Supplementary Material for "Keep Your Heads Held High Boys!: Examining the Relationship between the Proud Boys' Online Discourse and Offline Activities"

Catie Snow Bailard, Rebekah Tromble, Wei Zhong, Federico Bianchi, Pedram Hosseini, & David Broniatowski

Table of Contents

Proud Boys channels dataset on Telegram

Our dataset includes a total of 92 unique Proud Boys-affiliated public channels. Telegram allows channel owners to change channel username but not the channel ID, therefore, we observed the same channel with different usernames during the data collection process. In an effort to not function as a multiplier, for more information about these channels (i.e. the metadata of these channels as of July 2022, including the unique identification number, username, title, current count of subscribers, and biography), please email the authors.

Figure A1. Timeline of Proud Boys Channels on Telegram

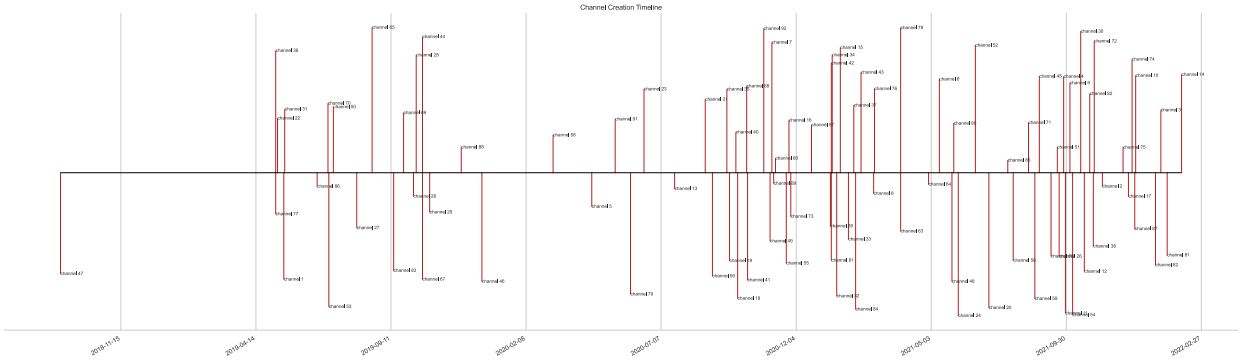
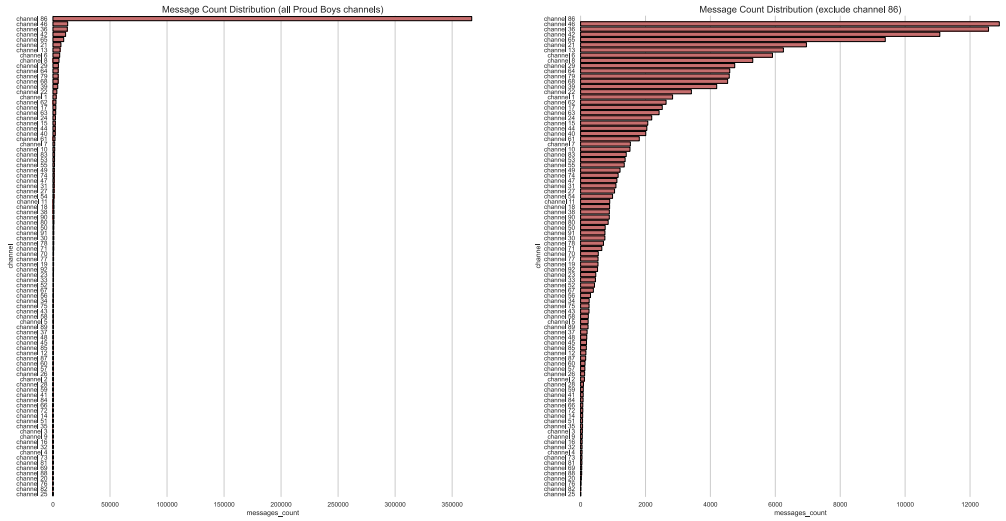


Figure A2. Number of Text Messages per Channel



Next, we display the number of text messages per channel in Figures A2. Among all the channels, “channel 86” stands out due to its higher volume of messages. To provide a

comprehensive view, we have separated the analysis into two subfigures in Figure A2: the first includes “channel 86”, while the second excludes it.

Finally, we provide a comprehensive view of the number of text messages across all channels over time, as depicted in Figure A3. In addition, we apply a log transformation to better visualize the frequency and percentage of the three main frames in Proud Boys’ Telegram conversations. This transformation and the subsequent analysis are presented in Figure A4.

Figure A3. Number of Text Messages across All the Channels over Time

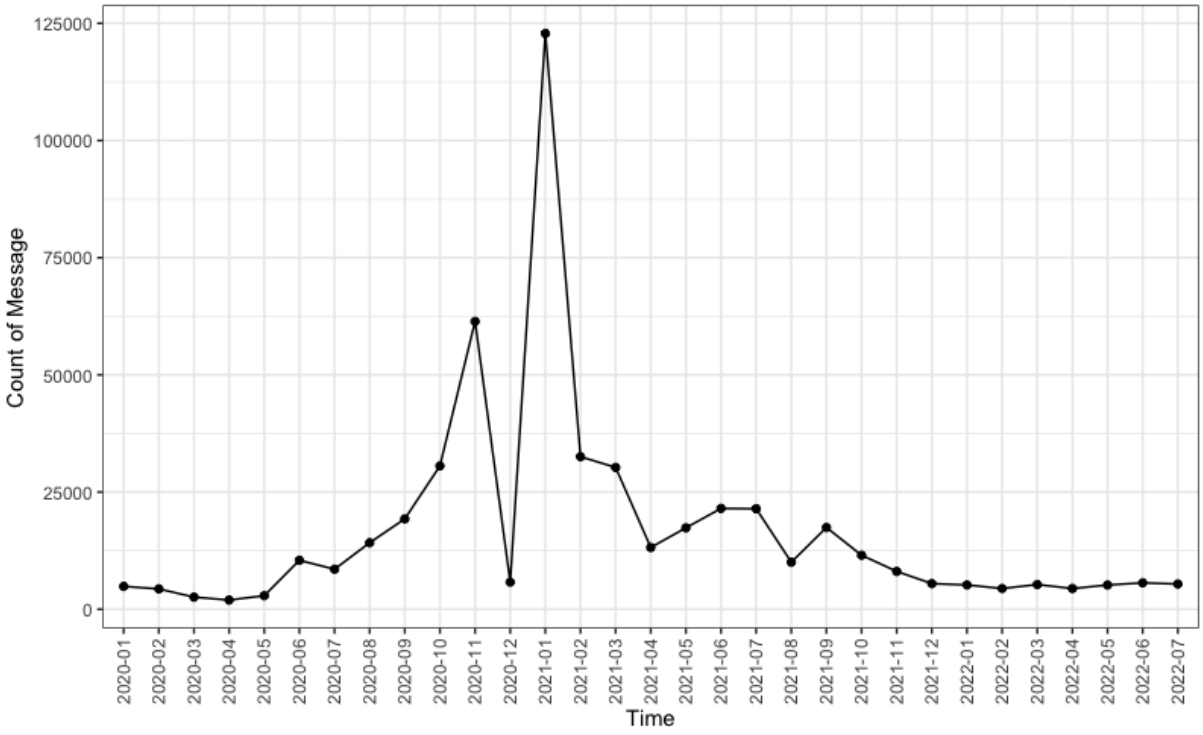
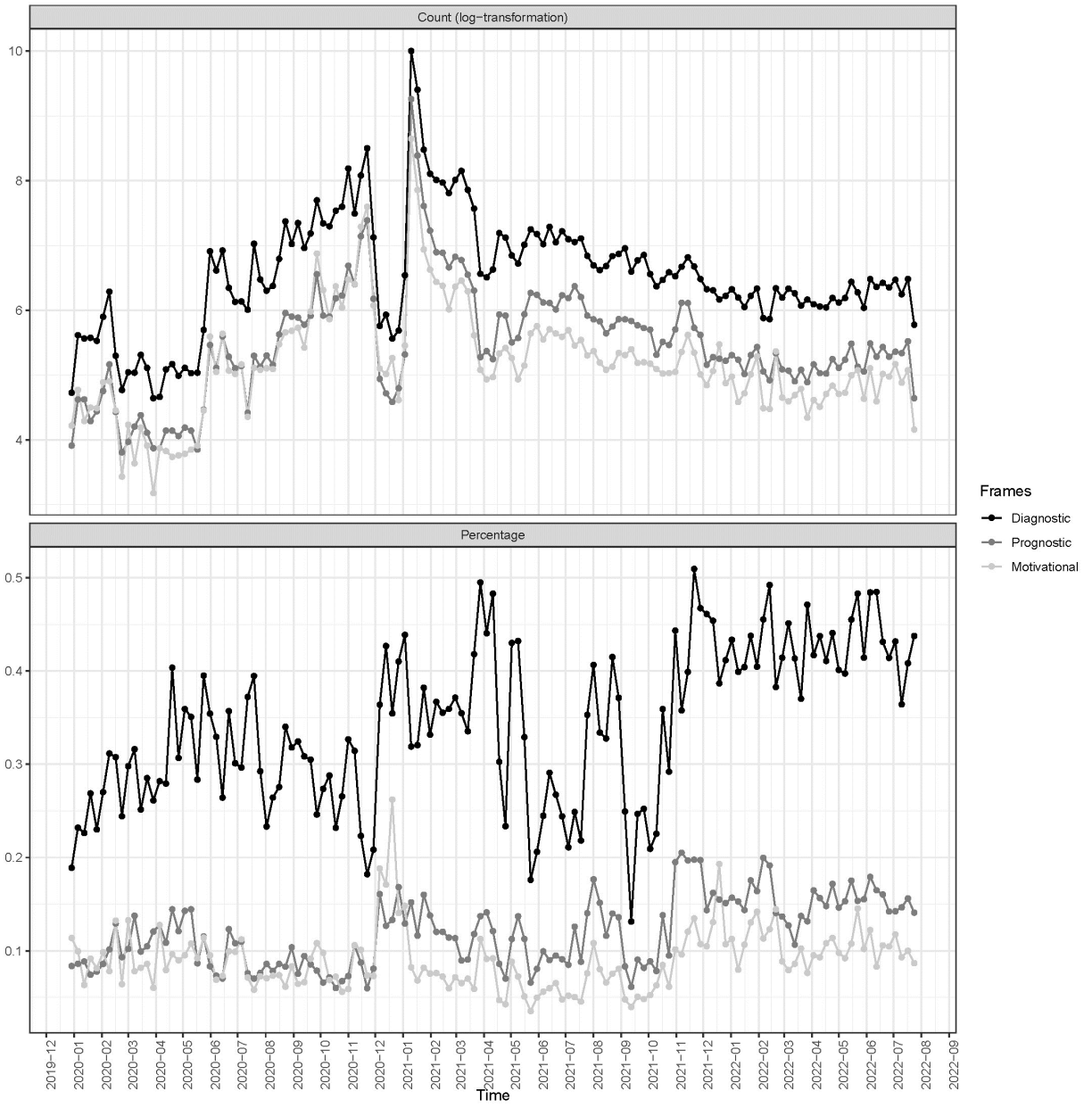


Figure A4. Frequency (log transformation) and Percentage of Collective Action Frames between January 2020 and July 2022



Collective action frames annotation framework

We provide descriptions and examples from our annotation framework for labeling the three main types of collective action frames – diagnostic (injustice and othering frame), prognostic, and motivational frames, and of the two diagnostic subtype frames (injustice and othering). For the detailed annotation framework, please contact the corresponding author (cbailard@gwu.edu).

Table A1. Description of Annotation Rules

Frame Type	Description	Examples
Diagnostic	Diagnostic frames identify a problem and/or attribute blame for a problem (Snow and Benford 1988, 200). Right-wing extremist groups often complain about censorship, threats to their racial or religious identity, the decline of certain social values or norms, societal changes that have emasculated or diminished men's traditional role in society, etc., but any text that identifies a problem constitutes a diagnostic frame.	e.g., "Everything they are doing it's like in every socialist state: standing above the law, they don't care about the people of that land, but have an own agenda, weaponizing the justice system, the intel agencies in their interest, censorship, fighting everyone who dare to oppose them. This is socialism"
Injustice	Injustice frames convey a sense of victimhood, in particular victimhood experienced by a group of people or by an individual because of their identity. An injustice frame might, for example, complain that someone/some group has been wronged, treated unfairly, slandered, or mischaracterized; has faced hypocrisy; has something that is rightfully theirs (e.g., choice, freedom, status, pride, culture, rights) taken away; is under threat; or has been (or might be) physically attacked.	e.g., "Affirmative action is reverse racism." "Our way of life is threatened." "They are trying to replace Christianity." "We can no longer tell our truth."
Othering	Othering frames cast aspersions on an identity-based out-group. They offer a negative form of collective identification. They often function to increase the perceived threat out-group poses, dehumanize its members, and/or amplify the perceived incompatibility between the in-group and out-group(s) and their ways of life.	e.g., "Maxine Waters exemplifies all the pathologies in black women today."
Prognostic	Prognostic frames "suggest solutions" to a problem and "identify strategies, tactics, and	e.g., "Let's hit the streets!" "Men should act like men again." "It's

	targets" for addressing a problem (Snow and Benford 1988, 201). Prognostic frames are signaled by words such as "need," "should," and "must." Directives also constitute prognostic frames.	time for us to take our country back." "Follow this link and learn more about our cause."
Motivational	Motivational frames provide "a rationale for action," or a "vocabulary of motives" for mobilizing (Snow & Benford, 1988, p. 202). Motivational frames serve to boost morale, pride, and/or a sense of belonging. They identify shared values, principles, priorities, norms, and/or characteristics (e.g., "Western culture", Christian values, manliness/masculinity, pro-guns, white, not (just) white); express pride in the actions of the group, its members, or affiliates; point to allies (e.g., Donald Trump) in a way that suggests strength and likelihood of success; or point to the group's legitimacy, their "right" to belong/form a group/do these things.	"POYB" ("Proud of You Boys" – a common abbreviation or hashtag used by the Proud Boys); "Uhuru" (the Swahili word for "freedom" – used as a rallying cry by the Proud Boys); "Donald Trump has our backs." "Our leader John Smith out there showing them how it's done." "We're not just white." (Helps define what the group is and isn't.)

Table A2. Examples of Collective Action Frames in Telegram Posts

Frame	Telegram Post Examples
Diagnostic *Subframes:* Injustice	"Trumps attempt to take down the deep state and expose the corrupt actors pulling the strings from the top has resulted in the country on the brink of a civil war. There is no more peaceful discourse. We watched American cities burn for 5 years while left-wing extremists coaxed red blooded Americans out of their peaceful lives to bring us to where we are now. It was all part of a plan. They intended to inflict damage and tear this country to shreds. Why you ask, what's the endgame? This is the end game. The end of America. " "WE THE PEOPLE. Deplatforming, arresting, doxxing. All these things the left is doing to us is fueling the fire. If Biden wants his "unity" why doesn't he come out and publicly say it? Why doesn't he call for peace and understanding? I'll tell you why. Because he doesn't fucking care about unity. He cares about silencing people who disagree with there socialist agenda. So here we are. 70+million Americans considered "enemies of the state" all because they are turning our country into something we have sacrificed thousands of American lives fighting against for decades. If this isn't the definition of clown world I don't know what is. Be prepared. Be ready. They are already coming after people who disagree."

Othering	"Gaggle of disgusting mutts and shitlibs try to shout down brave pro-white activists in Texas at WLM rally yesterday. This is what these 80-IQ mutts and shitlibs always do: shout down opposition with emotional psychobabble. Black/POC fragility on full display. When POCs aren't the center of attention they chimp out and throw a tantrum. They can't allow white people any room to voice support of our race because it undermines their ethno-narcissistic victim charade."
Prognostic	"If you don't share this post I don't ever want to hear y'all say you're fighting back against this oppressive government. Rufio is being wrongfully convicted along with other ProudBoys. We need all your help!" "Be ready for unwanted visitors. Give them zero information. Just shut the door. Or don't answer it." "This is what we are up against as a nation. We are in jeopardy! Every freedom loving citizen unite while you still have the choice!"
Motivational	"The Proud Boys are a force for good, I've seen it many times over the years, any group of people who organize to expose and fight wrongdoings which have come about because of NWO pushes, are always labelled the enemy, as the saying goes; Repeat a lie often enough and it becomes the truth, and that's the very tactics used, aided and abetted by MSM and politicians, crushing, infiltrating, anything they can with the sole purpose of silencing, the ordinary decent citizens stand with you's lads, keep your heads held high and never surrender!"

Results of additional statistical tests referenced in manuscript

Table A3. Results of Granger Causality Tests of Violent Events and Injustice and Othering Subframes (Figure 3)

Violent Events							
Percent of posts containing frame				*Percent of posts containing frame*			
Granger cause ->		chi2	Prob>chi2	Granger cause ->		chi2	Prob>chi2
Injustice	Violent Events	5.26	.07*	Othering	Violent Events	4.93	0.09*
Prognostic	Violent Events	2.97	0.23	Prognostic	Violent Events	3.86	0.15
Motivational	Violent Events	7.83	0.02**	Motivational	Violent Events	6.48	0.04**
Violent Events	Injustice	3.5	0.17	Violent Events	Othering	2.98	0.23
Violent Events	Prognostic	3.14	0.21	Violent Events	Prognostic	4.33	0.12
Violent Events	Motivational	0.04	0.98	Violent Events	Motivational	0.11	0.95

Table A4. Results of Granger Causality Tests of Non-violent Events and Collective Action Frames excluding Non-U.S. Channels (Footnote 11)

Non-Violent Protests							
Percent of posts containing frame				*Number (logged) of posts containing frame*			
Granger cause ->		chi2	Prob>chi2	Granger cause ->		chi2	Prob>chi2
Diagnostic	Non-Violent Protests	2.67	.26	Diagnostic	Non-Violent Protests	2.35	.31
Prognostic	Non-Violent Protests	1.41	.5	Prognostic	Non-Violent Protests	1.05	.59
Motivational	Non-Violent Protests	.03	.99	Motivational	Non-Violent Protests	.47	.79
Non-Violent Protests	Diagnostic	3.09	.21	Non-Violent Protests	Diagnostic	8.63	.01***
Non-Violent Protests	Prognostic	2.55	.28	Non-Violent Protests	Prognostic	8.59	.01***
Non-Violent Protests	Motivational	15.3	<.001***	Non-Violent Protests	Motivational	2.18	.34

Note: Significance-levels indicated as *p<0.10, **p<0.05, ***p<0.01.

Table A5. Results of Granger Causality Tests of Violent Events and Collective Action Frames excluding Non-U.S. Channels (Footnote 4)

Violent Events							
Percent of posts containing frame				*Number (logged) of posts containing frame*			
Granger cause ->		chi2	Prob>chi2	Granger cause ->		chi2	Prob>chi2
Diagnostic	Violent Events	7.34	0.03**	Diagnostic	Violent Events	2.7	.26
Prognostic	Violent Events	.11	.95	Prognostic	Violent Events	1.46	.48
Motivational	Violent Events	5.55	.06*	Motivational	Violent Events	3.56	.17
Violent Events	Diagnostic	1.73	.42	Violent Events	Diagnostic	.68	.71
Violent Events	Prognostic	2.03	.36	Violent Events	Prognostic	.82	.66
Violent Events	Motivational	.25	.89	Violent Events	Motivational	1.87	.39

Note: Significance-levels indicated as *p<0.10, **p<0.05, ***p<0.01.

Tables of the full specification of IRF figures in main document

Table A6. Table for Figure 2. IRF Plot of Non-violent Events on Percentage of Motivational Frames

Week	Lower 95%-CI	OIRF	Upper 95%-CI
0	-0.008	-0.004	0.000
1	-0.002	0.003	0.008
2	-0.002	0.003	0.008
3	-0.001	0.004	0.008
4	-0.001	0.003	0.006
5	-0.001	0.002	0.005
6	-0.001	0.002	0.004

Table A7. Table for Figure 3. IRF Plots of Violent Events and Percentage of Diagnostic, Motivational, and Prognostic Frames.

	Diagnostic			Motivational			Prognostic		
Week	Lower CI	OIRF	Upper CI	Lower CI	OIRF	Upper CI	Lower CI	OIRF	Upper CI
0	0	0	0	0	0	0	0	0	0
1	0.041	0.217	0.394	-0.155	0.015	0.186	-0.173	-0.005	0.162
2	-0.206	-0.063	0.080	0.065	0.215	0.365	-0.105	0.047	0.199
3	-0.095	0.009	0.114	0.033	0.139	0.246	-0.098	0.005	0.108
4	-0.089	-0.011	0.067	0.016	0.101	0.185	-0.091	0.000	0.091
5	-0.077	-0.015	0.048	0.005	0.074	0.144	-0.094	-0.017	0.060
6	-0.064	-0.013	0.038	-0.006	0.049	0.104	-0.089	-0.020	0.049

Table A8. Table for Figure 4. IRF Plots of Violent Events and Percentage of Injustice and Othering Frames

	Injustice			Othering		
Week	Lower CI	OIRF	Upper CI	Lower CI	OIRF	Upper CI
0	0	0	0	0	0	0
1	0.044	0.222	0.400	0.216	0.044	0.388
2	-0.151	-0.002	0.147	-0.006	-0.155	0.144
3	-0.076	0.024	0.123	0.089	-0.015	0.193
4	-0.048	0.017	0.082	0.016	-0.069	0.101
5	-0.050	0.003	0.055	0.017	-0.054	0.089
6	-0.041	0.000	0.042	0.005	-0.054	0.063

Table A9. Table for Figure 5. IRF Plot of Non-violent Events and Violent Events

Week	Lower 95%-CI	OIRF	Upper 95%-CI
0	0	0	0
1	0.013	0.190	0.367
2	-0.112	0.066	0.244
3	-0.276	-0.100	0.077
4	0.064	0.240	0.415
5	-0.021	0.089	0.199
6	-0.081	0.019	0.119

VAR tables for Granger Causality tests and IRFs reported in main document

Tables A10 & A11. VAR Results for Granger Causality Tests of Non-Violent Protests (Table 1)

Table A10	(1)	(2)	(3)	(4)
VARIABLES	**non_violent_ protest**	**diagnostic**	**prognostic**	**motivational**
L.non_violent_ protest	0.249***	0.00300	9.31e-05	0.00233**
	(0.0881)	(0.00246)	(0.000989)	(0.00102)
L2.non_violent_ protest	0.150*	-0.00102	0.000666	0.000508
	(0.0907)	(0.00253)	(0.00102)	(0.00105)
L.diagnostic	-1.361	0.314***	0.0472	0.0124
	(3.606)	(0.101)	(0.0405)	(0.0416)
L2.diagnostic	0.965	0.136	-0.0743*	-0.00818
	(3.478)	(0.0972)	(0.0390)	(0.0401)
L.prognostic	3.056	0.850***	0.578***	0.248**
	(9.082)	(0.254)	(0.102)	(0.105)
L2.prognostic	-7.510	-0.0932	0.248**	-0.194*
	(9.408)	(0.263)	(0.106)	(0.108)
L.motivational	-7.261	0.105	0.0410	0.529***
	(8.231)	(0.230)	(0.0924)	(0.0949)
L2.motivational	12.06	-0.0687	-0.00344	0.113
	(8.145)	(0.228)	(0.0914)	(0.0939)
Constant	1.539*	0.0917***	0.0255***	0.0180*
	(0.864)	(0.0242)	(0.00970)	(0.00997)
Observations	132	132	132	132
Note: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1				

Table A11.	(1)	(2)	(3)	(4)
VARIABLES	**non_violent_ protest**	**log_number_ diagnostic**	**log_number_ prognostic**	**log_number_ motivational**
L.non_violent_ protest	0.270***	-0.0169	-0.0226	0.00301
	(0.0922)	(0.0210)	(0.0219)	(0.0209)
L2.non_violent _protest	0.0868	0.00604	0.0177	0.0199
	(0.0927)	(0.0211)	(0.0220)	(0.0210)
L.log_number_ diagnostic	-0.868	0.771***	0.461*	0.306
	(1.060)	(0.242)	(0.252)	(0.240)
L2.log_number _diagnostic	0.830	0.00241	-0.240	0.0108
	(1.039)	(0.237)	(0.247)	(0.235)
L.log_number_ prognostic	0.808	0.226	0.555**	0.226
	(1.060)	(0.242)	(0.252)	(0.240)
L2.log_number _prognostic	-1.240	-0.287	0.00155	-0.471*
	(1.064)	(0.243)	(0.253)	(0.241)
L.log_number_ motivational	-0.265	0.0645	0.00423	0.397**
	(0.824)	(0.188)	(0.196)	(0.187)
L2.log_number _motivational	1.011	0.0765	0.0429	0.326*
	(0.808)	(0.184)	(0.192)	(0.183)
Constant	0.128	1.122***	0.757**	0.662*
	(1.545)	(0.353)	(0.367)	(0.350)
Observations	132	132	132	132
Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1				

Tables A12 & A13. VAR Results for Granger Causality Tests of Violent Events (Table 2)

Table A12	(1)	(2)	(3)	(4)
VARIABLES	**violent_event**	**diagnostic**	**prognostic**	**motivational**
L.violent_event	0.144*	-0.00708	-0.00325*	0.000370
	(0.0864)	(0.00465)	(0.00185)	(0.00197)
L2.violent_event	0.0477	0.00353	0.000315	0.000187
	(0.0868)	(0.00467)	(0.00186)	(0.00198)
L.diagnostic	3.946**	0.273***	0.0338	0.00958
	(1.899)	(0.102)	(0.0408)	(0.0434)
L2.diagnostic	-3.978**	0.164*	-0.0676*	-0.0133
	(1.837)	(0.0988)	(0.0394)	(0.0420)
L.prognostic	-0.554	1.013***	0.613***	0.275**
	(4.816)	(0.259)	(0.103)	(0.110)
L2.prognostic	-4.984	-0.213	0.230**	-0.189
	(5.079)	(0.273)	(0.109)	(0.116)
L.motivational	0.724	0.0329	0.0515	0.514***
	(4.111)	(0.221)	(0.0882)	(0.0939)
L2.motivational	9.604**	0.00194	0.00209	0.103
	(4.118)	(0.221)	(0.0884)	(0.0941)
Constant	0.0980	0.0967***	0.0272***	0.0251**
	(0.433)	(0.0233)	(0.00930)	(0.00990)
Observations	132	132	132	132
Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1				

Table A13.	(1)	(2)	(3)	(4)
VARIABLES	**violent_event**	**log_number_ diagnostic**	**log_number_ prognostic**	**log_number_ motivational**
L.violent_event	0.212**	0.0129	0.0202	0.0487
	(0.101)	(0.0433)	(0.0452)	(0.0428)
L2.violent_event	-0.0132	-0.0146	-0.0100	-0.0146
	(0.0938)	(0.0402)	(0.0420)	(0.0398)
L.log_number_ diagnostic	0.264	0.785***	0.478*	0.353
	(0.576)	(0.247)	(0.258)	(0.244)
L2.log_number_diagnostic	-0.859	-0.0121	-0.249	-0.0272
	(0.573)	(0.246)	(0.257)	(0.243)
L.log_number_ prognostic	-0.477	0.161	0.455*	0.113
	(0.580)	(0.249)	(0.260)	(0.246)
L2.log_number_prognostic	0.327	-0.233	0.0855	-0.351
	(0.583)	(0.250)	(0.261)	(0.247)
L.log_number_ motivational	-0.117	0.0800	0.0352	0.416**
	(0.438)	(0.188)	(0.196)	(0.186)
L2.log_number_motivational	0.990**	0.0693	0.0191	0.300*
	(0.428)	(0.183)	(0.191)	(0.181)
Constant	0.543	1.118***	0.746**	0.632*
	(0.823)	(0.353)	(0.368)	(0.349)
Observations	132	132	132	132
Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1				

Tables A14 & A15. VAR Results for Granger Causality Tests of Injustice & Othering (Table A3)

Table A14	(1)	(2)	(3)	(4)
VARIABLES	**violent_event**	**injustice_level2**	**prognostic**	**motivational**
L.violent_event	0.138	-0.00460*	-0.00324*	0.000283
	(0.0868)	(0.00261)	(0.00185)	(0.00197)
L2.violent_event	0.0674	0.00249	0.000172	0.000185
	(0.0875)	(0.00263)	(0.00186)	(0.00199)
L.injustice_level2	8.038**	0.155	0.0193	-0.00518
	(3.691)	(0.111)	(0.0786)	(0.0839)
L2.injustice_level2	-3.952	0.00708	-0.147*	-0.0188
	(3.640)	(0.109)	(0.0775)	(0.0828)
L.prognostic	-2.191	0.724***	0.632***	0.288**
	(5.090)	(0.153)	(0.108)	(0.116)
L2.prognostic	-7.375	-0.0494	0.279**	-0.185
	(5.512)	(0.165)	(0.117)	(0.125)
L.motivational	0.215	0.0242	0.0569	0.519***
	(4.184)	(0.126)	(0.0891)	(0.0951)
L2.motivational	9.684**	0.0299	0.00876	0.102
	(4.177)	(0.125)	(0.0890)	(0.0950)
Constant	0.0816	0.0228**	0.0229***	0.0245***
	(0.349)	(0.0105)	(0.00744)	(0.00793)
Observations	132	132	132	132
Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1				

Table A15	(1)	(2)	(3)	(4)
VARIABLES	**violent_event**	**othering**	**prognostic**	**motivational**
L.violent_event	0.134	-0.00331	-0.00390**	0.000626
	(0.0869)	(0.00202)	(0.00187)	(0.00197)
L2.violent_event	0.0665	0.00169	0.000813	9.02e-05
	(0.0864)	(0.00200)	(0.00186)	(0.00196)
L.othering	8.589**	0.262***	-0.0525	0.0413
	(3.977)	(0.0923)	(0.0856)	(0.0900)
L2.othering	-5.716	0.323***	0.00567	-0.0846
	(4.005)	(0.0929)	(0.0863)	(0.0907)
L.prognostic	0.638	0.236**	0.660***	0.265***
	(4.541)	(0.105)	(0.0978)	(0.103)
L2.prognostic	-7.132	-0.111	0.146	-0.171
	(4.614)	(0.107)	(0.0994)	(0.104)
L.motivational	0.450	0.0704	0.0821	0.511***
	(4.183)	(0.0970)	(0.0901)	(0.0947)
L2.motivational	8.922**	0.0143	-0.0185	0.117
	(4.194)	(0.0973)	(0.0903)	(0.0950)
Constant	-0.0427	0.0260***	0.0247***	0.0268***
	(0.402)	(0.00932)	(0.00865)	(0.00910)
Observations	132	132	132	132
Note: Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1				

Additional Details on Computational Methodology

For this analysis, we use DeBERTa v3 (large version), a cutting-edge model that is in the top leaderboards for text benchmarks: https://super.gluebenchmark.com/leaderboard.

We randomly divided the datasets into the following spits (80/10/10) for training, validation, and testing. This setup adheres to the widely-accepted standard for training models: the model is trained on the training set and evaluated during training using the validation set (multiple times per epoch). The model we use is the one that yielded the best validation score (considering the F1 score).

We report only the results computed on the test set (i.e., the data set that is completely unseen to the model).

Labels

The model is trained to predict all five labels together (since some labels co-occur, we want the model to use shared information to improve predictions). Thus, given in input a text, the model will generate 5 predictions, one for each label.

Parameters

We trained different DeBERTa models with different learning rates [1e-5, 5e-6, 8e-6, 9e-6, 5e-5]. These values come from the DeBERTa paper, see Table 10 - these are the same parameters used to fine-tune the model on other classification tasks. The results in the following section are from the "best model" at validation time.

The other parameters:

- batch_size = 8
- gradient_accumulation = 8
- eval_steps = 100
- weight_decay = 0.01
- training_epochs = 6
- warmup_steps = 50
- early_stopping = 3

Messages are truncated if longer than 200 tokens. In our dataset, the (average + 3 standard deviations) of the length is ~160 tokens.

Global Results

Table A16. Global Results on the Test Set

	P	R	F1	Support
Diagnostic	0.83	0.88	0.85	457
Prognostic	0.84	0.86	0.85	224
Motivational	0.81	0.75	0.78	204
Othering	0.78	0.74	0.76	189
Injustice	0.82	0.74	0.78	272
micro avg	0.82	0.81	0.81	1346
macro avg	0.82	0.79	0.80	1346
weighted avg	0.82	0.81	0.81	1346
samples avg	0.48	0.46	0.46	1346

Accuracy of Multi-Label Classifier

The accuracy score in the multi-label setting is around 0.68. This means that for 68% of all the messages we are able to predict all 5 classes without any error. This metric takes into account that, for each label, we have different error rates and these combine together when we look at the annotations for a single message. (For comparison's sake, the results of the same type of analysis employing a random Bernoulli of p=0.5 generates an accuracy rate of 0.028.)

Table A17. Results for Diagnostic Label on the Test Set

	P	R	F1	Support
0	0.92	0.89	0.91	762
1	0.83	0.88	0.85	457
macro avg	0.88	0.88	0.88	1219
weighted avg	0.89	0.89	0.89	1219

Table A18. Results for Prognostic Label on the Test Set

	P	R	F1	Support
0	0.97	0.96	0.97	995

1	0.84	0.86	0.85	224
macro avg	0.90	0.91	0.91	1219
weighted avg	0.94	0.94	0.94	1219

Table A19. Results for Motivational Label on the Test Set

	P	**R**	**F1**	**Support**
0	0.95	0.96	0.96	1015
1	0.81	0.75	0.78	204
macro avg	0.88	0.86	0.87	1219
weighted avg	0.93	0.93	0.93	1219

Table A20. Results for Othering Label on the Test Set

	P	**R**	**F1**	**Support**
0	0.95	0.96	0.96	1030
1	0.78	0.74	0.76	189
macro avg	0.87	0.85	0.86	1219
weighted avg	0.93	0.93	0.93	1219

Table A21. Results for Injustice Label on the Test Set

	P	**R**	**F1**	**Support**
0	0.93	0.95	0.94	947
1	0.82	0.74	0.78	272

macro avg	0.87	0.84	0.86	1219
weighted avg	0.90	0.90	0.90	1219

Confusion Matrices

Figure A5. Diagnostic

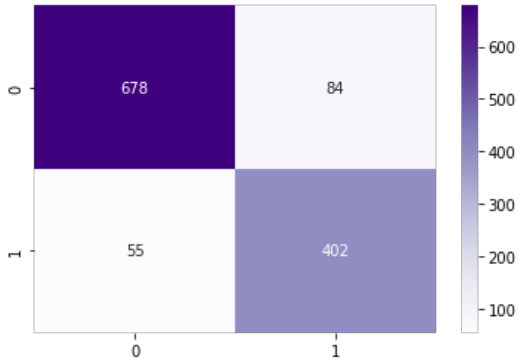


Figure A6. Prognostic

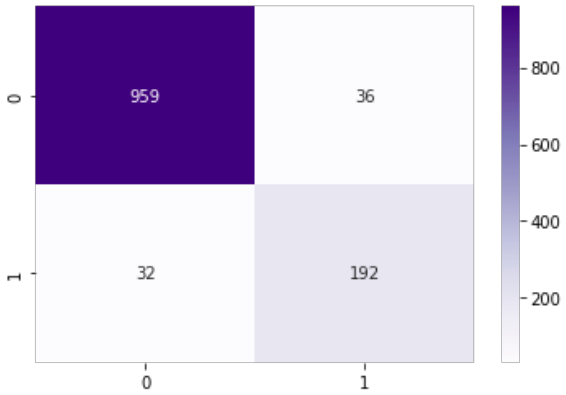


Figure A7. Motivational

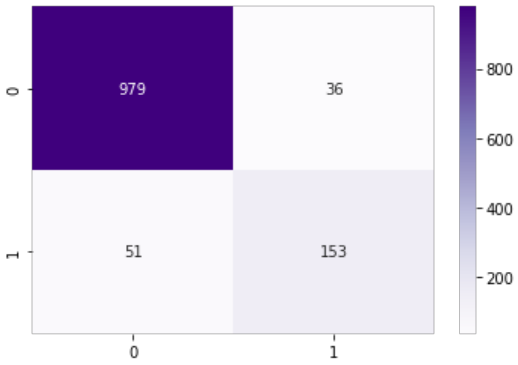


Figure A8. Othering

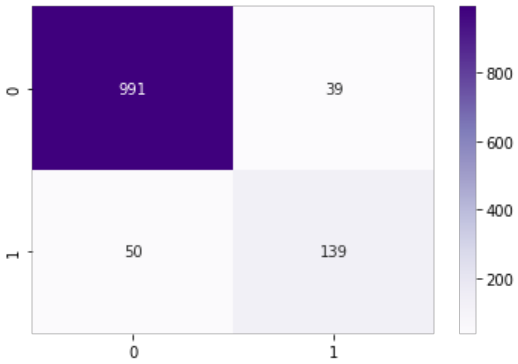


Figure A9. Injustice

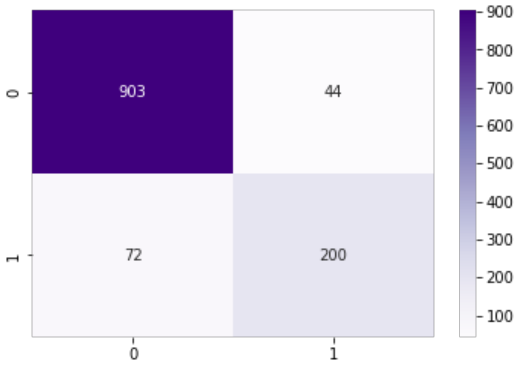


Figure A10. 8x8 Confusion Matrix

(Note: order of labels is: diagnostic, prognostic, motivational)