnout [99]. Our results suggest that community moderators can encourage the use of personal narratives to attract and retain new users into

the community. This can be done through norm-setting by highlighting exemplar comments containing personal narratives and by explicitly noting their usage in the community rules.

Attracting and retaining users is only one part of the puzzle for content moderators. While newcomers often provide fresh ideas and help stimulate discussions in newer directions [100], their contributions may not always align with the norms of the communities [101]. Indeed, in our analysis, political communities perceive the contributions of the PDIs less positively than those of the more politically active users. However, we find that the gap in community perceptions (based on karma score) of politically disinclined and more politically active users reduces when PDIs employ personal narratives. This may be achieved by encouraging introductory posts by newcomers to share their personal stories. Further, given that personal narratives also generally score higher than other comments, this early positive feedback may provide additional encouragement to newcomers to productively engage in the community. As Kraut et al. [99] note, "when newcomers have friendly interactions with existing community members soon after joining a community, they are more likely to stay longer and contribute more."

While these results are promising, their potential to drive actionable change can be fully unlocked only when combined with implementable strategies that enhance political discourse in online communities. Social media platforms, in particular, hold the power to amplify the visibility of personal narratives. For example, Bail highlights how algorithms can be optimized to prioritize content with broad resonance, rather than divisive or controversy-driven engagement metrics [12]. Similarly, platforms could adapt their algorithms to elevate content featuring personal narratives, helping to foster a more inviting environment that encourages participation by a wider range of users. Alternately, platforms can leverage gamification, such as badges for first-time contributors who employ personal narratives, to encourage disinclined users to share personal narratives in political discussion spaces.

Finally, political scientists have flagged concerns about how online political discussions are being dominated by users who are heavily politically involved, particularly how this often leads to a hostile partisan environment and misperceptions about polarization [11]. In an ideal scenario, a broader cross-section of individuals with differing political involvement participate in these discussions. Our results suggest that encouraging the use of personal narratives in political discussions likely attracts and retains users who typically avoid politics, advancing the normatively positive goal of broadening political participation online.

6.2 Cross-Partisan Political Discourse

This work evaluates the use and impact of personal narratives in naturally occurring online cross-partisan discussions. Our findings complement some of the results from experimental research. First, we find that personal narratives are used less frequently when engaging with opposing partisans than when engaging with copartisans. When they are used in cross-

partisan settings, replies with personal narratives by right-leaning users appear to be viewed more favourably by left-leaning communities than their other replies. We find no such effect for personal narratives told by left-leaning users. We also find that using personal stories has no effect on the likelihood of receiving a reply or on the hostility of the reply received from the original commenter.

Consistent with our prediction, we find that personal narratives are more likely to be used with copartisans than opposing partisans. As a result, the overall influence of personal narratives on cross-partisan political discussions may be rather diminished, with fewer people exposed to stories from opposing political perspectives. While we cannot identify the mechanisms behind this phenomenon we speculate that individuals are more likely to share such narratives with a more supportive audience [102].

We find that right-leaning users in left-leaning communities get more of a bump when using a personal narrative compared to left-leaning users in right-leaning communities. This result can be interpreted in two ways: right-leaning users tell better stories than left-leaning users or right-leaning communities do not appreciate stories by opposing partisans as much as left-leaning communities do. There are reasons to suspect both factors might be at play. Research suggests that stories told by conservatives elites are relatively straightforward stories that follow familiar plotlines which appears to strongly resonate with the American public [103]. Stories told by liberal elites are complex and unfamiliar, and do not appear to have the same effect [103]. On the other hand, stories are often evaluated by who is telling the stories [104]. A widespread perception of Democrats as being coastal elites may also affect how their stories are received. As Polleta [17] notes, "it is difficult to think of eggheads as folksy raconteurs. So Democratic stories are more likely to be heard as strategic, manipulative, or plain awkward ... than are Republican stories." Given that ordinary supporters look for cues from their party leaders [105], and often draw on the same set of deep stories [56], it is possible that similar forces may be at work both in terms of the perception and the kinds of everyday stories that left- and right-leaning users convey.

We also find that the hostility toward personal narratives do not differ significantly from responses to other types of comments. Seen in another way, replies to personal narratives appear to be just as hostile as replies to other content (about 16% of the replies are toxic). While this may be a *null* result, it raises important concerns about using personal narratives in political discussions. Through sharing personal stories, individuals narratively construct their identities and establish a sense of self [85]. Hostility directed toward these narratives may be received differently than hostility toward general arguments, since such hostility attacks are likely to feel more personal, potentially leading to a chilling effect on open, personal expression.

Taken together, these findings accentuate how personal narratives can act as a highly effective tool for overcoming common obstacles faced by online political communities—namely, limited participation from politically disinclined groups and ineffective communication across partisan

divides. Encouraging their use could help create more inclusive, engaging, and productive political discourse spaces, which ultimately benefit both community longevity and democratic dialogue more broadly.

6.3 Limitations and Future Work

While our study demonstrates the potential of personal narratives to engage politically disinclined individuals and improve cross-partisan communication, we recognize several limitations that must be considered. Firstly, our research draws exclusively from Reddit comments in political subreddits, which may not fully encapsulate broader online political discourse due to demographic biases inherent to the platform. Secondly, the dataset's temporal constraints—limited specifically to the period of 2020-2021—align with the COVID-19 pandemic and heightened political polarization. This unique historical context could have amplified narrative trends that may differ significantly under less extreme or more typical societal conditions.

Additionally, our operationalization of political activity, based solely on users' subreddit engagement, might lack sufficient complexity, potentially impacting the accuracy and generalizability of our findings. There is also the possibility that external personal events prompt users to both engage politically and share personal experiences simultaneously, suggesting that the tendency to use personal narratives could be issue-specific rather than broadly applicable across various topics. Furthermore, our analysis was limited to assessing hostility in replies to personal narratives, yet other deliberative ideals, such as reciprocity and reason-giving, are equally critical to evaluating the holistic effectiveness of narrative contributions in political discussions [106].

Moreover, our examination focused exclusively on text-based personal narratives, neglecting the various multimedia affordances of social media platforms. Visual narratives conveyed through images and videos could potentially offer more compelling and engaging content. This highlights another dimension not captured in our current analysis.

While our study primarily highlights the beneficial effects of personal narratives, we also acknowledge their potential misuse in political discourse. Personal narratives can sometimes be strategically employed to mislead audiences and propagate misinformation [54]. Our research did not explore this negative aspect of personal narratives, and future research should investigate how these narratives might contribute to misinformation and polarization. Additionally, the specific content of personal narratives was not analyzed in our work, leaving another valuable dimension for subsequent exploration.

Several avenues for future research arise from these limitations. Comparative analyses across diverse social media platforms, such as Twitter and Facebook, would reveal whether narrative patterns observed on Reddit are consistent or distinct elsewhere. Longitudinal studies could examine the persistence and sustainability of engagement fostered through personal narratives.

Furthermore, future research could explore the tangible impacts of personal narratives on political outcomes, such as voter turnout and activism, providing deeper insights into their efficacy in driving political participation.

By addressing these limitations and exploring suggested research directions, future studies can significantly expand our understanding of personal narratives' role in political discourse. It remains essential to investigate whether the observed positive impacts of personal narratives are generalizable across different online communities and political contexts. Ultimately, we hope this study encourages further exploration of personal narratives as a means of bridging the gap between political apathy and active participation, thereby promoting more inclusive and accessible political discussions.

Bibliography

[1] W. P. Eveland Jr and T. Thomson, "Is it talking, thinking, or both? a lagged dependent variable model of discussion effects on political knowledge," *Journal of communication*, vol. 56, no. 3, pp. 523–542, 2006.

[2] J. N. Cappella, V. Price, and L. Nir, "Argument repertoire as a reliable and valid measure of opinion quality: Electronic dialogue during campaign 2000," *Political Communication*, vol. 19, no. 1, pp. 73–93, 2002.

[3] D. C. Mutz, "Cross-cutting social networks: Testing democratic theory in practice," *American Political Science Review*, vol. 96, no. 1, pp. 111–126, 2002.

[4] L. R. Jacobs, F. L. Cook, and M. X. D. Carpini, *Talking together: Public deliberation and political participation in America*. University of Chicago Press, 2009.

[5] V. Price, J. N. Cappella, and L. Nir, "Does disagreement contribute to more deliberative opinion?" *Political communication*, vol. 19, no. 1, pp. 95–112, 2002.

[6] M. S. Levendusky and D. A. Stecula, *We need to talk: how cross-party dialogue reduces affective polarization*. Cambridge University Press, 2021.

[7] J. Kim, R. O. Wyatt, and E. Katz, "News, talk, opinion, participation: The part played by conversation in deliberative democracy," *Political communication*, vol. 16, no. 4, pp. 361–385, 1999.

[8] M. Jurkowitz and A. Mitchell, "A sore subject: Almost half of americans have stopped talking politics with someone," *Pew Research Center*, 2020.

[9] C. McClain, "70% of U.S. social media users never or rarely post or share about political, social issues," 2021, accessed: 2024-12-03. [Online]. Available: https://www.pewresearch.org/short-reads/2021/05/04/70-of-u-s-social-media-users-never-or-rarely-post-or-share-about-political-social-issues/

[10] M. Anderson and D. Quinn, "46% of us social media users say they are 'worn out'by political posts and discussions," 2019.

[11] Y. Krupnikov and J. B. Ryan, *The other divide.* Cambridge University Press, 2022.

[12] C. Bail, *Breaking the social media prism: How to make our platforms less polarizing.* Princeton University Press, 2022.

[13] M. S. Levendusky and N. Malhotra, "(mis) perceptions of partisan polarization in the american public," *Public Opinion Quarterly*, vol. 80, no. S1, pp. 378–391, 2016.

[14] T. N. Carlson and J. E. Settle, *What goes without saying.* Cambridge University Press, 2022.

[15] F. Justwan, B. Baumgaertner, J. E. Carlisle, A. K. Clark, and M. Clark, "Social media echo chambers and satisfaction with democracy among democrats and republicans in the aftermath of the 2016 us elections," *Journal of Elections, Public Opinion and Parties*, vol. 28, no. 4, pp. 424–442, 2018.

[16] I. M. Young, *Inclusion and democracy.* Oxford University Press, USA, 2002.

[17] F. Polletta, *It was like a fever: Storytelling in protest and politics.* University of Chicago Press, 2009.

[18] M. C. Green and T. C. Brock, "The role of transportation in the persuasiveness of public narratives." *Journal of personality and social psychology*, vol. 79, no. 5, p. 701, 2000.

[19] A. De Graaf, H. Hoeken, J. Sanders, and J. W. Beentjes, "Identification as a mechanism of narrative persuasion," *Communication research*, vol. 39, no. 6, pp. 802–823, 2012.

[20] E. Kubin, C. Puryear, C. Schein, and K. Gray, "Personal experiences bridge moral and political divides better than facts," *Proceedings of the National Academy of Sciences*, vol. 118, no. 6, p. e2008389118, 2021.

[21] D. Hagmann, J. A. Minson, and C. H. Tinsley, "Personal narratives build trust across ideological divides." *Journal of Applied Psychology*, 2024.

[22] J. L. Kalla and D. E. Broockman, "Which narrative strategies durably reduce prejudice? evidence from field and survey experiments supporting the efficacy of perspective-getting," *American Journal of Political Science*, vol. 67, no. 1, pp. 185–204, 2023.

[23] F. Polletta, "Characters in political storytelling," *Storytelling, Self, Society*, vol. 11, no. 1, pp. 34–55, 2015.

[24] A. De Fina, *Identity in Narrative: A Study of Immigrant Discourse.* John Benjamins Publishing, 2003, vol. 3.

[25] F. Polletta, "Contending stories: Narrative in social movements," *Qualitative sociology*, vol. 21, pp. 419–446, 1998.

[26] S. Iyengar, Y. Lelkes, M. Levendusky, N. Malhotra, and S. J. Westwood, "The origins and consequences of affective polarization in the united states," *Annual review of political science*, vol. 22, no. 1, pp. 129–146, 2019.

[27] J. E. Settle, *Frenemies: How social media polarizes America*. Cambridge University Press, 2018.

[28] S. Rathje, J. J. Van Bavel, and S. Van Der Linden, "Out-group animosity drives engagement on social media," *Proceedings of the National Academy of Sciences*, vol. 118, no. 26, p. e2024292118, 2021.

[29] F. Polletta and J. Callahan, "Deep stories, nostalgia narratives, and fake news: Storytelling in the trump era," *Politics of meaning/meaning of politics: Cultural sociology of the 2016 US presidential election*, pp. 55–73, 2019.

[30] N. Eliasoph, *Avoiding politics: How Americans produce apathy in everyday life*. Cambridge University Press, 1998.

[31] E. Noelle-Neumann, "The spiral of silence a theory of public opinion," *Journal of communication*, vol. 24, no. 2, pp. 43–51, 1974.

[32] S. G. Ulbig and C. L. Funk, "Conflict avoidance and political participation," *Political Behavior*, vol. 21, pp. 265–282, 1999.

[33] M. V. Hibbing, M. Ritchie, and M. R. Anderson, "Personality and political discussion," *Political Behavior*, vol. 33, pp. 601–624, 2011.

[34] J. Matthes, "Do hostile opinion environments harm political participation? the moderating role of generalized social trust," *International Journal of Public Opinion Research*, vol. 25, no. 1, pp. 23–42, 2013.

[35] A. S. Gerber, G. A. Huber, D. Doherty, and C. M. Dowling, "Disagreement and the avoidance of political discussion: Aggregate relationships and differences across personality traits," *American Journal of Political Science*, vol. 56, no. 4, pp. 849–874, 2012.

[36] C. Sunstein, *Republic. com*. Princeton University Press, 2001.

[37] B. Berelson, *Voting: A study of opinion formation in a presidential campaign*. University of Chicago Press, 1954.

[38] M. Patterson and K. R. Monroe, "Narrative in political science," *Annual review of political science*, vol. 1, no. 1, pp. 315–331, 1998.

[39] C. K. Riessman, "Analysis of personal narratives," *Inside interviewing: New lenses, new concerns*, pp. 331–346, 2003.

[40] R. G. Jones Jr and B. M. Calafell, "Contesting neoliberalism through critical pedagogy, intersectional reflexivity, and personal narrative: Queer tales of academia," *Journal of homosexuality*, vol. 59, no. 7, pp. 957–981, 2012.

[41] C. Stivers, "Reflections on the role of personal narrative in social science," 1993.

[42] J. A. Robinson, "Personal narratives reconsidered," *The journal of American folklore*, vol. 94, no. 371, pp. 58–85, 1981.

[43] A. Shuman, "Entitlement and empathy in personal narrative," *Narrative inquiry*, vol. 16, no. 1, pp. 148–155, 2006.

[44] E. E. Robinson, "Sharing stories: The role of personal narratives in community mobilization," *Humanity & Society*, vol. 40, no. 4, pp. 442–461, 2016.

[45] S. Akerman, "75personal narratives and the creation of a political voice," in *Words and Wounds: Narratives of Exile*. Oxford University Press, 08 2019. [Online]. Available: https://doi.org/10.1093/oso/9780190851712.003.0004

[46] I. M. Young, *Inclusion and democracy*. Oxford University Press, USA, 2000.

[47] A. Georgakopoulou, "Sharing the moment as small stories: The interplay between practices & affordances in the social media-curation of lives," *Narrative Inquiry*, vol. 27, no. 2, pp. 311–333, 2017.

[48] Z. Papacharissi, *A networked self and platforms, stories, connections*. Routledge New York, 2018.

[49] R. E. Page, *Stories and social media: Identities and interaction*. Routledge, 2013.

[50] A. De Fina, "Storytelling and audience reactions in social media," *Language in Society*, vol. 45, no. 4, pp. 473–498, 2016.

[51] A. Georgakopoulou, S. Iversen, and C. Stage, *Quantified Storytelling: A Narrative Analysis of Metrics on Social Media*, 1st ed., ser. Palgrave Pivot. Cham, Switzerland: Palgrave Macmillan, 2020.

[52] M. Wojcieszak and N. Kim, "How to improve attitudes toward disliked groups: The effects of narrative versus numerical evidence on political persuasion," *Communication Research*, vol. 43, no. 6, pp. 785–809, 2016.

[53] L. W. Black, "Deliberation, storytelling, and dialogic moments," *Communication Theory*, vol. 18, no. 1, pp. 93–116, 2008.

[54] J. J. Van Bavel, D. A. Reinero, V. Spring, E. A. Harris, and A. Duke, "Speaking my truth: Why personal experiences can bridge divides but mislead," *Proceedings of the National Academy of Sciences*, vol. 118, no. 8, p. e2100280118, 2021.

[55] S. Prochaska, K. Duskin, Z. Kharazian, C. Minow, S. Blucker, S. Venuto, J. D. West, and K. Starbird, "Mobilizing manufactured reality: How participatory disinformation shaped deep stories to catalyze action during the 2020 us presidential election," *Proceedings of the ACM on human-computer interaction*, vol. 7, no. CSCW1, pp. 1–39, 2023.

[56] A. R. Hochschild, *Strangers in their own land: Anger and mourning on the American right*. The New Press, 2018.

[57] F. W. Mayer, *Narrative Politics: Stories and Collective Action*. Oxford University Press, 2014.

[58] F. Polletta and J. M. Jasper, "Collective identity and social movements," *Annual review of Sociology*, vol. 27, no. 1, pp. 283–305, 2001.

[59] P. Gerbaudo, "From individual affectedness to collective identity: personal testimony campaigns on social media and the logic of collection," *New Media & Society*, vol. 26, no. 8, pp. 4904–4921, 2024.

[60] T. Gaudette, R. Scrivens, G. Davies, and R. Frank, "Upvoting extremism: Collective identity formation and the extreme right on reddit," *New Media & Society*, vol. 23, no. 12, pp. 3491–3508, 2021.

[61] E. Khazraee and A. N. Novak, "Digitally mediated protest: Social media affordances for collective identity construction," *Social Media+ Society*, vol. 4, no. 1, p. 2056305118765740, 2018.

[62] S. Iyengar and S. J. Westwood, "Fear and loathing across party lines: New evidence on group polarization," *American journal of political science*, vol. 59, no. 3, pp. 690–707, 2015.

[63] F. Polletta and J. Lee, "Is telling stories good for democracy? rhetoric in public deliberation after 9/11," *American sociological review*, vol. 71, no. 5, pp. 699–721, 2006.

[64] A. Gordon and R. Swanson, "Identifying personal stories in millions of weblog entries," in *Third international conference on weblogs and social media, data challenge workshop, San Jose, CA*, vol. 46, 2009, pp. 16–23.

[65] W. Yao and R. Huang, "Temporal event knowledge acquisition via identifying narratives," *arXiv preprint arXiv:1805.10956*, 2018.

[66] B. Ceran, R. Karad, A. Mandvekar, S. R. Corman, and H. Davulcu, "A semantic triplet based story classifier," in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE, 2012, pp. 573–580.

[67] J. Eisenberg and M. Finlayson, "A simpler and more generalizable story detector using verb and character features," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, pp. 2708–2715.

[68] A. Dirkson, S. Verberne, W. Kraaij, A. Jorge, R. Campos, A. Jatowt, and S. Bhatia, "Narrative detection in online patient communities." in *Text2Story@ ECIR*, 2019, pp. 21–28.

[69] A. Ganti, E. A. H. Hussein, S. Wilson, Z. Ma, and X. Zhao, "Narrative style and the spread of health misinformation on twitter," in *Findings of the Association for Computational Linguistics: EMNLP 2023*, 2023, pp. 4266–4282.

[70] M. Antoniak, J. Mire, M. Sap, E. Ash, and A. Piper, "Where do people tell stories online? story detection across online communities," *arXiv preprint arXiv:2311.09675*, 2023.

[71] N. Falk and G. Lapesa, "Reports of personal experiences and stories in argumentation: datasets and analysis," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, S. Muresan, P. Nakov, and A. Villavicencio, Eds. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 5530–5553. [Online]. Available: https://aclanthology.org/2022.acl-long.379

[72] J. Baumgartner, S. Zannettou, B. Keegan, M. Squire, and J. Blackburn, "The pushshift reddit dataset," in *Proceedings of the international AAAI conference on web and social media*, vol. 14, 2020, pp. 830–839.

[73] A. Rajadesingan, C. Budak, and P. Resnick, "Political discussion is abundant in non-political subreddits (and less toxic)," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 15, 2021, pp. 525–536.

[74] P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau, "Tweeting from left to right: Is online political communication more than an echo chamber?" *Psychological science*, vol. 26, no. 10, pp. 1531–1542, 2015.

[75] R. M. Bond and M. D. Sweitzer, "Political homophily in a large-scale online communication network," *Communication Research*, vol. 49, no. 1, pp. 93–115, 2022.

[76] A. Lai, M. A. Brown, J. Bisbee, J. A. Tucker, J. Nagler, and R. Bonneau, "Estimating the ideology of political youtube videos," *Political Analysis*, pp. 1–16, 2022.

[77] R. Bond and S. Messing, "Quantifying social media's political space: Estimating ideology from publicly revealed preferences on facebook," *American Political Science Review*, vol. 109, no. 1, pp. 62–78, 2015.

[78] M. Greenacre, *Correspondence analysis in practice*. chapman and hall/crc, 2017.

[79] E. Wulczyn, N. Thain, and L. Dixon, "Ex machina: Personal attacks seen at scale," in *Proceedings of the 26th international conference on world wide web*, 2017, pp. 1391–1399.

[80] D. Oh and J. Downey, "Does algorithmic content moderation promote democratic discourse? radical democratic critique of toxic language ai," *Information, Communication & Society*, pp. 1–20, 2024.

[81] P. Rossini, "Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk," *Communication Research*, vol. 49, no. 3, pp. 399–425, 2022.

[82] A. Rajadesingan, P. Resnick, and C. Budak, "Quick, community-specific learning: How distinctive toxicity norms are maintained in political subreddits," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, 2020, pp. 557–568.

[83] W. Labov, *Language in the inner city: Studies in the Black English vernacular*. University of Pennsylvania Press, 1972, no. 3.

[84] A. Georgakopoulou, "Small stories research: A narrative paradigm for the analysis of social media," *The SAGE handbook of social media research methods. London: SAGE*, pp. 240–252, 2017.

[85] M. Bamberg and A. Georgakopoulou, "Small stories as a new perspective in narrative and identity analysis," 2008.

[86] V. Sanh, "Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019.

[87] M. N. Rizve, K. Duarte, Y. S. Rawat, and M. Shah, "In defense of pseudo-labeling: An uncertainty-aware pseudo-label selection framework for semi-supervised learning," *arXiv preprint arXiv:2101.06329*, 2021.

[88] I. Loshchilov, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.

[89] H. Lea Gaydos, "Understanding personal narratives: an approach to practice," *Journal of advanced nursing*, vol. 49, no. 3, pp. 254–259, 2005.

[90] J. Rappaport, "Narrative studies, personal stories, and identity transformation in the mutual help context," *The journal of applied behavioral science*, vol. 29, no. 2, pp. 239–256, 1993.

[91] N. J. Cox, "Stata tip 96: Cube roots," *The stata journal*, vol. 11, no. 1, pp. 149–154, 2011.

[92] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of statistical software*, vol. 67, pp. 1–48, 2015.

[93] R. Lenth *et al.*, "emmeans: Estimated marginal means, aka least-squares means. r package version 1.7. 2," 2022.

[94] A. Rajadesingan, C. Duran, P. Resnick, and C. Budak, "'walking into a fire hoping you don't catch': Strategies and designs to facilitate cross-partisan online discussions," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–30, 2021.

[95] A. Barak and O. Gluck-Ofri, "Degree and reciprocity of self-disclosure in online forums," *CyberPsychology & Behavior*, vol. 10, no. 3, pp. 407–417, 2007.

[96] A. McQueen and M. W. Kreuter, "Women's cognitive and affective reactions to breast cancer survivor stories: a structural equation analysis," *Patient education and counseling*, vol. 81, pp. S15–S21, 2010.

[97] M. Miller-Day and M. L. Hecht, "Narrative means to preventative ends: A narrative engagement framework for designing prevention interventions," *Health communication*, vol. 28, no. 7, pp. 657–670, 2013.

[98] E. Groenendyk, Y. Krupnikov, J. B. RYAN, and E. C. Connors, "Selecting out of "politics": The self-fulfilling role of conflict expectation," *American Political Science Review*, pp. 1–16, 2024.

[99] R. Kraut, *Building Successful Online Communities: Evidence-based Social Design*. MIT Press, 2012.

[100] J. M. Levine, H.-S. Choi, and R. L. Moreland, "Newcomer innovation in work teams," *Group creativity: Innovation through collaboration*, pp. 202–224, 2003.

[101] S. Kiesler, R. Kraut, P. Resnick, and A. Kittur, "Regulating behavior in online communities," *Building successful online communities: Evidence-based social design*, vol. 1, pp. 4–2, 2012.

[102] N. Andalibi, P. Ozturk, and A. Forte, "Sensitive self-disclosures, responses, and social support on instagram: The case of# depression," in *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 2017, pp. 1485–1500.

[103] G. Lakoff, *The political mind: A cognitive scientist's guide to your brain and its politics*. Penguin, 2008.

[104] L. Gilmore, *Tainted witness: Why we doubt what women say about their lives*. Columbia University Press, 2017.

[105] M. Bisgaard and R. Slothuus, "Partisan elites as culprits? how party cues shape partisan perceptual gaps," *American Journal of Political Science*, vol. 62, no. 2, pp. 456–469, 2018.

[106] A. Bächtiger, J. S. Dryzek, J. Mansbridge, and M. Warren, "Deliberative democracy," *The Oxford handbook of deliberative democracy*, pp. 1–34, 2018.

Chapter 7

Appendix

7.1 Politically Disinclined Individuals

Table 7.1: Regression coefficients for H1

	Dependent variable:
	Personal narrative or not
User type (vs most active users)	
Least politically active	0.286^{***} (0.018)
Political newcomers	0.791^{***} (0.017)
Control variables	
Is top-level comment	-0.215^{***} (0.015)
COVID-19 subreddit	1.426^{***} (0.090)
Post score (cuberoot)	0.008^{***} (0.001)
Constant	-3.214^{***} (0.045)
Observations	360,000
Log Likelihood	−88,302.770
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.2: Regression coefficients for H2

	Dependent variable:
	Participation next month
Main effects	
Used personal narrative	0.261^{***} (0.058)
Political newcomers (vs least active)	1.179^{***} (0.019)
Control variables	
Is top-level comment	0.022 (0.018)
COVID-19 subreddit	0.378^{***} (0.139)
Post score (cuberoot)	0.017^{***} (0.001)
Comment score (cuberoot)	0.013^{***} (0.003)
Interaction term	
Used personal narrative × newcomers	-0.115^{*} (0.067)
Constant	-4.638^{***} (0.099)
Observations	480,000
Log Likelihood	−66,425.730
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.3: Regression coefficients for H3

	Dependent variable:
	User active next month
Main effects	
Used personal narrative	0.388^{***} (0.015)
Political newcomers (vs least active)	-0.299^{***} (0.006)
Interaction term	
Used personal narrative × newcomers	0.090^{***} (0.017)
Constant	-1.294^{***} (0.037)
Observations	949,413
Log Likelihood	−459,251.000
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.4: Regression coefficients for H4 and H5

	Dependent variable:
	Transformed comment score
User type (vs most active users)	
Least politically active	-0.292^{***} (0.005)
Political newcomers	-0.367^{***} (0.006)
Interaction terms	
Used personal narrative × least active	0.118^{***} (0.022)
Used personal narrative × newcomers	0.135^{***} (0.021)
Control variables	
Is top-level comment	0.078^{***} (0.005)
COVID-19 subreddit	0.195^{**} (0.042)
Post score (cuberoot)	-0.002^{***} (0.000)
Constant	1.379^{***} (0.024)
Observations	360,000
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.5: Contrasts for H4

Comment type	Comparisons	Estimate	SE	z-ratio	p-value
Personal narrative	Most politically active - least politically active	0.1733	0.02171	7.980	< .0001
	Most politically active - political newcomers	0.2530	0.01992	12.699	< .0001
Other comments	Most politically active - least politically active	0.2920	0.00540	54.074	< .0001
	Most politically active - political newcomers	0.3665	0.00545	67.192	< .0001

7.2 Cross-Partisan Discussions

Table 7.6: Regression co-efficients for model evaluating H1

	Dependent variable:
	Personal narrative or not
Interaction type (vs left-left interaction)	
Right-right interaction	−0.091** (0.036)
Left-right interaction	−0.309*** (0.027)
Right-right interaction	−0.253*** (0.032)
Control variables	
log(parent length)	0.246*** (0.008)
Discussion_level	−0.323*** (0.020)
Parent score (cuberoot)	0.036*** (0.003)
Copartisan community	−0.010 (0.028)
Constant	−4.296*** (0.067)
Observations	400,000
Log Likelihood	−67,010.800
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.7: Regression co-efficients for model evaluating H2a and H2b

	Dependent variable:
	Karma score
Personal narrative indicator	0.061** (0.030)
left-right interaction (vs right-left interaction)	0.031 (0.033)
Control variables	
Post score (log)	0.006*** (0.001)
Parent score (cuberoot)	0.040*** (0.001)
discussion_level	0.166*** (0.008)
Interaction term	
Personal narrative indicator:left-right interaction	0.023 (0.034)
Constant	0.954*** (0.028)
Observations	88,140
Log Likelihood	−71,057.300
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.8: Regression co-efficients for model evaluating H3

	Dependent variable:
	Reply present or not
Personal narrative indicator	0.069* (0.041)
left-right interaction (vs right-left interaction)	0.060*** (0.012)
Control variables	
Post score (log)	0.002 (0.005)
Parent score (cuberoot)	−0.003 (0.003)
Child score (cuberoot)	−0.012 (0.008)
discussion_level	0.014 (0.015)
Copartisan community	−0.013 (0.012)
Interaction term	
Personal narrative indicator:left-right interaction	−0.061 (0.055)
Constant	−0.163*** (0.018)
Observations	200,000
Log Likelihood	−92,155.410
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 7.9: Regression co-efficients for model evaluating RQ1

	Dependent variable:
	Toxic reply or not
Personal narrative indicator	0.053 (0.080)
left-right interaction (vs right-left interaction)	0.034 (0.024)
Control variables	
Post score (log)	0.005 (0.009)
Parent score (cuberoot)	−0.002 (0.006)
Child score (cuberoot)	0.008 (0.015)
discussion_level	0.012 (0.030)
Copartisan community	0.009 (0.025)
Interaction term	
Personal narrative indicator :left-right interaction (vs right-left interaction)	−0.128 (0.110)
Constant	−1.693*** (0.037)
Observations	92,764
Log Likelihood	−27,412.340
Note:	*p<0.1; **p<0.05; ***p<0.01