

# Social Media Participation in an Activist Movement for Racial Equality

Munmun De Choudhury<sup>†</sup>, Shagun Jhaver<sup>†</sup>, Benjamin Sugar<sup>†</sup>, Ingmar Weber<sup>§</sup>

<sup>†</sup> Georgia Institute of Technology, <sup>§</sup> Qatar Computing Research Institute, HBKU  
{munmund, jhaver.shagun, bsugar}@gatech.edu, iweber@qf.org.qa

## Abstract

From the Arab Spring to the Occupy Movement, social media has been instrumental in driving and supporting socio-political movements throughout the world. In this paper, we present one of the first social media investigations of an activist movement around racial discrimination and police violence, known as “Black Lives Matter”. Considering Twitter as a sensor for the broader community’s perception of the events related to the movement, we study participation over time, the geographical differences in this participation, and its relationship to protests that unfolded on the ground. We find evidence for continued participation across four temporally separated events related to the movement, with notable changes in engagement and language over time. We also find that participants from regions of historically high rates of black victimization due to police violence tend to express greater negativity and make more references to loss of life. Finally, we observe that social media attributes of affect, behavior and language can predict future protest participation on the ground. We discuss the role of social media in enabling collective action around this unique movement and how social media platforms may help understand perceptions on a socially contested and sensitive issue like race.

## Introduction

Racial inequality in the criminal justice system has been a consistent issue of socio-political significance in the United States. Empirical and official data, although scarce, indicates that over half of those killed by police in recent years have been black or latino (Parker, Onyekwuluje, and Murty 1995; Klinger 2011). Criminological theories such as “broken windows” have been argued to be behind the disproportionate rate of fatal police encounters, and broader perceptions of injustice in black communities (Goldkamp 1976; McKinley and Baker 2014).

In the last few years, conversations around the criminal justice system’s contributions to racial disparity against blacks have gathered renewed national attention. Following the acquittal of George Zimmerman in the death by shooting of black teen Trayvon Martin in Florida in 2013, three black community organizers, Alicia Garza, Patrisse Cullors-Brignac, and Opal Tometi started an activist movement with the use of the hashtag #BlackLivesMatter on

social media<sup>1</sup>. The Black Lives Matter (BLM) movement grew into a social juggernaut following the 2014 deaths of Michael Brown in Ferguson and Eric Garner in New York City (Bonilla and Rosa 2015). Over time, BLM has expanded its fight beyond racial police violence to situate itself as “an ideological and political intervention” (Garza 2014; Bonilla and Rosa 2015) that strives to end systemic presence of racial inequality against blacks.

Social media, especially Twitter, due to its pervasiveness and adoption, has provided the fundamental infrastructure to this activist movement. Cullors-Brignac, one of the co-founders of the movement, reported to the CNN<sup>2</sup>: “*Because of social media we reach people in the smallest corners of America. We are plucking at a cord that has not been plucked forever. There is a network and a hashtag to gather around. It is powerful to be in alignment with our own people.*” Similarly, and in contrast to the Civil Rights Movement (1954-1968), activist DeRay McKesson noted<sup>3</sup>: “*The tools that we have to organize and to resist are fundamentally different than anything that’s existed before in black struggle.*”

While BLM began online, the organization has since branched out into chapters in 31 cities and held protests, rallies, and boycotts across US and internationally<sup>2</sup>. However, little is understood regarding the relationship between the expanse of the movement online and its growth in the offline world, where both the demonstrations and the conditions they are in response to, take place. There is a lack of sufficient empirical evidence of how the social media activity around BLM relates to the historical discriminatory police actions. In this paper, we leverage the role of social media as a powerful lens to unpack the complex nuances of thoughts, opinions and sentiments that characterize this notable activist movement in different parts of the country.

Understanding community-wide expressions and thoughts of solidarity and agitation around as sensitive an issue as racial violence has typically been challenging. The Civil Rights movement, the Black Power movement, and the Black Feminist movement all provide us with a historical context around which most such investigations have been

<sup>1</sup>[https://en.wikipedia.org/wiki/Black\\_Lives\\_Matter](https://en.wikipedia.org/wiki/Black_Lives_Matter)

<sup>2</sup><http://www.cnn.com/2015/12/28/us/black-lives-matter-evolution/>

<sup>3</sup><http://www.wired.com/2015/10/how-black-lives-matter-uses-social-media-to-fight-the-power/>

based (Bridges and Crutchfield 1988; Sigelman et al. 1997; Bonilla-Silva 2006). However their scale and scope have been limited due to difficulty in gathering reliable data. Our work allows us to examine with computational rigor and large-scale data, spatio-temporal patterns of these perceptions manifested in social media within the context of BLM. We address the following research questions:

- *RQ 1:* What are the temporal characteristics of social media participation in the BLM movement?
- *RQ 2:* How do engagement and linguistic attributes of this participation manifest geographically, specifically in regions of high historical police violence against blacks?
- *RQ 3:* In what ways do these social media attributes of engagement and language around BLM relate to protests that unfolded on the ground?

Our results show that social media reflects the evolution of the BLM movement—the movement has kept on gaining newcomer attention in large volumes as different events unfolded between 2014 and 2015. We also find considerable continued participation, associated with increased social orientation over time. Significant geographical differences further characterize engagement and linguistic expression in this movement; states of high police violence exhibit greater negative affect and references to death and loss. Finally, we show that engagement and linguistic attributes gleaned from Twitter around BLM can predict well the size of the protests that commenced throughout the country. Emergence of a collective identity and, somewhat surprisingly, lowered anger and anxiety are observed to be prominent predictors of future protests.

We situate our findings in the context of collective action and activism around social movements. We also discuss the role of social media as a sensor for quantifying discourse around sensitive topics like race and societal violence.

## Background and Prior Work

### Race and Police Violence

Considerable prior literature, especially in criminology and sociology, has investigated the relationship between racial minorities and law enforcement agencies (Goldkamp 1976; Sampson and Lauritsen 1997; Sigelman et al. 1997). Broadly, these studies try to either explain the reasons for racial disproportionality in police violence (see (Kennedy 1998) for a comprehensive review), or study the impact on the affected communities (Bridges and Crutchfield 1988). Our work differs from these in that we examine the expressions of individuals in different parts of the country around the issues of racial conflict and misconduct subject to law enforcement, as observed through the lens of social media.

Further, the different BLM protests that commenced in different parts of the US and the world in response to racial police violence are, in many ways, unique compared to prior activism on racial inequality, such as the Civil Rights Movement (Kennedy 1998). The BLM protests were highly decentralized but coordinated, without any formalized hierarchical structure, and were often led by different groups of people in geographically disparate locations. The protests

were also influenced by a sequence of several events of police brutality, instead of one single incident. These distinctions, on their own right, render examining this movement through the lens of social media valuable.

### Social Movements

As a social movement, BLM is related to recent movements such as the Indignados in Spain, the Occupy movement in the US, and even the Arab Spring. All of these made heavy use of online social media for mobilization, and have been extensively studied (Lotan et al. 2011; Eltantawy and Wiest 2011; González-Bailón et al. 2011; Tufekci and Wilson 2012; Gleason 2013; Borge-Holthoefer et al. 2015). The bulk of this existing research has focused on characterizing the social and temporal dynamics as observed on social media, identifying influencers and modeling influence, and assessing different roles adopted by participants. To our knowledge, there is limited work examining a socio-politically contested activist movement springing from racial inequality, (i.e., BLM), as well as exploring the relationship of such a movement to offline events.

Nevertheless, our work is motivated by several investigations undertaken in this prior work. Studying the Occupy movement, Conover et al. (2013) explored participation from individuals who continued to engage in Twitter discourse around the movement. They found that while initial online participation on Twitter tended to come from highly-interconnected users with pre-existing interests, the users seemed to have lost interest in later phases. Bastos et al. (2015) examined whether online mobilization can be predictive of onsite protests and report mixed results across different instances (also see (Weber, Garimella, and Batayneh 2013)). Similarly, Varol et al. (2014) studied the Gezi Park movement and found Twitter activity to mirror geographic cues, and individual behavior to be affected by offline political events. Part of our temporal analysis also looks at links between offline events and online mobilization, thus contributing to an under-explored dimension in quantitative work on social movements and social media.

Topically, close to our work is the recent work of Olteanu et al. (2016), who studied the demographics of Twitter users who used the hashtag #BlackLivesMatter and found blacks to be more engaged with the hashtag. Though this work examines an offline context of the movement using demographics, our work differs in that we use statistics about police violence as well as about the protests as offline data, and include nuanced linguistic analysis for the social media data.

### Communal Coping

Societal racial inequality in general and incidents of police violence in particular, along with the ensuing protests can be viewed as a collective upheaval that triggers a coping mechanism. Previous research in crisis informatics has demonstrated the role of social media as a lens to understand how society copes with crises and how communities leverage these tools for collective action (Vieweg et al. 2010; Mark et al. 2012; Starbird and Palen 2012; De Choudhury, Monroy-Hernandez, and Mark 2014). A number of different kinds of crises events have been examined —

from natural ones: earthquakes and floods, to man-made ones: wars and terrorist attacks (Palen and Vieweg 2008; Cheong and Lee 2011; Mark et al. 2012; Glasgow, Fink, and Boyd-Graber 2014). It is known that traumatic events like these are often followed by social sharing, seeking of social support, changes in social interactions, and an increased collective orientation (Pennebaker, Mayne, and Francis 1997). Language analysis has been observed to be particularly valuable in studying communities in crises; there is an increased concern with social dynamics in the wake of these events that is reflected in frequent references to others, and a higher rate of second-person, third-person, and first-person plural pronoun usage (Cohn, Mehl, and Pennebaker 2004).

## Data and Methods

### Social Media Data

We collected data from Twitter around the BLM activist movement in four phases, corresponding to four major related events. An outline of the major events around which our data was collected is given in Table 2; these events were compiled from the BLM Wikipedia page<sup>1</sup>. Specifically, we utilized these events to identify relevant and major hashtags around which Twitter discourse on the movement commenced. Then we adopted a snowball approach to gather additional hashtags that co-occurred with the ones directly linked to the events. All of our datasets were collected using the Twitter streaming API.

Dataset	Posts	Users	Time Range
Ferguson I	13,260,158	2,868,315	8/08/14-8/27/14
Ferguson II	10,181,369	2,301,962	11/11/14-12/10/14
NYC	884,593	309,946	12/17/14-1/19/15
Baltimore	4,644,058	1,311,973	4/12/15-5/06/15

Table 1: Descriptive statistics of our Twitter data.

(1) Our first dataset (Ferguson I) spans between Aug 8 and Aug 27, 2014. The dataset was obtained via the single hashtag #ferguson . As can be observed in Table 2, this dataset aligned with the timeline of events around the shooting and killing of Michael Brown by police officer Darren Wilson.

(2) Our second dataset (Ferguson II) contains posts between Nov 11 and Dec 10, 2014, and was also obtained using #ferguson as the filtering hashtag. It captured posts made during the protests occurring in response to a grand jury decision that did not indict the police officer Darren Wilson.

(3) Our third dataset (NYC) included posts made between Dec 19, 2014 and Jan 19, 2015 around two counter protests that arose following the incidents in Ferguson. The first was from the police supporters who felt police officers were being unfairly depicted, and the realities of the danger or police work were not being recognized. These protests were captured on Twitter with the hashtag #BlueLivesMatter. Then, on Dec 20, 2014, New York police officers Liu and Ramos were shot and murdered by Brinsley; Brinsley claimed the murders to be in response to the “pro-police” protests. To capture these events in NYC, we also collected posts with the hashtags #BlackLivesMatter and #AllLivesMatter.

Ferguson I
Aug. 9, 2014: Michael Brown is killed by police officer Darren Wilson
Aug. 14, 2014: President Obama releases statement on events
Aug. 18, 2014: National Guard is ordered to Ferguson
Aug. 20, 2014: Police release name of shooting officer (Darren Wilson) and surveillance tape. Protesters come to Ferguson from across the nation
Aug 25, 2014: Michael Brown’s funeral
Ferguson II
Nov. 17, 2014: Governor of Missouri declares state of emergency
Nov. 24, 2014: Grand jury decides not to indict Wilson
Nov. 25, 2014: protests erupt in 170 cities across the U.S.
NYC
Dec. 19, 2014: “Pro-police” counter protests
Dec. 20, 2014: Ismaaiyl Abdullah Brinsley kills two on-duty NYPD officers Wenjian Liu and Rafael Ramos
Dec. 21, 2014: Sergeants Benevolent Assoc. blames Mayor DeBlasio on Twitter, Blue Lives Matter tweets call for Mayor’s resignation
Dec. 27, 2014: Funeral for Rafael Ramos, thousands of people attend
Jan. 4, 2015: Funeral for Wenjian Liu
Baltimore
April 19, 2015: Freddie Gray dies from injuries during an arrest on Apr 12
April 27, 2015: An image circulates on social media urging high schoolers to gather for a “purge”
April 27, 2015: Police pre-empt protests at the high school with riot gear
April 27, 2015: Governor declares emergency, activates National Guard
April 29, 2015: Solidarity protests reported in at least 10 other cities
May 3, 2015: Curfew is lifted

Table 2: Timeline of events captured in our Twitter Data.

(4) Finally, our last dataset (Baltimore) between Apr 12 and May 6, 2015 consisted of posts shared on Twitter during the demonstrations that followed the death of Freddie Gray while in police custody for allegedly possessing an illegal switchblade. In addition to using the hashtags as in NYC, we added #Baltimore, #BaltimoreRiots, #BaltimoreUprising, and #FreddieGray.

There were two conditions in which we supplemented our datasets by using a service called Topsy, which has an archive of every post on Twitter since 2009. The first was in capturing posts on any dates relevant to the sequence of events that were missed due to the inherent unpredictability of when the event would begin. The second was in correcting for a program error which resulted in noticeably less posts for NYC after Dec 25. To avoid over representation, we crawled the service such that the yield would be equivalent to the Twitter API. Table 1 gives basic statistics of the four datasets; Figure 1 gives distribution of the number of Twitter posts and unique users over time.

Following this data collection, we inferred US state-level location information from each Twitter post. We utilized the user reported location string which is known to yield more geographically-mapped posts than relying on the small amount of GPS-tagged tweets (Hecht and Stephens 2014). We made use of the Nominatim library (<http://wiki.openstreetmap.org/wiki/Nominatim>) for this task. We were able to infer state level location information for 3,100,132 posts in Ferguson I (23.4%), 2,165,630 (21.3%) in Ferguson II, 228,189 (25.8%) in NYC, and 1,079,068 (23.2%) in Baltimore.

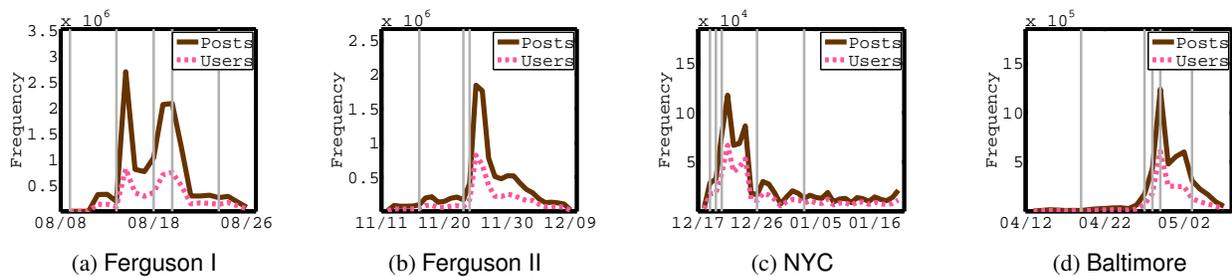


Figure 1: Post and user distribution over time. Vertical gray lines indicate important events (ref. Table 2).

### Police Shooting Data

Next, we obtained data on deaths attributed to police shootings. We utilized a police shooting dataset made available by Fatal Encounters (FE: <http://www.fatalencounters.org/>). FE includes information on just over 10,000 records of police killings since January 1, 2000. As of June 15, 2015, 85% percent of the data has been submitted by paid researchers, and all data submitted by volunteers is verified twice against published media reports. Each record in the FE database includes details about the location, time and cause of police shooting incident and race of the person being shot.

**Rate of Police Killing of Blacks Index (PK).** Combining FE data and state-level Census population data (<http://www.census.gov/popest/>), we calculated a per-state index we refer to as the rate of police killing of blacks (PK): defined as the ratio between the number of black citizens killed by a police officer within a given state (per million) to the total black population within that state (per million). Figure 2(a) gives a state-level representation of this index.

### Protest Data

Finally, we collected data on the number of individuals who participated in protests and demonstrations that took place in the US relating to BLM. As a starting point, we referred to a website Elephrame (<https://elephrame.com/textbook/protests>), which provides a compilation of BLM related protests between July 2014 and December 2015; they also report the number of individuals who participated in each of the reported protests, as well as a link to a source.

Using various sources of event timelines (e.g., Wikipedia), and comprehensive searches on Google for demonstrations reported for each day of the four datasets, we cross-checked and expanded this compilation. In cases where the reported volumes were inexact (e.g., “hundreds of ...”), we defaulted to the lowest possible definition. When multiple articles had conflicting reports for the same demonstration, we used the mean of all reported volumes. This data collection gave us 30,371 individuals involved in protest events during the time period of Ferguson I, 40,165 during Ferguson II, 63,744 corresponding to NYC and 24,370 during Baltimore.

**Protest Volume of Participating Individuals (PV).** Using this data, we define a daily measure we refer to as protest volume of participating individuals (PV). It is given as the

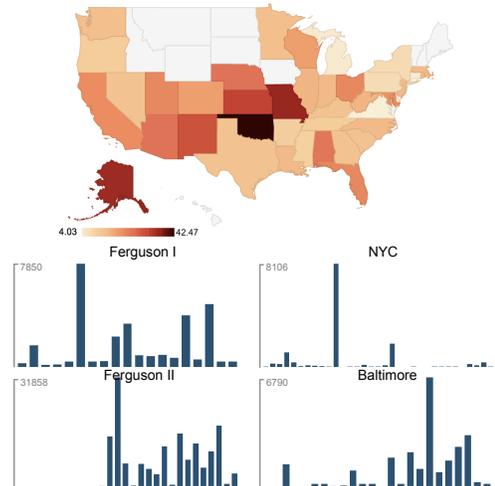


Figure 2: (a) (Top) Rate of Police Killing Index (PK) visualized over US states. MT, ID, ND, SD, WY, NE, NH, VT, ME and HI had no reported data. (b) (Bottom) Protest Volume (PV) in terms of total number of individuals involved ( $y$ -axis) over time ( $x$ -axis).

total number of individuals involved in all BLM protests on a certain day throughout the US.

### Measures

We now define a number of measures that are used in our investigations. These measures are based on prior work and capture aspects of affective expression, linguistic style, behavior, interpersonal interaction and psychological state of individuals from content shared on social media.

Our first measures are activity and engagement attributes. We utilized several content sharing, interpersonal interaction and information propagation related indicators as measures in this category: number of posts shared, number of unique users, number of retweets, number of @-reply posts, and number of link-bearing posts.

Focusing on the actual nature of shared content in posts, we considered two measures of affect: positive affect (PA), and negative affect (NA), and four other measures of emotional expression: *anger*, *anxiety*, *sadness*, and *swear*. These measures are computed using the psycholinguistic lexicon LIWC (Chung and Pennebaker 2007).

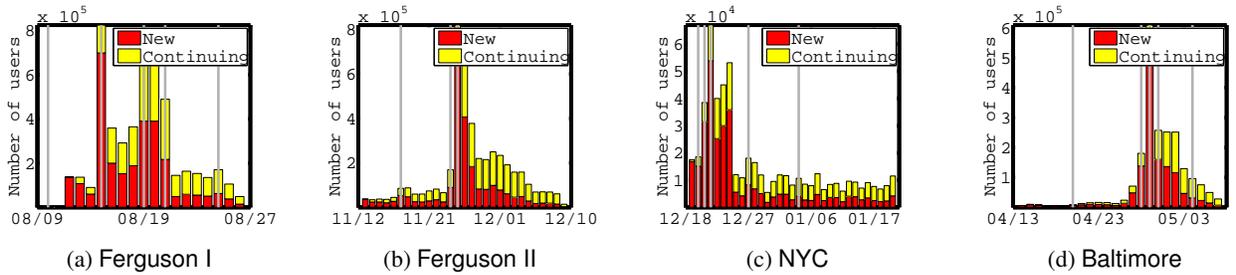


Figure 3: Distribution of new and continuing users over time. Vertical gray lines indicate important events (ref. Table 2).

We further used LIWC to define cognitive measures, comprising *cognitive mech*, *discrepancies*, and *negation*. Perception attributes consisted of the LIWC categories *death*, *see*, *hear*, *feel*, and *percept*. Next, we considered three measures of social orientation: *social*, *family*, and *friends*. Interpersonal awareness was assessed based on the frequency of usage of *1st person singular*, *1st person plural*, *2nd person*, and *3rd person* pronouns. Finally, we used two types of measures of psychological distancing (Cohn, Mehl, and Pennebaker 2004): Temporal references in Twitter content measured based on the use of *past*, *present*, and *future* tenses, and function word use based on the occurrence of *verbs*, *adverbs*, *articles*, *prepositions*, and *conjunctions*.

## Results

### RQ 1: Evolution of Participation

Per RQ 1, we begin by examining the evolution of participation of Twitter users across the four major events.

**New and Continuing Users.** First, we assess the growth of the community involved in discourse on this activist movement. We define two kinds of users over time—new users and continuing users. On a given day  $t_i$ , new users are those individuals for whom  $t_i$  was the first day they were observed to share a post in the entirety of our four datasets. In contrast, continuing users are those who shared a post on day  $t_i$ , and have at least one other post on a day  $t_m$  prior to  $t_i$  ( $m < i$ ).

Figure 3 shows the proportion of new and continuing users in the four datasets. The proportion of continuing users gradually increases over time in each case (mean Pearson correlation coefficient between proportion of continuing users and timestamp is  $r = .78$ ;  $p < .05$ ), indicating that many individuals who post initially, continued to participate as subsequent events unfold. The proportion of new users exhibit peaks during times of important events where the overall volume of posting is also high (correlation between post volume and proportion of new users over time is  $r = .15$ ;  $p < .05$ ; ref. Figure 1): indicating that those who would typically not post on these topics may be motivated to do so in the context of a notable event. This also aligns with prior literature that indicates that the audience composition changes during times of higher volumes and becomes more “mainstream”, while periods of low activity volume consist of mostly a “niche” audience (Weber and Jaimes 2010).

To extend this discussion, given the considerable gaps between the time periods of the datasets, can we analyze par-

ticipation from the new and continuing users from one event to another? To answer this, we computed the proportion of individuals who participated in event  $e_i$  (e.g., Ferguson I) and then continued to participate in the event  $e_j$  immediately succeeding it,  $j = i + 1$  (e.g., Ferguson II). We found this proportion to be 36.1% ( $\pm 23.6\%$ ), showing considerable continued participation across the events.

	$\mu$ (%)	$\sigma$ (%)		$\mu$ (%)	$\sigma$ (%)
<b>Affective attributes</b>			family	2.21	1.24
positive affect	-0.14	0.81	friends	5.59	5.76
negative affect	-62.15	10.47	<b>Interpersonal awareness</b>		
anger	-7.79	6.56	1st p. singular	-14.52	3.98
anxiety	-1.00	0.96	1st p. plural	11.15	3.15
sadness	-2.21	1.20	2nd pp.	5.35	0.84
swear	-6.59	2.07	3rd pp.	0.45	1.27
<b>Cognitive attributes</b>			<b>Psychological distancing</b>		
cognitive mech	1.30	3.10	<i>Temporal references</i>		
discrepancies	0.21	1.61	past tense	-3.71	2.64
negation	0.56	1.56	present tense	-1.56	0.81
<b>Perception attributes</b>			future tense	10.53	6.67
percept	1.32	2.08	<i>Function words</i>		
see	0.82	1.55	article	1.82	3.16
hear	0.98	0.61	adverbs	1.02	7.71
feel	1.16	1.95	verbs	0.84	2.22
death	-21.23	3.66	preposition	0.41	1.95
<b>Social orientation</b>			conjunction	0.31	1.97
social	12.58	3.64			

Table 3: Temporal changes in different measures.

**Temporal Change.** Next, we examine whether users exhibit any changes in their engagement and linguistic attributes on Twitter with continued participation over time. For a given social media measure (e.g., NA) and a given user, we define a *temporal change metric* as the percentage change in the measure in the user’s posts on day  $t_j$  compared to that on the day immediately prior,  $t_i$ , where  $j = i + 1$ .

Table 3 gives the mean temporal changes (and their standard deviations) for all our social media measures aggregated across the four datasets. We observe NA, *death*, *1st p. singular*, *anger* and *swear* to show the largest decreases (6.59–62.15% decrease;  $p < 10^{-5}$  based on Wilcoxon rank sum tests); while *social*, *1st p. plural*, *friends* and *2nd pp.* indicate significant increases over time as users continue to post in the context of this activist movement (5.35–14.52% increase;  $p < .0001$  based on Wilcoxon rank sum tests).

Decreasing measures of *NA*, *death*, *1st p. singular*, *anger* and *swear* indicate that initial postings of users may reflect more of a personal narrative or opinion, including heightened negativity, exasperation, displeasure, and references to the loss of lives of black people. However, given the growth of the BLM movement over time, greater awareness of and reference to one’s social environment (use of *cognition*, *perception* and *function words*), in combination with the increase in *future tense* may indicate a shift of focus from the personal reaction of the current state, to a sense of empowerment. The manifestation of a collective identity in participation over time is lent credence by the observation of a higher usage of *1st p. plural pronouns* and increased social orientation (via higher use of *2nd pp.* and *social* and *friends* words). These observations align with prior literature wherein communities experiencing trauma, like wars and disasters, expand “the circle of we” over time, promoting and deepening solidarity and a sense of collectivity (Mark et al. 2012).

## RQ 2: Contrasting Geographical Patterns

Next, we direct our attention to ascertain in what ways content and expression on Twitter, as reflected through the different measures, differ across different geographical regions.

First we investigate: *in what ways are the manifested Twitter measurements associated with the rates of police killings (PK) in different states?* To examine these multivariate effects, we adopt a regression modeling approach, where we use the state-level rate of police killings of Blacks index (PK) as the dependent variable, and per-state normalized Twitter measures averaged over the four datasets as independent variables in a Poisson regression model. This model is appropriate for contexts in which the dependent variable is a count-like quantity that cannot take on negative values (case of PK), and that has significant skew (conditional variance nearly equal to or greater than the conditional mean). Note, a regression modeling approach is appropriate here since we intend to examine multiple pairwise correlations between PK and Twitter derived measures. We caution that the approach does not imply that the measures causally affect PK in different states.

We summarize the performance of this model in Table 4. The model is assessed in a number of ways — pseudo- $R^2$ , log likelihood and  $\chi^2$ -test statistic of the log likelihood ratio. We find that the Twitter derived state-level measures are well-suited to characterize the dependent variable PK. Specifically, our model is able to account for more than 61.8% of variance in the PK data, with a log likelihood of 141.7, that significantly improves over a null model based on a  $\chi^2$  test (LR  $\chi^2 = 22.86, p < 10^{-4}$ ).

Next, we examine the effects of specific measures in accounting for state-level PK.

**Observation 1:** *High negativity manifests in high PK states.*

We find that the affective measures *NA*, *PA*, *anxiety*, *swear* show high  $\beta$  values, thus high explanatory power for PK. This shows individuals posting from states with high PK may be engaging over Twitter to express their relatively higher negative perceptions, reactions and feelings of the existing racial violence context.

	$\beta$	[95% conf. interval]		$p$
Activity and engagement				
# @-replies	1.6386	0.939	4.216	*
Affective attributes				
PA	-20.278	-37.63	-11.07	***
NA	28.54	16.306	33.39	***
anger	9.757	7.5289	12.043	***
anxiety	14.071	10.369	21.512	***
sadness	13.268	7.84	25.376	***
swear	12.101	9.4505	14.752	***
Cognitive attributes				
cognitive mech	-1.1461	-5.981	-0.688	*
negation	-7.1799	-15.847	-3.4876	**
Perception attributes				
see	10.489	3.722	22.699	***
hear	10.27	7.52	18.06	***
feel	4.2771	2.7379	14.292	**
death	40.036	22.271	52.34	***
Social orientation				
social	-7.4167	-11.643	-5.81	**
family	-5.5384	-16.059	-1.982	**
friends	-4.8271	-10.897	-1.2425	***
Interpersonal awareness				
1st p. singular	39.213	18.93	58.505	***
1st p. plural	-12.3689	-16.225	-5.4877	***
2nd p.	4.6663	1.227	14.559	**
Psychological distancing				
Temporal references				
past tense	4.4593	1.24	9.16	**
present tense	3.8796	1.368	14.127	*
Function words				
article	-2.7275	-9.641	-0.1859	**
adverbs	-4.2942	-14.168	-1.5798	**
conjunction	-3.1092	-7.832	-0.6138	**
pseudo $R^2 = .618$ , LL = 141.7; LR $\chi^2 = 22.86, p < 10^{-4}$				

Table 4: Summary of a Poisson regression model with PK in states as dependent variable. Significance is estimated following Bonferroni correction:  $*\alpha = .05/34$ ;  $**\alpha = .01/34$ ;  $***\alpha = .001/34$ .

**Observation 2:** *Increased self-preoccupation and low social orientation are observable in high PK states.*

High use of *1st p. singular*, lower use of *1st p. plural* words and low social orientation words (*social*, *friends*, *family*) are associated with higher PK in states. These patterns are known to indicate heightened self-attentional focus and greater detachment from the social realm (Cohn, Mehl, and Pennebaker 2004). We conjecture that individuals in regions of greater racial police violence may be resorting to social media to share their personal experiences and opinions about this topic, explaining this observation.

**Observation 3:** *Low cognition, but high personal accounts of perceived incidents, along with attribution to death are observable in high PK states.*

We also observe reduced use of cognitive attributes (*cognitive mech*) and higher use of perception attributes like (*see*, *hear*, *feel*) to be associated with posts from high PK states. These observations are known to be associated with language that depicts personal and first-hand accounts of real

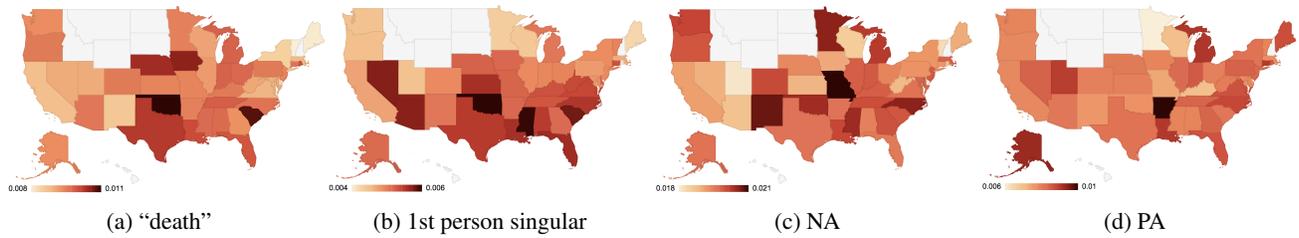


Figure 4: State-level measures shown for four measures with the highest  $\beta$  weights in Table 4.

world happenings, events and experiences. Further, greater use of *death* words is positively associated with states of high PK. While Ferguson, New York, and Baltimore are just a few of the states with high profile cases, the greater use of *death* words across other high PK states may highlight the resonance of the loss of black lives subject to law enforcement countrywide; a launching point of BLM discourse.

**Observation 4:** *High psychological distancing is observable in high PK states.*

Posts from states with high PK also show high psychological distancing as observable from the use of temporal reference and function words. Use of high *past* and *present* tense words indicate tendency to recollect prior experiences and events, as well as focus on the here and now (Pennebaker, Mayne, and Francis 1997). Lower use of *articles* and *adverbs* are associated with a personal narrative writing style (Cohn, Mehl, and Pennebaker 2004). Together, this aligns with our observation above that individuals in high PK states may be using Twitter to express their personal thoughts on the topic of racial conflict.

**Observation 5:** *High interpersonal interaction is observable in high PK states.*

While most activity and engagement measures are non-significant predictors of PK across states, we find that the number of @-replies is. It is known that during societal crises and upheavals, communities bond and engage in interpersonal exchange (Glasgow, Fink, and Boyd-Graber 2014). Potentially, Twitter users in high PK states, despite high self-focus, tend to interact with others to seek and provide psychosocial support around issues of racial inequality.

*What are some of the states where the above observations are manifested?* For this purpose, we show state-wise differences across the four most significant measures (i.e., with highest  $\beta$  weights, and highlighted in Table 4) in Figure 4. We find notable “signatures” in manifested use of *death* words (high in the south and the midwest)<sup>4</sup>, *1st p. singular* (highest in the south), *NA* (most states in the south and midwest show high levels) and *PA* across states (most states show low positivity). In an aggregated sense, all of the first three measures are the highest in the states of the south ( $\mu = .014$ ), followed by the midwestern states ( $\mu = .006$ ). The Clifford, Richardson, and Hémon, or CRH test, a method that corrects traditional  $p$ -value calculation by

<sup>4</sup>State-region assignments are made based on Census definitions: [http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us\\_regdiv.pdf](http://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf)

taking into account spatial auto-correlation in data, yields statistically significant differences across the four different regions;  $F = -6.4$ ;  $p < 10^{-3}$ . We note that in the south and the midwest, PK is also found to be the highest, relative to the other parts of the country – PK in the south ( $\mu = 11.55$ ), in midwest ( $\mu = 13.66$ ) (ref. Figure 2). These patterns summarize the geographical tone of the discourse that unfolded on Twitter, and are also found to bear significant relationship with the reported PK in different states.

### RQ 3: Predictors of Protest Volume

Finally, in RQ 3, we focus on understanding how Twitter activity and expression, captured by the different measures, relates to and predicts the actual protests and demonstrations that commenced throughout the country. For this, we employ a negative binomial regression modeling approach, due to the presence of over dispersed count data (PV) and due to sufficiently large number of samples. We consider the daily values of the countrywide protest volumes as the dependent variable (PV) (say, PV on  $t_i$ ), and the *previous day* ( $t_{i-1}$ ) normalized averages of the different measures as independent variables, concatenated by time over the four datasets.

We summarize this model’s fit in Table 5. The model yields high pseudo  $R^2 = .42$  with significance at  $p < 10^{-6}$ , indicating that the Twitter derived daily measures bear considerable power in explaining more than 41% of variance in the PV data. The log likelihood of this model is found to be 826.1, an improvement over a null model (intercept-only model) based on a  $\chi^2$  test: the log likelihood ratio is  $\chi^2 = 81.24$ ,  $p < 10^{-6}$ , on 35 degrees of freedom. We next discuss the different significant variables in this model.

**Observation 1:** *Greater levels of activity and engagement are associated with high future daily protest volume (PV).*

We observe that greater volumes of posts, @-replies, retweets and posts bearing links demonstrate higher explanatory power in estimating PV over the next day. These measures indicate greater social awareness, interaction and information sharing and dissemination practices via social media, which may play a role in motivating greater number of individuals to participate in the BLM protests that happened throughout the country.

**Observation 2:** *High NA and sadness but low anger and anxiety are positively associated with high PV in the future.*

Among the affective attributes, we observe that higher *NA* and *sadness* but lower *anger* and *anxiety* are associated with greater number of individuals protesting next day. This may

	$\beta$	[95% conf. interval]		$p$
[intercept]	1.453	0.795	2.110	***
Activity and engagement				
# posts	7.010	1.244	15.264	***
# @-replies	6.635	4.258	10.528	***
# retweets	2.727	0.4844	3.138	***
# posts w/ link	1.831	0.2574	2.236	**
Affective attributes				
PA	-0.2985	-1.104	-0.1072	*
NA	21.74	12.624	36.103	***
anger	-0.7518	-0.957	-0.4533	**
anxiety	-3.9124	-6.3191	-1.5057	***
sadness	1.7109	0.8238	3.4021	**
swear	-2.0366	-4.5444	-1.4712	***
Cognitive attributes				
discrepancies	1.0604	0.4118	2.3327	***
negation	2.8294	0.97	4.3113	***
Perception attributes				
hear	0.9437	0.675	2.587	**
feel	1.0153	0.5229	2.0923	**
death	-15.144	-24.726	-5.437	**
Social orientation				
social	0.2811	0.0922	1.485	*
family	1.8024	1.1568	4.961	***
friends	1.9103	1.0357	5.2563	***
Interpersonal awareness				
1st p. singular	-4.3174	-14.178	-1.5428	***
1st p. plural	2.0803	1.7301	5.8907	***
2nd p.	0.8822	0.3311	2.7957	**
Psychological distancing				
Temporal references				
past tense	-1.1051	-2.3229	-0.1127	**
present tense	0.3832	0.1706	1.9382	*
future tense	0.2803	0.0301	2.8627	*
Function words				
article	-0.3403	-1.6725	-0.1924	*
verbs	-0.3344	-1.54	-0.1709	*
pseudo $R^2 = .422$ , LL = 826.1; LR $\chi^2 = 81.24$ , $p < -10^6$				

Table 5: Summary of negative binomial regression with daily protest volume (PV) as dependent variable. Significance is estimated following Bonferroni correction: \* $\alpha = .05/34$ ; \*\* $\alpha = .01/34$ ; \*\*\* $\alpha = .001/34$ .

indicate that while a negative attitude towards racial violence due to law enforcement remains unchanged, anger and anxiety may be replaced with empowerment derived from collective action and identity.

**Observation 3:** *High cognitive and perceptual processing and attribution to death characterize high PV in the future.*

Our regression model further showed more complex cognition and perception in Twitter content to predict high PV next day. This is known to indicate heightened awareness of one’s social environment and more objective psychological expression (Pennebaker, Mayne, and Francis 1997). In our case, these may be related to people’s reflection of thoughts around police violence, misconduct and mistreatment of minority communities (the greater use of *death* words further bolsters this observation). These, in turn, are also likely to be factors positively linked to larger participation in protests.

**Observation 4:** *Greater social orientation and depiction of collective identities characterize high PV in the future.*

Next, we find that higher *social*, *friends*, *family* and *1st p. plural* words indicate more participation in the next day protests. Several media organizations have reported these protests to be solidarity protests bringing to the fore the racial challenges experienced by the black community. Through the social orientation words, we conjecture individuals to share social well-being impressions and aspirations to end victimization of blacks (Garza 2014). Further, a characteristic of societal crises is the eventual emergence of a collective identity and group cohesion among those directly or indirectly affected, and use of *1st p. plural* words in language are known to be indicative of such an identity (Cohn, Mehl, and Pennebaker 2004). Likely, protest participation also involves the manifestation of collective perceptions of a social and racial challenge, and one’s identification with a greater community of people—hence the observed link between the above measures and PV.

**Observation 5:** *Increased futuristic inclination, greater use of categorical language and abstract information processing are associated with high PV in the future.*

Days of high protest volume are preceded by greater use of *future* tense and certain function words. The function words indicate complex categorical thinking, implying concrete information sharing about a topic (Chung and Pennebaker 2007). Many of the protests were issue-focused, and about fact-sharing and coordination action, often in response to a recent incident or development (e.g., in Table 2 events on Aug 20 2014, Dec 19 2014 and Apr 29 2015). It is thus consistent that a similar form of social media discourse preceded these demonstrations, and perhaps even garnered increased online support.

Finally, we describe the predictive power of the negative binomial regression model in assessing next day PV values (day  $t_i$ ) based on current day Twitter measures ( $t_{i-1}$ ). We employ  $k$ -fold cross validation ( $k = 5$ ) and train and test several models: “Activity, Engagement”, “Affective attributes”, “Cognition, Perception attributes”, “Social orientation”, “Interpersonal awareness” and “Psychological distancing”. These models allow us to independently examine how forms of Twitter activity and expression predict offline protests. We also train and test a final model that uses all of our measures (referred to as “All”). We additionally include two auto-regressive baseline models: the first one, a “Constant model” where we utilize PV on day  $t_i$  to predict the mean value of PV over all time; and a moving average (MA) model with a 1 day lag (“Next day MA”), where we predict the PV value on day  $t_i$  using PV on the day before, i.e.,  $t_{i-1}$ . These baseline models allow us to examine whether the models that use Twitter measures provide us additional predictive power over time series relationships in PV.

We summarize the mean predictive performance of the various models across all five cross validation folds in Table 6. We observe that the “All” model performs the best (81% estimates within 20% of true PV values), followed by the model that uses the affective attributes. Compared to the two baseline models, we show significant improvements in

	RMSE	MAPE	SMAPE (%)	Correct @ $\leq 20\%$ (%)
Constant model	9496.5	84.63	34.59	42
Next day MA	8379.2	70.68	26.35	48
Activity, Engagement	7155.4	57.26	20.40	59
Affective Attributes	5775.6	36.05	8.34	74
Cognition, Perception	6100.4	38.04	9.64	71
Social Orientation	6532.5	54.15	15.39	62
Interpers. Awareness	6311.3	42.57	10.43	69
Psychological Dist.	6519.9	49.74	12.56	65
All	5528.1	32.62	6.37	81

Table 6: Performance metrics of predicting daily PV. Here (1) RMSE is root mean squared error; (2) MAPE is median absolute percentage error; (3) SMAPE is symmetric mean absolute percentage error; and (4) Correct @  $\leq 20\%$  is the percent of PV estimates within 20% of the true values.

performance with our “All” model (33-39% improvement in number of predicted PV within 20% of true values).

## Discussion

### Implications

Historians of the 1960s Civil Rights Movement view how the media of the time helped establish a “new common sense” about race in America (Kennedy 1998). In 2015, with the plethora of available social media platforms, any individual loosely or tightly supporting the Black Lives Matter movement can seek to call attention to issues of racialized policing, and the vulnerability that black people experience in general. As prior research indicates (Lotan et al. 2011; Tufekci and Wilson 2012), one of the most powerful attributes of social media platforms has been their ability to bring the voices of the masses to the fore during times of societal and political upheavals and our findings indicate that the Black Lives Matter activist movement is no exception.

**Collectivism.** Our results demonstrate that while notable events may have triggered many individuals to engage in cursory or one-time discourse on the various issues of the Black Lives Matter activist movement, some individuals remained involved in the social media conversations over a long period and across temporally spread-out events. This indicates that Twitter emerged as an important platform of discourse and reflection for many individuals, allowing them to share stories, find common ground and agitate for police and government reform around racial issues. Further, we observed that continued participation is associated with lowered negativity and anger, and with emergent collective identity and reduced psychological distancing over time. This indicates people’s desire to organize collective action and to socially connect, support, cope and engage with each other as a community, as though they have experienced “collective abuse” (Mark et al. 2012), that transcends the specific incidents of police brutality.

**Role of Historical Police Violence.** We also found that the historical rate of police killings of blacks was linked to the way people responded on social media around this activist movement. This illustrates that the expression and evolution of new social movements on social media are shaped not just

by isolated contemporary phenomena, but also by historical racial discrimination and long-standing perceptions of race and law enforcement. Our results also show how social media may be providing an alternate channel for discussing race related issues, and for sharing community response to loss of life in different parts of the country.

**Digital Activism.** Finally, we also found that participation in future protests was associated with a spike in the intensity of social media conversations, as well as an increase in negative affect and sadness, heightened cognitive and perceptual processing and manifestation of a collective agency. While it is challenging to answer whether this observed digital activism is indeed the same real world activism it predicts, our analysis may help augment traditional survey-based efforts to assess community solidarity and collectivism in the light of protests and social movements. Our observations also suggest how face-to-face and online forms of activism work in interrelated and aggregative ways towards helping drive social and political change. As protestor Johnetta Elzie puts it (Bonilla and Rosa 2015): “[...] *Then I saw Brown’s body laying out there, and I said, Damn, they did it again! [...] I’m not just going to tweet about it from the comfort of my bed. So I went down there.*”

**Community Sensing.** The nature of behavioral patterns that we gleaned from Twitter in different parts of the country can provide new insights to authorities and policymakers to understand issues of public unrest, and to identify opinions and expressions on a sensitive topic like race, at a scale and scope not possible through conventional means such as surveys. The attributes of participation as we observed can also empower activists to better understand community growth and mobilization over time, and how the activities of the movement are being recognized by the online population. Finally, our results can help researchers derive theories about socio-psychological responses to discriminatory police violence.

### Limitations and Future Work

There are limitations to our approach and findings. We cannot claim to have captured the complete engagement around this movement online or in the physical world. Additionally, we cannot make causal claims from our analysis. While greater Twitter activity was positively associated with heightened protest volumes, we did not ascertain whether the individuals posting on Twitter were also the ones who were involved in the protests themselves. While new media platforms have proven to be a powerful aid in facilitating activism offline, in most cases, demonstrations are organized by groups and individuals in the physical world first. Therefore, understanding how activity on social media translates into real world mobilization will require understanding the corresponding offline social structures.

## Conclusion

We provided some of the first empirical insights into social media discourse on the Black Lives Matter movement. Our results on over 28M Twitter posts show continued participation in the conversation around this movement. To the best

of our knowledge, this is the first work to explore how historical racial disparities in police killings in different parts of the country are associated with the way people feel and express themselves on social media in the context of an activist movement. Another important finding of our work is that activism on social media predicted future protests and demonstrations that commenced on the streets throughout the country. As observed in other social movements like the Occupy and Arab Spring, we observed BLM participation on social media to indicate an emergent collective identity. Our work extends the literature on social movements and the role of social media in collective action.

## Acknowledgements

We acknowledge the efforts of Molly Loyd, Gregory Coleman, Kimberly Lamke, and Ed Summers in compiling the Ferguson Twitter datasets. We also thank Alexandra Olteanu for insightful discussions. De Choudhury was partly supported through an NIH grant # 1R01GM11269701.

## References

- Bastos, M. T.; Mercea, D.; and Charpentier, A. 2015. Tents, tweets, and events: The interplay between ongoing protests and social media. *Journal of Communication* 65:320–350.
- Bonilla, Y., and Rosa, J. 2015. #ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the united states. *American Ethnologist* 42(1):4–17.
- Bonilla-Silva, E. 2006. *Racism without racists: Color-blind racism and the persistence of racial inequality in the United States*. Rowman & Littlefield Publishers.
- Borge-Holthoefer, J.; Magdy, W.; Darwish, K.; and Weber, I. 2015. Content and network dynamics behind egyptian political polarization on twitter. In *Proc. CSCW*, 700–711. ACM.
- Bridges, G., and Crutchfield, R. 1988. Law, social standing and racial disparities in imprisonment. *Social Forces* 66(3):699–724.
- Cheong, M., and Lee, V. C. 2011. A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via twitter. *Information Systems Frontiers* 13(1):45–59.
- Chung, C., and Pennebaker, J. W. 2007. The psychological functions of function words. *Social communication* 343–359.
- Cohn, M. A.; Mehl, M. R.; and Pennebaker, J. W. 2004. Linguistic markers of psychological change surrounding september 11, 2001. *Psychological science* 15(10):687–693.
- Conover, M. D.; Ferrara, E.; Menczer, F.; and Flammini, A. 2013. The digital evolution of occupy wall street. *PLOS ONE* 8(5):e64679.
- De Choudhury, M.; Monroy-Hernandez, A.; and Mark, G. 2014. Narco emotions: affect and desensitization in social media during the mexican drug war. In *CHI*, 3563–3572.
- Eltantawy, N., and Wiest, J. B. 2011. The arab spring— social media in the egyptian revolution: reconsidering resource mobilization theory. *International Journal of Communication* 5:18.
- Garza, A. 2014. A herstory of the black lives matter movement. *Black Lives Matter*.
- Glasgow, K.; Fink, C.; and Boyd-Graber, J. 2014. Our grief is unspeakable: Measuring the community impact of a tragedy. In *ICWSM*.
- Gleason, B. 2013. # occupy wall street: Exploring informal learning about a social movement on twitter. *American Behavioral Scientist* 0002764213479372.
- Goldkamp, J. S. 1976. Minorities as victims of police shootings: Interpretations of racial disproportionality and police use of deadly force. *The Justice System Journal* 2(2):169–183.
- González-Bailón, S.; Borge-Holthoefer, J.; Rivero, A.; and Moreno, Y. 2011. The dynamics of protest recruitment through an online network. *Scientific reports* 1.
- Hecht, B., and Stephens, M. 2014. A tale of cities: Urban biases in volunteered geographic information. In *ICWSM*.
- Kennedy, R. 1998. *Race, crime, and the law*. Vintage.
- Klinger, D. A. 2011. On the problems and promise of research on lethal police violence: A research note. *Homicide Studies* 1088767911430861.
- Lotan, G.; Graeff, E.; Ananny, M.; Gaffney, D.; Pearce, I.; et al. 2011. The arab spring— the revolutions were tweeted: Information flows during the 2011 tunisian and egyptian revolutions. *International journal of communication* 5:31.
- Mark, G.; Bagdouri, M.; Palen, L.; Martin, J.; Al-Ani, B.; and Anderson, K. 2012. Blogs as a collective war diary. In *CSCW*, 37–46.
- McKinley, J., and Baker, A. 2014. Grand jury system, with exceptions, favors the police in fatalities. <http://www.nytimes.com/2014/12/08/nyregion/grand-juries-seldom-charge-police-officers-in-fatal-actions.html>.
- Olteanu, A.; Weber, I.; and Gatica-Perez, D. 2016. Characterizing the demographics behind the #blacklivesmatter movement. In *OSSM*. <http://arxiv.org/abs/1512.05671>.
- Palen, L., and Vieweg, S. 2008. The emergence of online widescale interaction in unexpected events: assistance, alliance & retreat. In *CSCW*, 117–126.
- Parker, K. D.; Onyekwuluje, A. B.; and Murty, K. S. 1995. African americans’ attitudes toward the local police: A multivariate analysis. *Journal of Black Studies* 25(3):396–409.
- Pennebaker, J. W.; Mayne, T. J.; and Francis, M. E. 1997. Linguistic predictors of adaptive bereavement. *Journal of personality and social psychology* 72(4):863.
- Sampson, R. J., and Lauritsen, J. L. 1997. Racial and ethnic disparities in crime and criminal justice in the united states. *Crime and Justice* 311–374.
- Sigelman, L.; Welch, S.; Bledsoe, T.; and Combs, M. 1997. Police brutality and public perceptions of racial discrimination: A tale of two beatings. *Political Research Quarterly* 50(4):777–791.
- Starbird, K., and Palen, L. 2012. (how) will the revolution be retweeted?: information diffusion and the 2011 egyptian uprising. In *CSCW*, 7–16.
- Tufekci, Z., and Wilson, C. 2012. Social media and the decision to participate in political protest: Observations from tahrir square. *Journal of Communication* 62(2):363–379.
- Varol, O.; Ferrara, E.; Ogan, C. L.; Menczer, F.; and Flammini, A. 2014. Evolution of online user behavior during a social upheaval. In *Proc. 2014 ACM WebSci*, 81–90. ACM.
- Vieweg, S.; Hughes, A. L.; Starbird, K.; and Palen, L. 2010. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *CHI*, 1079–1088.
- Weber, I., and Jaimes, A. 2010. Demographic information flows. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, 1521–1524. ACM.
- Weber, I.; Garimella, V. R. K.; and Batayneh, A. 2013. Secular vs. islamist polarization in egypt on twitter. In *ASONAM*, 290–297.