Images that Matter:
Online Protests and the Mobilizing Role of Pictures

Andreu Casas
New York University

Nora Webb Williams
University of Washington

Accepted for publication at *Political Research Quarterly*

Abstract

Do images affect online political mobilization? If so, how? These questions are of fundamental importance to scholars of social movements, contentious politics, and political behavior generally. However, little prior work has systematically addressed the role of images in mobilizing online participation in social movements. We first confirm that images have a positive mobilizing effect in the context of online protest activity. We then argue that images are mobilizing because they trigger stronger emotional reactions than text. Building on existing political psychology models we theorize that images evoking enthusiasm, anger, and fear should be particularly mobilizing, while sadness should be demobilizing. We test the argument through a study of Twitter activity related to a Black Lives Matter protest. We find that both images in general and some of the proposed emotional attributes (enthusiasm and fear) contribute to online participation. The results hold when controlling for alternative theoretical mechanisms for why images should be mobilizing, as well as for the presence of frequent image features. Our paper thus provides evidence supporting the broad argument that images increase the likelihood of a protest to spread online while also teasing out the mechanisms at play in a new media environment.

1 Introduction

Do images affect online political mobilization? If so, how? People today are bombarded with more images than ever before in human history. However, despite small-N studies and experimental research demonstrating the effects of images on political information seeking,[1] issue-framing,[2] voting preferences,[3] political attitudes,[4] and even on compliance with authoritarian regimes,[5] there is still little work systematically addressing the role of images in mobilizing participation in protests and social movements online,[6] nor are there studies that have leveraged large, digitized corpora of real world protest images. Moreover, those works that study the more general political effects of images tend to rely on clear experimental treatments,[7] while real political images from everyday individuals are messy and often vary on multiple dimensions, making large-N observational studies a must.

In this paper, we attempt to fill these gaps in the literature by presenting and testing a set of hypotheses derived from a specific mechanism pathway for why images might affect social movement mobilizations. We first confirm that, as expected from prior research, images increase participation in the context of online mobilization. Beyond this main effect, we suggest that images are particularly mobilizing because they generate stronger emotional reactions than text. Emotions such as enthusiasm, anger, and fear are known to positively affect participation in a wide range of political contexts. We argue that images evoking these three emotions are likely to be strongly correlated with higher levels of protest participation. We test the argument with a new, large-N dataset related to the Black Lives Matter (BLM) movement. We track the online spread of general support in Twitter for BLM and a specific BLM protest, ShutdownA14, that occurred on April 14, 2015.

The idea that images might matter to social movements like BLM is not new. The Civil Rights movement in the United States and recent contentious events such as the Arab Spring became known for its powerful mobilizing images.[8] The issue with studying

[1] Ryan 2012

[2] Corrigall-Brown and Wilkes 2012; Rohlinger and Klein 2012

[3] Rosenberg et al. 1986; Todorov et al. 2005

[4] Grabe and Bucy 2009; Wright and Citirn 2011; Dahmen 2012

[5] Bush et al. 2016

[6] See Kharroub and Bas 2015 for a preliminary attempt. See also Bas and Grabe 2016 for a study of how images affect participation in other types of political behavior.

[7] e.g. Ryan 2012

[8] Raiford 2007; Howard and Hussain 2013

these cases, however, is the biasing selection effect of only looking at potentially rare cases. Our challenge is to study images without knowing in advance whether any of them will have an effect on behavior. We chose a case prior to mobilization to see which images, if any, explained subsequent variation in the spread of the given protest and support for its associated social movement.

Our paper also speaks to the urgency of studying images now, in the current new media environment. With the rise of mobile phones with cameras, the ability of almost everyone to share images from or related to a protest has become an important consideration for scholars.[9] Today small or emerging social movements such as BLM can rely on thousands of participants to take pictures "from the trenches"[10] and immediately share them.

We focus our efforts on the effect of online image sharing on online social movement mobilization. We acknowledge that mobilization in the offline arena is equally, if not more, important for social movements. Organizations today use hybrid offline and online tactics to achieve their goals.[11] And while some scholars are skeptical of the role of online activism,[12] others find that online participation is an increasingly important tool to recruit committed supporters,[13] increase protest turnout,[14] and to set the media and political agendas.[15]

We operationalize online mobilization as retweets of movement-related messages on Twitter. While retweeting, like most forms of online mobilization, is a relatively low cost form of participation, it is still has value to social movement organizers. A simple action such as a retweet of a protest message can help drive public, media, and political attention.[16] We refer to this form of mobilization as generating movement "attention," or an increase in the amount of public discussion about the movement. Retweeting can also help a movement see who its new supporters are. If many new Twitter users are engaging with a movement through its messages and hashtags, this is a signal that movement themes are impacting, and being impacted by, a broader audience. We refer to this form of mobilization as generating movement "diffusion," or the spread of online support to new members. We discuss the nuances of treating retweeting as participation in more

[9]cf Howard and Hussain 2013; Webb Williams 2015
[10]Payne 1998
[11]Chadwick 2011; Bimber, Flanagin, and Stohl 2012; Theocharis et al. 2015
[12]Morozov 2011
[13]Hsiao and Yang 2018
[14]De Choudhury et al. 2016
[15]Freelon, C. McIlwain, and M. Clark 2016; Casas, Davesa, and Congosto 2016
[16]Freelon, C. McIlwain, and M. Clark 2016

detail below.

The contributions of this paper are fourfold. First, we confirm an image effect in the case of a specific BLM protest. Second, we derive specific theoretical mechanisms that might make certain types of images more effective at mobilizing participants. Third, we test these hypotheses using a large-N observational dataset of tweets containing protest keywords and hashtags from April 13 to April 20 2015, along with all of the images included in those tweets. The dataset includes approximately 150,000 tweets and 9,500 manually labeled images. Finally, we add to the available body of knowledge regarding the BLM movement and the means by which the movement has spread.

2 Theoretical Framework and Expectations

2.1 Social media messaging as mobilization

Treating retweets as participation in a social movement links our research to a wider conversation about the nature of mobilization in the digital age. Many scholars have debated the value of social media for social movements.[17] Internet technologies have greatly lowered the costs of connecting with potential supporters,[18] and although some argue that networked technologies have given place to cheap and meaningless forms of engagement,[19] another line of research clearly shows that this is often not the case. In particular, this literature argues that today, simply sharing social media messages should be seen as a form of participation that helps social movements achieve their intermediate and longer-term goals.[20]

Social movements are groups protesting against a *status quo* and seeking some sort of social change.[21] In order to achieve such goal, they use a wide range of tactics,[22] such as public campaigning[23] and litigation strategies.[24] Raising awareness about a movement's claims is often a necessary condition preceding social change, and as a result, tactics pursuing agenda setting objectives are seen as crucial by social movements and organiza-

[17] Notably Morozov 2011; Bennett and Segerberg 2013; Bimber, Flanagin, and Stohl 2012; Earl and Kimport 2011; Castells 2012; Gonzalez-Bailon et al. 2011; Howard and Hussain 2013; Kharroub and Bas 2015

[18] Olson 1965; Bimber, Flanagin, and Stohl 2012; Lupia and Sin 2003

[19] c.f. the "slacktivism" concept of Morozov 2011

[20] Gonzalez-Bailon et al. 2011; Theocharis et al. 2015; Barberá et al. 2015

[21] Tarrow 2011

[22] McAdam 2000

[23] Benford and Snow 2000

[24] McCann 2006

tions.[25] In the current digital media environment, protest movements use a repertoire of offline but also online tactics to influence the public, media, and political agenda.[26]

In this context, a high volume of social media messages (original and shared) can raise awareness about a protesting group in three ways. First, it helps set public agendas by exposing the movement's claims to new audiences and by keeping engaged those who already care about the issue. For example, Gonzalez-Bailon et al. show how retweet chains allowed the variety of claims of the *Indignados* movement to spread to a massive amount of users,[27] and Barbera et al. show how even retweets from peripheral users (users with very few followers engaged with a given protest movement) were responsible for the growth of online discussions related to the "United for Global Change" demonstration and the Gezi Park protest.[28]

Second, sharing social media messages related to a protest helps a movement get media attention and coverage. This is of particular importance in cases where mainstream media is initially reticent to cover a given movement.[29] Penney and Dadas describe how retweet chains by early Occupy Wall Street protesters helped the movement attract mainstream media attention.[30] In the context of the expansion of BLM, Freelon et al. also noted that simple retweets of protest messages helped drive media coverage.[31] Retweeting helps amplify a tweet or a hashtag, even if the retweeter is doing so ironically or without intending to indicate support for the tweet.

Finally, retweeting messages associated with a protest also helps the movement to raise awareness about its claims by recruiting new future offline protesters. The more public support a movement receives in an offline protest, the more likely the movement is to get political attention and foster social change.[32] Recent research shows that, after controlling for a battery of alternative explanations, a larger number of tweets and retweets related to the BLM movement were closely associated with higher participation rates in posterior demonstrations.[33] By studying numerous protests related to the movement, De Choudhury et al. found that, on average, each retweet in time t was associated with 15

[25] Baumgartner, De Boef, and Boydstun 2008
[26] Chadwick 2011
[27] Gonzalez-Bailon et al. 2011
[28] Barberá et al. 2015
[29] Gitlin 1980; Penney and Dadas 2013
[30] Penney and Dadas 2013
[31] Freelon, C. McIlwain, and M. Clark 2016
[32] Walgrave and Vliegenthart 2012
[33] De Choudhury et al. 2016

more protesters on the streets in time $t + 1$.

Moreover, although the participation costs of sharing a message are not as high as physically attending a protest, there are still some reputation costs attached to it: people can lose ties with friends in their network[34] and they open themselves to criticism from people in their immediate social environment.[35] In sum, based on this observation and on the rich line of research described above, we consider social media messaging as a meaningful way to help a social movement succeed, and we align with previous studies and treat retweets as a form of political participation worth studying.[36] Despite the existence of this rich literature debating online participation in social movements, there has been less consideration of the role of images in mobilizing online action. We now turn to discussing those works that do exist to situate our project.

2.2 Existing research on images and politics

Images are a central part of our lives but political scientists have traditionally paid little attention to how they affect social and political processes.[37] In particular, literature studying the effect of images on social protest mobilization is scarce.[38] The existing literature mostly focuses on the effect that images have on issue-framing, political attitudes, and participation in non-protest related political activities, such as voting.

In the last few years a growing body of literature has paid more attention to how mass media images play a particular issue-framing role for protests and social movements. Corrigall-Brown and Wilkes study newspaper images of a collective action in Canada to conclude that, whereas textual content confirmed the "protest paradigm," protesters were just as likely as government authorities to receive visual coverage.[39] In another study, Rohlinger and Klein look at how different news sources covered several abortion-related protests and find that visual content is very similar across outlets and events.[40]

Another line of research studies the role of images in shaping political attitudes. Wright and Citrin test if hypotheses derived from the common in-group identity model still hold

[34] Boyd, Golder, and Lotan 2010

[35] Griffiths 2018

[36] For a recent overview of the debate and useful typology of retweets as participation, see Dutceac Segesten and Bossetta 2017

[37] Grabe and Bucy 2009; Corrigall-Brown and Wilkes 2012

[38] See Kharroub and Bas 2015

[39] Corrigall-Brown and Wilkes 2012

[40] Rohlinger and Klein 2012

in an experimental setting when using images as treatments. Participants hold more positive view towards immigrant population holding an American than a Mexican flag.[41] In another study, Powell et al. perform an experiment studying individual-level framing effects and find that images shape people's opinions and behavioral intentions more than similar textual content.[42]

Finally, research also finds that images affect political participation in activities other than social protests. A large body of literature demonstrates how images of political candidates affect viewers' evaluations of those candidates and their voting preferences. Todorov et al. for example ask people to evaluate pairs of candidates competing for United States Congress seats based on their visual appearance. Candidates who subjects rated as more competent after glancing at candidate photos often matched the candidate who actually won the electoral seat.[43]

Regarding the relationship between images and participation in mass contentious activities, although many authors suggest that images are important in explaining social protest mobilizations,[44] they do not advance clear theoretical expectations nor empirical tests. In the next section we build on prior work from the fields of visual communication and cognitive psychology to explain the theoretical underpinnings of a positive image effect expectation in the context of online protest mobilization.

2.3 The General Image Effect

The logic of a general image effect on attitudes and behaviors is well established in the existing literature. First, individuals principally learn about the reality surrounding them through experience, and images act as quasi-experiences that trigger a faster and more efficient learning process.[45] Learning from visuals takes place in a specialized part of the brain (the visual cortex) whereas no such specialized area exists for text processing, making learning from text a much more consuming task.[46] Moreover, people are more capable of structuring information learned from visuals and what they learn is more likely to affect their consciousness. Messaris and Abraham point out that because images have an 'analogical' quality (they resemble real life), it is easier for people to 'index' and later

[41] Wright and Citirn 2011
[42] Powell et al. 2015
[43] Todorov et al. 2005
[44] Castells 2012; Bennett and Segerberg 2013; Howard and Hussain 2013
[45] Barry 1997; Gazzaniga 1998; Graber 1996; Grabe and Bucy 2009; Kraidy 2012
[46] Grabe and Bucy 2009

access visually-learned information;[47] and Grabe and Bucy point out that since the visual cortex is in the part of the brain where thinking takes place, the neocortex, visually-learned information has a significant impact on social cognition.[48] This brings Grabe and Bucy to argue that "visual experience remains the most dominant form of learning... visual processing is central to building synaptic connections and ultimately forms the basis of extended awareness,"[49] and Graber to state that "human brains extract valuable information from audiovisuals more quickly and more easily than from purely verbal information."[50]

A large collective action literature portrays information costs as playing a key role in determining the failure or success of mobilizing efforts: people need to know about the existence, costs, and benefits of a mobilization before deciding whether to support it or not.[51] The more rapidly they can process pro-mobilization information, the more likely they are to join the action, underscoring the potential importance of images in the context of protest mobilization.

Second, images trigger stronger emotional reactions than written or spoken information.[52] An extensive literature argues that individual emotional responses are important to understand social mobilization and political participation in general.[53] Marcus et al. show that when exposed to new information, individuals feel first and think second: emotions motivate information-seeking and participation in political processes such as elections.[54] Social movement scholars also argue that emotions generate moral shocks that become motives for mobilization,[55] Existing visual communication literature argues that images are "especially powerful in transmitting realism and emotional appeal"[56] and that "because visual are processed via emotional pathways in the brain, they are inherently affect laden."[57] Thus, the existing literature not only suggests that emotions are important for social mobilization but that images play a key role in generating strong emotional reactions.

[47] Messaris and Abraham 2001
[48] Grabe and Bucy 2009
[49] Grabe and Bucy 2009, 13
[50] Graber 1996
[51] Downs 1957; Olson 1965
[52] Graber 1996; Grabe and Bucy 2009; Barry 1997
[53] J. Jasper 1998; George E. Marcus, Neuman, and M. MacKuen 2000; Goodwin, J. M. Jasper, and Polletta 2004; Flam and King 2005; Goodwin and J. M. Jasper 2006; Valentino et al. 2011; Papacharissi 2014
[54] George E. Marcus, Neuman, and M. MacKuen 2000; Valentino et al. 2011; Ryan 2012
[55] J. Jasper 1998; Papacharissi 2014
[56] Graber 2009
[57] Grabe and Bucy 2009, 8

Although there is a clear expectation of a positive effect based on the cognitive findings described above, to our knowledge the general image effect has not been studied in the context of online protest participation. Our initial empirical hypothesis is therefore a basic test of the image effect in this context:

$\mathbf{H}_1$ (*General Image Effect*): Compared to protest messages without images, messages with images will attract more online attention and recruit more new online participants.

2.4 Images and the Emotional Mechanisms of Online Protest Participation

A key contribution of this paper is to go beyond the general image effect hypothesized above. We wish to better understand the reasons why images are more likely to foster online political participation. To contribute to this goal, we turn our attention to the emotions images evoke. As mentioned above, one of the main reason why images are such a powerful form of communication is because they trigger stronger emotional reactions than their potential textual counterfactual;[58] which means that people are more likely to pay attention to, think about, and react to messages with emotion-evoking images. However, relative to text, all images are potentially more emotionally-triggering. So the question remains: which emotions matter most?

Although some scholars theorize about the effects of a long list of emotions,[59] empirical research shows that groups of emotions covary at high rates and only some can be empirically distinguished. We build on a well established model in political psychology (the *affective intelligence* model)[60] to argue that three empirically distinct emotions have the most mobilizing potential: enthusiasm, anger, and fear.

The affective intelligence model argues that people's emotions are governed by a dual emotional systems: the dispositional and the surveillance systems. The 'dispositional system' is in charge of governing how people feel about issues they deem relevant. Two key emotions are part of this system: anger and enthusiasm. *Anger* "emerges in situations when people are threatened or find obstacles blocking their path to reward"[61] and it

[58] Graber 1996; Barry 1997; Grabe and Bucy 2009

[59] J. Jasper 1998

[60] George E. Marcus, Neuman, and M. MacKuen 2000; Valentino et al. 2011; George E. Marcus, Neuman, and M. B. MacKuen 2017

[61] Brader and George E Marcus 2013, 179

motivates individuals to act in order to find a solution to the threat or to remove the existing obstacle.[62] Individuals experience *enthusiasm* "when the system receives positive feedback about a pursuit, namely when rewards appear within reach, are getting closer, or have been attained."[63] Similar to anger, enthusiasm also might boost participation because there is a desire to achieve certain goals.

On the other hand, the 'surveillance system' is an emotional system in charge of scanning the environment for potential threats. Fear (or anxiety)[64] is in charge of this emotional dimension and it often triggers a reflective process capable of mobilizing those who care about an issue as well as new audiences; it increases the likelihood that people will reconsider their believes, seek further information, and mobilize on new issues.

In a study of electoral turnout in Presidental elections in the United States from 1980 to 2004, Valentinto et al. found that people feeling enthusiasm, anger, and fear during the campaign were more likely to engage in "cheap" forms of participation comparable to the ones under study here.[65] Marcus et al. also found that people feeling these same emotions were also more likely to actively seek information about campaigns.[66]

Turning to the specific role of anger, in the context of social movement participation Van Zomeren and colleagues found in a set of experiments that students feeling angry were more likely to join demonstrations against an increment of tuition fees in the Netherlands.[67] In a similar study, Sturmer and Simon found anger to have a very similar effect on protest mobilization in Germany.[68] We aim to study how images triggering anger affected online mobilization in the context a BLM protest. Although the specific image features triggering anger are not our main focus, we expect images reflecting racial discrimination and injustice to be relevant motivations for anger. Based on the described literature, when exposed to these types of images, we expect people to be more likely to engage with the movement online and so we advance the following hypothesis:

$\mathbf{H}_2$ (*Anger*): Messages with images that generate anger will attract more online attention

[62] Valentino et al. 2011; Brader and George E Marcus 2013

[63] Brader and George E Marcus 2013, 175

[64] Although fear and anxiety can be theoretically distinguished, empirical evidence show that they are highly correlated and difficult to distinguish in practice George E. Marcus, Neuman, and M. MacKuen 2000; Brader 2005 For this reason in this paper we use the Valentino et al. 2011 approach and we treat fear and anxiety interchangeably.

[65] Valentino et al. 2011

[66] George E. Marcus, Neuman, and M. MacKuen 2000

[67] Zomeren, Spears, Fischer, et al. 2004

[68] Stürmer and Simon 2009

and recruit more new online participants.

The literature on social movements and emotions also argues that enthusiasm should encourage people to support and join a movement.[69] In relation to the fight for racial equality in the United States, Jasper highlights that enthusiasm-evoking speeches by movement leaders in the 1906s played a key role in keeping people engaged with all movement actions.[70] Research on the role of enthusiasm in fostering movement participation is scarce, but work on other types of political behavior point in this direction. Apart from Marcus et al.'s and Valentino et al's. work mentioned earlier,[71] in an experimental setting Brader finds that political ads triggering enthusiasm motivate participation and strengthen party loyalty.[72] As it happened during the Civil Rights movements,[73] we expect numerous images related to the BLM protest under study will show people taking action that will likely evoke enthusiasm. Based on the described literature we hypothesize the following:

$\mathbf{H}_3$ (*Enthusiasm*): Messages with images that generate enthusiasm will attract more online attention and recruit more new online participants.

The social movements literature also highlights the mobilizing role of fear.[74] As it happens with enthusiasm, empirical research on social movement participation triggered by fear-evoking information is also scarce; but as Marcus et al. and Brader point out,[75] fear-evoking information is likely to make people devote more attention to an issue and also to act on it. In the context of the BLM movement, we expect numerous images to evoke fear, particularly images related to police violence. Building on this literature, we expect these images to capture people's attention and to motivate them to participate in spreading the word about the movement:

$\mathbf{H}_4$ (*Fear*): Messages with images that generate fear will attract more online attention

[69] Castells 2012
[70] J. Jasper 1998
[71] George E. Marcus, Neuman, and M. MacKuen 2000; Valentino et al. 2011
[72] Brader 2005
[73] Raiford 2007
[74] J. Jasper 1998; Kemper 2001
[75] George E. Marcus, Neuman, and M. MacKuen 2000; Brader 2005

and recruit more new online participants.

As a final consideration, a recent study on the online diffusion of the BLM movements finds a fourth emotion, sadness, to have a some mobilizing effect.[76] The finding does not align with existing political psychology literature, which argues that sadness should be demobilizing: "[it] motivates withdrawal and more effortful processing of information, encouraging individuals to accept the loss, reflect on their situation, and change goals and plans accordingly."[77] We incorporate sadness into our analysis to adjudicate between these contradictory findings and to illustrate the point that not all emotions (nor all images) are mobilizing. Numerous BLM-related images could trigger this emotion, such as images of victims of police violence. Building on existing political psychology research we propose the following hypothesis:

$\mathbf{H}_5$ (*Sadness*): Messages with images that generate sadness will attract less online attention and recruit fewer new online participants.

3 Research Design

In setting out to test the above hypotheses, we faced four fundamental research design challenges: 1) case selection; 2) measuring online social movement mobilization; 3) treating images as data; and 4) making valid causal claims.

To assuage case selection concerns, we chose our case in advance of the protest event. In the spring of 2015, we learned of an upcoming BLM action against police brutality, called ShutdownA14, which would be held on April 14, 2015. We decided to track this case on Twitter without knowing in advance if any images would be spread online during our established protest window of April 13-20, 2015. ShutdownA14 was organized by a coalition of activist groups, including the Stop Mass Incarceration Network and Hands Up United. Actions took place on the national level with demonstrations in numerous cities. The organizing groups not only called for a mobilization on the streets but also coordinated an online social media campaign. To promote the movement, organizing materials

[76] De Choudhury et al. 2016
[77] Brader and George E Marcus 2013, 176-177

asked people to share messages about the protest and its goals by using specific hashtags and keywords such as #shutdownA14, #policebrutality, and #murderbypolice. In addition, organizing materials and tweets about the protest often included #blacklivesmatter, highlighting the crossover between the April 14 protest and the broader BLM movement.

As previously mentioned, we use two operationalizations of online participation in our analysis: attention (number of retweets of ShutdownA14 and BLM related tweets) and diffusion (number of ShutdownA14 messages retweeted by new ShutdownA14 users). We care about attention because it is a necessary condition for a movement to exist or to succeed at setting policy agendas.[78] We study diffusion because it is key for social movements in order to achieve larger support and be more likely to set media and political agendas.[79] Diffusion here is therefore conceptually equivalent to online recruitment into the action. Given the narrow time frame of our observational data collection, we were unable to analyze the number of new recruits to BLM due to April 14 images, as individuals may have been active online in the movement long before ShutdownA14. This is why we focus our diffusion analysis on the April 14 protest.

Our next challenge was in treating images as data. While computer programs have become more adept at categorizing images, the level of detail and emotional response data that we required from the collected ShutdownA14 images necessitated human coding. We worked with both university undergraduates and Mechanical Turk workers to label the roughly 9,500 unique images collected over the course of the ShutdownA14 protest. Each image was manually labeled on each of the hypothesized emotion mechanisms, so that we could analyze each mechanism in isolation by controlling for the remaining mechanisms.

Finally, we recognize the difficulty of making causal claims for our hypotheses and analyses. One contribution of this paper is our attempt to describe important patterns in human behavior using messy, real-world data and events, which necessarily poses a challenge for causal research. We attempt to rule out alternative explanations using a four-fold strategy.

First, we control for three alternative mechanisms through which images may have a direct impact on online protest participation: '*disgust*', '*expectation of success*' and '*social collective identity*'. Recent political psychology literature has been exploring the effect of

[78]Kingdon 1984; Baumgartner, De Boef, and Boydstun 2008; Casas, Davesa, and Congosto 2016
[79]Barberá et al. 2015; De Choudhury et al. 2016; Freelon, C. McIlwain, and M. Clark 2016

disgust on political attitudes towards policies such as health[80] and homelessness.[81] The effects of disgust on behavior are much less clear but possible.

Existing literature also suggests that people with a material interest in joining a collective action will do so if their action is likely to make a contribution.[82] Joining a small movement may not be rational when it is unlikely to succeed, but this may change when the expectations of success increase.[83] Images of protesting crowds may encourage participation because they generate enthusiasm but also because they increase people's perception of the movement's potential for success.

Another set of scholars argue that a sense of collective identity also creates motivations for individuals to join collective actions[84] In constructing and connecting purposes, and thus in building motives for others to join the movement, symbols may play a very important role.[85] We aim to rule out the effect of these three alternative mechanisms by controlling for whether images evoke disgust and for the presence of protesting crowds and symbols of collective identity in the images.

Second, to assuage concerns that the text of the tweets are driving results, in Appendix D we check whether our findings hold when controlling for the topics in the text messages. Third, to rule out the possibility that other salient image features, such as the level of informativeness, have a direct impact on protest participation, in Appendix E we check whether the presence of some salient image features (e.g. the presence of text and/or police in the image) have an effect on our main findings. Finally, we also control in all our models for characteristics of the users sending the tweets that are known to impact engagement.

4 Data & Measurement

Our data are Twitter messages related to the BLM movement and to the ShutdownA14 protest. We used hashtags promoted by the groups organizing the demonstrations and a similar set of keywords to identify which messages were about the protest. The hashtags and keywords came from the websites of the main organizing groups. From April 13 to

[80] Clifford and Wendell 2016
[81] Clifford and Piston 2016
[82] e.g Downs 1957; Olson 1965
[83] Klandermans 1984; Kuran 1997; Kharroub and Bas 2015
[84] cf Polletta and J. M. Jasper 2001; Tajfel 1982; Zomeren, Spears, and Leach 2008
[85] Kharroub and Bas 2015, 7

April 20, we collected all Twitter messages containing the hashtags and keywords in Table 1 using the Twitter Streaming API.

Table 1: List of Hashtags and Keywords Used to Collect the Tweets

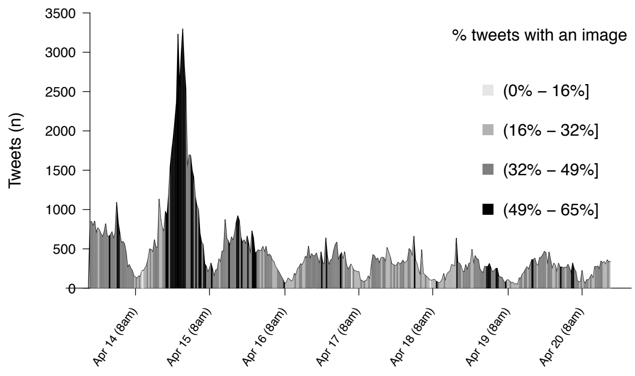
ShutdownA14	BLM	
#shutdownA14	murder by police	mass incarceration
shutdownA14	killer cops	police murder
	stop business as usual	stolenlives
	massincarceration	stolen lives
	#policebrutality	#stolenlives
	#blacklivesmatter	black lives

We looked at this particular case and both BLM and ShutdownA14 messages because it allowed us to test the effect that images have on both attention and diffusion. We obtained a data set with 150,324 tweets sent by 67,484 unique users; 26.8% of the messages were related to the specific ShutdownA14 protest, and about 43.2% of all messages contained an image.

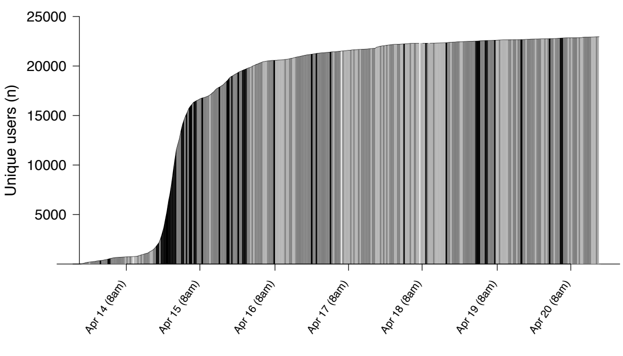
Figure 1 displays a general summary of our data over time by dividing the tweets into periods of 30 minutes. The first panel shows the percent of BLM and ShutdownA14 tweets in a given time period with an image and the total number of tweets for that period. We see a general trend, in that there seems to be a congruence between high concentrations of messages with images and larger numbers of protest-related tweets. The second panel shows the percent of ShutdownA14 tweets in a given time period with an image and the number of new tweeters for that period (displayed as cumulative unique users, where the slope shows the rate of recruitment of new users). Again, we see a general trend where a high concentration of images appears to track with a spike in the number of new tweeters.

Figure 1: Percentage of posts over time with images and (a) total tweets and (b) number of new tweeters

(a) Number of tweets (attention)



(b) Number of new tweeters (diffusion)



One concern with modeling attention and diffusion using the aggregated, 30-minute time-break data as shown in Figure 1 is that we cannot be sure that posters are responding to having seen images shared by friends on Twitter. Our analysis strategy reduces this concern by focusing on retweets.

4.1 Main Variables

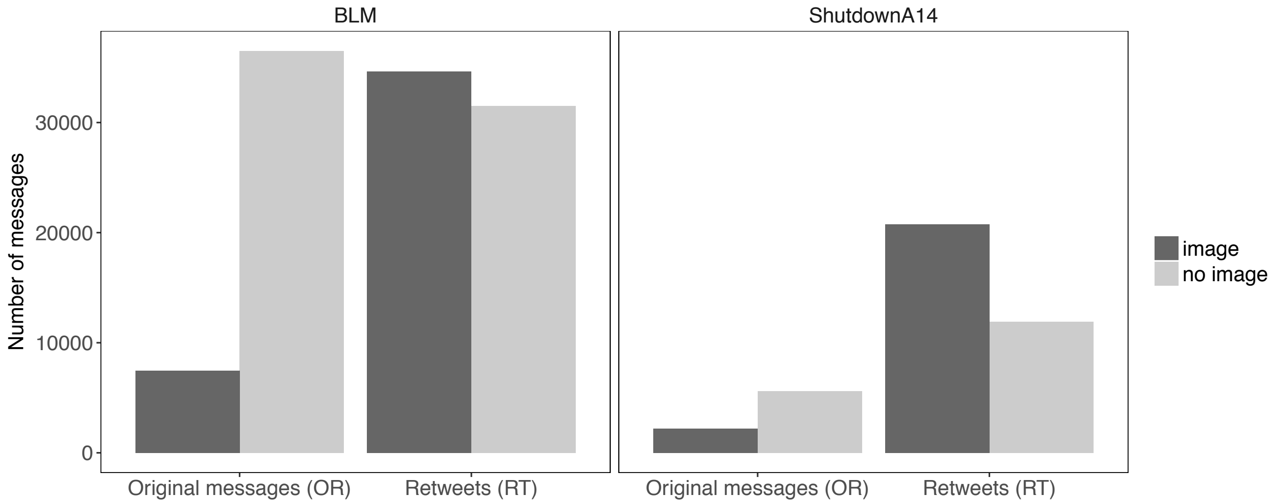
We split the Twitter messages into original tweets and retweets. We then linked retweets to their original tweet in order to count how many times an original tweet was subsequently retweeted. The count of retweets is our measure of movement *attention*, where the original tweets could include either BLM or ShutdownA14 hashtags and keywords. To measure movement *diffusion*, we consider only the ShutdownA14 tweets, for reasons described above. We check which retweeters of a given original message had never before tweeted about the ShutdownA14 protest. If their first tweet in regards to the protest was a retweet, we count them as an individual to whom the original tweet diffused the protest.

Our first key explanatory variable of interest is whether or not an original tweet contained an image (*image*). This is a binary variable derived from the Twitter data available for each original tweet. Figure 2 provides an overview of the data. The plot on the left provides information about messages in our dataset that contain at least one of the BLM hashtags from Table 1 while the plot on the right provides information for messages containing only ShutdownA14 hashtags. Both plots show similar trends: as expected, the number of original messages is smaller than the number of retweets. A majority of the original messages do not include an image, but most of the retweets are of original messages that do include an image.

4.2 Mechanism Variables: Images As Data

To address our mechanism hypotheses, we required information about each particular image, not simply the number of individual tweets with images. A week after the protest, we wrote a computer program to collect all of the images that were present in the tweets, using the image links provided by the Twitter Streaming API. Some tweets had the same image under a different link, so we first identified which images were the same, following a three-step procedure (see Appendix A). After collecting and matching all images, we ended up with a dataset of 9,458 unique images.

Figure 2: Number of tweets about the overall BLM movement and number of tweets only about the ShutdownA14 protest. Each panel shows original messages *versus* retweets, split based on whether the message contained an image.



The next step was to manually label each of the 9,458 images for the presence of our hypothesized mechanisms and alternative explanations. We needed to know how much anger, enthusiasm, fear, sadness, and disgust each imaged evoked (*emotional* mechanisms), and whether a protest (*expectation of success* mechanism) and/or a symbol (*social collective identity* mechanism) were present in the image. We had two main concerns during this labeling process. First, we wanted to make sure that the labels for the top 1,000 most-tweeted images[86] (949 after removing duplicates) were reliable since these would strongly influence the analysis: the distribution of the images was right skewed, with a few images being highly tweeted and the rest being tweeted only once or a couple times.

Second, for modeling purposes we needed to give to each unique image one score per emotion. However, emotions are subjective and the same image might trigger different emotions in different people. We wanted to make sure that the emotion scores for the most influential images were the result of multiple emotional reactions, and that on average different people reacted with similar emotional intensity to these images.

We mitigated these concerns by having 5 people label each of the top 1,000 images (two undergraduate research assistants and three Mechanical Turk workers) and by limiting to 100 the number of images a single Mechanical Turk annotator could label. This meant that the large and diverse pool of people who participated in the labeling process (a total

[86]These 'most-tweeted images' are the pictures that were present in the largest number of original tweets and retweets. To calculate this frequency we first matched all unique images in our dataset.

of 1,259 Mechanical Turk annotators, see Appendix B for demographics of the annotators) were each assigned up to 100 images to assess. The remaining images (n = 8,509) were labeled only once by individuals from the large Mechanical Turk pool.

Annotators indicated the extent to which an image evoked each of the five emotions (this generated five 0-10 scores per image), and whether each image had a protest (a binary indicator) or a symbol (a binary indicator). We then built seven image-level mechanism variables. *Anger*, *Enthusiasm*, *Fear*, *Sadness*, and *Disgust* are each continuous variables ranging from 0 to 10 addressing the *emotional* mechanisms ($H_{2,3,4,5}$). *Protest* is a binary variable addressing the alternative *expectation of success* mechanism and *Symbol* is a binary variable addressing the alternative *collective identity* mechanism. The emotional score for each emotion on each image is the average of the values given by all five annotators. For the top 1,000 images, we considered an image as having a protest or a symbol if at least one of the five annotators indicated the presence of these elements.

A copy of the labeling protocol form, two examples of labeled images, and a summary table for these seven mechanism variables plus the controls can be found in Appendix A. Appendix B contains interrater reliability measures for the two undergraduate research assistants, showing that on average they coded the same images as evoking similar emotional intensities.

Figure 3 contains four examples of images that annotators indicated as evoking high levels of anger, enthusiasm, fear, and sadness. These and additional examples available in Appendix E are a good representation of the type of images that triggered each of the emotions under study. As expected, anger was mainly triggered by images illustrating scenarios that one can see as unfair, enthusiasm by images of people taking action, fear by images of police action and brutality, and sadness by pictures memorializing people killed by police (see Appendix E for more on associations between image features and evoked emotions).

Finally, we matched each unique image to all original messages containing that picture. We also extracted from each original message user level characteristics that are known to have an impact on participation.[87] For each sender and message we extracted the number of followers, the number of friends, the number of previous tweets about the protest, and the time of day a tweet was sent. Table 2 provides a brief description of all of the study variables. The final output is a dataset with information about each original BLM and

[87] Gonzalez-Bailon et al. 2011; Barberá et al. 2015

Figure 3: Examples of images significantly evoking the four hypothesized emotions

(a) Anger

(b) Enthusiasm

(c) Fear

(d) Sadness

ShutdownA14 message that we use to model the image effects.

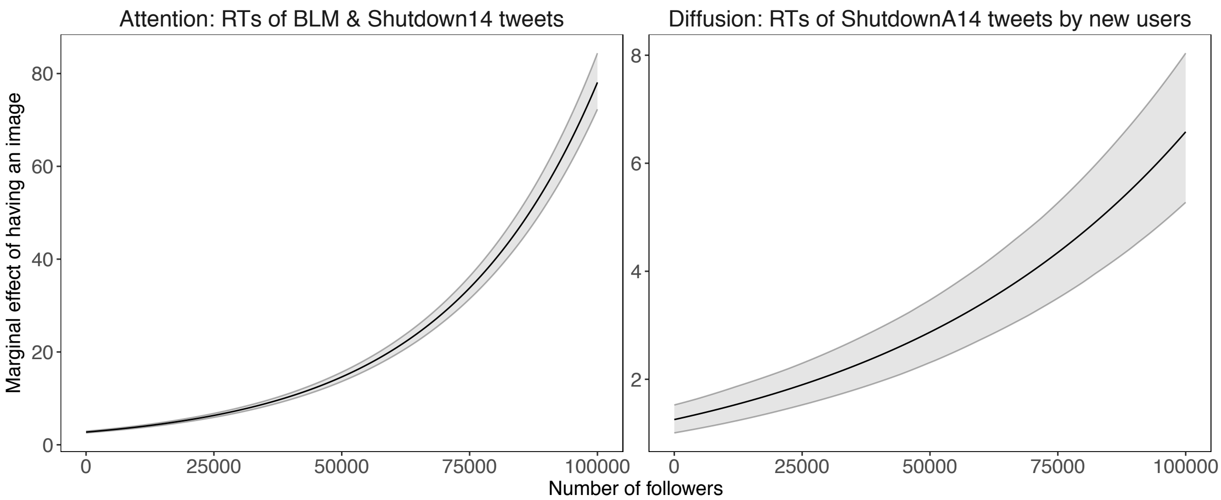
Table 2: Study variable descriptions

Variable	**Description** (Unit of Analysis = Original Tweet)
	Outcome Variables
BLM and A14 tweets (attention)	Number of retweets for tweets mentioning any of BLM hashtags/keywords from Table 1
A14 new users (diffusion)	Number of retweets from users mentioning the A14 hashtags/keywords for the first time
	Explanatory Variables
Image	Whether or not the tweet contains an image
Anger	Average anger score evoked by the image (0-10)
Enthusiasm	Average enthusiasm score evoked by the image (0-10)
Fear	Average fear score evoked by the image (0-10)
Sadness	Average sadness score evoked by the image (0-10)
Disgust	Average disgust score evoked by the image (0-10)
Symbol	Whether or not the image contains a symbol
Protest	Whether or not the image is of a street protest
	Control Variables
Number of followers	Number of followers of original tweeter
Number of friends	Number of friends of original tweeter
Number of previous tweets	Number of previous tweets by the original tweeter in the dataset
Time	6-class categorical variable (each class is a 4-hour break)

5 Modeling and Results

We use negative binomial models to predict how our factors of interest affect the number of times an original tweet is retweeted. We first model the *attention* to the overall BLM movement (Model 1, with 49,345 original messages) and the *diffusion* of the ShutdownA14 action (Model 2, with 7,502 original messages). In both of these basic models the independent variable of interest is whether or not an original tweet contained an image. We then control for the number of followers, number of friends, number of previous tweets, and time. The regression table with the results for these models can be found in Appendix C. Our interpretation focuses on the marginal effects of our hypothesized explanatory variables.

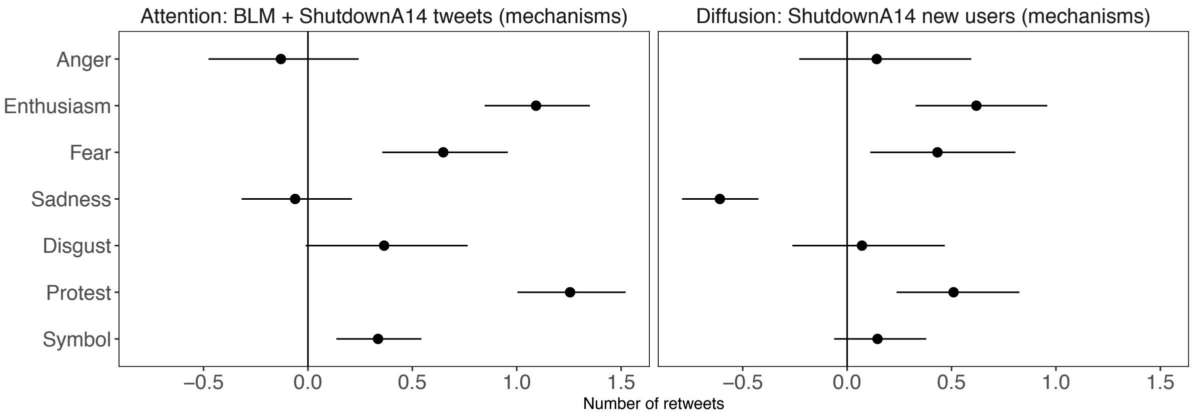
Figure 4: Marginal effect of an original tweet having an image versus not having an image on the number of retweets (on the left) and number of retweets by new users (on the right). Marginal effect shown over a selected range of number of followers.



The results of the two basic multivariate analyses shown in Figure 4 are consistent with the well established *General Image Effect* hypothesis (H_1): the likelihood of a protest tweet to diffuse to new recruits and to get more attention increases if the tweet contains an image. Using the general BLM data, we find that for users with few followers (e.g. 1,000), including an image with an original message means getting approximately three more retweets than they would have if they had not included an image (holding all else constant at the mean). Using the specific ShutdownA14 tweets, and again considering an original tweeter with 1,000 followers, we find that tweets with images on average recruit one more new retweeter than tweets without images. The marginal effect is higher for users with a larger number of followers. For example, a BLM message with an image from a hypothetical user with 75,000 followers would get about 35 more retweets compared to a tweet without an image. Original ShutdownA14 messages from the same user would get 4 more retweets from people messaging about the protest for the first time if the tweet has an image.

The next step is to test to what extent the hypothesized mechanisms explain why images related to a protest increase attention and diffusion. We estimate two new negative binomial models (Models 3 and 4) only using information from original messages that had an image (8,706 original tweets in Model 3 and 2,078 original tweets in Model 4). In this case we include all of the mechanism variables (Anger, Fear, Sadness, Disgust, Enthusiasm, Protest, Symbol) while keeping the same controls (Number of Followers, Number of Friends, Number of Previous Tweets, and Time).

Figure 5: Predicting attention and diffusion by image mechanisms (Negative Binomial Models)*



**Standardized coefficients (the effect of a variable moving from its mean to 1 standard deviation above)*

The results, shown in Figure 5, are supportive of most of the hypothesized mechanisms.[88] The coefficient for *anger* (H_2) is negative in the attention model, contradicting our initial hypothesis, but in neither model are the coefficients statistically significant. We observe as expected that, all else equal, an increase in the amount of *enthusiasm* (H_3) and *fear* an image evokes increases both attention and diffusion. In both models the mobilizing effect of enthusiasm is larger, on average, than the effect of fear. Also as expected, we find images evoking *sadness* to be mainly demobilizing (H_4): statically significant effects in the model predicting diffusion, and negative although not significant effect in the model predicting attention. This contradicts the findings of a recent paper studying the diffusion

[88]See Appendix C for the regression tables.

of BLM in social media[89] but aligns with the existing political psychology research.[90]

Figure 6: Predicting attention to BLM over range of evoked emotions

Figure 7: Predicting diffusion of ShutdownA14 over range of evoked emotions

Figures 6 and 7 highlight the differential effects of the emotions evoked by images. In Figure 6 we see that as the amount of fear increases, attention also increases. Holding all of the other variables at their means, increasing the anger evoked by an image from 0 to 10 increases the predicted number of retweets by about 3. A similar change in enthusiasm has an effect of about 5 more retweets. Increasing the amount of anger in an image decreases the attention a tweet receives, though the size of the effect is slight. For sadness, there is no substantial or significant effect over the range of evoked emotion.

Similar trends appear in Figure 7. Increasing anger seems to have essentially no effect on the diffusion of the movement, while increasing the amount of enthusiasm or fear evoked does increase diffusion. An increase in both enthusiasm and fear from 0 to 10, holding all else at the mean, increases the predicted number of new user retweets by approximately 4. Increasing sadness decreases the diffusion to new protest tweeters, with a decrease of

[89] De Choudhury et al. 2016

[90] Brader and George E Marcus 2013

about 1 new user retweet over the range of sadness. These findings demonstrate the varied effects of images on mobilization based on the emotional responses they generate. The same findings hold when controlling for the topics in the text of the messages (Appendix D) and for the presence of other salient image features (Appendix E).

As for the proposed alternative mechanisms, some of them do have a positive effect on mobilization. Images containing protesting crowds ('expectation of success' mechanism: 'Protest' coefficient in Figure 5) are associated with higher levels of online protest attention and diffusion. For a Twitter user with 1,000 followers, holding all else constant, an image with protesting crowds translates into 5 more retweets and 2 more retweets from new users. Images with symbols ('collective social identity' mechanism: 'Symbol' coefficient in Figure 5) have a positive effect on attention and diffusion but the effect is only statistically significant in the attention model. For a similar user with 1,000 followers, retweeting a symbol of collective social identity such as the American flag translates on average into 1 more retweet. Finally, the alternative emotional 'disgust' mechanism is associated with higher attention and diffusion but the effect is not statistically significant in either of the models.

5.1 Limitations and Alternative Explanations

The findings above have clear statistical and substantive significance. That said, while a strength of this work is its leveraging of a large corpus of real world images and responses to those images, our causal conclusions rest on the validity of our labels and our ability to control for alternative explanations. We find our results to be in line with prior work that uses experimental research designs to assess the impact of emotions, but we would encourage future research in a variety of settings on the effects of these messy images, which often vary on multiple dimensions.

We acknowledge the particular complications that arise when testing for the effects of emotions. Even in experimental designs, it can take extensive pretesting to conclude that a given treatment does in fact elicit the intended emotion, especially if a researcher is intending to elicit a single emotion (e.g. anger) without also triggering similar feelings (e.g. anxiety).[91] Our approach in attempting to isolate causal emotions mechanisms has been to control for other emotions and for characteristics of the tweet and original tweeter.

[91] Albertson and Gadarian 2016

In addition, we control for two theoretically derived alternative mechanism pathways. First, to account for possible *expectation of success* effects, we control for whether or not the picture contains a large protesting crowd. Second, to rule out possible *social collective identity* effects, we control for whether or not the picture contains a social symbol. Additional controls for image and text features were added for the robustness checks in Appendices D and E.

We also note that our findings rest on the accuracy of our annotator labels. We have self-reported emotional reactions to images, without confirmation of reactions from, for example, physiological data. Our image annotators could be reporting the emotion that they think they are supposed to feel, as opposed to what they actually feel. While our interrater reliability scores are acceptable for what is objective about a picture (e.g. whether it is of a large street protest), they are less strong for labels which could be considered subjective (e.g. emotional reactions, or whether the image contains a social identity symbol).[92]. While we suspect that it is precisely the subjective nature of these features that drives the variation in responses, this is a fruitful area for future research. It may be, for example, that demographic characteristics of our labelers predicts the variation we observe in emotional responses. That is, there may be an additional alternative explanation relating to an in-group identity which we are not capturing.[93]

Despite our efforts, we are unable to obviate every potential weakness in our research design. Well-designed randomized experiments, following the best practices of Albertson and Gadarian (2016) and using images either from our dataset or manipulated by researchers to vary only one emotional component (while holding other features such as group size and symbols constant), would provide an excellent complement to the presented findings.

6 Discussion and Conclusion

Despite the prevalence of images in modern life and the prior literature on the importance of images in swaying political opinions and behavior, very little research has leveraged large quantities of observational data to test the role of images in mobilizing political activism. In this paper we test the general image effect theory in the context of a Black Lives Matter

[92]See Appendix B, Table 3
[93]Wright and Citirn 2011

protest. We then specify and test a set of emotional mechanisms explaining why images might increase the likelihood of a protest to receive more attention and to diffuse to new participants online. We test these hypotheses using observational data, including roughly 150,000 tweets and 9,500 unique images from a Black Lives Matter (BLM) protest that took place in April 2015: ShutdownA14.

We argue that images are particularly mobilizing because, compared to text, they trigger stronger emotional reactions. The literature on political and social movement participation argues some emotions can play a mobilizing role. We build on the *affective intelligence* model to argue images evoking anger, enthusiasm, and fear should be associated with higher levels of online protest attention and diffusion, whereas those triggering sadness should discourage participation.

In line with the theoretical predictions, we find that in the context of the ShutdownA14 BLM protest promoted on Twitter, messages with images were more likely to be retweeted and were more likely to receive retweets from individuals who had not previously tweeted about the protest. Images evoking enthusiasm increased attention and diffusion, as did images evoking fear. Images triggering sadness appeared to depress attention and diffusion, while the effects of anger were imprecise. The findings hold after accounting for alternative theoretical mechanisms as to why images should be mobilizing, and after controlling for a wide range of sender and message level features.

Given that we test these theories on only one movement and one specific protest, a note of caution about generalizing from these results is warranted. Scholars have noted that the Black Lives Matter movement has some unique features, including a lack of highly centralized organizing and the heightened role of social media in increasing support for the issues in the absence of more traditional media coverage.[94] Social media behavior for BLM may be fundamentally different from other movements. Image features that mobilize Black Lives Matter supporters may differ from features that mobilize other groups and, images aside, the underlying dynamics of retweeting in the movement may differ from other social movement cases. These results should be validated by further research, especially work that includes a larger variety of protest movements and protest incidents. In addition, retweeting is not the only form of online mobilization of interest to scholars, and future work should compare the retweeting image effects we find to the effects of images on other participation measures, such as signing petitions, donating to a cause, or participating in

[94]Freelon, C. D. McIlwain, and M. D. Clark 2016

offline protests.

Our study contributes to broad and increasingly relevant discussions of collective action in the age of social media.[95] The ability to send and receive images via social media is a transformative force in social organizing, allowing groups and individuals to circumvent traditional mass media channels. Crucially, images have historically helped marginalized populations put their interests on the public agenda, and the explosion of images via social media may serve to amplify these voices.[96] Our study of Black Lives Matter thus illuminates some of the important intersections of organizing, social media, and the mobilizing power of images.

[95] Bennett and Segerberg 2013; Bimber, Flanagin, and Stohl 2012
[96] Freelon, C. McIlwain, and M. Clark 2016

References

Albertson, Bethany and Shana Kushner Gadarian (2016). "Did that scare you? Tips on creating emotion in experimental subjects". In: *Political Analysis* 24.4, pp. 485–491.

Barberá, Pablo et al. (2015). "The Critical Periphery in the Growth of Social Protests". In: *PLOS ONE* 10.11. Ed. by Stephen C Pratt, e0143611.

Barry, Ann Marie (1997). *Visual Intelligence: Perception, Image, and Manipulation in Visual Communication*. Albany, NY: SUNY Press.

Bas, Ozen and Maria Elizabeth Grabe (2016). "Personalized News and Participatory Intent: How Emotional Displays of Everyday Citizens Promote Political Involvement". In: *American Behavioral Scientist* 60.14, pp. 1719–1736.

Baumgartner, Frank, Suzanna De Boef, and Amber Boydstun (2008). *The Decline of the Death Penalty and the Discovery of Innocence*. Cambridge: Cambridge University Press.

Benford, Robert D. and David A. Snow (2000). "Framing Processes and Social Movements: An Overview and Assessment". In: *Annual Review of Sociology* 26.1, pp. 611–639.

Bennett, W Lance and Alexandra Segerberg (2013). *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. New York: Cambridge University Press.

Bimber, Bruce A, Andrew J Flanagin, and Cynthia Stohl (2012). *Collective action in organizations: interaction and engagement in an era of technological change*. Communication, society and politics. New York: Cambridge University Press.

Boyd, Danah, Scott Golder, and Gilad Lotan (2010). "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter". In: *HICSS-43. IEEE*. Kauai, HI.

Brader, Ted (2005). "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions". In: *American Journal of Political Science* 49.2, pp. 388–405.

Brader, Ted and George E Marcus (2013). "Emotion and Political Psychology". In: *The Oxford Handbook of Political Psychology*. Ed. by Leonie Huddy, David O. Sears, and Jack S. Levy. Oxford: Oxford University Press.

Bush, Sarah Sunn et al. (2016). "The Effects of Authoritarian Iconography: An Experimental Test". In: *Comparative Political Studies*.

Casas, Andreu, Ferran Davesa, and Mariluz Congosto (2016). "The media coverage of a 'connective action': The interaction between the Spanish Indignados and the mass media". In: *Revista Española de Investigaciones Sociológicas* 155, pp. 73–96.

Castells, Manuel (2012). *Networks of outrage and hope: social movements in the Internet Age*. Cambridge, UK ; Malden, MA: Polity Press.

Chadwick, Andrew (2011). "The Political Information Cycle in a Hybrid News System: The British Prime Minister and the 'Bullygate' Affair". In: *The International Journal of Press/Politics* 16.1, pp. 3–29.

Clifford, Scott and Spencer Piston (2016). "Explaining Public Ambivalence about Homelessness Policy: The Role of Disgust". In: *Political Behavior*.

Clifford, Scott and Dane G. Wendell (2016). "How Disgust Influences Health Purity Attitudes". In: *Political Behavior* 38.1, pp. 155–178.

Corrigall-Brown, Catherine and Rima Wilkes (2012). "Picturing Protest: The Visual Framing of Collective Action by First Nations in Canada". In: *American Behavioral Scientist* 56.2, pp. 223–243.

Dahmen, Nicole S (2012). "Photographic Framing in the Stem Cell Debate". In: *American Behavioral Scientist* 56.2, pp. 189–203.

De Choudhury, Munmun et al. (2016). "Social Media Participation in an Activist Movement for Racial Equality". In: *Proceedings of the 10th International AAAI Conference on Weblogs and Social Media*. Cologne, Germany.

Downs, Anthony (1957). *An Economic Theory of Democracy*. New York: Harper & Row.

Dutceac Segesten, Anamaria and Michael Bossetta (2017). "A typology of political participation online: How citizens used Twitter to mobilize during the

2015 British general elections". In: *Information Communication and Society* 20.11, pp. 1625–1643. URL: https://doi.org/10.1080/1369118X.2016.1252413.

Earl, Jennifer and Katrina Kimport (2011). *Digitally Enabled Social Change: Activism in the Internet Age*. Cambridge, MA: MIT Press.

Flam, Helena and Debra King, eds. (2005). *Emotions and Social Movements*. London: Routledge.

Freelon, Deen, Charlton D McIlwain, and Meredith D Clark (2016). *Beyond the Hashtags*. Tech. rep. Washington, DC: Center for Media & Social Impact, School of Communication, American University.

Freelon, Deen, Charlton McIlwain, and Meredith Clark (2016). "Quantifying the power and consequences of social media protest". In: *New Media & Society*.

Gazzaniga, Michael S. (1998). *The Mind's Past*. Berkely and Los Angeles, CA: University of California Press.

Gitlin, Todd (1980). *The whole world is watching: Mass media in the making and unmaking of the new left*. Berkeley: University of California Press.

Gonzalez-Bailon, Sandra et al. (2011). "The Dynamics of Protest Recruitment through an Online Network". In: *Scientific Reports* 1, p. 197.

Goodwin, Jeff and James M Jasper (2006). "Emotions and Social Movements". In: *Handbook of the Sociology of Emotions*. Ed. by Jan E Stets and Jonathan H. Turner. Boston: Springer, pp. 611–635.

Goodwin, Jeff, James M Jasper, and Francesca Polletta (2004). "Emotional Dimensions of Social Movements". In: *The Blackwell Companion to Social Movements*. Ed. by David A Snow, Sarah A Soule, and Hanspeter Kriesi. Malden: Blackwell, pp. 413–432.

Grabe, Maria Elizabeth and Erik Page Bucy (2009). *Image bite politics: news and the visual framing of elections*. Oxford; New York: Oxford University Press.

Graber, Doris A (1996). "Say it with Pictures". In: *The Annals of the American Academy of Political and Social Science* 546.The Media and Politics, pp. 85–96.

Graber, Doris A (2009). *Mass Media and American Politics*. Washington, DC: CQ Press.

Griffiths, James (2018). *Trump says he didn't know UK far-right group were 'horrible racists'*. URL: https://www.cnn.com/2018/01/26/politics/trump-britain-first-apology-intl/index.html.

Howard, Philip N. and Muzammil M. Hussain (2013). *Democracy's Fourth Wave?: Digital Media and the Arab Spring*. New York, NY: Oxford University Press.

Hsiao, Yuan and Yunkang Yang (2018). "Commitment in the cloud? Social media participation in the sunflower movement". In: *Information, Communication & Society*, pp. 1–18.

Jasper, James (1998). "The Emotions for Protest: Affective and Reactive Emotions In and Around Social Movements". In: *Sociological Forum* 13.3, pp. 397–424.

Kemper, Theodore D. (2001). "Passionate Politics: Emotions and Social Movements". In: *Passionate Politics: Emotions and Social Movements*. Ed. by Jeff Goodwin, James M. Jasper, and Francesca Polletta. Chicago: The University of Chicago Press, pp. 58–73.

Kharroub, Tamara and Ozen Bas (2015). "Social media and protests: An examination of Twitter images of the 2011 Egyptian revolution". In: *New Media & Society*. URL: http://nms.sagepub.com/cgi/doi/10.1177/1461444815571914.

Kingdon, John W. (1984). *Agendas, Alternatives, and Public Policies*. Boston: Little Brown.

Klandermans, Bert (1984). "Mobilization and Participation: Social-Psychological Expansions of Resource Mobilization Theory". In: *American Sociological Review* 49.5, p. 583.

Kraidy, Ute (2012). "Digital Media and Education: Cognitive Impact of Information Visualization". In: *Journal of Educational Media* 27.3, pp. 95–106.

Kuran, Timur (1997). *Private truths, public lies: the social consequences of preference falsification*. 1st Harvar. Cambridge, Mass: Harvard University Press.

Lupia, Arthur and Gisela Sin (2003). "Which Public Goods are Endangered?: How Evolving Communication Technologies Affect The Logic of Collective Action". In: *Public Choice* 117.3, pp. 315–331.

Marcus, George E., W. Russell Neuman, and Michael MacKuen (2000). *Affective Intelligence and Political Judgement*. Chicago and London: University of Chicago Press.

Marcus, George E., W. Russell Neuman, and Michael B. MacKuen (2017). "Measuring Emotional Response: Comparing Alternative Approaches to Measurement". In: *Political Science Research and Methods* 5.4, pp. 733–754.

McAdam, Doug (2000). "Culture and Social Movements". In: *Culture and Politics: A Reader*. Ed. by Lane Crothers and Charles Lockhart. New York: Palgrave Macmillan US, pp. 253–268.

McCann, Michael (2006). "Law and Social Movements: Contemporary Perspectives". In: *Annual Review of Law and Social Science* 2.1, pp. 17–38.

Messaris, Paul and Limus Abraham (2001). "The Role of Images in Framing News Stories". In: *Framing Public LIfe: Perspectives on Media and Our Understanding of the Social World*. Ed. by Reese Stephen D., Gandy Oscar H., and August E Grant. Mahwah, New Jersey: Lawrence Erlbaum Associates Plublishers, pp. 215–226.

Morozov, Evgeny (2011). *The net delusion: The dark side of Internet freedom*. New York: PublicAffairs.

Olson, Mancur (1965). *The Logic of Collective Action: Public Goods and the Theory of Goups*. Cambridge, MA: Harvard University Press.

Papacharissi, Zizi (2014). *Affective Publics: Sentiment, Technology, and Politics*. Oxford: Oxford University Press.

Payne, Charles (1998). "Debating the Civil Rights Movement: The View from the Trenches". In: *Debating the Civil Rights Movement, 1945-1968*. Ed. by Steven Lawson and Charles Payne. Lanham, MD: Rowman and Littlefield.

Penney, Joel and Caroline Dadas (2013). "(Re)Tweeting in the service of protest: Digital composition and circulation in the Occupy Wall Street movement". In: *New Media & Society2* 16.1, pp. 74–90.

Polletta, Francesca and James M Jasper (2001). "Collective Identity and Social Movements". In: *Annual Review of Sociology* 27, pp. 283–305.

Powell, Tom et al. (2015). "A Clearer Picture: The Contribution of Visuals and Text to Framing Effects". In: *Journal of Communication* 65.6, pp. 997–1017.

Raiford, Laigh (2007). "World Together: SNCC and Photography of the Civil Rights Movement". In: *American Quarterly* 59.4, pp. 1129–1157.

Rohlinger, Deana A and Jesse Klein (2012). "Visual Landscapes and the Abortion Issue". In: *American Behavioral Scientist* 56.2, pp. 172–188.

Rosenberg, Shawn W. et al. (1986). "The Image and the Vote: The Effect of Candidate Presentation on Voter Preference". In: *American Journal of Political Science* 30.1, pp. 108–127.

Ryan, Timothy J. (2012). "What makes us click? Demonstrating incentives for angry discourse with digital-age field experiments". In: *Journal of Politics* 74.4, pp. 1138–1152.

Stürmer, Stefan and Bernd Simon (2009). "Pathways to Collective Protest: Calculation, Identification, or Emotion? A Critical Analysis of the Role of Group-Based Anger in Social Movement Participation". In: *Journal of Social Issues* 65.4, pp. 681–705.

Tajfel, Henri (1982). "Social Psychology of Intergroup Relations". In: *Annual Review of Psychology* 33.1, pp. 1–39.

Tarrow, Sidney (2011). *Power in Movement: Social Movements and Contentious Politics*. Cambridge: Cambridge University Press.

Theocharis, Yannis et al. (2015). "Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements". In: *Information Communication and Society* 18.2, pp. 202–220.

Todorov, Alexander et al. (2005). “Inferences of Competence from Faces Predict Election Outcomes”. In: *Science* 308.5728, pp. 1623–1626.

Valentino, Nicholas A et al. (2011). “Election Night’s Alright for Fighting: The Role of Emotions in Political Participation”. In: *The Journal of Politics* 73.01, pp. 156–170.

Walgrave, Stefaan and Rens Vliegenthart (2012). “The Complex Agenda-Setting Power of Protest: Demonstrations, Media, Parliament, Government, and Legislation in Belgium, 1993-2000”. In: *Mobilization: An International Quarterly* 17.2, pp. 129–156.

Webb Williams, Nora (2015). “Observing protest: Media use and student involvement on 7 April 2010”. In: *Central Asian Survey* 34.3, pp. 373–389.

Wright, Matthew and Jack Citirn (2011). “Saved by the Stars and Stripes? Images of Protest, Salience of Threat, and Immigration Attitudes”. In: *American Politics Research* 39.2, pp. 323–343.

Zomeren, Martijn van, Russell Spears, Agneta H. Fischer, et al. (2004). “Put Your Money Where Your Mouth Is! Explaining Collective Action Tendencies Through Group-Based Anger and Group Efficacy”. In: *Journal of Personality and Social Psychology* 87.5, pp. 649–664.

Zomeren, Martijn van, Russell Spears, and Colin Wayne Leach (2008). “Exploring psychological mechanisms of collective action: Does relevance of group identity influence how people cope with collective disadvantage?” In: *British Journal of Social Psychology* 47.2, pp. 353–372.

A Appendix: Image Labeling Procedures and Summary Statistics Table

This appendix describes the deduplication procedure we used for images and discusses the manual image labeling steps in greater details. Some tweets had the same image under a different link, so before studying them we first identified which images were the same. We did that in three different steps. First we looked for which messages shared an image stored in the same URL. Second we wrote a computer program to identify which images were very similar. As a result we obtained a list of images that were the same but also a list of images that were potentially the same. In the third step two annotators manually revised the second group and indicated which were exactly the same. During this last step we found some images that were pictures of the same scene but from different angles or from slightly different times. We decided to code those as unique images.

Table 1 presents the questions used to manually label images. Figures 1 and 2 show two sample images with their labeling scores. Table 2 is a summary statistics table for the key model covariates. For our top 1,000 images, two research assistants tagged images as being of a street protest or not, and as having a collective identity symbol or not. For the remaining images, we labeled an image of being as a street protest if an annotator indicated that there were more than ten people present in the image and at least one protest sign or slogan. We labeled an image as having a collective symbol if at least one annotator believed one was present in the image.

Table 1: Labeling Form for Images

Variable	Question	Options
sign_slogan	Is there a protest sign or slogan in the picture? (e.g. Black Lives Matter; Hands Up, Don't Shoot!)	(0,1)
symbol	Is there a symbol in the picture? (e.g. flags, logos)	(0,1)
anger	How much anger does the image incite in you? If none, select 0.	(0, 1, ..., 10)
fear	How much fear does the image incite in you? If none, select 0.	(0, 1, ..., 10)
disgust	How much disgust the image incite in you? If none, select 0.	(0, 1, ..., 10)
sadness	How much sadness does the image incite in you? If none, select 0.	(0, 1, ..., 10)
enthusiasm	How much enthusiasm does the image incite in you? If none, select 0.	(0, 1, ..., 10)

Figure 1: The Most Tweeted Image During the April 14 Protest

Research staff labeled this image with average emotion scores of: anger: 2, fear: 1, disgust: 2, sadness: 3, enthusiasm: 1. They indicated that it was not a protest, and that it had no symbols.

Figure 2: The Fifth Most Tweeted Image During the April 14 Protest

Research staff labeled this image with average emotion scores of: anger: 2, fear: 1.5, disgust: 1, sadness: 1, enthusiasm: 2.5. They indicated that it was a protest, but that it had no symbols.

Table 2: Key Variable Summary Statistics

Variable	Minimum	Maximum	Mean	SD
Image	0	1	0.19	0.40
Protest*	0	1	0.03	0.18
Symbol*	0	1	0.02	0.14
Anger*	0	10	1.75	2.70
Disgust*	0	10	1.74	2.79
Enthusiasm*	0	10	1.51	2.41
Fear*	0	10	1.05	2.04
Sadness*	0	10	1.93	2.84
Number of Followers	0	5540545	4692.23	59339.65
Number of Friends	0	350644	1425.84	5198.41
Previous Tweets	0	1815	54.29	179.21

*For these variables, we provide summary statistics for the messages that have an image. The statistics for the other variables are based on the whole sample of original messages.

B Appendix: Interrater Reliability, Evidence of Stable Emotions Labeling, and Turker Demographics

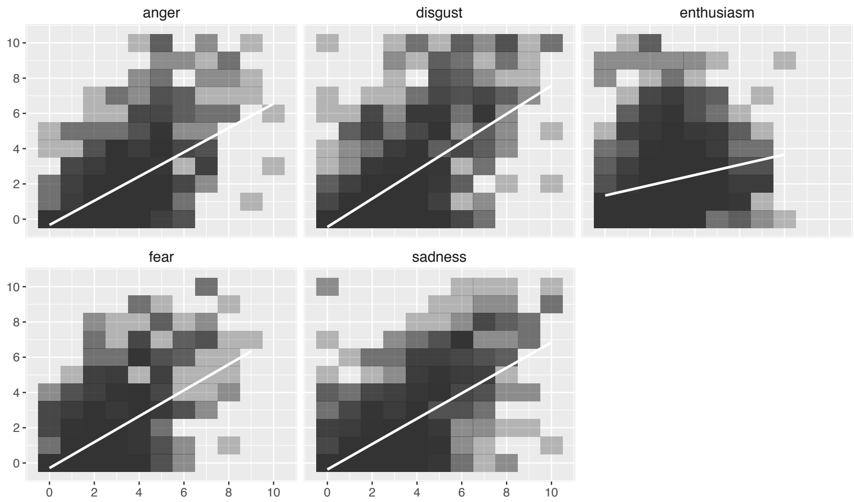
In this appendix we address interrater reliability concerns. Table 3 presents Cohen's Kappa or one-way intraclass correlation coefficients for each of the seven independent variables of interest. These values were generated based on the ratings generated by our two undergraduate coders on the top 1000 most-tweeted images. The raters had generally good agreement, with the lowest agreement for the symbol and enthusiasm labels.

Table 3: Interrater Reliability Measures

Variable	Interrater Reliability	Cohen's Kappa or one-way intraclass correlation coefficient (ICC)
Symbol	0.23	Kappa
Protest	0.78	Kappa
Anger	0.46	ICC
Fear	0.48	ICC
Disgust	0.55	ICC
Sadness	0.54	ICC
Enthusiasm	0.19	ICC

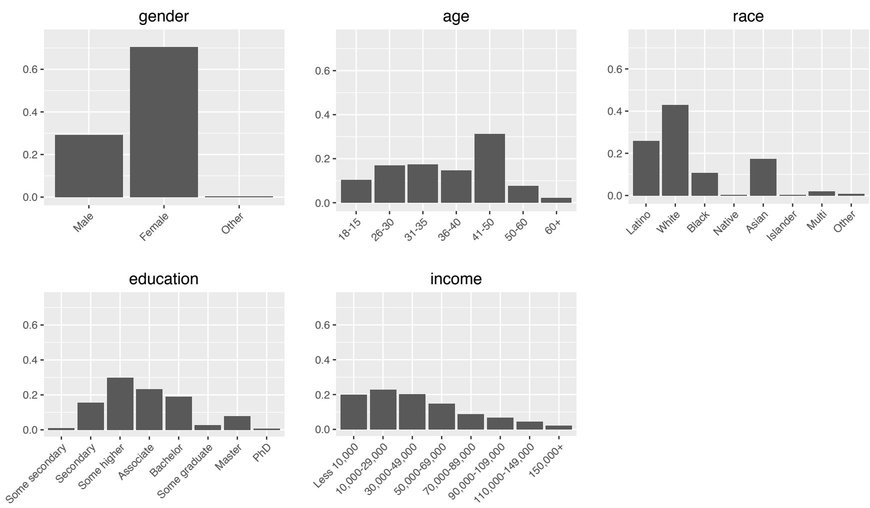
The emotions portion of labeling is particularly important for our purposes. Although emotions are subjective, and we expected a wide range of emotional responses, on average the top 1,000 images (which account for more than 50% of the messages with images) triggered very similar emotions to different people. When modeling the data we give each unique image a single score per emotion (on a 0-10 point scale). Each image has been labeled by five different people and for each image and emotion we averaged the scores given by the five individuals. In a first iteration two research assistants labeled the top 1,000 images. We had weekly meetings with them during the labeling process, they were aware of the substance and goals of the project, and they helped us improve other parts of the labeling form. Figure 3 shows the correlation between the emotional scores given by the two research assistants to the same images. The correlation is very strong in all the cases. The *enthusiasm* score shows the weakest correlation but it is still strong.

Figure 3: Correlation between the emotion scores given by 2 research assistants to the same images (top 1,000 images)



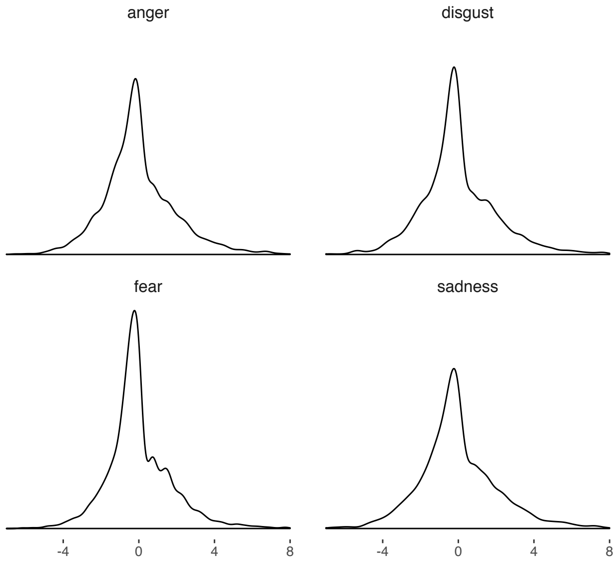
In a second iteration we used Mechanical Turk (MT) to label the top 1,000 images three more times. We decided to do so not only to have more emotion labels per image but also to get scores from people with different backgrounds, since our two research assistants were both undergraduate students, male, and white. We set it up so that only MT workers from the United States could participate and we also set it up so that workers could label more than one image but never the same image twice. Figure 4 presents summary statistics for the 1,259 MT workers that participated in the labeling process. The figure shows how workers had a very diverse background.

Figure 4: Summary of the socio demographic characteristics of Mechanical Turk workers that labeled the images



To see whether people gave very different emotion scores to the same images, for each image and emotion we calculated the average score given by the five annotators (the two research assistants and three people from MT), and then for each of the five scores we calculated the difference between them and the average score. Figure 5 shows again that the same images triggered very similar emotions in different people, with most individual scores being around 1 or 2 points from the five-scores average. *Enthusiasm* is again the emotion that presents the most variation.

Figure 5: Distribution of the difference between emotions scores for the same top 1,000 images



C Appendix: Regression Analysis Results Table

	Attention		Diffusion	
	(Basic)	(Mechanism)	(Basic)	(Mechanism)
Image	1.68982*	-	1.51099*	-
	(0.07496)	-	(0.19128)	-
Anger	-	-0.01258	-	0.03302
	-	(0.05133)	-	(0.09402)
Enthusiasm	-	0.09965*	-	0.14805*
	-	(0.03006)	-	(0.05829)
Fear	-	0.07293	-	0.13137
	-	(0.04744)	-	(0.08795)
Sadness	-	-0.00559	-	-0.196*
	-	(0.0353)	-	(0.06565)
Disgust	-	0.0308	-	0.01541
	-	(0.04753)	-	(0.08637)
Protest	-	0.58714*	-	0.59756*
	-	(0.15666)	-	(0.27813)
Symbol	-	0.2464	-	0.27148
	-	(0.21536)	-	(0.37851)
Number of Followers	0.00003*	0.00002*	0.00002*	0*
	(0)	(0)	(0)	(0)
Number Previous Tweets	-0.00136*	-0.00127*	-0.00319*	-0.00692*
	(0.00021)	(0.00046)	(0.00112)	(0.00229)
Number of Friends	0.00004*	0.00004*	0.0001*	0.00003
	(0.00001)	(0.00002)	(0.00002)	(0.00006)
Time (t2)	-0.37681*	-0.88402*	-0.90594*	-1.4849*
	(0.12506)	(0.29657)	(0.42199)	(0.71093)
Time (t3)	-0.4106*	-0.73863*	1.2054*	-0.75196
	(0.14518)	(0.3462)	(0.47383)	(0.86792)
Time (t4)	-0.29262*	-0.59474*	-0.3746	-0.59322
	(0.10221)	(0.24731)	(0.39159)	(0.64909)
Time (t5)	-0.27161*	-0.49169*	0.30034	0.33694
	(0.09395)	(0.21096)	(0.28022)	(0.41588)
Time (t6)	0.10589	-0.24391	1.1195*	0.61577
	(0.08981)	(0.20221)	(0.23906)	(0.37724)
Constant	-0.31865*	1.2157*	-1.5364*	-0.1109
	(0.07033)	(0.17538)	(0.20791)	(0.37247)
Original Tweets (n)	49,345	8,706	7,502	2,078
AIC	113,204	36,478	11,462	5,152.6

Note: $^*p < 0.5$

D Appendix: Robustness Check, controlling for the text of the messages

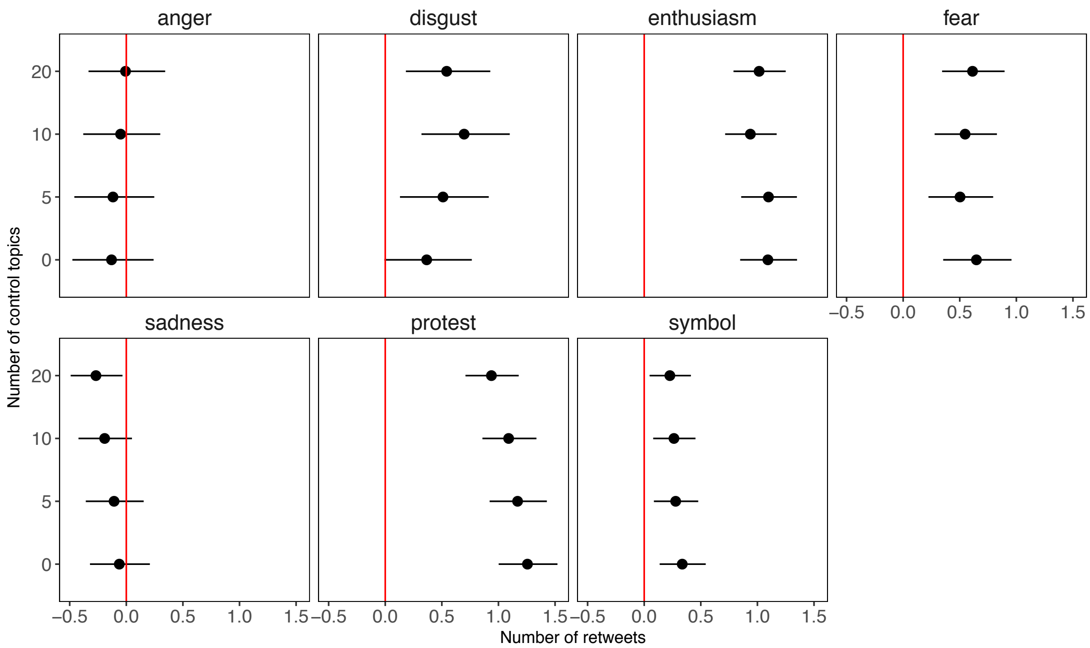
Are the results reported in the paper a mere function of the tweets' text instead of image effects? We do not believe this is the case for two main reasons.

First, the protest was about a very specific issue, police brutality against African American citizens in the United States, and so we expect most messages to be related to a narrow set of claims and their text to be very similar; and second, after exploring a large number of messages, we observed that most texts were very short (e.g. they only contained a hashtag such as #blacklivesmatter), corroborating our low textual variation expectation.

Nevertheless, to rule out this potential concern, in this Appendix we add different textual controls to Model the mechanism attention model (Model 3 in the paper) as follows.[1] First we pre-process the text of the messages by removing urls, mentions to users, punctuation, numbers, stopwords, and by stemming all remaining words. Then we fit 3 Latent Dirichlet Allocation models with a varying number of topics: $k = \{5, 10, 20\}$. Our goal is not to perfectly capture specific topics and/or frames of interest, but to group messages that contain very similar words. This is why instead of choosing a specific number of topics, we check whether our findings hold when controlling for the semantic content of the message, and for different number of topics.

To do so, after estimating each topic model, we re-run Model 3 by including message-level covariates indicating the probability of a message to belong to each of the $k-1$ topics. For example, the first time we include 4 new message-level variables to the model, the probability that a given messages belongs to topic 1, 2, 3, and 4 of a 5-topic model (we exclude one topic probability to avoid perfect colinearity issues). The second time we include 14 variables, and so on.

Figure 6: Key coefficients of interest when controlling for the textual content of the message, and for a different number of potential textual topics



[1] We manually checked all messages used to estimate Model 4 and the textual variation was extremely low. For this reason we do not replicate it using these textual controls. A very large percentage for example only had the hashtag #shutdownA14.

Figure 6 shows the key coefficients of interest across these different versions of Model 3. For each variable we can see the original coefficient at the bottom (number of topics = 0) and the coefficient when controlling for 20 topics at the top. We see that the findings are robust to all textual controls, with the single exception of the sadness coefficient, which goes from not having a statistically significant effect to have a significant negative effect on attention when controlling for 20 topics. If this were to be the true sadness effect, it would actually corroborate our hypothesis H_4 about the demobilizing effect of images evoking sadnes. In the paper we can only corroborate this hypothesis as it relates to the diffusion model (Model 4). The coefficient for disgust, one of the alternative mechanisms for which we had no clear expectation, seems to have a positive effect on online mobilization in models controlling for 5 to 20 topics. However, this new effect does not invalidate our findings for the anger, enthusiasm and fear mechanisms, which are the main focus of the paper.

E Appendix: Image Features and Emotions

In this Appendix we explore the presence and mobilizing power of image features beyond those described the main text. We have a two-fold objective here: to have a better understanding of what particular elements in an image are associated with each emotion under study, and to check the extent to which other image features are responsible for the variation currently explained by our key mobilizing mechanisms. The first objective is essentially descriptive, presenting an initial take on how the content of an image might evoke a specific emotion. The second serves as a robustness check for the presented models, showing the stability of our mechanism coefficients when controlling for specific image features.

As discussed in the main manuscript, we rated how much anger, fear, disgust, sadness, and enthusiasm each image in our dataset evoked (see Appendix B for a description of the rating procedure) and found, after controlling for other alternative explanations, that images triggering fear and enthusiasm were associated with higher levels of online protest attention and diffusion, and that images evoking sadness were correlated with lower diffusion levels. However, due to space limitations we were not able to provide detailed information of what images triggering each emotion look like. We address this shortcoming here by exploring the top 1,000 images in our dataset in more depth (these images account for about 50% of all messages with images).[2]

As a first step, we looked through these top images to see if we detected any particularly common or potentially salient features. Nine particular features stood out. A substantive number of images had or were: (1) text, (2) a head-shot of someone killed by police, (3) police forces, (4) police brutality scenes, (5) the singer John Legend, (6) injured people, (7) die-ins, (8) cartoon, and (9) protest advertisement. We then asked two research assistants to go through these top images again and label them for the presence of these 9 features. Figure 7 shows the proportion of the top 1,000 images in which these image features are present. For example, about 25% of the images had some text in it, and between 20 and 25% had a headshot of someone killed by police.

One might expect that some of these new image features could be correlated with the emotions the images evoke. We illustrate such correlations in Figure 8. To highlight some examples, images evoking anger and disgust tend to have scenes of police brutality and injured people, those triggering enthusiasm are more likely to have protest advertisement, and fearful and sad images are likely to show police forces, police brutality, and injured people. Moreover, images with text and cartoons appear to be more likely to trigger stronger emotions in general. They seem to be precise and "informative" and to generate more clear and stronger reactions. In Figure 10 you can see the top 5 images (among the top 1,000) with the highest fear, sadness, enthusiasm, anger, and disgust score; and you can see how these different image features are present in them.

We believe that these new image features impact protest attention and diffusion mainly through the theoretical mechanisms presented in the paper and that they do not challenge the core findings of the study. That is, the presence of police violence affect mobilization via the emotions-evoking channel, not a separate causal pathway. To check this assumption we proceeded as follows.

[2] To determine the most retweeted images, we first removed duplicate images from the dataset and then we mapped each unique image to all tweets and retweets. We study here the 1,000 images that were most often present in the messages in our dataset.

Figure 7: In how many top images were these new image features present?

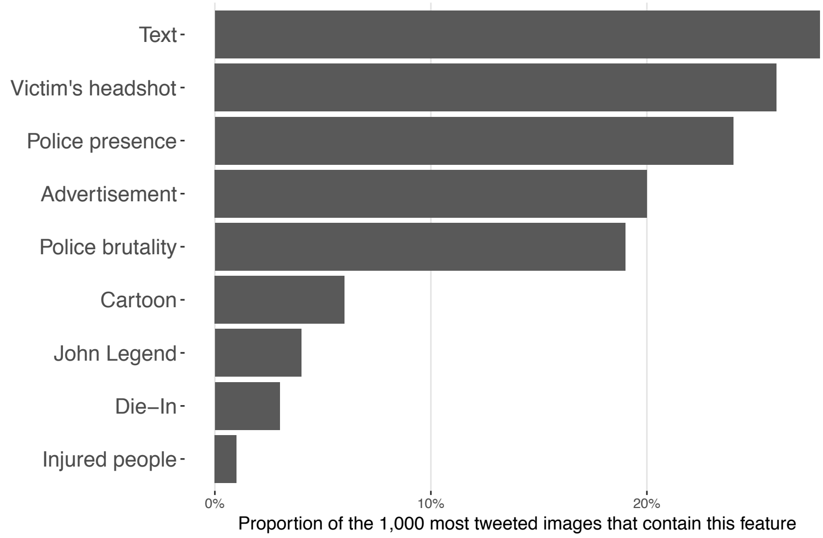
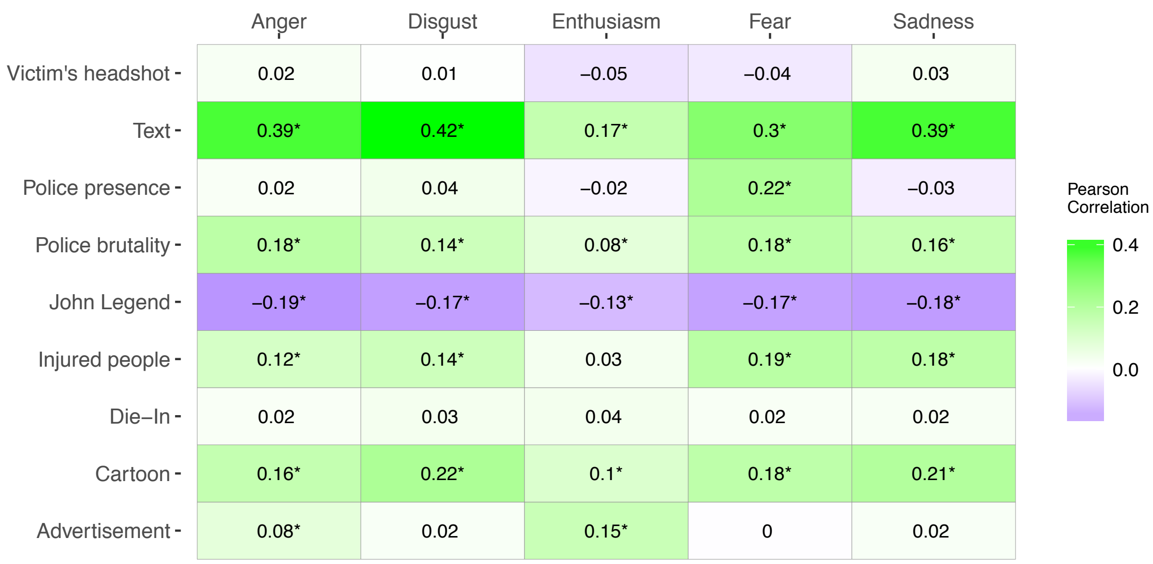


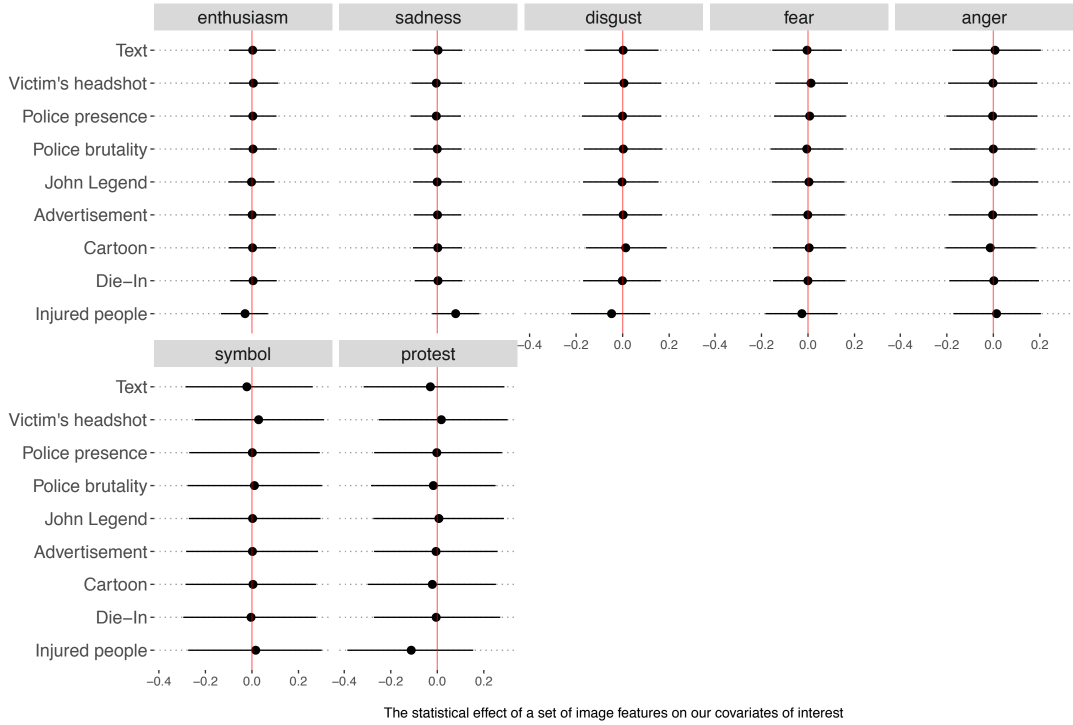
Figure 8: Pearson correlation between a set of image features and the emotions image evoke



First, we selected from our dataset all the original messages containing one of the top 1,000 images (for which we have labels for these extra image features): a total of 1,312 messages (out of the 49,345 original messages in the full sample used in the manuscript). Then we fitted to these messages the same negative binomial model we fit in the paper when estimating the effect of our key image mechanisms on protest attention (Model 3). This is a biased sample and so the estimated effects are not of particular interest. However, we can still use this sample to see how adding new features to the model affect the estimated effects for the mobilizing mechanisms originally theorized. To do so we fitted the same model 9 more times but each time we added one of the new image features we explore in this Appendix. We then checked if adding these extra features significantly affected the explanatory power of our key covariates.

The results in Figure 9 show that these image features either affect the

Figure 9: Do these new image features take away the explanatory power of the theorized mobilizing mechanisms?



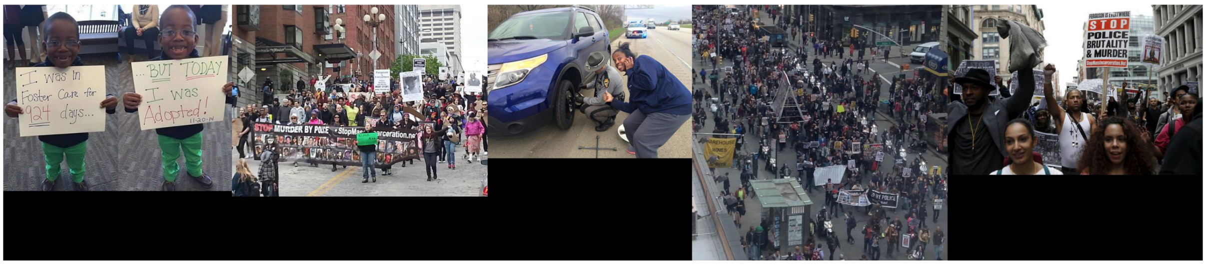
outcome through our mechanisms of interest, or that they have an independent effect on attention that does not invalidate the mobilizing power of the mechanisms we initially theorized. Each label on the left is a new feature we added into the model, and each label in the facet titles indicate one of our theorized mechanisms. Each dot represents how adding that new feature into the model changed the coefficient for that specific mechanism (with a 95% bootstrapped confidence interval). We see that each of these new features have a null effect on the estimated mobilizing effect of the key mechanisms (though some effects have wide confidence bands). Only the injured people feature seems to slightly take away the negative effect of sadness on protest attention, though the confidence interval still crosses zero.

Figure 10: Top 5 images (among the top 1,000) with the highest fear, sadness, enthusiasm, anger, and disgust scores

Fear

Sadness

Enthusiasm



Anger

Disgust