Framing Social Movements on Social Media: Unpacking Diagnostic, Prognostic, and Motivational Strategies

Anonymous WiNLP submission

Introduction Social movements use social media to draw attention to their cause, disseminate information, coordinate offline action, and build collective identity (Harlow, 2012; Jost et al., 2018). Social media further provides a space for activists, bystanders and opponents to collectively construct and contest narratives of their movement via *collective action framing* (Snow et al., 1988). Focusing on Twitter (X) messages from social movements surrounding three issues (guns, immigration, and LGBTQ rights), we create a codebook, annotated dataset, and computational models to detect diagnostic (problem identification and attribution), prognostic (proposed solutions and tactics) and motivational (calls to action) framing strategies. We conduct an in-depth unsupervised linguistic analysis of each strategy and uncover cross-movement similarities in associations between framing and linguistic features such as personal pronouns. We then analyze frame variation across a broad set of sociocultural and interactional contexts.

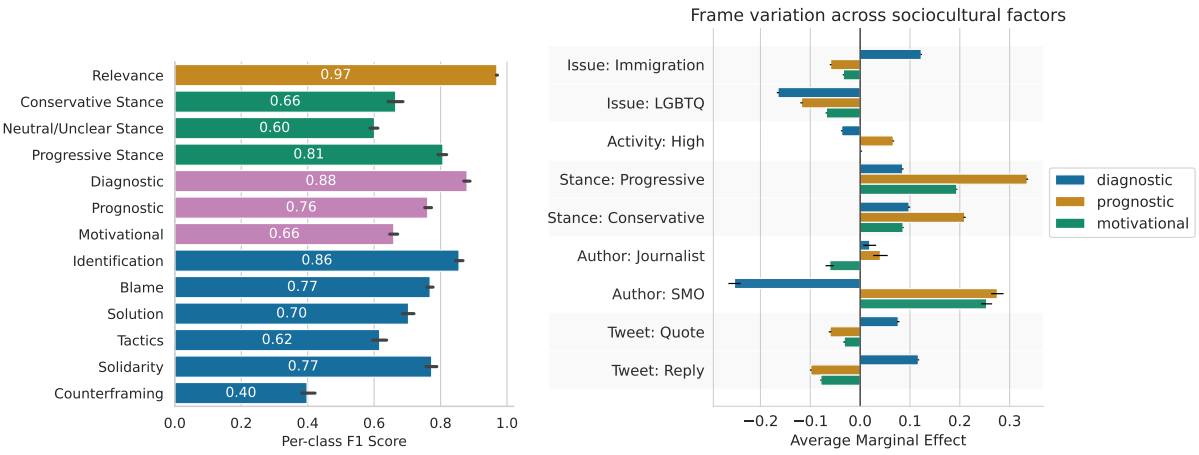
Background Social movements are sustained efforts to enact or hinder social and political changes (Jasper, 2014). *Collective action framing* is the dynamic, interactive process of constructing and negotiating shared meaning of a movement (Snow et al., 1986) and is accomplished by engaging with three *core framing tasks*: *diagnostic*, *prognostic*, and *motivational* framing (Snow et al., 1988). Social movement scholars recognize the need for large-scale, empirical and comparative framing research in addition to in-depth interpretive case studies (Snow et al., 2014), but such endeavors have been limited by the relative lack of sociologically-grounded computational approaches. In contrast, NLP framing research is typically grounded in political communication, with studies focusing on detecting issue-generic policy frames in mass media (what aspects of political issues are emphasized, e.g. *economic* or *morality*) (Boydstun et al., 2013; Card et al., 2015). Several papers detect issue-specific frames for gun violence (Liu et al., 2019) and immigration (Mendelsohn et al., 2021). While we investigate these issue areas, our conceptualization of framing differs greatly from the political communication perspective adopted in prior work.

Table 1: Annotation typology and codebook descriptions for stance, core framing tasks, and frame elements

Category	Sub-Category	Brief Codebook Description
Stance	Stance	Would you guess that the author has a progressive, conservative, or neutral/unclear attitude towards the specified issue?
Diagnostic	Identification	Does this message identify a social or political problem? *Ex: homophobia, school shootings, family separation at the border*
	Blame	Does this message assign blame for a societal problem? *Ex: to the government, corporations, socioeconomic systems*
Prognostic	Solutions	Does this message propose solutions for a societal problem? *Ex: changes in policies, political leaders, or societal norms*
	Tactics	Does this message discuss strategies or tactics for achieving a movement's goals? *Ex: protests, boycotts, petitions, contacting politicians*
	Solidarity	Does this message express support or solidarity for a movement? *Ex: celebrating a movement, honoring activists, raising visibility*
	Counterframing	Does this message explicitly challenge arguments made by the opposing side?
Motivational	Motivational	Does this message try to convince readers to join, participate in, or support a social movement through calls to action?

Data Our dataset comes from Bozarth and Budak (2022) and contains 1.85M tweets from 2018-2019 covering three issue areas: *guns*, *immigration*, and *LGBTQ rights*. We developed and iteratively refined a codebook based on theoretical definitions (Benford and Snow, 2000) and existing codebooks (Phadke et al., 2018). We code tweets for *relevance*, *stance*, and *diagnostic*, *prognostic*, and *motivational* framing strategies based on the presence of sub-categories, or *frame elements* (Table 1). Two authors annotated 750 tweets and obtained sufficient inter-annotator agreement (Krippendorff's $\alpha \geq 0.75$). The two authors proceeded with independent annotation until a total of 6,000 tweets were labeled.

Classifying framing strategies We operationalize our taxonomy as a set of four classification problems: (1) binary relevance, (2) multi-class stance, (3) multi-label core framing tasks, and (4) multi-label frame elements. We first finetune a RoBERTa-base model on the full corpus, and separately train the finetuned



(a) F1-scores for each category

(b) Effects of sociocultural factors on frame usage

model for each classification tasks. Figure (a) shows F1-scores for each category. We exclude categories with F1-scores below 0.6 (i.e. *counterframing*), and infer all other categories in the full corpus.

Linguistic properties of core framing tasks We identify linguistic features associated with each core framing task by calculating the log-odds ratio with informative Dirichlet prior within each issue area (Monroe et al., 2008). Patterns in **pronoun person marking** suggest boundary framing processes that delineate protagonists and antagonists in conflict (Snow et al., 1986). For all issue areas, 3rd person pronouns (e.g., *they, he*) are most associated with diagnostic framing, 1st person pronouns (e.g., *we, our*) with prognostic framing, and 2nd person pronouns (e.g., *you*, *your*) with motivational framing. **Moral language** is prominent: adjectives associated with diagnostic framing often express moral disapproval, such as *sick*, *disgusting*, *cruel*, and *evil*. Verbs associated with prognostic framing include deontic modals such as *need*, *should*, and *must*, which signal moral obligation.

Variation across sociocultural contexts We address how framing varies across three movement-level factors (issue area, stance, and high vs. average offline protest levels) and two message-level factors: author role (social movement organization (SMO), journalist, or other/public) and tweet type (original/broadcast tweets), quote tweets, and replies). We fit a series of logistic regression models with sociocultural factors as independent variables and core framing task presence as dependent variables.[1]

We highlight several findings with more results shown in Figure (b). First, we observe surprisingly large variation across author roles: relative to journalists and the public, SMOs are much less likely to use diagnostic framing, and much more likely to use prognostic and motivational framing. Prior work has questioned and debated the relevance of SMOs in the digital age (Earl, 2015), but these patterns suggest that SMOs still play a unique role in the online social movement ecosystem. Second, there is considerable variation across tweet types: compared to broadcast tweets, quote tweets and replies are more likely to contain diagnostic framing and less likely to contain prognostic or motivational framing. This finding emphasizes that social media meaning-making occurs not through one-sided messaging, but through conversations, with each kind of interaction offering a unique contribution to the broader discourse.

Conclusion We demonstrate the utility of synthesizing sociological theory and NLP methods for social movement content analysis. Our comprehensive descriptive analysis lays foundations for future research to address a broad range of causal questions, such as the effect of collective action framing on social movement success. Moreover, our interdisciplinary approach enables us to conduct the empirical comparative work lacking in extant social movement framing scholarship. Snow et al. (2014) argue that, "empirical investigations of framing hold the potential to influence activists' practice toward greater efficacy in mobilizing recruits and gaining media attention", suggesting that our work can hold direct implications for activist practices and strategies.

[1]We use the following reference levels: *guns* for issue area, *neutral/unclear* for stance, *average* for protest level, *other/public* for author role, and *broadcast* for tweet type.

References

Robert D. Benford and David A. Snow. 2000. Framing Processes and Social Movements: An Overview and Assessment. *Annual Review of Sociology*, 26:611–639. Publisher: Annual Reviews.

Amber E Boydstun, Justin H Gross, Philip Resnik, and Noah A Smith. 2013. Identifying media frames and frame dynamics within and across policy issues.

Lia Bozarth and Ceren Budak. 2022. Keyword expansion techniques for mining social movement data on social media. *EPJ Data Science*, 11(1):30.

Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The Media Frames Corpus: Annotations of Frames Across Issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 438–444, Beijing, China. Association for Computational Linguistics.

Jennifer Earl. 2015. The Future of Social Movement Organizations: The Waning Dominance of SMOs Online. *American Behavioral Scientist*, 59(1):35–52. Publisher: SAGE Publications Inc.

Summer Harlow. 2012. Social media and social movements: Facebook and an online Guatemalan justice movement that moved offline. *New Media & Society*, 14(2):225–243. Publisher: SAGE Publications.

James M. Jasper. 2014. *Protest: A Cultural Introduction to Social Movements*. John Wiley & Sons. Google-Books-ID: iu7cBAAAQBAJ.

John T. Jost, Pablo Barberá, Richard Bonneau, Melanie Langer, Megan Metzger, Jonathan Nagler, Joanna Sterling, and Joshua A. Tucker. 2018. How Social Media Facilitates Political Protest: Information, Motivation, and Social Networks. *Political Psychology*, 39(S1):85–118. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/pops.12478.

Siyi Liu, Lei Guo, Kate Mays, Margrit Betke, and Derry Tanti Wijaya. 2019. Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 504–514, Hong Kong, China. Association for Computational Linguistics.

Julia Mendelsohn, Ceren Budak, and David Jurgens. 2021. Modeling framing in immigration discourse on social media. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2219–2263.

Burt L Monroe, Michael P Colaresi, and Kevin M Quinn. 2008. Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403.

Shruti Phadke, Jonathan Lloyd, James Hawdon, Mattia Samory, and Tanushree Mitra. 2018. Framing Hate with Hate Frames: Designing the Codebook. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 201–204, Jersey City NJ USA. ACM.

David Snow, Robert Benford, Holly McCammon, Lyndi Hewitt, and Scott Fitzgerald. 2014. The Emergence, Development, and Future of the Framing Perspective: 25+ Years Since "Frame Alignment". *Mobilization: An International Quarterly*, 19(1):23–46.

David A Snow, Robert D Benford, et al. 1988. Ideology, frame resonance, and participant mobilization. *International social movement research*, 1(1):197–217.

David A. Snow, E. Burke Rochford, Steven K. Worden, and Robert D. Benford. 1986. Frame Alignment Processes, Micromobilization, and Movement Participation. *American Sociological Review*, 51(4):464–481. Publisher: [American Sociological Association, Sage Publications, Inc.].