Motivation Science

Digital Dissent: An Analysis of the Motivational Contents of Tweets From an Occupy Wall Street Demonstration

Melanie Langer, John T. Jost, Richard Bonneau, Megan MacDuffee Metzger, Sharareh Noorbaloochi, and Duncan Penfold-Brown

Online First Publication, February 1, 2018. http://dx.doi.org/10.1037/mot0000084

Motivation Science

2333-8113/18/$12.00 http://dx.doi.org/10.1037/mot0000084

Digital Dissent: An Analysis of the Motivational Contents of Tweets From an Occupy Wall Street Demonstration

Melanie Langer, John T. Jost, Richard Bonneau, Megan MacDuffee Metzger, Sharareh Noorbaloochi, and Duncan Penfold-Brown
New York University

Social scientific models of protest activity emphasize instrumental motives associated with rational self-interest and beliefs about group efficacy and symbolic motives associated with social identification and anger at perceived injustice. Ideological processes are typically neglected, despite the fact that protest movements occur in a sociopolitical context in which some people are motivated to maintain the status quo, whereas others are motivated to challenge it. To investigate the role of ideology and other social psychological processes in protest participation, we used manual and machine-learning methods to analyze the contents of 23,810 tweets sent on the day of the May Day 2012 Occupy Wall Street demonstration along with an additional 664,937 tweets (sent by 8,244 unique users) during the 2-week lead-up to the demonstration. Results revealed that social identification and liberal ideology were significant independent predictors of protest participation. The effect of social identification was mediated by the expression of collective efficacy, justice concerns, ideological themes, and positive emotion. The effect of liberalism was mediated by the expression of ideological themes, but conservatives were more likely to express ideological backlash against Occupy Wall Street than liberals were to express ideological support for the movement or demonstration. The expression of self-interest and anger was either negatively related or unrelated to protest participation. This work illustrates the promise (and challenge) of using automated methods to analyze new, ecologically valid data sources for studying protest activity and its motivational underpinnings—thereby informing strategic campaigns that employ collective action tactics.

Keywords: collective action, protest, social identification, political ideology, justice concerns

> "Haven't you heard, it's a battle of words?"
> The poster bearer cried.
> —Pink Floyd, "Us and Them"

Sociologists and political scientists recognize that the decision made by citizens about whether to engage in political protest is a complicated one. One major focus of theoretical and empirical attention pertains to the utility of the decision to the individual, as emphasized in rational choice perspectives (Downs, 1957; Finkel, Muller, & Opp, 1989; Kuran, 1991; Marwell & Oliver, 1993; Oberschall, 1973; Olson,

Melanie Langer, Department of Psychology, New York University; John T. Jost, Department of Psychology, Center for Data Science, and Department of Politics, New York University; Richard Bonneau, Center for Data Science, Departments of Computer Science and Biology, New York University; Megan MacDuffee Metzger, Department of Politics, New York University; Sharareh Noorbaloochi, Department of Psychology, New York University; Duncan Penfold-Brown, Departments of Computer Science and Biology, New York University.

This work was supported by the INSPIRE program of the National Science Foundation (Awards SES-1248077 and SES-1248077-001) as well as New York University's Research Investment Fund (RIF) and Global Institute for Advanced Study (GIAS). Data were collected by the NYU Social Media and Political Participation Laboratory (https://wp.nyu.edu/smapp/), where John T. Jost and Richard Bonneau are Principal Investigators along with Jonathan Nagler and Joshua Tucker. Parts of this research program were presented at annual meetings of the International Society of Political Psychology (ISPP) in San Diego, CA, and Warsaw, Poland. We are grateful for the feedback we received there as well as incisive comments by Vivienne Badaan, Sarah DiMuccio, Shahrzad Goudarzi, Mark Hoffarth, Anni Sternisko, Maarten van Bezouw, Gregory Webster, and Rex Wright on a draft of this article.

Correspondence concerning this article should be addressed to John T. Jost, New York University, 6 Washington Place, New York, NY 10003. E-mail: john.jost@nyu.edu

1965; Riker & Ordeshook, 1968; Tilly, 1978; Useem, 1998). The decision to protest, in this view, is the result of a subjective calculation involving the value of potential gains if the protest succeeds, the costs of participation, the probabilities that these costs and benefits will be realized, and a judgment of how one's own participation will affect these probabilities. The individual's sense of efficacy (Bandura, 1997)—especially collective efficacy (Tausch et al., 2011), or the expectation that the group can succeed—has also been identified as a key factor in such calculations.

As noted by Muller and Opp (1986), the costs and benefits of participation may enter differently into such calculations. Whereas the benefits of a successful protest are reaped by all, those who participate generally incur extra costs—such as missing work, risking arrest, exposing oneself to violence, and so on—relative to those who do not. Moreover, as with the so-called "voter's paradox" (Aldrich, 1993), a single individual's participation is unlikely to be the deciding factor in the success of a protest. As a result, even if a successful protest would have enormous value for an individual, a "free rider" might reap the benefits of protest while avoiding the costs of participation (and without jeopardizing the protest's chance of success), thus making a rational decision to stay home.

And yet people throughout history have sometimes chosen to incur the costs of participation in protest (Gurr, 1970; Klandermans & van Stekelenburg, 2013; Tilly, 1978). Social psychological theories have therefore moved beyond models that focus exclusively on issues of rational choice to consider additional motives for protest participation. The perception of injustice is considered to be a necessary prerequisite, and so is anger—or moral outrage—in response to injustice (Barbalet, 1998; Goodwin & Jasper, 2006; Jost et al., 2012; Kawakami & Dion, 1995; Stürmer & Simon, 2009; van Zomeren, Postmes, & Spears, 2008; van Zomeren, Spears, Fischer, & Leach, 2004; Wakslak, Jost, Tyler, & Chen, 2007). Indeed, feelings of moral outrage are posited to motivate collective action independent of any rational calculations about the likelihood of securing particular gains.

Social identification with fellow protesters is also assumed to play a key role, insofar as the individual comes to understand unjust circumstances as affecting not only him or her, but also the social group as a whole (Drury & Reicher, 2009; Jost et al., 2012; Kelly & Breinlinger, 1996; Klandermans, 1997; McGarty, Thomas, Lala, Smith, & Bliuc, 2014; Subasic, Reynolds, & Turner, 2008). Viewing oneself as a member of a social group allows for assessments of collective efficacy, or the sense that the group is capable of achieving its goals. Just as beliefs about self-efficacy are instrumental in shaping individual behavior, beliefs about collective efficacy affect the individual's decision to engage in collective action (Bandura, 1997; Tausch et al., 2011; Van Zomeren, Leach, & Spears, 2012).

Van Zomeren et al. (2012) highlighted the fact that there are rational and emotional pathways to protest and refer to potential protesters as "passionate economists." Following Lazarus (1991, 2001), they construed protest as an approach-related coping behavior rooted in processes of social identification and relative deprivation. They emphasized the role of anger, the importance of identifying disadvantages as unfair, and placing blame on an external agent. Van Zomeren and colleagues also posit that a collective response to disadvantage produces an assessment of coping potential, or group efficacy. They note that motivational factors are mutually reinforcing, so that protest participation strengthens group identification, and stronger identification reinforces confidence in the group's efficacy and the likelihood of protesting further. These assumptions are built into the Social Identity Model of Collective Action (SIMCA), which was proposed by van Zomeren et al. (2008) to specify the ways in which the effects of social identification are mediated by anger at perceived injustice and beliefs about collective efficacy (see Figure 1).

In social psychology, SIMCA has become the most influential model of collective action. Although it is a very useful model, Jost, Becker, Osborne, and Badaan (2017) argued that SIMCA neglects overtly ideological factors.[1] This neglect is probably due to the fact that, in

[1] Klandermans and van Stekelenburg (2013) cited "ideological motivation" as a predictor of collective action (p. 786), but they did not elaborate on the role of ideology, nor did they distinguish between system-justifying and system-challenging ideologies or consider the possibility that social movements on the left and right may be inspired by qualitatively different goals.

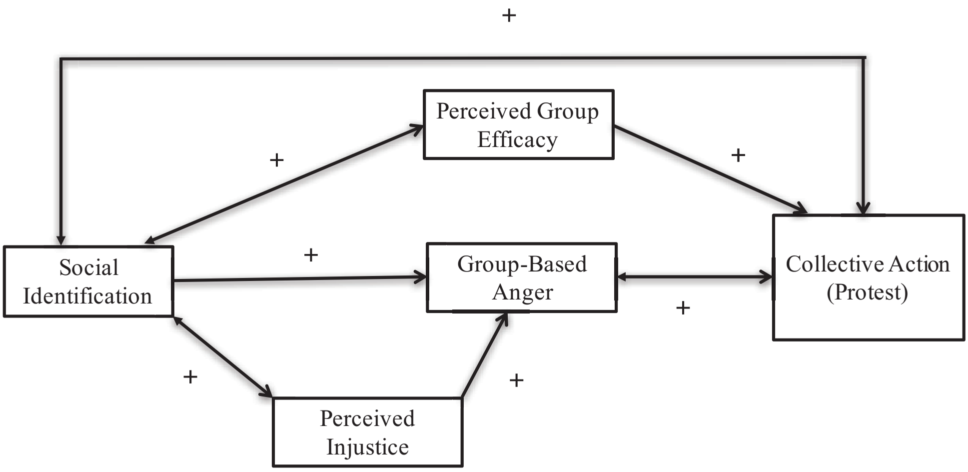

Figure 1. Schematic illustration of the Social Identity Model of Collective Action (SIMCA; see Van Zomeren et al., 2008, p. 521) The variable of group-based anger was added by Jost, Becker, Osborne, and Badaan (2017).

accordance with social identity theory (Tajfel & Turner, 1979), it conceptualizes protest movements purely in terms of ingroup/outgroup dynamics. The intergroup level of analysis is extremely important, but it does not enable us to tell the "whole story," so to speak, of social protest. As McGarty et al. (2014) pointed out, "the debate about abortion is a genuine intergroup conflict, but it is not a conflict between men and women, or between Christians and non-Christians, but between groups based around contrasting pro-Choice and pro-Life positions" (p. 729). It is, in other words, an irreducibly *ideological* conflict (see also Jost et al., 2017).

From our perspective, existing social psychological models—including SIMCA—fail to appreciate the fact that the decision to participate (or withhold participation) in protest is not just like every other instance of group behavior. Rather, it is an inherently ideological decision—not only because it involves conflicting beliefs, values, and opinions (as well as identities), but also because it occurs in a societal context in which some people are motivated to defend and bolster the existing regime, whereas others are motivated to challenge and oppose it. Often—but not always—these groups of political actors turn out to be "conservatives" (or, since at least the time of the French Revolution, "rightists") and "progressives" (or "leftists"), respectively (Jost, 2006; Jost, Nosek, & Gosling, 2008).

When the problem is posed in these terms, one sees that a complete social psychological account of collective action must incorporate not only interpersonal and intergroup processes associated with relative deprivation and social identification but also social structural and ideological processes such as those associated with the phenomenon of system justification, defined as the motivational tendency to defend, bolster, and justify aspects of the societal status quo (Jost, Banaji, & Nosek, 2004). Among other things, system justification theory can help to specify when individuals and groups will—and, just as importantly, will *not*—experience moral outrage (Wakslak et al., 2007) and whether moral outrage is directed at defenders or challengers of the status quo (Rudman, Moss-Racusin, Glick, & Phelan, 2012).

Several studies demonstrate that system justification motivation undermines support for progressive forms of protest, such as demonstrations associated with the feminist and Occupy Wall Street movements (Becker & Wright, 2011; Jost et al., 2012; Osborne & Sibley, 2013), while enhancing support for conservative protests such as those organized by the Tea Party movement (Hennes, Nam, Stern, & Jost, 2012). System justification motivation also inspires opposition to or backlash against progressive activists and others who are seen as challenging the societal status quo (Diekman & Goodfriend, 2007; O'Brien & Crandall, 2005; Yeung, Kay, & Peach, 2014). An analysis of ideological conflict in terms of the opposition between system-justifying goals to legitimize the status quo and system-challenging goals to delegitimize the status quo leads to the conclusion that social movements of the right and left may be inspired by qualitatively different motivational concerns (Hennes et al., 2012; Jost, 2006).

To incorporate these ideological factors, which are missing from SIMCA, Jost et al. (2017) proposed an integrative model in which group identification and ideological motives both contribute independently to two very different types of protest activity, namely system-justifying and system-challenging types of protest. As in SIMCA, mediating variables include beliefs about group efficacy and anger at perceived injustice. From a system justification perspective, however, it is important to recognize that emotions may be elicited by and directed not only at individuals and groups, but also overarching social systems (Solak, Jost, Sümer, & Clore, 2012). Whereas the effects of group efficacy and group-based anger on protest are likely to depend upon the relative social status of the social group, system-based anger is likely to be positively associated with system-challenging protest but negatively associated with system-justifying protest. A somewhat simplified version of Jost et al.'s (2017) model is illustrated in Figure 2.

In the present study, we sought to explore parts of this integrative theoretical model, which incorporates elements of rational, emotional, and ideological motivations to understand participation in political protest. Specifically, we used a large dataset gleaned from user-generated social media content to analyze sentiments expressed on Twitter concerning an Occupy Wall Street demonstration that took place in New York City on May Day 2012. The dataset includes Twitter messages sent by those

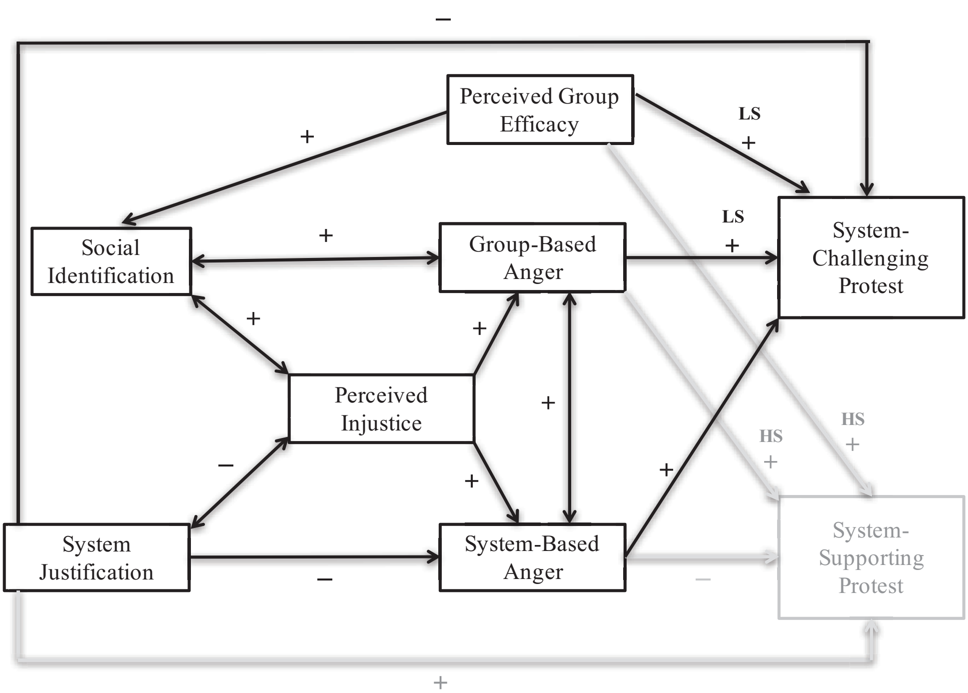

Figure 2. Illustration of Jost et al.'s (2017) integrative model suggesting that group identification and ideology contribute independently (and differentially) to system-supporting and system-challenging types of protest activity. Pathways in gray were not investigated in the present study. HS = High Status Group; LS = Low Supporting Group.

who participated in the demonstration as well as those who supported it but did not participate and those who objected to it.

To capture rational motivations, we coded tweets for content related to self-interest and collective efficacy. To measure emotion, we coded expressions of positive emotion in addition to anger, because several theorists have proposed that positive emotions (such as pride and hope) may also play an important role in inspiring collective action (e.g., Bar-Tal, Halperin, & de Rivera, 2007; Goodwin & Jasper, 2006). Unfortunately, the methods we employed did not facilitate the drawing of clear distinctions between the expression of group-based and system-based emotions, so there is some ambiguity when it comes to interpreting the effects of emotional variables.

To gauge ideological motivations, we used follower-based network methods to estimate the ideological positions of individual users and also coded messages for justice-related themes and evaluations of the system. Because the Occupy Wall Street movement involved system-challenging forms of protest, we were unable to investigate all aspects of the theoretical model specified in Figure 2, which also takes into account system-supporting forms of protest. To some extent, however, we were able to observe instances of system justification in the form of backlash against the protestors.

The model that we tested is illustrated schematically in Figure 3. We hypothesized that—in addition to social identification with the Occupy Wall Street demonstration—liberal (or leftist) ideology would predict participation in the Occupy Wall Street demonstration held on May Day 2012. We also hypothesized that the social and psychological variables highlighted by existing models of collective action—such as self-interest, perceptions of collective efficacy, anger, and positive emotion—would mediate the effects of social identification and liberalism on participation in protest. In an effort to incorporate ideological processes that had been neglected in previous studies of collective action, we also hypothesized that the expression of justice concerns and ideological themes (such as criticism or affirmation of the existing social system) would mediate the effects of group identification and liberalism on protest participation. Thus, the present study complements existing research on the role of social media in promoting participation in the Occupy Wall

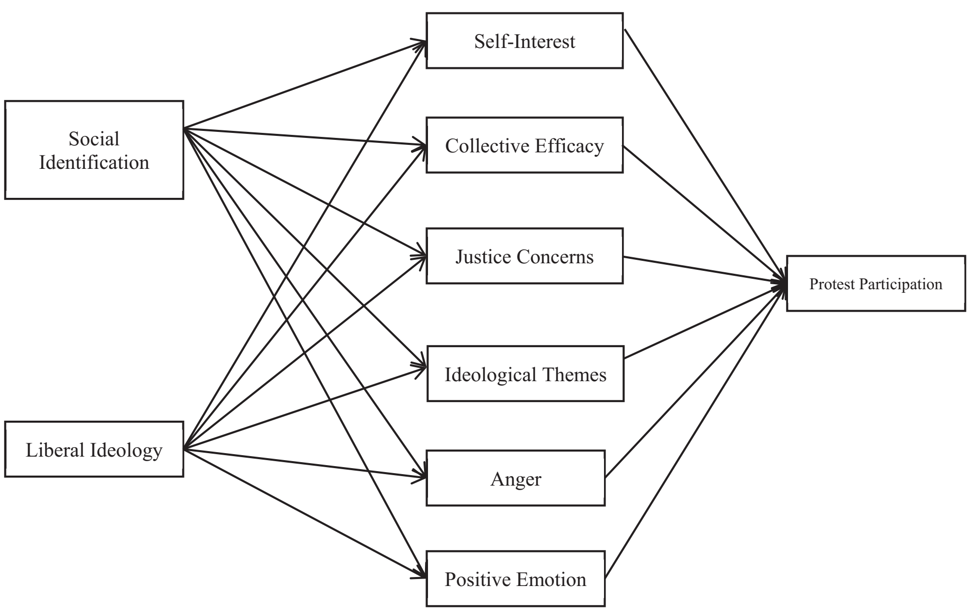

Figure 3. Schematic illustration of a path model in which the effects of social identification and liberal ideology on protest participation are mediated by the expression of self-interest, collective efficacy, justice concerns, ideological themes, anger, and positive emotion.

Street movement—including especially valuable studies carried out by Smith, Gavin, and Sharp (2015) and Theocharis, Lowe, van Deth, and García-Albacete (2015)—but it also breaks new ground by incorporating ideological and emotional processes in addition to processes of identity formation and the spread of logistical information.

Method

Data Collection

We collected 23,810 tweets from 12:19 PM May 1 to 12:00 AM May 2, 2012 EDT by passing a predetermined set of keywords through Twitter's public Streaming API via the "track" parameter (https://dev.twitter.com/streaming/overview). We selected the keywords "ows," "union sq," "union square," and "mayday" after reviewing Twitter early in the day to identify common references to the Occupy Wall Street May Day (mayday) protest in New York. The data set contains formal as well as copied and pasted retweets. We stored all tweets in a MongoDB database.

To further investigate the role of psychological factors leading up to the protest, we later back-fetched an additional 664,937 tweets from users included in the original data set. We gathered tweets two weeks prior to the protest (from May 17 to April 30, 2012) using Twitter's User Timeline endpoint (https://dev.twitter.com/rest/reference/get/statuses/user_timeline). It is possible that some users deleted tweets after sending (and before we collected) them, in which case they would be excluded from our data set. We did not remove abbreviations, hyperlinks, or hashtags from any of the tweets prior to coding them.

Tweet Coding

Students and research assistants from the psychology and politics departments at New York University manually coded a subset of tweets. Because of resource constraints, manual coding was carried out in two rounds. For the first round, we selected a random subset of 7,705 tweets and instructed judges to respond *yes* (scored as 1), *no* (0), or *don't know* (missing value) to the eight questions listed in Table 1. Every tweet was coded by 2 or 3 judges. The analysis was conducted at the user level of analysis (rather than the tweet level) to transfer coding information from multiple tweets to a single actor.

Table 1
Coding Questions for Variable Creation

Coding variable	Question
Protest participation	Does this tweet indicate that the author participated, is participating, or will participate in the Occupy Wall Street protests?
Non-participation (used for data cleaning only)	Does this tweet explicitly indicate that the tweeter DID NOT or WILL NOT participate in the Occupy Wall Street protests?
Social identification	Does this tweet evoke social identification with Occupy Wall Street, or a feeling of membership in the group organizing or participating in the protest?
Self-interest	Does this tweet appeal to individual or collective self-interest in any way (e.g., costs or benefits to the group or individual of political participation)?
Collective efficacy	Does this tweet appeal to a shared or collective sense of efficacy—that it is possible to make a difference (e.g., "we can do it")?
Justice concerns	Does this tweet mention concerns about fairness, morality, social justice, poverty, deprivation, or exploitation?
Ideological themes	Does this tweet draw explicitly on ideological themes (i.e., such as liberal or conservative opinions), or on criticism or affirmation of the U.S. (or global) political or economic systems, etc.?
Emotion (used in conjunction with LIWC, see anger and positive emotion below)	Does this tweet contain any form of positive or negative emotional expression (e.g., sadness, anger, fear, disgust, joy, enthusiasm, contentment, etc.)?

Note. LIWC = Linguistic Inquiry and Word Count.

For the dependent variable of protest participation, judges sought to determine whether any of the tweets indicated that the social media user participated, was participating, or would participate in the Occupy Wall Street protest on May 1. If a single tweet indicated that the user participated (or intended to participate) in the protest, we applied this information to all of his or her tweets and therefore assigned a value of 1 for the variable of participation in such cases. Judges also indicated whether any of the tweets explicitly indicated that the user did not or would not participate in the protest, but responses to this item were only used to identify a possible conflict with the preceding judgment.

The remaining six questions listed in Table 1 were used to estimate the independent variables, namely social and psychological antecedents of protest participation. Specifically, judges were asked whether each tweet included mentions of group identification, self-interest, collective efficacy, justice concerns, ideological themes, and emotional expression. For these variables, we computed the proportion of each user's tweets that were judged to contain specific content (see Figure 4 for distribution and Table 2 for frequency of number of tweets sent by each user). For instance, if a user tweeted twice and one of those tweets received a *yes* coding for "group identification" and the other received a *no* coding, s/he would receive a score of .5 for the group identification variable. If a tweet received a *don't know* response from one judge and a nonmissing value (*yes* or *no*) from the other coder, we used only the nonmissing value.

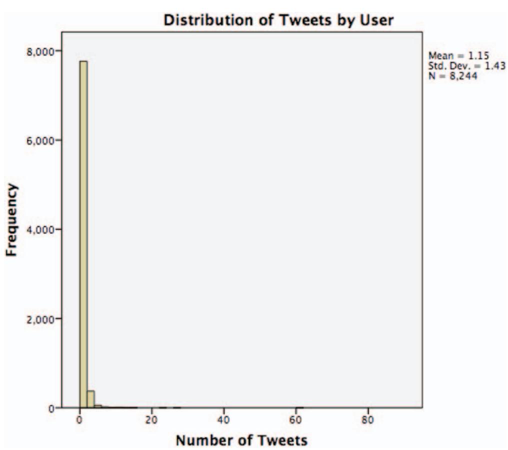

Figure 4. Distribution of number of tweets sent by user (user "tweet counts") for hand-coded dataset. $N = 8{,}244$. See the online article for the color version of this figure.

Table 2
Frequency of Number of Tweets Sent by User (User "Tweet Counts") for Hand-Coded Dataset

Tweet count	Frequency	Percent
1	7,764	94.2
2	293	3.6
3	80	1.0
4	36	.4
5	20	.2
6	12	.1
7	7	.1
8	6	.1
9	2	.0
10	8	.1
11	4	.0
12	1	.0
13	2	.0
14	4	.0
15	1	.0
22	1	.0
26	1	.0
60	1	.0
90	1	.0

Note. $N = 8{,}244$.

In the second round of manual coding, a new group of volunteers coded the same 7,705 tweets coded in round one so that we could obtain adequate intercoder reliability estimates. Once again, every tweet was coded by 2 or 3 judges. We also asked judges to code an additional subset of 7,000 unique tweets—2,500 of which were selected randomly. The remaining 4,500 were selected in accordance with active learning techniques (Settles, 2012). Based on the first round of coding, we observed that positive classes for the variables were much rarer than negative, leaving us with an imbalanced data set, which poses problems for machine learning methods. For example, tweets indicating participation in the OWS protest comprised only 3% of all tweets coded in round one, whereas the negative class, nonparticipation, accounted for 97%. To achieve a better class balance, we trained models predicting the outcome of all nine coding variables based on manually coded tweets from the first round and selected 500 previously uncoded tweets predicted to be from the minority class. This iterative, active learning approach provided a more balanced set of manually coded tweets, which improved the machine-learning technique used subsequently.

Specifically, we trained algorithmic models based on the manually coded data from the two rounds to analyze the contents of all of the remaining tweets in the sample. Throughout the article, we refer to tweets coded by humans as "manually coded," and the tweets coded automatically as "machine-coded." Thus, we were able to estimate the effects of each psychological variable on protest participation based upon a very large set of tweets, with only a small subset coded manually.

Data Cleaning and Inter-Coder Reliability

To ensure that the manually coded data set was clean, we employed several heuristics to remove poor coders and poor codings. Of the 21,284 initial tweet codings, we removed 4,263 tweets that were considered spam and 380 tweets that received *yes* codings for two logically incompatible categories: "Does this tweet indicate that the author participated, is participating, or will participate in the Occupy Wall Street protests?" and "Does this tweet explicitly indicate that the tweeter DID NOT or WILL NOT participate in the Occupy Wall Street protests?"

We defined intercoder agreement in terms of the probability that two or more coders provided the same response to a given question. According to this metric, we achieved an average intercoder reliability of 80.1% after removing a single poor judge who exhibited an agreement rate of less than 32% with the other judges. After implementing these data cleaning methods, we were left with a final data set of 9,452 manually coded tweets (see Table 3 for response frequencies at the tweet level as well as sample tweets) from 8,244 unique Twitter users (see Table 4 for response frequencies at the user level of analysis).

Variable Creation With Machine Coding of Data

As noted above, we used a machine-learning algorithm trained on the manually coded dataset to code the 14,394 tweets from the initial data-set that had not been manually coded. For the machine-coded portion of the dataset, we first aggregated the tweet text for each user and used models to code the aggregated text. Whereas individual tweets are less confusing and easier to interpret for human coders, user-aggregated tweet texts provide more efficient and accurate estimates using machine-coding techniques, because problems associated with sparse text data are attenuated. Using this method of aggregation, the set of 14,394 tweets was reduced to 12,450 machine-coded user texts.

We trained logistic regression models to predict whether the aggregated texts contained each type of psychological content described in Tables 1 and 3. Models were trained on "bag-of-words" representations of manually coded tweets, and were optimized for AUROC (area under receiver operating curve) using either L1 or L2 shrinkage parameters, and a grid search over possible model parameter values to find values that would give best AUROC over three-fold cross-validation. Optimizing for AUROC provided an intuitive understanding of performance and true-positive versus false-positive prediction rates, by allowing us to maximize the probability of predicting true positives as positive while minimizing false positive rates. The predicted values for a given user's level of social identification, self-interest, collective efficacy, ideological themes, and justice concerns is the predicted positive probability—that is, the probability that the user's text contains the specific type of content (e.g., an expression of social identification). These results are summarized in Table 5.

Anger and Positive Emotion

To obtain estimates for emotional variables, we processed the text from manual and machine-coded datasets using Linguistic Inquiry and Word Count software (LIWC; see Pennebaker, Booth, & Francis, 2007). We calculated the proportion of a user's tweets containing words that, according to LIWC's dictionaries, expressed anger or positive emotion, and that had also received a *yes* coding for "emotion" (defined as .5 or higher predicted positive probability for "emotion"). The manual "emotion" codings enabled us to verify the dictionary-based codings.

Table 3
Response Frequencies for Manually Coded Tweets (at Tweet Level) With Sample Tweets

Variable	Coding	Frequency	Sample tweets
Protest participation	Yes	8.90%	1. Our May Day march was a huge success! We are now in Union Sq #ows @99PicketLines 2. In an #occupywallstreet protest down Broadway Ave in NYC! http://t.co/akGu7f4d 3. I am standing under a giant blue tarp in the largest march I've ever seen. #OWS #m1gs #m1nyc
	No	91.10%	1. In defiance of #OWS May Day request, I went to work, went shopping and sat in on a Pace Psych class just for the hell of it. @OccupyWallSt 2. Sorry I just don't get this ows stuff. You get what you vote for. You want changes then vote everyone out and start fresh 3. If your still apart of Occupy wall st. You either have no life, no ability to think for yourself, a wannabe or attention seeker or an idiot
Social identification	Yes	11.88%	1. Good luck & a safe night to all my brothers & sisters at #OccupyWallStreet tonight. We are all the 99% 2. WHOSE STREETS??? OUR STREETS #OCCUPIED #OWS 3. @OccupyWallSt: We are unstoppable another world is possible #mayday #m1gs #occupywallstreet"
	No	85.47%	1. #OWS you don't represent me nor the rest of this city. GTFO and let us go about our lives 2. I don't protest. I vote. #occupy 3. The Occupy idiots are nothing but stooges - left brainwashed and ignorant by the liberal infiltration of our educational system
Self-interest	Yes	6.95%	1. Failed capitalism is desperately cannibalizing everything we love: education, health care, old people. Make it stop. #concise #ows #mayday 2. Preach it #TaxCheat RT@cbrangel "#MayDay reminds us work together as 1 nation to make the American Dream a reality for everyone." #ows #m1gs 3. "Wall Street is a symbol to the indifference of the suffering of the people!" As the countdown to strike begins. #M1NYC #OWS #M1GS
	No	91.34%	1. Tens of Thousands Take Broadway in NYC for #Occupy Wall Street Mayday - YouTube http://t.co/rmEvtCgG 2. #OccupyWallStreet on 34th http://t.co/Ks0y8P0W 3. nyclu: Another march leading from Madison sq pk to union sq. pk #occupywallstreet #ows NYCLU
Collective efficacy	Yes	4.88%	1. Direct action is similar to voting, In fact it actually makes a difference! #ows #mayday #m1gs 2. nyclu: just came from #ows #occupywallstreet where we handed out know your rights cards to occupiers NYCLU 3. Many thousands marching down Broadway at 10th Street now. Numbers just growing by the hour #winning #ows #MayDay
	No	93.10%	1. May Day. NYC. 2012 #M1NYC http://t.co/SfEiAAYK 2. Occupy protest outside Grand Central by Krystyl Baldwin - http://t.co/kxe415er 3. City Room: Morning Buzz: A Day Of Occupy Protests http://t.co/gootpHir

(*table continues*)

Table 3 (*continued*)

Variable	Coding	Frequency	Sample tweets
Justice concerns	Yes	9.48%	1. Tell Pres. @BarackObama: Don't give Wall Street crooks a "get out of jail free" card. http://t.co/dAiAYWIC #p2 #ows #ffraud 2. #NOW says #EnoughRUSH it's time to #StopRush. End corporate-funded #hatespeech. #Occupy #WOOD #Grandrapids http://t.co/5dhSCpSE 3. Today was awesome. Thousands of people in the streets, calling for an end to capitalism, war, and racism. #mayday #p2 #OWS #socialism
	No	88.87%	1. The Occupy Wall Street movement really needs to end. . . 2. Stop occupying Wall Street and occupy a job! #mfs 3. Dear Occupy Wall Street, You people are annoying. . . Go get a job and shut up. Sincerely, Normal people with lives
Ideological themes	Yes	11.64%	1. The news coming in from #OWS is loud and clear. "Business as usual" just won't do. The economic order of the day has to change! 2. Wealth distribution, fairness, corruption, banks - reasons for #OccupyWallStreet occupation - http://t.co/RKqO8qWB RT 3. EVERY single Worker performs a Necessary job. Do they Not deserve to be paid enuf to LIVE ON after providing 40 Hours of PRODUCTIVITY? #OWS
	No	86.38%	1. Occupy Wall Street May Day Protests in New York http://t.co/KFZr0VgK 2. Occupy takes May Day protests to streets: NEW YORK (Reuters) - Occupy Wall Street protesters massed outside bank . . . http://t.co/yQ7nazAq 3. OWS occupying Broadway http://t.co/nWS0vbLk
Anger and positive emotion	Yes	33.46%	Anger 1. #Occupy We have the right to assemble! Our rights are being violated! To all my followers, I'd like (cont) 2. OCCUPY! Lovin the protest, 'They say come back, we say fight back!' @AngellyneK 3. Peaceful protest, chants, fellow NY's together! NYPD tried to run us over, but these are our streets! #May1 #OWS http://t.co/0UKrmnNn Positive Emotion 1. Love the diversity, energy, and just how many people are here for #ows #mayday at Union Square #M1NYC http://t.co/I1jHlIUs 2. I love #occupywallstreet 3. Everywhere I go it's hugs & smiles. This is a revolution of love. #M1NYC
	No	63.58%	1. Whatever you feel about #OWS, you're getting a teach-in on the new American police state. (via @alexanderchee) 2. People in the streets at union square NYC. #ows 3. @JonathanHoenig If OWS only had a real understanding of what they are asking for.. Just cant wrap my head around the mentality

Twitter User Ideology Point Estimates

We drew precomputed user ideology point estimates ranging from liberal (−2) to conservative (2) for a subset of users included in Barberá's (2015) large-scale study of Twitter users' ideological orientations. This well-validated measure is based on follower relation-

Table 4
Response Frequencies for Manually Coded Tweets (at Individual User Level of Analysis)

Variable	Response category	Response frequency
Protest participation	Yes	11.8%
	No	88.2%
Social identification	Yes	18.1%
	No	81.9%
Self-interest	Yes	11.8%
	No	88.2%
Collective efficacy	Yes	8.3%
	No	91.7%
Justice concerns	Yes	14.4%
	No	85.6%
Ideological themes	Yes	17.3%
	No	82.7%
Anger	Yes	12.1%
	No	87.9%
Positive emotion	Yes	22.3%
	No	77.7%

Note. N = 8,244 unique Twitter users.

ships using political annotations of well-known political actors as seed data to estimate individual users' ideology in a sample of 30 million U.S. Twitter users. We reverse-scored user ideology to range from conservative (−2) to liberal (2), so that a positive relationship would indicate that liberals were more likely to participate in the demonstration (see Figure 5 for the distributions of user ideology in the three datasets we analyzed).

Back-Fetching Tweet Data From the Preceding Two Weeks

To explore whether the psychological processes during the lead-up to the protest—when individuals may have been considering whether to protest—differed from those that occurred on the day of the protest, we applied the same methodology to a set of tweets gathered from the days preceding the May Day protest. Analyzing these back-fetched tweets allowed us to examine the same psychological variables (measured earlier) for users who tweeted during the protest itself. Using the Twitter API we obtained a set of 664,937 tweets sent by the users included in the manually coded data set from the period two weeks prior to May 1. We applied machine-learning models to all back-fetched tweets, without filtering for Occupy-related hashtags. Because we trained the models to estimate psychological variables of interest in Occupy-related tweets, relevance to Occupy Wall Street was inherent in each statistical model.

Coding and Variable Creation for Prior Data

We coded back-fetched tweets using the same models that were applied to the machine-coded dataset on the day of the protest, except that estimates for anger and positive emotion were obtained by processing the text with LIWC, and user ideology estimates were again taken from Barberá's (2015) database.

Results

Analysis of Manually Coded Tweets

Correlation. Bivariate correlations among all manually coded variables are listed in Table 6.

Regression. We first investigated the effects of the independent and mediating variables on protest participation using the manually coded dataset taken from the day of the protest using logistic regression (see Table 7). We entered all predictor variables in a single step, so results reported are adjusting for all other predictors in the model. In keeping with theoretical expectations, results revealed that social identification was associated with participation in the protest, $b = 2.26$, $SE = .11$, $Wald\ \chi^2(1) = 430.46$, $p < .001$. People who sent tweets containing themes of social identification

Table 5
Percentage of Tweets in Training Data Predicted by Machine Models to Contain Expressions of Protest Participation, Social Identification, Self-Interest, Collective Efficacy, Ideological Themes, and Justice Concerns

Type of content	Percent positive	Percent negative
Protest participation	11.78%	88.22%
Social identification	17.24%	82.76%
Self-interest	11.27%	88.73%
Collective efficacy	7.84%	92.16%
Ideological themes	16.74%	83.26%
Justice concerns	13.65%	86.35%

Note. Ns range from 8,138 to 8,244.

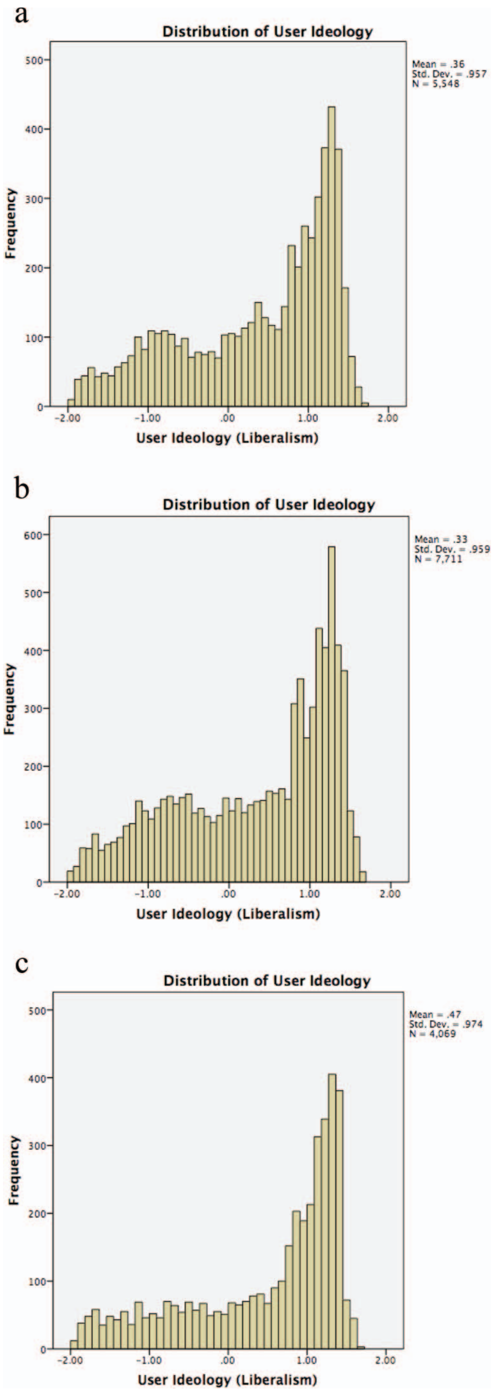

Figure 5. Distribution of user ideology for the (a) hand-coded dataset (N = 5,548), (b) machine-coded dataset from the day of the protest (N = 7,711), and (c) machine-coded dataset from the two weeks prior to the protest (N = 4,069) See the online article for the color version of this figure.

with the Occupy Wall Street movement were more likely to participate in the protest than people who did not. Using an odds ratio as a measure of effect size, we see that the increase in the odds of an individual participating in the protest for each unit increase in social identification is Exp (B) = 9.58, 95% CI [7.74, 11.86]. Political ideology also predicted participation, b = .64, SE = .07, $Wald$ $\chi^2(1)$ = 89.03, $p <$ .001, such that individuals who were more liberal (or less conservative) were more likely to participate in the demonstration, Exp (B) = 1.88, 95% CI [1.64, 2.15].

Unexpectedly, participation in the protest was negatively associated with the expression of justice concerns, $b = -.41$, SE = .16, $Wald$ $\chi^2(1)$ = 6.35, p = .012, Exp (B) = .67, 95% CI [.49, .93], ideological themes, $b = -1.15$, SE = .17, $Wald$ $\chi^2(1)$ = 46.67, $p <$ .001, Exp (B) = .31, 95% CI [.23, .44], and positive emotion, $b = -.29$, SE = .12, $Wald$ $\chi^2(1)$ = 6.08, p = .014, Exp (B) = .74, 95% CI [.60, .91]. That is, users who expressed justice concerns, ideological themes, and positive emotion were *less* likely to participate in the protest. The effect of self-interest on protest participation was nonsignificant, b = .20, SE = .16, $Wald$ $\chi^2(1)$ = 1.69, p = .194, as were the effects of collective efficacy, b = .30, SE = .16, $Wald$ $\chi^2(1)$ = 3.46, p = .063, and anger, $b = -.10$, SE = .17, $Wald$ $\chi^2(1)$ = .37, p = .546.

Path analysis. To investigate the hypothesis that social identification and liberal ideology motivate participation in protest through other psychological variables—such as self-interest, collective efficacy, justice concerns, ideological themes, anger, and positive emotion—we conducted a path analysis using MPlus 6 (Muthén & Muthén, 1998). Specifically, we examined whether direct associations between (a) social identification and political ideology and (b) protest participation were mediated by (c) self-interest, collective efficacy, justice concerns, ideological themes, anger, and positive emotion (see Figure 6 and Table 8).

We observed that social identification was indeed positively associated with collective efficacy, b = .29, SE = .01, β = .40, Z = 22.20, $p <$ .001, the expression of justice concerns, b = .31, SE = .01, β = .34, Z = 22.14, $p <$ .001, ideological themes, b = .23, SE = .01, β = .24, Z = 17.10, $p <$ .001, and positive emotion, b = .29, SE = .02, β = .26,

Table 6
Correlations Involving Motivational Factors and Protest Participation in Manually Coded Tweets

Variable	1	2	3	4	5	6	7	8
(1) Protest participation								
(2) Social identification	.391**							
(3) Liberal ideology	.216**	.262**						
(4) Self-interest	.092**	.305**	.115**					
(5) Collective efficacy	.155**	.421**	.181**	.508**				
(6) Justice concerns	.044**	.350**	.120**	.440**	.452**			
(7) Ideological themes	−.046**	.198**	−.100**	.296**	.308**	.516**		
(8) Anger	−.024*	.007	−.123**	.038**	−.032**	.053**	.061**	
(9) Positive emotion	.039**	.249**	.016	.286**	.306**	.297**	.286**	.156**

Note. *N*s range from 5,548 to 8,244.
* $p < .05$. ** $p < .01$.

$Z = 19.68$, $p < .001$. That is, adjusting for all other variables in the model, people who expressed identification with the Occupy movement were more likely to mention collective efficacy, issues of justice, ideology, and positive emotion. As predicted, collective efficacy was positively associated with protest participation, $b = .09$, $SE = .04$, $\beta = .04$, $Z = 2.00$, $p < .05$.

Interestingly, liberal ideology was negatively associated with the expression of ideological themes, $b = -.06$, $SE = .01$, $\beta = -.15$, $Z = -11.26$, $p < .001$. That is, conservatives were more likely than liberals to express ideological sentiments when tweeting about this demonstration. The use of ideological language was negatively related to protest participation, $b = -.15$, $SE = .02$, $\beta = -.08$, $Z = -7.46$, $p < .001$, and so was the expression of justice concerns, $b = -.14$, $SE = .03$, $\beta = -.07$, $Z = -5.17$, $p < .001$, and the communication of positive emotion, $b = -.04$, $SE = .02$, $\beta = -.03$, $Z = -2.14$, $p < .05$ (see Figure 6).

Overall, the model explained 18% of the variance in protest participation. When collective efficacy, justice concerns, ideological themes, and positive emotion were included in the model, the direct effects of group identification on participation remained significant, $b = .69$, $SE = .03$, $\beta = .40$, $Z = 23.10$, $p < .001$. Likewise, when the expression of ideological themes was included in the model, the direct effect of liberal ideology on participation remained significant, $b = .07$, $SE = .01$, $\beta = .10$, $Z = 10.19$, $p < .001$.

Finally, we tested indirect effects using a bootstrapping analysis and observed that the effect of social identification on protest participation were partially mediated by collective efficacy 95% CI [.003, .027], justice con-

Table 7
Logistic Regression Predicting User's Protest Participation on the Basis of Manually Coded Tweets Sent on the Day of the Protest

Variable	*B*	Exp (*B*)	*SE*	95% CI for Exp (*B*) Lower	Upper
Social identification	2.26***	9.58	.11	7.74	11.86
Liberal political ideology	.64***	1.88	.07	1.64	2.15
Self-interest	.20	1.22	.16	.90	1.65
Collective efficacy	.30	1.33	.16	.97	1.82
Justice concerns	−.41**	.67	.16	.49	.93
Ideological themes	−1.15***	.31	.17	.23	.44
Anger	−.10	.88	.17	.64	1.22
Positive emotion	−.29**	.74	.12	.60	.91

Note. $N = 8{,}244$.
** $p < .01$. *** $p < .001$.

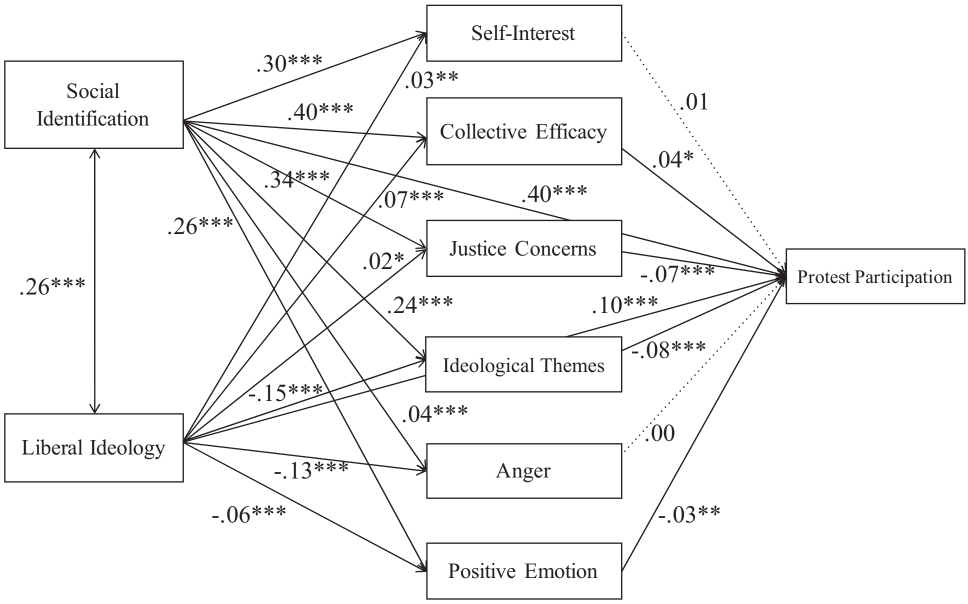

Figure 6. Path model illustrating the effects of social identification and liberal ideology on protest participation, mediated by the social and psychological variables of self-interest, perceptions of collective efficacy, justice concerns, ideological themes, anger, and positive emotion. Saturated model: * $p < .05$. ** $p < .01$. *** $p < .001$.

cerns 95% CI [−.034, −.017], ideological themes 95% CI [−.024, −.015], and positive emotion 95% CI [−.013, −.002]. The effect of liberal ideology on protest participation was partially mediated by the expression of ideological themes 95% CI [.009, .015]. Because 0 was excluded from these unstandardized 95% confidence intervals, we consider these mediation effects to be significant (see Table 8).

Analysis of Machine-Coded Tweets

Turning now to the corpus of tweets that were machine-coded (rather than manually coded by research assistants), we investigated parallel relationships involving social and psychological variables and protest participation using correlation (see Table 9 for mean values and standard deviations and Table 10 for correlations) and linear regression, again with all variables entered in a single step (see Table 11). The independent and mediating variables explained a significant proportion of variance in protest participation, $R^2 = .47$, $F(8, 7{,}702) = 838.05$, $p < .001$. As hypothesized, social identification predicted participation in the protest, $b = .94$, $SE = .01$, $t(7{,}702) = 67.81$, $p < .001$, 95% CI [.91, .96], such that people who expressed stronger identification with the Occupy Wall Street movement were more likely to participate in the protest. Political ideology also positively predicted participation, $b = .01$, $SE = .001$, $t(7{,}702) = 6.40$, $p < .001$, 95% CI [.01, .01]; individuals who were more liberal were more likely to participate in the demonstration.

Consistent with the foregoing results, participation in the protest was negatively associated with the expression of self-interest, $b = -.27$, $SE = .02$, $t(7{,}702) = -11.48$, $p = <.001$, 95% CI [−.32, −.22], justice concerns, $b = -.06$, $SE = .01$, $t(7{,}702) = -4.85$, $p < .001$, 95% CI [−.08, −.03], ideological themes, $b = -.05$, $SE = .01$, $t(7{,}702) = -7.96$, $p < .001$, 95% CI [−.07, −.04], anger, $b = -.01$, $SE = .003$, $t(7{,}702) = -2.39$, $p = .017$, 95% CI [−.01, .00], and positive emotion, $b = -.02$, $SE = .003$, $t(7{,}702) = -6.29$, $p < .001$, 95% CI [−.02, −.01]. In this model, collective efficacy was also negatively related to participation, $b =$

Table 8
Direct and Indirect Effects of Social Identification, Liberal Ideology, Self-Interest, Collective Efficacy, Justice Concerns, Ideological Themes, Anger, and Positive Emotion

	b	β
Direct effects		
Social identification → Protest participation	.69***	.40***
Liberal ideology → Protest participation	.07***	.10***
Self-interest → Protest participation	.01	.01
Collective efficacy → Protest participation	.09*	.04*
Justice concerns → Protest participation	−.14***	−.07***
Ideological themes → Protest participation	−.15***	−.08***
Anger → Protest participation	.00	.00
Positive emotion → Protest participation	−.04*	−.03*
Social identification → Self-interest	.25***	.30***
Liberal ideology → Self-interest	.01**	.03**
Social identification → Collective efficacy	.29***	.40***
Liberal ideology → Collective efficacy	.02***	.07***
Social identification → Justice concerns	.31***	.34***
Liberal ideology → Justice concerns	.01*	.02*
Social identification → Ideological themes	.23***	.24***
Liberal ideology → Ideological themes	−.06***	−.15***
Social identification → Anger	.03***	.04***
Liberal ideology → Anger	−.04***	−.13***
Social identification → Positive emotion	.29***	.26***
Liberal ideology → Positive emotion	−.02***	−.06***
Indirect effects		
Social identification → Self-interest → Protest participation	{−.005, .009}	
Social identification → Efficacy → Protest participation	**{.003, .027}**	
Social identification → Justice → Protest participation	**{−.034, −.017}**	
Social identification → Ideological themes → Protest participation	**{−.024, −.015}**	
Social identification → Anger → Protest participation	{−.001, .001}	
Social identification → Positive emotion → Protest participation	**{−.013, −.002}**	
Liberal ideology → Self-interest → Protest participation	{−.001, .001}	
Liberal ideology → Collective efficacy → Protest participation	{.000, .005}	
Liberal ideology → Justice concerns → Protest participation	{−.003, .000}	
Liberal ideology → Ideological themes → Protest participation	**{.009, .015}**	
Liberal ideology → Anger → Protest participation	{−.002, .002}	
Liberal ideology → Positive emotion → Protest participation	{.000, .003}	
R^2	18.0%	

Note. R^2 signifies proportion of variance in protest participation explained by the saturated model. $N = 5{,}391$. The bold values are for 95% Confidence Intervals that do not include zero. So they are significant at $p < .05$.
* $p < .05$. ** $p < .01$. *** $p < .001$.

−.22, $SE = .02$, $t(7{,}702) = -11.88$, $p = <.001$, 95% CI [−.25, −.18] in this model. Thus, people who tweeted more about these topics were *less* likely to participate in the May Day demonstration, in comparison with people who tweeted less about them.

Analysis of Tweets Sent During the Preceding Two Weeks

When we analyzed the corpus of back-fetched tweets sent during the two weeks preceding the demonstration, a linear regression with all predictors entered in a single step revealed that the independent and mediating variables explained a significant proportion of variance in protest participation, $R^2 = .41$, $F(8, 4060) = 348.90$, $p < .001$ (see Table 12 for correlations and Table 13 for regression results).

As before, social identification predicted participation in the protest, $b = .53$, $SE = .02$, $t(4{,}060) = 32.61$, $p < .001$; people who ex-

Table 9
Mean Values and Standard Deviations for Machine Coded Tweets

Variable	Mean	*SD*
Protest participation	.30	.38
Social identification	.34	.38
Liberal ideology	.47	.97
Self-interest	.48	.42
Collective efficacy	.23	.37
Justice concerns	.48	.45
Ideological themes	.54	.46
Anger	.09	.09
Positive emotion	.33	.21

Note. $N = 4{,}069$.

pressed stronger identification with the Occupy Wall Street movement during the two weeks prior to the demonstration were more likely to turn out for it. Political ideology also predicted participation once again, $b = .04$, $SE = .01$, $t(4{,}060) = 6.81$, $p < .001$. People who were more liberal were more likely to attend the demonstration.

Consistent with the results obtained for tweets sent on the day of the protest, individuals who (during the two weeks prior to the demonstration) expressed ideological themes, $b = -.21$, $SE = .01$, $t(4{,}060) = -14.97$, $p < .001$, anger, $b = -.26$, $SE = .06$, $t(4{,}060) = -4.73$, $p < .001$, and positive emotion, $b = -.20$, $SE = .02$, $t(4{,}060) = -.11$, $p < .001$, were less likely to participate. Those who expressed a stronger sense of collective efficacy were also less likely to attend, $b = -.04$, $SE = .02$, $t(4{,}060) = -2.53$, $p = <.05$, possibly because they felt that their participation was unnecessary—or because they were encouraging others to take their place. There were no significant effects of the expression of self-interest, $b = -.01$, $SE = .02$, $t(4{,}060) = -.45$, $p = .653$, or justice concerns, $b = .02$, $SE = .01$, $t(4{,}060) = 1.55$, $p = .121$, on protest participation.

Accounting for the Expression of Backlash Against Occupy Wall Street

To account for the fact that there were social media users in these datasets who were included because they mentioned Occupy Wall Street but were criticizing or attacking the movement (or the demonstration), we ran two sets of additional sensitivity analyses. In one of these analyses we adjusted statistically for the expression of anti-Occupy sentiment, and in the other we filtered on pro-Occupy sentiment—using only tweets that were coded as containing sentiment in favor of the movement or demonstration.

For the manually coded tweets, the results of these sensitivity analyses were consistent with the findings summarized above. For machine-coded tweets sent the day of the protest, we observed that the effect of justice concerns became a nonsignificant (rather than negative) predictor of participation after adjusting statistically for anti-Occupy sentiment. When we filtered on pro-Occupy sentiment, however, the expression of justice concerns was again a negative predictor of participation. When we adjusted statistically for anti-Occupy sentiment, anger remained a negative predictor of partici-

Table 10
Correlations Among Motivational Factors and Protest Participation in Machine Coded Tweets Sent on the Day of the Demonstration

Variable	1	2	3	4	5	6	7	8
(1) Protest participation								
(2) Social identification	.613**							
(3) Liberal ideology	.232**	.203**						
(4) Self-interest	.190**	.562**	.009					
(5) Collective efficacy	.290**	.651**	.110**	.591**				
(6) Justice concerns	.088**	.425**	.006	.588**	.456**			
(7) Ideological themes	−.115**	.084**	−.247**	.338**	.180**	.516**		
(8) Anger	−.059**	.020*	−.124**	.162**	.015	.148**	.165**	
(9) Positive emotion	−.005	.175**	−.105**	.272**	.163**	.178**	.192**	.209**

Note. $N = 12{,}540$.
* Correlation is significant at the .05 level (2-tailed). ** Correlation is significant at the .01 level (2-tailed).

Table 11
Linear Regression Predicting User Participation in Machine-Coded Tweets From the Day of the Protest

Variable	b	SE	β	95% CI for B Lower	95% CI for B Upper
Social identification	.94***	.01	.81	.91	.96
Liberal ideology	.01***	.00	.06	.01	.01
Self-interest	−.27***	.02	−.14	−.32	−.22
Collective efficacy	−.22***	.02	−.14	−.25	−.18
Justice concerns	−.06***	.01	−.06	−.08	−.03
Ideological themes	−.05***	.01	−.08	−.07	−.04
Anger	−.01*	.00	−.021	−.01	−.00
Positive emotion	−.02***	.00	−.06	−.02	−.01
R^2	47.00%				

Note. N =7,711.
* $p < .05$. *** $p < .001$.

pation, but this effect became nonsignificant when we filtered on pro-Occupy sentiment.

When we reanalyzed the corpus of machine-coded tweets sent during the two weeks prior to the protest, the effect of positive emotion became nonsignificant when we adjusted for anti-Occupy sentiment, but it remained a negative predictor of participation when we filtered on pro-Occupy sentiment. When we adjusted for anti-Occupy sentiment, self-interest was a positive predictor of participation, but this effect was nonsignificant when we filtered on pro-Occupy sentiment. Finally, when we adjusted statistically for anti-Occupy sentiment, the expression of justice concerns was negatively associated with participation, but it was positively associated with participation when we filtered on pro-Occupy sentiment.

Thus, we observed some suppression effects when analyzing the three datasets. For instance, justice concerns were positively correlated with protest participation in the hand-coded and machine-coded datasets, but they were negatively associated with protest participation when entered into multiple regressions. In addition, self-interest was positively correlated with protest participation in the machine-coded dataset, but it was a negative predictor of participation in the regression analysis. In two of the datasets, group efficacy was positively correlated with protest participation, but it was a negative predictor after adjusting for other factors. Finally, positive emotion was positively correlated with protest participation in one dataset, but in all three datasets it was negatively associ-

Table 12
Correlations Involving Motivational Factors and Protest Participation in Tweets Sent Two Weeks Prior to the Protest

Variable	1	2	3	4	5	6	7	8
(1) Protest participation								
(2) Social identification	.513**							
(3) Liberal ideology	.338**	.273**						
(4) Self-interest	−.007	.305**	−.096**					
(5) Collective efficacy	.249**	.618**	.185**	.394**				
(6) Justice concerns	−.053**	.246**	−.095**	.471**	.339**			
(7) Ideological themes	−.268**	.023	−.329**	.448**	.159**	.605**		
(8) Anger	−.184**	−.02	−.202**	.295**	.032*	.199**	.255**	
(9) Positive emotion	−.060**	.184**	−.142**	.370**	.179**	.081**	.163**	.319**

Note. N = 6,285.
* Correlation is significant at the .05 level (2-tailed). ** Correlation is significant at the .01 level (2-tailed).

Table 13
Linear Regression Predicting User Participation in Machine-Coded Tweets Sent Two Weeks Prior to the Protest

Variable	b	SE	β	95% CI for B Lower	Upper
Social identification	.53***	.02	.53	.50	.56
Liberal ideology	.04***	.01	.09	.03	.05
Self-interest	−.01	.02	−.01	−.04	.02
Collective efficacy	−.04*	.02	−.04	−.08	−.01
Justice concerns	.02	.01	.03	−.01	.05
Ideological themes	−.21***	.01	−.25	−.23	−.18
Anger	−.26***	.06	−.07	−.37	−.15
Positive emotion	−.20***	.02	−.11	−.25	−.16
R^2	41.0%				

Note. N =4,069.
* $p < .05$. *** $p < .001$.

ated with protest participation in multiple regression analyses.

Overall, then, we did not find that taking into account whether social media users expressed pro- versus anti-OWS sentiment did much to clarify the role of communicating (vs. not communicating) ideological themes. The suppression effects are potentially interesting but do not alter the general motivational portrait of Occupy Wall Street protestors that emerges from our study. Nevertheless, it may be useful to keep these reversals in mind when conducting or interpreting other findings pertaining to collective action and to pursue them more directly and systematically in future research.

Discussion

As hypothesized, social identification and liberal ideology were robust predictors of participation in a May Day demonstration organized by the Occupy Wall Street movement. The association between social identification and protest participation was partially mediated by the expression of collective efficacy, justice concerns, ideological themes, and positive emotion. This pattern of mediation is broadly consistent with the Social Identity Model of Collective Action (SIMCA), which posits that identification with a given social group fosters participation in collective action by strengthening the perception of oneself as part of a group with a shared mission (McGarty, Bliuc, Thomas, & Bongiorno, 2009), an emphasis on unjust circumstances that must be righted through joint action on behalf of one's group (Van Zomeren et al., 2012), and a sense of collective efficacy (Tausch et al., 2011). It is also consistent with the results of a study by Theocharis et al. (2015), which suggested that social media was used in the case of Occupy Wall Street and other protest movements to foster a sense of community (and to provide logistical information) rather than to broadcast calls for participation per se (see also Jost et al., in press).

The effect of liberalism on protest participation was partially mediated by the expression of ideological themes. This finding highlights the importance of beliefs about the legitimacy (or illegitimacy) of the social system—variables that are often overlooked in social psychological models of collective action (Jost et al., 2017). In this context, the expression of ideological themes and positive emotion were negatively (rather than positively) associated with protest participation in the three corpuses of tweets that we analyzed. That is, political ideology manifested itself more explicitly in terms of system-justifying forms of *backlash*—protest against the protestors (see also Diekman & Goodfriend, 2007; O'Brien & Crandall, 2005; Rudman et al., 2012; Yeung et al., 2014)—than in terms of system-challenging forms of protest against the status quo. In the two weeks prior to the demonstration, the expression of anger was also a negative predictor of participation. Although these findings were somewhat unexpected, they are consistent with the notion that—whereas decisions to participate in nondisruptive action are often made spontaneously on the basis of anger—decisions to participate in disruptive action (such as protest) are more likely to be guided by group commitment and strategic considerations (Jost et al., 2012; Tausch et al., 2011).

In the human- and machine-coded tweets sent on the day of the demonstration, we observed that justice concerns were negatively associated with participation. This effect does not seem to be attributable to conservative backlash, because it remained significant in sensitivity analyses that accounted for sentiment directed at the Occupy movement. In the two weeks prior to the demonstration, we also observed that the expression of efficacy was a negative predictor of participation. It is possible that more geographically distant supporters of the Occupy movement were especially likely to emphasize justice-related themes and to promote collective efficacy as a way of encouraging others to take part in the demonstration, whereas local supporters were focused on more pragmatic considerations. Such an interpretation would be broadly consistent with construal level theory, which suggests that abstract ideas loom larger from a distance, whereas more concrete concerns take center stage in the "here-and-now" (Ledgerwood, Trope, & Chaiken, 2010).

In terms of methodological contributions, this work illustrates the promise (as well as challenges) of applying machine-learning techniques to analyzing new data sources, such as social media messages, to study protest activity. The possibility of using a relatively limited amount of hand-coded data to train models that can analyze message contents as nuanced and complex as these psychological variables affirms the value and viability of using social media platforms as real-time windows into the motivations of would-be protesters. Importantly, we obtained very similar results when we compared messages sent during the two-week lead-up to the protest to those sent on the day of the protest. Even after taking into account the results of sensitivity analyses, the overall pattern of results is largely consistent, suggesting that the motivational antecedents of protest in this case were relatively stable over the 2-week period.

The goal of the present research was to explore the social and psychological factors that motivate participation in political protest, by focusing on messages written by potential protestors (as well as nonprotestors). In future research, it would be useful to analyze the contents of social media messages *viewed* by actual and potential protestors as well; this might shed some light on what types of persuasive appeals lead people to participate in protest. Although we observed that the use of language pertaining to self-interest and emotion was unrelated to protest participation, exposure to these types of messages may nonetheless motivate participation. It is also possible that messages received by individual users would shape—wittingly or unwittingly—the messages that they subsequently transmit on the same topic. Insofar as at least some of the motivational antecedents of support for protest movements on the political left and right may diverge (Hennes et al., 2012; Jost et al., 2017), it would be extremely useful to extend the types of analyses we have conducted in the case of Occupy Wall Street to illuminate the antecedents of participation in a conservative (or right-wing) demonstration.

References

Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science, 37,* 246–278. http://dx.doi.org/10.2307/2111531

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Barbalet, J. (1998). *Emotion, social theory, and social structure: A macrosociological approach*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511488740

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis, 23,* 76–91. http://dx.doi.org/10.1093/pan/mpu011

Bar-Tal, B., Halperin, E., & de Rivera, J. (2007). Collective emotions in conflict situations: Societal implications. *Journal of Social Issues, 63,* 441–460. http://dx.doi.org/10.1111/j.1540-4560.2007.00518.x

Becker, J. C., & Wright, S. C. (2011). Yet another dark side of chivalry: Benevolent sexism undermines and hostile sexism motivates collective action for social change. *Journal of Personality and Social Psychology, 101,* 62–77. http://dx.doi.org/10.1037/a0022615

Diekman, A. B., & Goodfriend, W. (2007). The good and bad of social change: Ambivalence toward activist groups. *Social Justice Research, 20,* 401–417. http://dx.doi.org/10.1007/s11211-007-0050-z

Downs, A. (1957). *An economic theory of democracy*. New York, NY: Harper & Row.

Drury, J., & Reicher, S. D. (2009). Collective psychological empowerment as a model of social change: Researching crowds. *Journal of Social Issues, 65,* 707–725. http://dx.doi.org/10.1111/j.1540-4560.2009.01622.x

Finkel, S. E., Muller, E. N., & Opp, K. D. (1989). Personal influence, collective rationality, and mass political action. *The American Political Science Review, 83,* 885–903. http://dx.doi.org/10.2307/1962065

Goodwin, J., & Jasper, J. (2006). Emotions and social movements. In J. Stets, & J. Turner, (Eds.) *Handbook of the sociology of emotions* (pp. 611–635). New York, NY: Springer. http://dx.doi.org/10.1007/978-0-387-30715-2_27

Gurr, T. (1970). *Why men rebel*. Princeton, NJ: Princeton University Press.

Hennes, E. P., Nam, H. H., Stern, C., & Jost, J. T. (2012). Not all ideologies are created equal: Epistemic, existential, and relational needs predict system-justifying attitudes. *Social Cognition, 30,* 669–688. http://dx.doi.org/10.1521/soco.2012.30.6.669

Jost, J. T. (2006). The end of the end of ideology. *American Psychologist, 61,* 651–670. http://dx.doi.org/10.1037/0003-066X.61.7.651

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology, 25,* 881–919. http://dx.doi.org/10.1111/j.1467-9221.2004.00402.x

Jost, J. T., Barberá, P., Bonneau, R., Metzger, M. M., Nagler, J., Sterling, J., & Tucker, J. (in press). How social media facilitates political protest: Information, motivation, and social networks. *Advances in Political Psychology*.

Jost, J. T., Becker, J., Osborne, D., & Badaan, V. (2017). Missing in (collective) action: Ideology, system justification, and the motivational antecedents of protest behavior. *Current Directions in Psychological Science, 26,* 99–108. http://dx.doi.org/10.1177/0963721417690633

Jost, J. T., Chaikalis-Petritsis, V., Abrams, D., Sidanius, J., van der Toorn, J., & Bratt, C. (2012). Why men (and women) do and don't rebel: Effects of system justification on willingness to protest. *Personality and Social Psychology Bulletin, 38,* 197–208. http://dx.doi.org/10.1177/0146167211422544

Jost, J. T., Nosek, B. A., & Gosling, S. D. (2008). Ideology: Its resurgence in social, personality, and political psychology. *Perspectives on Psychological Science, 3,* 126–136. http://dx.doi.org/10.1111/j.1745-6916.2008.00070.x

Kawakami, K., & Dion, K. L. (1995). Social identity and affect as determinants of collective action: Toward an integration of relative deprivation and social identity theories. *Theory & Psychology, 5,* 551–577. http://dx.doi.org/10.1177/0959354395054005

Kelly, C., & Breinlinger, S. (1996). *The social psychology of collective action*. Basingstoke, United Kingdom: Taylor & Francis.

Klandermans, B. (1997). *The social psychology of protest*. Oxford, United Kingdom: Blackwell.

Klandermans, B., & van Stekelenburg, J. (2013). Social movements and the dynamics of collective action. In L. Huddy, D. O. Sears, & J. S. Levy (Eds.), *The Oxford handbook of political psychology* (pp. 774–811). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199760107.013.0024

Kuran, T. (1991). Now out of never: The element of surprise in the East European revolution of 1989. *World Politics, 44,* 7–48. http://dx.doi.org/10.2307/2010422

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Lazarus, R. S. (2001). Relational meaning and discrete emotions. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion* (pp. 37–67). New York, NY: Oxford University Press.

Ledgerwood, A., Trope, Y., & Chaiken, S. (2010). Flexibility now, consistency later: Psychological distance and construal shape evaluative responding. *Journal of Personality and Social Psychology, 99,* 32–51. http://dx.doi.org/10.1037/a0019843

Marwell, G., & Oliver, P. (1993). *The critical mass in collective action: A micro-social theory*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511663765

McGarty, C., Bliuc, A. M., Thomas, E. F., & Bongiorno, R. (2009). Collective action as the material expression of opinion-based group membership. *Journal of Social Issues, 65,* 839–857. http://dx.doi.org/10.1111/j.1540-4560.2009.01627.x

McGarty, C., Thomas, E. F., Lala, G., Smith, L. G., & Bliuc, A. M. (2014). New technologies, new identities, and the growth of mass opposition in the Arab Spring. *Political Psychology, 35,* 725–740. http://dx.doi.org/10.1111/pops.12060

Muller, E., & Opp, K. D. (1986). Rational choice and rebellious collective action. *The American Political Science Review, 80,* 471–487. http://dx.doi.org/10.2307/1958269

Muthén, B., & Muthén, L. (1998). *Mplus user's guide*. Los Angeles, CA: Author.

Oberschall, A. (1973). *Social conflict and social movements*. Englewood Cliffs, NJ: Prentice Hall.

O'Brien, L. T., & Crandall, C. S. (2005). Perceiving self-interest: Power, ideology, and maintenance of the status quo. *Social Justice Research, 18,* 1–24. http://dx.doi.org/10.1007/s11211-005-3368-4

Olson, M. (1965). *Logic of collective action*. Cambridge, MA: Harvard University Press.

Osborne, D., & Sibley, C. G. (2013). Through rose-colored glasses: System-justifying beliefs dampen the effects of relative deprivation on well-being and political mobilization. *Personality and Social Psychology Bulletin, 39,* 991–1004. http://dx.doi.org/10.1177/0146167213487997

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. Austin, TX: Liwc.net.

Riker, W., & Ordeshook, P. (1968). A theory of the calculus of voting. *The American Political Science Review, 62,* 25–42. http://dx.doi.org/10.2307/1953324

Rudman, L. A., Moss-Racusin, C., Glick, P., & Phelan, J. (2012). Reactions to vanguards: Advances in backlash theory. *Advances in Experimental Social Psychology, 45,* 167–227. http://dx.doi.org/10.1016/B978-0-12-394286-9.00004-4

Settles, B. (2012). Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning, 6,* 1–114. http://dx.doi.org/10.2200/S00429ED1V01Y201207AIM018

Smith, L., Gavin, J., & Sharp, E. (2015). Social identity formation during the emergence of the Occupy movement. *European Journal of Social Psychology, 45,* 818–832. http://dx.doi.org/10.1002/ejsp.2150

Solak, N., Jost, J. T., Sümer, N., & Clore, G. L. (2012). Rage against the machine: The case for system-level emotions. *Social and Personality Psychology Compass, 6,* 674–690. http://dx.doi.org/10.1111/j.1751-9004.2012.00456.x

Stürmer, S., & Simon, B. (2009). Pathways to collective protest: Calculation, identification, or emotion? A critical analysis of the role of anger in social movement participation. *Journal of Social Issues, 65,* 681–705. http://dx.doi.org/10.1111/j.1540-4560.2009.01620.x

Subasic, E., Reynolds, K. J., & Turner, J. C. (2008). The political solidarity model of social change: Dynamics of self-categorization in intergroup power relations. *Personality and Social Psychology Review, 12,* 330–352. http://dx.doi.org/10.1177/1088868308323223

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tausch, N., Becker, J. C., Spears, R., Christ, O., Saab, R., Singh, P., & Siddiqui, R. N. (2011). Explaining radical group behavior: Developing emotion and efficacy routes to normative and nonnormative collective action. *Journal of Personality and Social Psychology, 101,* 129–148. http://dx.doi.org/10.1037/a0022728

Theocharis, Y., Lowe, W., van Deth, J. W., & García-Albacete, G. (2015). Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication & Society, 18,* 202–220. http://dx.doi.org/10.1080/1369118X.2014.948035

Tilly, C. (1978). *From mobilization to revolution*. New York, NY: Longman Higher Education.

Useem, B. (1998). Breakdown theories of collective action. *Annual Review of Sociology, 24,* 215–238. http://dx.doi.org/10.1146/annurev.soc.24.1.215

van Zomeren, M., Leach, C. W., & Spears, R. (2012). Protesters as "passionate economists": A dynamic dual pathway model of approach coping with collective disadvantage. *Personality and Social Psychology Review, 16,* 180–199. http://dx.doi.org/10.1177/1088868311430835

van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin, 134,* 504–535. http://dx.doi.org/10.1037/0033-2909.134.4.504

van Zomeren, M., Spears, R., Fischer, A. H., & Leach, C. W. (2004). Put your money where your mouth is! Explaining collective action tendencies through group-based anger and group efficacy. *Journal of Personality and Social Psychology, 87,* 649–664. http://dx.doi.org/10.1037/0022-3514.87.5.649

Wakslak, C. J., Jost, J. T., Tyler, T. R., & Chen, E. S. (2007). Moral outrage mediates the dampening effect of system justification on support for redistributive social policies. *Psychological Science, 18,* 267–274. http://dx.doi.org/10.1111/j.1467-9280.2007.01887.x

Yeung, A. W. Y., Kay, A. C., & Peach, J. M. (2014). Anti-feminist backlash: The role of system justification in the rejection of feminism. *Group Processes & Intergroup Relations, 17,* 474–484. http://dx.doi.org/10.1177/1368430213514121

Received July 11, 2017
Revision received October 15, 2017
Accepted October 18, 2017 ■