Selected Papers of #AoIR2022:
The 23rd Annual Conference of the
Association of Internet Researchers
Dublin, Ireland / 2-5 Nov 2022

UNCIVIL FOR CIVIL RIGHTS: A MACHINE LEARNING AND QUALITATIVE ANALYSIS OF INCIVILITY IN THE TWITTER-BASED CONVERSATION ABOUT BLACK LIVES MATTER

Rob Eschmann
Columbia University

Lei Guo
Fudan University

Jacob Groshek
Kansas State University

Philipe Copeland
Boston University

Alex Rochefort
Boston University

Introduction

The Movement for Black Lives, a social movement geared towards ending police violence against unarmed Blacks (among numerous other racial, gender, and economic calls for equity) is one of the most prominent social movements of the 21st Century, more commonly known by the #BlackLivesMatter hashtag.

Representative of a new form of social engagement, many view the Movement for Black Lives as being our most promising example of the way marginalized people have used technology to resist oppression using the tools of the oppressor (Brock, 2012; Hill, 2018).

However, the Internet remains a vast communicative forum where explicit expressions of racism are still common and often anonymous. This content and its underlying behavior are often explained by the online disinhibition effect (Lapidot-Lefler & Barak,

Suggested Citation (APA): Eschmann, R., Guo, L., Groshek, J., Copeland, P., and Rochefort, A. (2022, November). *Uncivil for civil rights: A machine learning and qualitative analysis of incivility in the Twitter-based conversation about Black Lives Matter*. Paper presented at AoIR 2022: The 23rd Annual Conference of the Association of Internet Researchers. Dublin, Ireland: AoIR. Retrieved from http://spir.aoir.org.

2012), which refers to the tendency for online communications to be more hostile and less civil than face-to-face communication.

Therefore, this paper seeks to clarify the overarching research question of: How does the online disinhibition effect shape communication and interaction about the Movement for Black Lives on social media?

Reviewing the Online Disinhibition Effect

While the online disinhibition effect was conceptualized before the development of contemporary social media platforms, much of the research encompassing the online disinhibition effect deals with the anonymity-related dimensions of online self-disclosure (Hollenbaugh & Everett, 2013) and harassment and bullying (Udris, 204; Wachs & Wright, 2018; Wright et al., 2019).

Unpacking Incivility on the Internet and Mobile Devices

Papacharissi (2004) defines incivility as “a set of behaviors that threaten democracy, deny people their personal freedoms, and stereotype social groups” (p. 267). Anderson & Huntington (2017) identify all of the following as examples of incivility: “Name-calling, mockery, negative character exaggerations through spin, or the attachment of negative emotion to another person” (p. 600).

Despite scholarly interest in civility, relatively little work has been done regarding civility on mobile devices. Mobile phones have provided expanded opportunity for disturbing forms of political dysfunction (Groshek & Cutino, 2016). This topic is the primary focus of the present research.

Methods

In carrying out this inquiry, public content from Twitter was collected through the Boston University Twitter Collection and Analysis Toolkit (BU-TCAT) that handshakes with the Twitter streaming API (Borra & Rieder, 2014) on an ongoing basis and that has archived over 375 million tweets to date (Groshek, 2014). While the BU-TCAT does not collect every tweet posted, this open-source software nonetheless captures customizable samples of public tweets and output has been shown to be generalizable to Twitter content more broadly (Gerlitz & Reider, 2013).

Using this interface, one year’s worth (5/1/15 to 5/1/2016) of tweets and accompanying metadata were downloaded for two distinct areas: (1) genetically modified organisms (GMO), and (2) black lives matter (BLM). The rationale in engaging these topics were simply on the basis that both were high on the agendas of mainstream and social media at the time of data collection, but they do not conceptually overlap explicitly in the political arena.

Using the BU-TCAT keyword search facility, GMO data comprised a total of 775,943 tweets. Data for the BLM consisted of 1,945,494 tweets. Only English tweets were considered in this study.

Coding for Incivility

One key dependent variable in this study is incivility. Previous work (e.g., Papacharissi, 2004) defined uncivil utterance as that containing (1) personal or inflammatory attacks, (2) threats, (3) vulgarities, abusive or foul language, (4) xenophobic or other hateful language or expressions, or (5) epithets or ethnic slurs, sentiments that are racist or bigoted, and/or disparaging on the basis of race/ethnicity or that assign stereotypes. Due to the large volume of datasets, we applied a dictionary-based text analysis approach and operationalized uncivil tweet as it contains one or more uncivil terms that are vulgar, abusive, foul, xenophobic, hateful, racist or bigoted.

To build a dictionary of an exhaustive list of uncivil terms, we started with the one used in Wang et al.'s (2014) study in which the authors collected cursing and uncivil words on social media from a comprehensive list of sources. New words and abbreviations were added to reflect the current trend of cursing on Twitter. A tweet will be coded as "uncivil" if it includes one of the terms provided in the updated dictionary.

The other key variable was extracted from a metadata field available through the BU-TCAT identified as "source" that made it possible for determining if a tweet originated from a mobile (phone or table) or fixed web (laptop or desktop PC) device. Tweets that were posted from other sources such as "Facebook" and "Hootsuite," which do not indicate mobile versus non-mobile devices, were removed from the analysis.

In total, 1,636,744 tweets were included in the final BLM dataset, representing 84.13% of the original dataset with all Twitter sources. For the GMO dataset, 479,921 tweets (61.85%) were included in the final analysis.

Unsupervised Learning for Coding Topics

To better understand how mobile and non-mobile users discussed GMO or BLM differently on Twitter, we applied an unsupervised topic modeling approach (Guo et al., 2016) to identify salient topics inherent in the tweets that originated from mobile and non-mobile devices, respectively. The unsupervised approach is based on the Latent Dirichlet Allocation (LDA) model (Blei, Ng, & Jordan, 2003), a statistical tool for automatically discovering latent thematic structure or topics.

Findings

Out of all the 1,636,744 tweets related to the subject of Black Lives Matter (see table 1), 1,241,618 originated from mobile devices, representing 75.87% of the population, and the rest were from non-mobile devices. Out of all the mobile-generated tweets, 80,577 contained uncivil expression, with the percentage (6.49%) about twice as much as in non-mobile tweets (3.76%).

As for the GMO dataset (see Table 2), 287,951 out of 479,921 tweets were generated from mobile devices, which account for 60.0% of all the tweets. The percentage of uncivil tweets appeared to be smaller compared with that in the BLM dataset. 2.41% of the tweets contained uncivil terms in mobile-generated tweets, while 1.03% of non-mobile tweets were considered uncivil.

Qualitative Analysis

Our findings yielded seven thematic codes, which can be seen in Table 3. While most of the themes showed up in similar percentages across the mobile and nonmobile datasets, two notable exceptions are 1) the Supporting BLM theme, which made up 22% of nonmobile but 34% of Mobile tweets, and 2) the Overtly racist / All Lives Matter theme, which was twice as likely to be used in the nonmobile dataset (26%) as the mobile dataset (13%).

Discussion

This study uses a combination of machine-learning and fine-tuned qualitative analysis to explore the online disinhibition effect in Twitter-based discourse around #BlackLivesMatter. We find that uncivil language is more common in conversations about Black Lives Matter than conversations about GMOs. In addition, this study breaks new ground in explicating the online disinhibition effect in the context of justice-oriented conversations and through different mediums (mobile vs nonmobile).

One noticeable pattern for both conversations is that people tended to share breaking news, conflicts, and content more relevant to their lives on the go, while choosing to use fixed devices when they had to discuss more serious issues and problems. Our findings may suggest that the mobile use of social media platforms not only stimulates more uncivil conservations, but also conversations that are sensational and superficial. Given most Twitter users are mobile, results of the study reinforced the belief that Twitter may not be an ideal platform for democratic discourse.

References

Anderson, A. A., & Huntington, H. E. (2017). Social Media, Science, and Attack Discourse: How Twitter Discussions of Climate Change Use Sarcasm and Incivility. *Science Communication, 39*(5), 598–620. https://doi.org/10.1177/1075547017735113

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Brock, A. (2012). From the Blackhand Side: Twitter as a Cultural Conversation. *Journal of Broadcasting & Electronic Media*, *56*(4), 529–549. https://doi.org/10.1080/08838151.2012.732147

Guo, L., Vargo, C., Pan, Z., Ding, W., & Ishwar, P. (2016). Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling. *Journalism & Mass Communication Quarterly*, *93*(2), 332–359.

Hill, M. L. (2018). “Thank You, Black Twitter”: State Violence, Digital Counterpublics, and Pedagogies of Resistance. *Urban Education*, *53*(2), 286–302. https://doi.org/10.1177/0042085917747124

Hollenbaugh, E., & Everett, M. (2013). The Effects of Anonymity on Self-Disclosure in Blogs: An Application of the Online Disinhibition Effect. *Journal of Computer-Mediated Communication,18*(3), 283-302.

Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, *28*(2), 434–443. https://doi.org/10.1016/j.chb.2011.10.014

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society, 6*, 259–283.

Udris, R. (2014). Cyberbullying among high school students in Japan: Development and validation of the Online Disinhibition Scale. *Computers in Human Behavior, 41*(C), 253-261.

Wachs, S., & Wright, M. (2018). Associations between Bystanders and Perpetrators of Online Hate: The Moderating Role of Toxic Online Disinhibition. *International Journal Of Environmental Research And Public Health, 15*(9).

Wang, W., Chen, L., Thirunarayan, K., & Sehth, A. P. (2014). Cursing in english on twitter. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. Presented at the ACM.

Wright, M. F., Harper, B. D., & Wachs, S. (2019). The associations between cyberbullying and callous-unemotional traits among adolescents: The moderating effect of online disinhibition. *Personality and Individual Differences*, *140*, 41–45. https://doi.org/10.1016/j.paid.2018.04.001

Table 1: “Topics” in Twitter Coverage of Black Lives Matter

	Mobile tweets: “Topic” and associated words	Label
1	activist, arrest, call, charlestonshoot, creat, ferguson, group, hate, hillari, love, media, protest, racism, racist, ralli, shit, support, tcot, terrorist, trump, white	Activism/protests (e.g., Charleston shooting, Ferguson unrest) + political candidates for the 2016 U.S. election
2	believ, can, cold, confirm, cop, forget, heart, just, justice, justiceformonroebird, knew, monroebird, peopl, realli, sandrabland, sayhernam, see, still, tell, wrong, wtf	The #SayHerName campaign
3	berni, beyonc, come, dai, death, great, kill, messag, movement, name, new, old, power, protest, real, rememb, sander, shut, stand, start, time, todai, video	Sanders and Beyoncé on BLM
4	alllivesmatt, black, fight, folk, fuck, know, life, like, live, matter, mean, peopl, rainforest, sai, save, think, try	Black lives matter vs. All lives matter
5	america, american, black, brutal, call, case, cop, even, happen, hear, help, import, man, mckinnei, murder, need, never, now, offic, pleas, polic, right, sad, shot, stop, thing, unarm	Police brutality against African American people
	Non-mobile tweets: “Topic” and associated words	**Label**
1	call, group, hate, know, leader, make, media, movement, never, obama, race, see, start, tcot, terrorist, thug, want, white, will	Obama on BLM
2	anoth, arrest, baltimor, black, blacktwitt, chicago, cop, death, kill, life, man, mckinnei, murder, offic, old, polic, rememb, shoot, shot, tamirric, women, year	Black Twitter + BLM protests in multiple places
3	america, american, can, charlestonshoot, dai, fight, happen, help, just, love, march, must, need, new, now, nyc, pleas, racism, racist, read, right, stand, take, thank, time, today, truth, video, watch	Fight against racism in America
4	activist, attack, berni, candid, clinton, feelthebern, ferguson, good, hillari, protest, ralli, right, sander, show, shut, support, talk, trump, video, vote	Political candidates for the 2016 U.S. election
5	alllivesmatt, believ, black, come, demand, folk, fuck, investing, justic, know, like, live, matter, mean, need, people, realli, sai, sandrabland, sayhernam, tell, think, time	Black lives matter vs. All lives matter + The #SayHerName campaign

Table 2: "Topics" in Twitter Coverage of GMO

	Mobile tweets: "Topic" and associated words	**Label**
1	antibiot, carb, enjoi, fat, final, free, gluten, low, meal, non, nongmo, organ, raw, settl, soi, vegan	Viral tweet: "Enjoying my vegan, gluten free, non GMO, soy free, antibiotics free, raw, organic, fat free, low carb meal."
2	act, bill block, boycott, bui, call, consum, countr, fight, food, help, know, label, law, live, mandator, product, right, righttoknow, state, stop, support, vote, want, will	House passes bill to prevent mandatory GMO food labeling
3	anti, beer, berni, big, don, even, farm, isi, kill, let, make, monsanto, need, peopl, poison, pro, protect, scienc, take, tell, world	Sanders on GMO, farming and Monsanto
4	can, cancer, caus, dai, eat, fact, feed, food, gaga, genet, get, ladi, like, love, modify, money, natur, oscar, safe, think, time, year	GMO's linkage to cancer
5	farmer, glyphos, health, just, million, monsanto, new, now, pesticide, plant, roundup, sai, seed, show, studi, talk	Health problems linked to Monsanto's roundup
	Non-mobile tweets: "Topic" and associated words	**Label**
1	america, big, bring, clean, cleanfood, food, gmofre, gpdb, healthi, isbad, know, let, look, non, now, offer, organ, partner, right, secret, sell, tpp, trade	The TPP and global trade of GMOs
2	act, bill , block, boycott, call, company, congress, dark, darkact, hous, know, label, labelgmo, law, mandator, new , pass, petit, product, protect, requir, righttoknow, sai, senat, sign, state, stop, support, tell, vote, will	House passes bill to prevent mandatory GMO food labeling
3	antibiot, can, carb, dai, don, eat, farm, fat, free, get, gluten, help, just, like, low, make, meal, non, nongmo, organ, pleas, raw, safe, salmon, soi, vegan, win	Viral tweet: "Enjoying my vegan, gluten free, non GMO, soy free, antibiotics free, raw, organic, fat free, low carb meal."
4	ban, cancer, caus, contamin, corn, crop, end, environ, farmer, genet, glyphos, health, modifi, new, pesticid0, plant, roundup, seed, studi, usda	GMO's linkage to cancer and other health and environmental problems
5	anti, berni, countri, hillari, kill, monsanto, need, peopl, poison, pro, scienc, take,.vaccin, world, year	Sanders and Clinton on GMO

Table 3 Themes in uncivil BLM conversation

Codes	Nonmobile		Mobile	
	# of tweets	% of tweets	# of tweets	% of tweets
Supporting BLM	236	22%	156	34%
Challenging BLM	140	13%	67	15%
Overtly racist / Anti-BLM	275	26%	61	13%
Popular Culture and Media / Fake Activism	28	3%	22	5%
Learning/Education about BLM	89	8%	50	11%
Solutions, Raising Awareness to BLM	52	5%	22	5%
Activism/BLM in Media	215	20%	78	17%
Clickbait / other unrelated	38	3%	0	0%
Total	1073	100.00%	456	100%