Demographic Representation and Collective Storytelling in the Me Too Twitter Hashtag Activism Movement

AARON MUELLER, Johns Hopkins University, USA
ZACH WOOD-DOUGHTY, Johns Hopkins University, USA
SILVIO AMIR, Johns Hopkins University, USA
MARK DREDZE, Johns Hopkins University, USA
ALICIA LYNN NOBLES, University of California San Diego, USA

The #MeToo movement on Twitter has drawn attention to the pervasive nature of sexual harassment and violence. While #MeToo has been praised for providing support for self-disclosures of harassment or violence and shifting societal response, it has also been criticized for exemplifying how women of color have been discounted for their historical contributions to and excluded from feminist movements. Through an analysis of over 600,000 tweets from over 256,000 unique users, we examine online #MeToo conversations across gender and racial/ethnic identities and the topics that each demographic emphasized. We found that tweets authored by white women were overrepresented in the movement compared to other demographics, aligning with criticism of unequal representation. We found that intersected identities contributed differing narratives to frame the movement, co-opted the movement to raise visibility in parallel ongoing movements, employed the same hashtags both critically and supportively, and revived and created new hashtags in response to pivotal moments. Notably, tweets authored by black women often expressed emotional support and were critical about differential treatment in the justice system and by police. In comparison, tweets authored by white women and men often highlighted sexual harassment and violence by public figures and weaved in more general political discussions. We discuss the implications of this work for digital activism research and design, including suggestions to raise visibility by those who were under-represented in this hashtag activism movement. **Content warning: this article discusses issues of sexual harassment and violence.**

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**; • **Social and professional topics → Gender**; **Race and ethnicity**.

Additional Key Words and Phrases: metoo, topic modeling, demographic inference

1 INTRODUCTION

Tarana Burke aimed to break through the silence surrounding sexual harassment and violence by initiating Me Too, an offline grassroots movement, in 2006. This movement focused on increasing

---

Authors' addresses: Aaron Mueller, amueller@jhu.edu, Johns Hopkins University, 3400 N. Charles St., Baltimore, Maryland, USA, 21218; Zach Wood-Doughty, zach@cs.jhu.edu, Johns Hopkins University, 3400 N. Charles St., Baltimore, Maryland, USA, 21218; Silvio Amir, samir@jhu.edu, Johns Hopkins University, Baltimore, Maryland, USA, 21218; Mark Dredze, mdredze@cs.jhu.edu, Johns Hopkins University, 3400 N. Charles St., Baltimore, Maryland, USA, 21218; Alicia Lynn Nobles, alnobles@ucsd.edu, University of California San Diego, 9500 Gilman Dr., La Jolla, California, USA, 92093.

---

resources and building a community of peer advocates for survivors of sexual violence—especially for low-wealth communities of black women—in part by engaging in storytelling and data sharing about sexual violence [45, 78]. Alyssa Milano elevated the visibility of the movement when she tweeted a request for solidarity of victims by asking people to simply reply with *"#MeToo"* [45] on October 15, 2017. This request sparked a viral hashtag activism movement across multiple social media platforms with victims responding in solidarity, many recounting personal experiences. This resulted in over half a million #MeToo tweets within the first 24 hours after Milano's tweet [77].

The online movement has been praised for opening the door to the long-held silence among victims of sexual harassment and violence, enabling people from disparate geographic regions and cultures to participate broadly in both collective storytelling online and collective action offline, shifting societal response to sexual harassment and violence. However, the online movement has been rightfully criticized of dismissal of the contributions to feminist movements from women of color, as well as generally under-representing the voices of women of color [19, 61, 80, 94]—a significant criticism, especially given the original founding of Me Too by a woman of color and the intent of Me Too to help marginalized communities of color [45, 78].

While the movement has drawn attention from scholars focusing on a myriad of its aspects, including the content and emotions expressed in tweets [67], how tweets revealed first-person accounts of sexual violence and received support [5, 74], how tweets communicated victimhood [90], and the role of reciprocal disclosure as a motivator for self-disclosure [40], less is known about "*who*" is represented in this online movement. That is, what voices contributed to the hashtag activism movement, and what issues were salient to these voices? Attempts to understand representation are currently limited to an examination of types of social support across binary gender during the first six months of the movement [49] or recalling the frequency of disclosure of sexual assault or abuse across binary gender, race, and ethnicity during the initial week of the movement [74].

Examining who participated in the #MeToo movement is challenging, but understanding representation is critical to effecting change that reaches all individuals, especially those that are most vulnerable and often excluded from discourse [53]. Moreover, critical race theory highlights the necessity of marginalized voices, which are best heard through storytelling of their lived experience [76]. Our study aims to examine large-scale patterns in representation across intersecting identities of gender, race, and ethnicity (i.e., the framework of intersectionality [31]) which have been critically important in historical movements [71, 79, 80]. We also aim to investigate topics salient to these groups through the following research questions.

RQ1: *Who actively tweeted* in the #MeToo Twitter hashtag activism movement? We automatically inferred perceived demographic identities including gender, race, ethnicity, and the intersection of these, for user accounts tweeting about #MeToo and related topics (e.g., sexual assault). Note that we are inferring perceived demographics, which do not necessarily align with self-identified demographics. We discuss these limitations in the paper.

RQ2: *What stories were shared* and how did they vary temporally and across inferred demographic identities? We conducted a topic model analysis to examine the stories and comments shared in #MeToo tweets, sharing edited tweets to contextualize these automatically discovered topics. We contextualize these stories by finding the most probable and most distinctive topics per-demographic. To analyze these topics of conversation across time, we also analyze which topics are most common at specific pivotal points during the movement.

RQ3: *What hashtags were associated* with this movement and how did they vary temporally and across inferred demographic identities? We perform a similar analysis to the previous topic analysis; namely, we observe which hashtags were most common overall in #MeToo and related tweets, as well as which hashtags were most common across demographic groups and at pivotal points in the movement.

1.1 Contributions

Our work extends previous studies that investigated how collaborative communication technologies were leveraged for digital activism. We examine how intersectional identities leveraged Twitter to participate in and contribute to the framing of the #MeToo hashtag activism movement. More specifically, rather than *a priori* selecting a particular construct to examine, we examine large-scale patterns in the topics discussed and hashtags used by the intersected identities of gender, race, and ethnicity during the first year of #MeToo—a first in examining contributions from intersected identities for this movement. We discuss how these identities contributed different narratives to frame the movement (e.g., critical narratives about differential treatment in the justice system and by police), co-opted the movement to raise visibility in parallel ongoing movements (e.g., #Brutality, #ClimateChange), employed the same hashtags both critically and supportively (#MAGA), and revived and created new hashtags (#BelieveSurvivors, #WhyIDidntReport) over the one-year period in response to pivotal moments. We also discuss the implications of our study for digital activism design and research including computing tools that may help under-represented voices gain visibility. Because our analysis employed automated demographic inference methods, we include a discussion on the ethical, privacy, and fairness implications of these tools for our analysis.

1.2 Ethics, Fairness, and Privacy

Given the sensitive nature of this study, we reflect on the ethical, fairness, and privacy implications of our work.

First, our study was exempted from review from our respective institutional review boards under 45 CFR § 46.104. Although the tweets in our study are public, to prevent re-identification outside of a tweet's intended audience, we adhere to data protections [13] including modifying quotes of non-organizations to avoid reverse identification [11]. We only include modified tweets to contextualize our computational results. Some of the content is sensitive, so we suggest caution for readers.

Second, several recent studies have exposed a lack of fairness for demographic inference methods applied to images [21] and text [18, 106]. We acknowledge the limitations of the demographic inference methods employed in this study and discuss these in upfront in this paper (see §2). Third, social media users have raised concerns about protecting autonomy from automatic demographic inference and suggest that increased transparency may help mediate a lack of trust [46]. We hope that our upfront transparency about the demographic inference methods we employed combined with the data protection methods described above can help ameliorate concerns about using these tools in this specific context.

We believe that this analysis presents an opportunity to examine large-scale patterns of representation in this online movement. While we do not want to draw final conclusions from our analysis given the limitations of the demographic inference methods that we employ, we hope that our results provide sufficient evidence to warrant further study including those inclusive of additional intersectional identities (e.g., non-binary gender, socioeconomic status, immigration status) that likely play a role in how people participate in digital activism.

Finally, the coauthors of this paper hold different racial, ethnic, and gender identities, which has helped to ensure thoughtful perspectives in the discussion of our work. However, we acknowledge that among the coauthors, we do not hold the identity of a woman of color—a critical voice in the Me Too and feminist movements.

2 BACKGROUND

2.1 Hashtag Activism

Hashtags, initially introduced by Chris Messina in 2007 as a way to designate groups on Twitter, have evolved into a way to categorize conversations, find like-minded people, and even become self-referential in everyday conversation [53]. While hashtags can amplify the reach of voices that were already mainstream in the public discourse (e.g., celebrities, those affiliated with elite institutions), hashtags now allow other voices that were formerly excluded and ignored to be heard in public discourse. Ordinary people can democratically share their narratives and engage with like-minded people rather than waiting on those with power and privilege to narrate.

Hashtag activism is a form of digital activism that uses a hashtag/hashtags to anchor the conversation(s) related to an online movement. This bottom-up approach affords a "crowd-sourced framing process" [35] that has traditionally been held by only the leaders of a social movement. This activism has extended beyond the online realm and into offline action (i.e., computer-supported collective action) by serving as a medium of collaborative communication to mobilize people, build support for resources, organize offline actions, and warn against danger [68, 93]. For example, the Occupy movement assembled support and built momentum for an offline movement [41]. The Arab Spring uprisings used hashtag activism to organize protests and avoid state actors [20]. Hashtag activism was used to build support for public health safety measures (e.g., promoting social distancing during the COVID-19 pandemic by "flattening the curve"). Most recently, hashtag activism was used to share individual narratives of individual- and state-sanctioned violence to coalesce into collective storytelling and renewed world-wide support for the Black Lives Matter movement [39].

Previous work has examined the role of collaborative communication technologies in digital activism including how the collaborative nature of Wikipedia functions as a knowledge building tool and collective memory of a collective action [98], how singular mentions of a hashtag on social media do indeed contribute to collective storytelling (as opposed to being considered spam) [90], and how offline activism groups have used collaborative communication technologies to democratically engage with the public [9]. We extend on these previous examinations of the role of collaborative communication technologies in digital activism to understand how intersectional identities are using collaborative communication technologies to engage in storytelling, effectively framing their perspective, in a hashtag activism movement.

2.2 #MeToo

Despite that 81% of U.S. women have experienced some form of sexual harassment or violence in their lifetime [55], gendered harassment and violence have historically been framed as individual problems (e.g., victim blaming to normalize harassment and violence). In part, this has served to keep stories and subsequently the scale of the social problem hidden [57, 70]. Feminist hashtag activism opened the door to challenging mainstream narratives about gendered violence online. In these spaces, women have shared their stories and found solidarity [104].

The Me Too movement was an activism movement led by women of color long before (more than 10 years before) it was endorsed by a white celebrity woman. This both increased visibility of the online #MeToo movement, but also nearly erased the history of it; many people believe the movement was started by Milano's tweet. The frequency of storytelling in the #MeToo social media postings by women across geographic areas, races, ethnicities, socioeconomic statuses, and political affiliations shifted the cultural narrative of gendered violence from being an individual problem to being a legitimate social problem. The collective storytelling of the movement has been credited with eliciting collective action and real change offline, ranging from removal of those committing

such misconduct from positions of power in their respective institutions to increased consciousness of and interest in the prevention of misconduct [24].

However, the #MeToo movement has also received criticism for initially ignoring the contributions from women of color in combating gendered violence and the movement not being centered on the marginalized voices that are more vulnerable to sexual harassment and violence, exemplifying how women of color have been historically excluded in feminist movements in the U.S. [19, 61, 80, 94]. For example, the credibility of women of color is discounted in investigations and prosecutions of sexual violence [63, 96], and women of color are subjected not only to gender bias (e.g., myth rapes [81]), but also racial bias (e.g., the intersected racial-gendered stereotype of an "angry black woman") [95].

In turn, scholars have analyzed how demographic groups contributed to the #MeToo movement. PettyJohn et al. (2019) qualitatively examined the content of one day of #HowIWillChange tweets (3,182 tweets), initiated to engage men, during the first week of #MeToo [86]. The researchers found that these tweets either expressed allyship by discussing strategies to dismantle rape culture (e.g., examining their role in toxic masculinty) or resistance, including open hostility (e.g., accusing men that were allies of being weak). Hosterman et al. (2018) examined the types of social support expressed by men and women (identified by manual review of their profile) in 2,782 tweets during the first six months of #MeToo [49]. The researchers found that informational support was the most common type of support expressed by both men and women—but as a percentage of all tweets, men shared slightly more informational support compared to women who shared slightly more emotional support. Modrek and Chakalov (2019) examined the frequency of disclosures of sexual abuse or assault expressed by singular identities (gender and, separately, race/ethnicity that were inferred using commercial marketing software) in 11,935 tweets during the first week of #MeToo [74]. The researchers found that 11% of the tweets disclosed an incidence and that these tweets were overwhelmingly authored by inferred white (89%) and female (90%) posters.

We extend prior work by examining intersectional identities of gender and race/ethnicity—an important lens [94] for viewing this movement—in a large sample (660,237 tweets from 256,650 unique users) of tweets during the first year of #MeToo. Rather than *a priori* examining a specific issue, our work uses topic modeling to broadly examine the content, or storytelling, of the tweets.

2.3 Demographic Inference on Social Media

Numerous recent studies have used publicly available social media data to investigate public health matters [14, 82, 85, 91]. However, social media posts in themselves, including tweets, present little contextualizing information. Thus, to align social media data with traditional data sources, demographic information can be useful (and is often essential for more controlled analyses). Demographics—in our case, gender, race, and ethnicity—have been used in a variety of studies to contextualize online conversations [6, 25, 50, 51, 66, 73]. Prior to the advent of demographic inference, data accessibility of demographics was a major hindrance to demographically-informed social media research [10].

Numerous approaches exist for demographic inference of social media including text-based [22, 26, 58, 72, 89, 102] and image-based approaches [65, 75]. It is widely acknowledged that both of these approaches have inherent biases and limitations. For example, known biases include, but are not limited to, a tendency for image-based approaches to incorrectly classify darker skin tones, especially for women [21], and for text-based approaches to display racial disparities [18]. Methods that are reliant on previous postings or social networks are potentially subject to insufficient historical data or prohibitively expensive calls to an API for a user's historical data [105]. Additionally, all of these approaches limit the inference of gender, race, and ethnicity to a binary classification (e.g., users are either men or women and users are limited to a singular racial/ethnic category).

This oversimplication overlooks the complexity of identities. First, binary classification of race and ethnicity overlooks that multiracial/ethnic identities can be comprised of a primary and secondary identity. Second, these approaches are incapable of identifying non-binary genders. Third, inferred identities may not accurately reflect their self-identified demographics. Finally, identities can be contextual, shifting across situations and settings; these approaches cannot capture that nuance.

We now upfront discuss limitations specific to the demographic inference methods employed in this study (details of the method itself can be found in the Methods section) and the potential impacts to our analysis. First, for gender, we employ a model that examines user names. The model achieves high to reasonable precision and recall for women and men, in part because of its reliance on character-level representations in text, which can identify features that are likely distinctive of gender (e.g., usage of a male vs female emoji). Nonetheless, its training data is limited to binary gender labels, and thus our model will misrepresent users who do not identify in binary terms. Additionally, our model may be biased by cultural naming norms, as its training data over-represents users in the United States. Given the high to reasonable performance, we suspect these biases do not impact the overall results of this study on a large scale.

Second, for race or ethnicity, we employ an ensemble of two models: one based on unigrams of a user's most recent tweets, and the other based on known audience demographics of popular Twitter accounts (e.g., websites like ESPN) that the user follows. This ensemble method results in high performance for white users and moderate performance for people of color. To overcome this lackluster performance, we (1) favor precision (i.e., correctly identifying the race or ethnicity) over recall (i.e., recalling all instances of a race or ethnicity) by (2) only labeling race or ethnicity for a particular user if both of these models are in agreement. Similar to the gender model, examining word-level content does allow us to pick up on key features potentially indicative of race (e.g., skin tone of emoji). We suspect that although we are not recovering every post authored by people of color, we are reasonably sure that our label is correct when we do. Both of the models do not randomly misclassify gender, race, or ethnicity, but instead misclassify on data that are more ambiguous and do not contain key identifying features. We suspect that this correlates with how users' identities are *perceived* on Twitter. For example, usage of skin tone and gendered emojis may also cause a human reader to infer the same identify of a user.

We acknowledge that these limitations and biases make our study imperfect; however, we believe that the ability to understand these patterns at a large-scale level warrants their usage despite these concerns. Our study investigates these conversations across intersectional identities over a longer time period than previous work, allowing for a more nuanced view of how demographic groups engaged in this event of hashtag activism.

3 DATA AND METHODS

3.1 Data

We retrieved one year of public tweets in English that contained the hashtags '#metoo', '#survivor', '#domesticviolence', '#abuse', '#whyididn(o)t', and '#myvoice' (as well as hashtags containing any of those character sequences as sub-strings) from October 1, 2017 to October 28, 2018 from the 1% Streaming Twitter API.[1] These hashtags were selected based on a survey of media reports covering the movement. Our dataset consists of 660,237 tweets from 256,650 unique users.[2]

---

[1] https://developer.twitter.com/en/docs/tweets/sample-realtime/
[2] Extrapolating from the 1% feed, this suggests that there were 66 million English Me Too related tweets in the first year.

3.2 RQ1: Who actively tweeted?

To infer gender (female, male)[3] for the users in our dataset, we employed NeuralGenderDemographer—a convolutional neural network model that classifies the binary gender of the user—from Demographer [105]. The model is trained on the characters of 58,046 Twitter usernames with gender identified through linked blogs. We evaluated the model on a gold-standard dataset consisting of Twitter users who explicitly self-reported their gender in a survey [87], finding the model achieves 90.9% precision and 83.2% recall for women, and 71.4% precision and 83.5% recall for men on this dataset.[4] The benefit of this character-level neural model is that it can learn patterns in character sub-sequences, which is beneficial for Twitter user names that contain special characters (like emojis) that may reveal a user's gender. More specific details of the model can be found in Wood-Doughty et al. (2018) [105].

To infer a *singular* race or ethnicity of the users in our dataset, we employ an ensemble of two methods. The first model, the EthSelfReportNeuralDemographer from Demographer [106], is a regularized logistic regression classifier that uses unigrams from the user's most recent 200 tweets to classify a user's race or ethnicity. The EthSelfReportNeuralDemographer model was trained on an automatically-collected corpus of users who self-reported their race or ethnicity in their Twitter user profile description. The second model, from Culotta et al. (2016) [32], is a regularized logistic regression classifier that infers a user's race or ethnicity by analyzing which popular Twitter accounts that the user follows (i.e., high traffic websites with known measurements of audience demographics that visit those website). Due to limitations in available training data, both models limit classifications to the most populous racial and ethnic groups in the U.S.: Asian, black, Hispanic, and non-Hispanic white. To avoid drawing conclusions from incorrectly-labeled demographics, we favor precision (i.e., correctly identifying the race/ethnicity) over recall (i.e., identifying all instances of a given race/ethnicity) by only labeling the race/ethnicity for users for which our two classifiers agreed; if our two classifiers did not agree, we labeled race/ethnicity as 'unknown'. We evaluated this ensemble method on a dataset of self-reported race/ethnicity data from Twitter users [106] achieving the following precisions and recalls: 20.0% precision and 32.4% recall for Asian users, 50.8% precision and 81.3% recall for black users, 52.2% precision and 19.4% recall for Hispanic users, and 93.3% precision and 87.3% recall for white users.

3.3 RQ2: What stories were shared?

We employed latent Dirichlet allocation (LDA) [17], a probabilistic topic modeling method that infers topics from distributions of words. This method identifies salient topics in a set of documents—in this case, the tweets in our dataset. Topic models have been used to discover salient points of discussion and to delineate different foci of conversations and documents in many domains, including in social media data [14, 15, 48, 60, 83, 88, 107].

We tokenized tweets using TweetTokenizer from NLTK [16], removed capitalization, and removed tweets with three or fewer words. We trained LDA topic models with 20, 30, 40, and 50 topics with a vocabulary consisting of the 20,000 most frequent tokens in our dataset. We ran the models for 2,000 iterations with the hyperparameter defaults defined in Paul and Dredze (2015) [84][5] and averaged the final 100 iterations to estimate the LDA parameters. Upon qualitative review of the topics of each of the topic models by the authors, we selected to move forward with the topic model with 40 topics because it produced the most qualitatively meaningful and cohesive topics.

---

[3]We use male/man and female/woman interchangeably in this paper to refer solely to socially-constructed gender and not physical sex, though we acknowledge that "male" and "female" tend to refer to physical sex in other literature.
[4]Precision refers to the proportion of predictions for a given category that were correct (i.e., exactness/quality), whereas recall refers to the number of members of a given category that were correctly classified (i.e., completeness/quantity).
[5]https://bitbucket.org/adrianbenton/sprite/

Then we assigned final labels for the topics by qualitatively reviewing the word distributions of the topic with randomly selected tweet examples. We excluded topics from our results that did not have any coherent semantic groupings (hence why we present fewer than 40 topics in the results); non-coherent groupings were found by locating topics whose top-10 tokens did not frequently co-occur in the same documents (as measured by per-topic normalized pointwise mutual information).

We describe our results in three ways. First, by the topic labels, words associated with each topic, and probability of each topic's occurrence in the dataset. The probability of a topic was calculated by counting the number of tokens occurring in the tweet dataset which are members of that topic, then normalizing by the number of tokens which occur in *any* topic in the dataset. This method allows us to account for tweets which contain discussions of multiple topics, and ensures that the probability of all topics will sum to 1. Second, by temporally examining the most probable topics during spikes in tweet volume. Third, we compare topics across intersectional demographic groups using both conditional probabilities and pointwise mutual information (PMI). The conditional probability of a topic characterizes the frequency of the topic by demographic group; this is calculated in the same way as overall topic probabilities, but here we restrict the dataset to tweets from a particular demographic group. In contrast, PMI characterizes how distinctive a topic is for each demographic group. Formally, PMI measures the deviation of the joint distribution (x,y) from the separate independent distributions of x and y. Here, $x = p(\text{demographic})$ and $y = p(\text{topic})$, resulting in a metric which roughly captures how distinctive a given topic is of a given demographic group.

$$\text{PMI}(x; y) = \log \frac{p(x, y)}{p(x)p(y)}$$

Finally, we provide modified quotes that exemplify the topics to contextualize the qualitative narrative of selected topics.

3.4 RQ3: What hashtags were associated?

Similarly to our topic analysis, we examine the overall probability of specific hashtags in our dataset. We then analyze and describe which hashtags are most commonly used among each intersectional demographic group using conditional probabilities, and which hashtags are more distinctive of each demographic group using PMI. We also examine the hashtags temporally in a similar nature to how we temporally investigate topics.

4 RESULTS

4.1 RQ1: Who actively tweeted?

Table 1 presents the number of users per-demographic, as well as the mean and standard deviation of tweets per-user across the inferred demographic groups (gender, race, and ethnicity).[6] Approximately 52% of Twitter users are women,[7] whereas women represented 53.32% of users in our dataset. Women also authored significantly more #MeToo tweets per-user (M=2.85, SD=19.01) than men (M=2.38, SD=19.64) ($p < .0001$).

---

[6]11,176 users were classified as unknown by our gender classification model. These tweets were included during topic modeling and overall hashtag probability calculations, but are otherwise excluded from analysis.

[7]As of Jan. 2018, 23% of men and 24% of women in the U.S. use Twitter [92]. The most recent U.S. census states that 50.9% of the U.S. population are women, whereas 49.1% are men [99]. We make a simplifying assumption that Twitter users are either men or women, and use these proportions to calculate that approximately 52% of Twitter users are women and 48% are men.

Table 1. The number of users, percentage of classifiable users, mean authored tweets, and standard deviation of authored tweets for each demographic identity.

Demographic	Users	Percentage	Mean	Stdev
Women	133,717	53.32%	2.85	19.01
Men	119,432	46.68%	2.38	19.64
White	99,601	59.65%	2.90	22.77
Black	42,164	25.25%	2.53	14.98
Asian	16,422	9.84%	2.36	25.36
Hispanic	8,781	5.26%	2.00	8.31
White Women	52,560	31.48%	3.24	21.67
White Men	47,041	28.17%	2.52	23.93
Black Women	22,397	13.41%	2.74	15.64
Black Men	19,767	11.84%	2.28	14.19
Asian Women	8,561	5.12%	2.41	24.49
Asian Men	7,861	4.71%	2.31	26.27
Hispanic Women	5,055	3.03%	2.06	7.48
Hispanic Men	3,726	2.23%	1.93	9.32

In our dataset, the largest racial group is non-Hispanic white, comprising approximately 60% of users inferred; compare to the U.S. population which is comprised of approximately 61% non-Hispanic white individuals.[8] Precision and recall for our ensemble method for the non-Hispanic white group was high (see Methods). Black users comprised 25% of our dataset, followed by Asian (10% of users), and Hispanic (5% of users). Our ensemble method achieves significantly lower precision and recall for people of color; thus, there are likely more users that participated in the #MeToo movement that were not recalled. We opted to favor precision over recall and use an ensemble method where both classifiers had to agree on a label which, in practice, likely translates to labels that would be consistent with how other users perceive racial or ethnic identity. There were significant differences in the number of tweets per-user between racial and ethnic groups ($p < .001$ in a one-way ANOVA test). Specifically, post-hoc analyses using Tukey's HSD [97] indicated that white users authored significantly more #MeToo tweets per-user (M=2.90, SD=22.77) than people of color—Asian (M=2.36, SD=25.36), black (M=2.53, SD=14.98), or Hispanic (M=2.00, SD=8.31) users ($p < .01$). Differences in the number of tweets per-user between Asian, black, and Hispanic groups were not significant ($p > .1$).

Given the differences between the tweeting patterns across gender, race, and ethnicity, it is unsurprising that the number of tweets per-user significantly differed between intersectional demographics as well ($p < .001$ in an ANOVA test). Specifically, the number of tweets per-user was significantly higher for white women (M=3.24, SD=21.67) than any other intersectional demographic ($p < .001$ in a series of Tukey *post-hoc* tests), while differences between any other two intersectional demographics was not significant ($p > .1$ in a series of Tukey *post-hoc* tests). This suggests that, despite the Me Too movement's origins among women of color, the viral Twitter #MeToo movement was primarily represented by white women.

---

[8]https://www.census.gov/quickfacts/fact/table/US/PST045218

Table 2. The topics discovered by topic modeling, probabilities of each topic in the dataset, and most probable tokens associated with each topic.

Topic	Prob.	Representative Tokens	Topic	Prob.	Representative Tokens
Love/support	.044	love good happy thank like family great best	Hollywood accusations/Weinstein	.032	weinstein hollywood harvey star accused alyssamilano porn woman
Trump insults	.043	realdonaldtrump f***ing f*** s*** trumprussia hell lying stupid man	Domestic violence/abuse survivors	.031	abuse violence domesticviolence survivors survivor awareness domestic victims trauma
MAGA	.041	metoo resist maga theresistance fbr resistance realdonaldtrump fbr party qanon trumpshutdown	Rape culture	.030	sexual assault harassment abuse rape victims violence culture predator misconduct
Article/thread sharing	.040	read thread story media article children body metoo piece	Wrestling/sports (unrelated)	.029	survivorseries wwe team raw vs nfl sdlive match players play
Sexual allegations	.039	sexual trump metoo assault women accused everyone telling congress allegations resign guilty	Animals' rights (unrelated)	.029	protectwildlife sign animals trophy animal petition dog hunting stop kill
Women's rights	.039	women rights human right justice due political civil means	Social media	.028	twitter tweet account facebook tweets accounts video followers social posted block
Promoting women's stories	.037	metoo women thank stories survivors speak story forward sharing believe brave support truth	Racism	.028	white black people sarah sanders hate serious racism racist face
U.S. president	.036	trump president america obama country american people history party united states	National security	.027	trump russia putin president security donald north korea saudi national russian
Gender & sexual violence	.035	sexually women assaulted men man mean reminder boys people sexual abused	Elections/voting	.026	vote election gop state democrats republicans house democratic candidate running november voters voting
#MeToo movement	.035	metoo movement women time timesup silence moment campaign taranaburke started hollywood	Politics in the media	.026	news fox cnn john msnbc president press white rally media kazweida fake kelly
Democrat politics	.034	clinton hillary bill obama 2018 bluesea1964 party democrats hillaryclinton	Kavanaugh hearings	.024	kavanaugh ford dr court brett supreme judge senate blasey christine believe believesurvivors senatorcollins fbi hearing jeffflake
Republican politics	.034	realdonaldtrump potus gop foxnews speakerryan tedlieu senatemajldr donaldtrumpjr seanhannity senategop	Police brutality	.021	police state arrested officer killed black woman news video copcrisis
News/media sharing	.033	report video story youtube hero abuser change via	Mueller investigation	.021	trump fbi mueller investigation cohen michael breaking russia campaign manafort comey attorney sessions lawyer doj
Legislative discussion	.033	tax million money pay bill gop cut americans health care ryan billion	Child abuse	.020	children sex child woman abuse rape parents wife kids matter families friend body church girlfriend family forcing
Sharing personal assault stories	.032	years ago old told raped days whyididntreport months man took today said last	School shootings	.015	school gun nra high students shooting moore guns kids parkland violence mass student schools florida teachers davidhogg111

4.2 RQ2: What stories were shared?

4.2.1 Topics in the General Dataset. Table 2 presents the topics identified by the topic model, as well as the probability of occurrence in our dataset and representative tokens for each topic.[9]

---

[9]The 'wrestling/sports' topic is from the 'survivor' hashtags: #SurvivorSeries refers to a wrestling event, and #SurvivorFinale refers to the game show Survivor. The 'Animals' rights' topic results from our search for 'abuse' hashtags, for we obtain #AnimalAbuse and #Abuse (the latter of which is often applied to animal abuse discussions). We removed any tweet in our dataset with the #SurvivorSeries, #SurvivorAU, #SurvivorFinale, and #AnimalAbuse hashtags before running all other quantitative analysis scripts or generating figures.

The topics covered a range of issues including disclosures of sexual harassment and violence, political discourse, supportive discourse, and discussions of public figures involved in high-profile sexual harassment and violence investigations. Below, we describe the ten most probable topics in the general dataset.

The most probable topic in the general dataset was 'love/support'. These tweets often expressed solidarity and emotional support for those who had experienced sexual harassment or violence and were participating in the online movement. Often these tweets were directed towards people who had disclosed an experience of sexual harassment or violence. For example,

> *"love you girl! thanks for sharing #theresistance"*

Of the ten most probable topics, three topics ('Trump insults' (second most probable), 'MAGA' (third most probable), 'U.S. president' (eighth most probable)) focused explicitly on President Trump, each with different emphasis. The 'Trump insults' topic consisted of tweets that primarily expressed criticisms, complaints, and jokes aimed at Trump. For example,

> *"I hate everything about this administration. HATE. EVERYTHING. F*** Trump, and f*** you if you voted for this"*

The 'MAGA' topic featured mentions of President Trump's campaign slogan of Make America Great Again. This topic initially appeared as if it would be a topic opposed to the #MeToo movement, but we found that this abbreviation was often employed subversively and ironically to call attention to a high-profile figure who has been accused, but not investigated, of sexual harassment and assault. For example,

> *"i believe the women that have accused @realDonaldTrump of sexual assault #DumpTrump #MAGA #TheResistance"*

> *"If you really want to #MAGA, start speaking up against these guys who abuse their power and #BelieveSurvivors."*

The 'U.S. president' topic included tweets that expressed discussions of the presidency in general, including discussions of former U.S. presidents (primarily Barack Obama and George W. Bush). For example,

> *"@realDonaldTrump – you are causing the divisiveness. At least when the Bush men were POTUS I respected them because they were doing SOMETHING to move America forward."*

The fourth most probable topic, 'Article/thread sharing', included tweets that primarily focused on providing informational support by sharing data, like statistics on the prevalence of sexual harassment and assault, and news (including reactions to the news). For example,

> *"Janelle Monae delivered an empowering message at the Grammys about the #MeToo movement and #TimesUp campaign. Read the full speech: <link>"*

> *"Important piece on the deep-rooted discrimination that women of color experience from the justice system as #MeToo survivors <link>"*

The fifth most probable topic, 'sexual allegations', included tweets that primarily focused on allegations and investigations of sexual harassment and violence against high-profile individuals. For example,

> *"The 16 women who accused Trump of sexual assault are telling their story in one video—please share!"*

The sixth most probable topic of 'Women's rights' included tweets that focused on historical and current feminist movements including women's suffrage, reproductive rights, and—most relevantly—sexual assault. For example,

> *"ofc the war on women keeps goin. how much longer we gonna have to fight like we did to vote? #believewomen"*

Similar to the most probable topic in the dataset ('love/support'), the seventh most probable topic in the general dataset, 'promoting women's stories', included tweets that explicitly provided emotional support for people who disclosed incidences of sexual harassment or violence either publicly or privately. For example,

> *"To all the women sharing stories of sexual assault and sexual harrassment, you are not alone. Thank you for your bravery. #metoo"*

Finally, the ninth and tenth most probable topics of 'gender and sexual violence' and '#MeToo movement' focused on discussions of gendered violence and how it is reinforced through cultural norms. The former often entailed sharing personal experiences or opinions about the extent and nature of sexual violence against women, as well as broader discussions of the causes that may have led to the #MeToo movement. The latter tended to include direct discussions of developments in the #MeToo movement, including personal #MeToo stories as well as articles specifically about the movement. For example,

> *"fr, stop teaching women to not get assaulted. Teach boys and men not to assault"*

> *"Evangelical women just joined #Metoo, and They're Urging Churches to Address Abuse"*

Although not the most probable topics, notable topics including 'racism' and police brutality' were present. Tweets that contained these topics featured discussions of institutional and systemic racism and bias that contribute to the increased vulnerability of women of color to sexual harassment and violence. These topics highlighted differential racial justice for victims of sexual harassment and violence in our criminal system while contextualizing the #MeToo movement with respect to the disparities in treatment across racial and ethnic identities. For example,

> *"It's cause we got to where we are because men, specifically white men, have been given most of the power in this country."*

> *""Whining" on Twitter helped us with targeted activism on two important fronts: police brutality and sexual assault. Neither would be a conversation without hashtags like #MeToo and #BLM. Bodycams and firing powerful men would not have been possible."*

4.2.2 Variations in Topics Temporally. Figure 1 presents the daily number of tweets by the gender of the tweet author. The start of the viral movement is clear on October 15, 2017. Approximately 68% of all original tweets—not retweets or replies—in the first 48 hours of the movement were authored by women (whereas 53% of all original tweets in the first year of the movement were authored by women.) The most probable topics during initial popularity are 'promoting women's stories' ($P = .077$), 'rape culture' ($P = .075$), 'gender and sexual violence' ($P = .075$), and 'Sexual allegations' ($P = .067$). These topics primarily focus on disclosures of personal experiences of sexual harassment or violence, providing emotional support for those who have experienced sexual harassment or violence, and sharing informational support including data about prevalence of sexual harassment and violence.

After the initial viral popularity of #MeToo on Twitter, tweet volume declined and remained fairly stable at just under 2,000 tweets per day; however, there were smaller spikes in December 2017 after TIME magazine named the #MeToo movement person of the year, and in January 2018 when Catherine Deneuve denounced the movement and soon after when Aziz Ansari was accused of sexual assault. There was a much larger spike in mid-to-late September 2018, when the #MeToo hashtag was revived again during the widely publicized accusations of sexual violence against
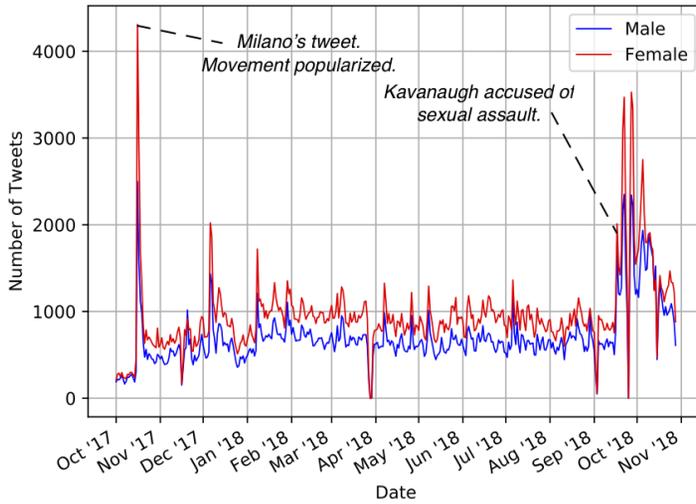
Fig. 1. Daily tweet volume by gender in our dataset.

Judge Brett Kavanaugh.[10] The most probable topics that were discussed during this latter spike (September 15 to October 7, 2018) included 'Kavanaugh hearings', 'Promoting women's stories', ($P = .057$), and 'sharing personal assault stories' ($P = .047$). For example,

> "#WhyIDidntReport The culture that discouraged Christine Blasey Ford from speaking out punishes women their entire careers."

4.2.3   Variations in Topics Across Demographic Groups. Figure 2 presents the conditional probabilities of each topic by demographic group (i.e., the probability of topic occurrence within a demographic group) and Table 3 presents the PMI of each of the topics by each demographic group (i.e., how distinct each topic is for each demographic group, with high PMI indicating the topic is more distinctive of that group). The tweets of all demographic groups discussed a mixture of personal (e.g., 'sharing personal assault stories') and political topics (e.g., 'democrat politics'). We first discuss our observations across race and ethnicity, then across gender, and finally across intersectional identities of gender, race, and ethnicity.

Comparing Topics Across Race and Ethnicity. There were some notable differences between topics expressed in tweets authored by people of color versus white users. For example, tweets authored by white users more often discussed a wider variety of topics (e.g., topics ranging from 'Mueller investigation' to 'sexual allegations/resignation requests' to 'elections/voting') compared to tweets authored by people of color, which generally focused on topics that centered around social support (e.g, on 'promoting women's stories', 'gender and sexual violence', and 'love/support'). Tweets authored by white users were slightly more likely to discuss political topics (e.g., the topics of 'Republican politics' and 'U.S. president') than those authored by people of color, who were slightly more likely to disclose personal incidences of sexual harassment and assault (e.g., 'sharing personal assault stories'), but we reiterate that this difference is relatively small. The topics of 'rape

---

[10]Brett Kavanaugh is a U.S. Supreme Court Justice whose U.S. Senate confirmation hearings took place in September to October 2018. During the hearings, he was accused of a sexual assault that took place while in high school. The issue became of national interest in the wake of the first year of to the #MeToo movement.
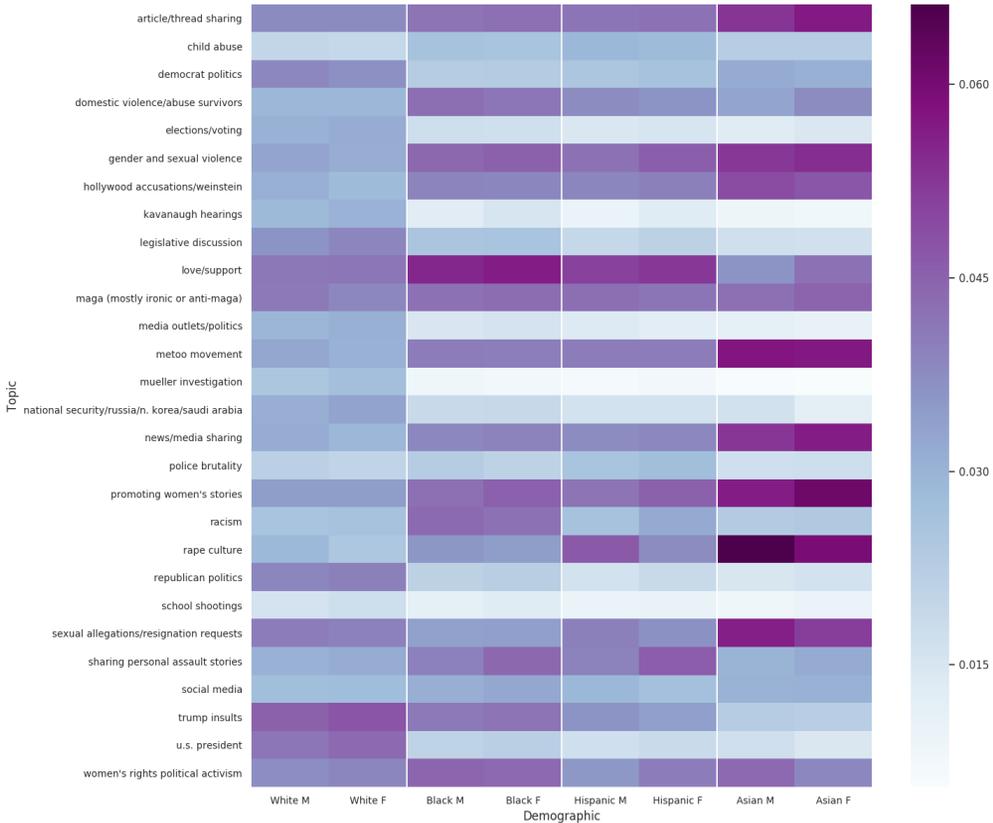
Fig. 2. Conditional probabilities of topic occurrence given intersectional demographics. Demographics are grouped first by race/ethnicity then gender along the x axis, as topics seemed to vary more markedly by race/ethnicity. F = female, M = male.

culture' and 'gender and sexual violence' were more likely among tweets authored by Hispanic and Asian users, specifically those authored by males. For example,

> *"Self respect is sorely needed when you have been sexually violated in our culture. Aggressors notice this weakness"* – Hispanic Male Author

Consistent with the observations above, the topic of 'love/support', which focuses on providing emotional support, is more likely among tweets authored by black users. For example,

> *"#Love is important. We'll all get through #MeToo too"* – Black Male Author

Comparing Topics Across Gender. We now observe topics expressed in tweets across gender. In general, the variety of topics discussed varied more markedly across race and ethnicity than it did across gender. At the surface level, this may be surprising because the #MeToo movement was primarily centered around women's experiences of sexual harassment and violence; thus, we expected a more noticeable divergence between topics discussed in tweets authored by females compared to males. Instead, we observe that the topics discussed in tweets are similar across gender within each race/ethnicity, but that males and females employ different language to discuss these

Table 3. Top five most distinctive topics for each demographic group as measured by pointwise mutual information.

Women			Men		
Black	**PMI**		**Black**	**PMI**	
racism	0.546		racism	0.585	
sharing personal assault stories	0.383		domestic violence/abuse survivors	0.396	
domestic violence/abuse survivors	0.357		child abuse	0.310	
gender and sexual violence	0.313		gender and sexual violence	0.267	
love/support	0.300		love/support	0.253	
White			**White**		
mueller investigation	0.404		mueller investigation	0.281	
kavanaugh hearings	0.367		kavanaugh hearings	0.277	
u.s. president	0.307		elections/voting	0.230	
media outlets/politics	0.289		media outlets/politics	0.228	
national security	0.286		u.s. president	0.225	
Hispanic			**Hispanic**		
sharing personal assault stories	0.446		rape culture	0.553	
child abuse	0.425		child abuse	0.464	
police brutality	0.340		police brutality	0.247	
gender and sexual violence	0.321		sharing personal assault stories	0.219	
rape culture	0.253		hollywood accusations/weinstein	0.213	
Asian			**Asian**		
rape culture	0.945		rape culture	1.089	
news/media sharing	0.731		metoo movement	0.692	
metoo movement	0.689		news/media sharing	0.627	
promoting women's stories	0.689		hollywood accusations/weinstein	0.577	
gender and sexual violence	0.590		promoting women's stories	0.562	

topics (e.g., they often discuss different perspectives within a topic). Consider, for example, the following tweets from the 'Kavanaugh hearings' topic:

> *""I survived. I was fine." This is the mantra of women everywhere #BelieveSurvivors #StopKavanaugh"* – White Female Author

> *"The drinking is important. if Brett was a churchgoer he wouldn't have a rep for binge drinking. character counts!"* – White Male Author

The former tweet (White F) contributes to the topic by sharing the pervasiveness of sexual harassment and violence, whereas the latter (White M) contributes to the topic by discussing the Kavanaugh's character. This exemplifies a noticeable difference that we observed across demographics: male authors tended to contribute to the conversation by framing the issues as a general public or political problem, whereas female authors tended to discuss them by drawing on and sharing their personal experiences.

Comparing Topics Across Intersectional Identities. We now observe topics expressed in tweets across intersectional identities of gender, race, and ethnicity. As shown in Table 3, tweets authored by white users (male and female) were more likely than tweets authored by other identities to relate the #MeToo movement to contemporary social issues and political events, such as school shootings, discussions on the voting process, and the Mueller investigation (as well as more directly related topics, such as the Kavanaugh hearings). In contrast, tweets authored by black females emphasized the broader presence of racial discrimination and differential treatment. For example,

> *"Bill Cosby was always gonna go down as a #MeToo sacrifice. They all guilty but the black men are going to jail"* – Black Female Author

"Are you saying she should put aside her blackness for the sake of the #MeToo movement? Are the 2 exclusive? Really?" – Black Female Author

"I will never understand how we can live in a time of #MeToo, school shootings, our president being a #felon, and there's nothing being done about any of it." – White Female Author

Additionally, tweets authored by black and Hispanic females tended to share personal experiences of sexual harassment and violence more often than their counterparts (white and Asian females). This is evidenced by the relatively high PMI in Table 3 for this topic among that demographic identities.

In general, tweets authored by males were more likely to discuss the accusations of sexual assault against high-profile individuals, such as Harvey Weinstein ('Hollywood accusations/Weinstein' topic). That said, the tweets authored by black males tended to discuss unfair treatment across racial and ethnic groups when they discussed Me Too, similarly to tweets authored by black female users. For example,

"black men always seen as rapists more than anyone else. took a whole damn hashtag to get people to notice that famous white men ain't all innocent" – Black Male Author

In a similar vain, tweets by Hispanic men and women focused slightly more on tying contemporary issues like police brutality against bodies of color with the #MeToo movement. For example,

"When is #MeToo coming for cops? officers keep sexually assaulting undocumented women and punishing people who point out anything wrong" – Hispanic Male Author

Notably we observed less difference in topics expressed in tweets of Asian females and males than other genders. Tweets authored by both Asian females and males expressed unification in topics including tending to share more news articles about Me Too than other demographics. For example,

"Interesting article about all these celebrities "fall from grace" lately: <link>" – Asian Male Author

4.3 RQ3: What hashtags were associated?

Hashtags can function as a link between a tweet and a larger conversation. Thus, in addition to analyzing the topics of the tweets in our datasets, we also investigate which hashtags are used and how they are employed across different demographic identities.

4.3.1 Hashtags in the General Dataset. The most probable hashtag in our dataset was #MeToo and the top 10 next most probable hashtags in our dataset, shown in Figure 3, include (in descending order) #WhyIDidntReport, #Survivor, #BelieveSurvivors, #TimesUp, #abuse, #TheResistance, #DomesticViolence, #Trump #MAGA, and #resist.

The five most probable hashtags in our dataset—#MeToo, #WhyIDidntReport, #Survivor, #BelieveSurvivors, and #TimesUp—each work hand-in-hand with the objectives of the #MeToo movement, anchoring multiple perspectives (#Survivor linking stories of surviving sexual harassment or violence and #WhyIDidntReport linking stories of people who did not report these incidents) and simultaneous movements (#BelieveSurvivors emphasizing historical tendencies for the burden of proof to be upon the victims of sexual harassment and violence and #TimesUp advocating for the end of workplace sexual harassment and violence). Other probable hashtags which are applicable more broadly include #Abuse and #DomesticViolence.
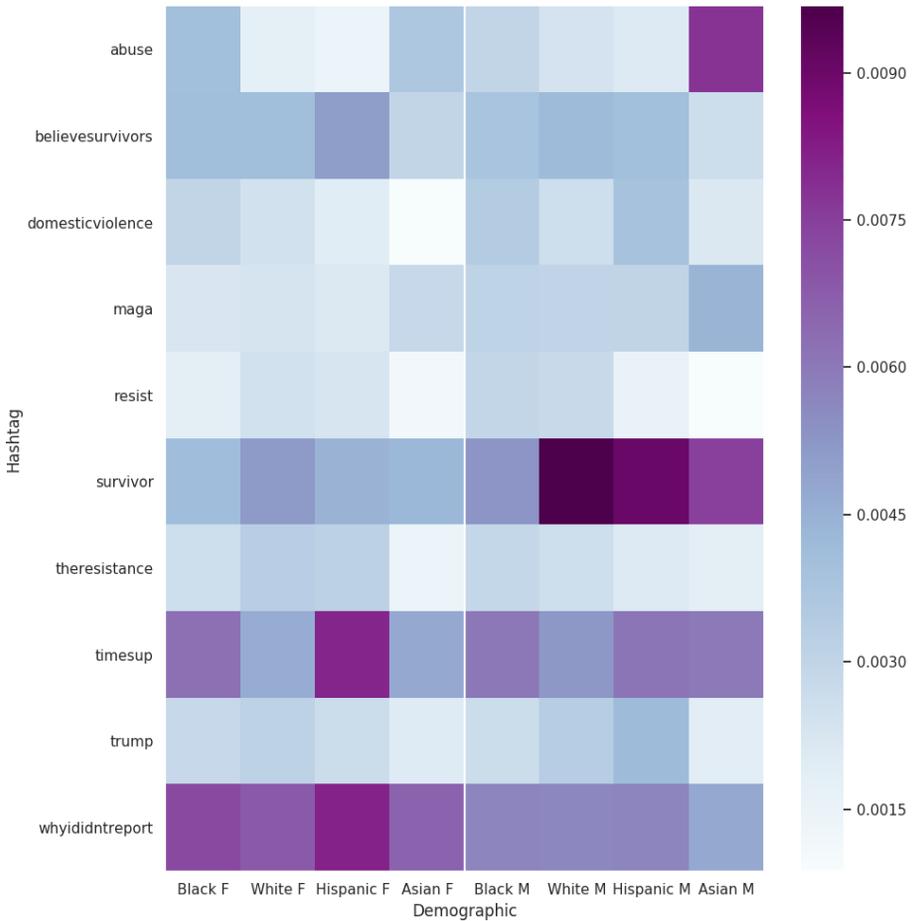
Fig. 3. Conditional probabilities of the top-10 most probable hashtags (excluding #MeToo) by intersectional demographic. Demographics are grouped by gender first and then race and ethnicity becauase hashtags seemed to vary more markedly by gender. F = female, M= male.

Discussions of politics are pervasive in our dataset, and this is reflected in the next most probable hashtags. #Trump and #MAGA are the 8th and 9th most probable hashtags in our dataset, respectively. We found that tweets incorporating this hashtag were frequently anti-Trump and tended to employ MAGA in a subversive way, consistent with our findings about the topic. Tweets containing #Trump and #MAGA often contained hashtags like #TheResistance and #FBRParty (Follow-Back Resistance Party), which are associated with opposition to the presidency. In a random subsample of 100 #MAGA tweets in our dataset, 37 were ironic and/or anti-Trump, and 12 were neutral/unclear.

4.3.2 Variations in Hashtags Temporally. During the initial viral popularity of the #MeToo movement in the first week after October 15, 2017, the most probable hashtags were #MeToo, #NoMore, #Survivor, #IBelieveYou, and #HimThough.

Table 4. Top five most distinctive hashtags of each demographic group as measured by pointwise mutual information.

Women			Men		
Black	PMI		Black	PMI	
EmbracingMyself	2.772		HumanTrafficking	2.490	
Friendship	2.423		Bullying	2.021	
Brutality	2.380		Suicide	1.690	
Law	2.136		StopChildAbuse	1.317	
Humanity	2.089		SexualAbuse	1.206	
White			White		
Leadership	1.242		Pray	1.968	
SexAbuseChat	1.128		Jesus	1.758	
ClimateChange	0.845		God	1.308	
Facebook	0.702		SubstanceAbuse	1.272	
GenderEquality	0.651		Recovery	1.092	
Hispanic			Hispanic		
WCW	1.813		LGBTQ	1.435	
BreakingNews	1.712		WCW	1.304	
BoycottNRA	1.608		Equality	1.282	
NoMore	1.494		IBelieveChristineBlaseyFord	1.259	
Kavanope	1.470		SESTA	1.216	
Asian			Asian		
TakePride	3.388		Brutality	4.155	
VoteThemOut2018	2.269		Friendship	4.018	
Veterans	2.066		Humanity	3.813	
Healing	1.934		HumanRight	3.470	
MosqueMetoo	1.649		Child	3.312	

During the latter spike related to the Kavanaugh hearings (defined here as September 15 to October 7, 2018), the top 5 most probable hashtags were #MeToo, #WhyIDidntReport, #BelieveSurvivors, #Kavanaugh, and #StopKavanaugh. These were thematically similar to the first spike, but with the notable addition of new hashtags (#Kavanaugh and #StopKavanaugh) directly related to the Kavanaugh hearings and hashtags originating (#WhyIDidntReport and #BelieveSurvivors) in response to support Professor Christine Blasey Ford and used broadly to share incidences of sexual harassment and violence that went unreported. As with the first spike, hashtags were generally consistent across demographics—this time, even more so than during the start of the movement. Indeed, none of the demographic groups had different top-5 hashtags during this time period. Thus, this moment was perhaps the most unified with respect to the content of its conversations across demographics, as there was a singular popular event around which users could direct their discussions.

4.3.3 Variations in Hashtags Across Demographic Groups. Figure 3 presents the conditional probabilities of the 10 most probable hashtags in our dataset expressed in tweets by each demographic group, and Table 4 presents the top five distinctive (as measured by PMI) hashtags expressed in tweets for each demographic group.

Comparing Hashtags Across Race and Ethnicity. Unlike topics—which were more consistent across gender, but varied more by race and ethnicity—the usage of the most probable hashtags in our dataset vary more by gender and are more consistent across race and ethnicity. However, we observed some notable variations in usage of hashtags across race and ethnicity. For example, as shown in Figure 3, #TimesUp is likely to be expressed in tweets authored by black and Hispanic users than among Asian or white users, and the same is true to a lesser extent for #WhyIDidntReport.

> *"The #TimesUp and #MeToo movement have given women even bigger and louder platforms to use to speak out about the way we are treated !!!"* – Black Author

> *"I reported it didn't matter I was dismissed & still get blamed #WhyIDidntReport"* – Hispanic Author

Comparing Hashtags Across Gender. As mentioned above, the top hashtags in our dataset were more vary more across gender. For example, as shown in Figure 3, the hashtags #BelieveSurvivors, #TimesUp, and #WhyIDidntReport were more probable among tweets authored by females than males, regardless of their race or ethnicity. Meanwhile, hashtags such as #Trump and #MAGA were more probable among tweets authored by males than females. Note, at the surface-level, #Survivor seems to be more likely to be expressed in tweets authored by males, but this is reflective of a difference in semantics. In a random sample of 100 #Survivor tweets for women, 59 referred to surviving sexual violence and 24 referred to the game show Survivor. In contrast, in a random sample of 100 #Survivor tweets for men, 64 refer to the game show and 22 refer to domestic and/or sexual abuse.

Comparing Hashtags Across Intersectional Identities. As shown in Figure 3, the hashtag #BelieveSurvivors was more probable among tweets authored by Hispanic females than any other demographic group. The hashtag #Abuse was more probable among tweets authored by Asian males primarily because the word "abuse" was often made into a hashtag in the main text of tweets. The higher probability of this hashtag among black and Asian women is for a similar reason. For example,

> *"The #MeToo campaign was so impactful because it applies to #women who shared their stories of #abuse, and so many more."* – Asian Male Author

When observing the most distinctive hashtags of each demographic group (see Table 4), we observe again that tweets authored by black women tended to relate the #MeToo movement to issues that disproportionately affect people of color (e.g., #Brutality is often used in tweets discussing police brutality) as well as emotional support, whether directed towards themselves (e.g., #EmbracingMyself) or others (#Friendship). For example,

> *"Thank you everyone for sharing your stories. Victims need a voice and to be respected! #MeToo #DomesticAbuse #EmbracingMyself"* - Black Female Author

Many of the top hashtags that are most distinctive in tweets authored by black males focused on other contemporary issues that notably affect young adults, including #Bullying, #Suicide, and #StopChildAbuse. For example,

> *"Crazy how many friends I have with #MeToo stories from so young. #StopChildAbuse"* - Black Male Author

Among tweets authored by white users, we observed more distinction in the usage of hashtags across gender. Tweets authored by white females were more likely to relate #MeToo to other contemporary issues, such as #ClimateChange and #GenderEquality. For example,

> *"For real? Adding to the list: #MuellerTime #ClimateChangeIsReal #Resistance #FBR #MyBodyMyChoice"* - White Female Author

Meanwhile, tweets authored by white males were more likely than other intersectional demographics to incorporate religious hashtags (#Pray, #Jesus, and #God). These tweets were not always directly related to Me Too; in fact, in a random subsample of 100 tweets with #Pray, 38 were related to recovery from addiction. These hashtags were sometimes used to express support for victims, and, more infrequently, #Pray was employed ironically.

> *"This #WhyIDidntReport is heartbreaking. Will #pray for victims who seek justice"* - White Male Author

> *"Gonna think and #Pray for #Kavanaugh during these challenging times. We won't miss you #IBelieveChristineBlaseyFord"* - White Male Author

Among tweets authored by Hispanic men and women, the most distinctive hashtags were all associated with current events and political developments.[11] The #SESTA hashtag refers to the Stop Enabling Sex Traffickers Act, which users in our dataset tended to be against. For example,

> *"we can't fix all the problems that led to #MeToo without treating our sex workers better. #Resist #SESTA"* – Hispanic Female Author

Many of the other hashtags with high PMI expressed in the tweets of Hispanic users were employed in the strings of tweets which expressed dissatisfaction about the current political administration. For example,

> *"until @realDonaldTrump is voted out I am Puerto Rican, I am a woman, I am #LGBTQ, I am black, I am #metoo, I am our environment."* - Hispanic Female Author

Finally, tweets authored by Asian females and males were more likely to use hashtags referring to contemporary political issues and discussions, such as #Brutality and #VoteThemOut2018.

5 DISCUSSION

In our analysis of the first year of the #Metoo movement on Twitter, we found notable differences in representation and storytelling across demographic groups. Below we contextualize these differences in the larger literature centered on representation, disclosure, and visibility on social media across gender, race, and ethnicity.

In our dataset, we found that the number of men and women actively tweeting in #MeToo were nearly equal. However, women authored significantly more tweets, especially during the first week of the movement. In general, white users authored significantly more tweets per-user and tweeted more compared to Asian, black, and Hispanic users. Indeed, Modrek and Chakalov (2019) [74] found that disclosures during the first week of #MeToo were primarily from tweets authored by white and female users. Many scholars have been critical that the #MeToo exemplified historical exclusion of women of color from feminist movements and our results support that despite the Me Too movement's origin to support women of color, the intersectional identity that tweeted the most was white female users.

The most probable topics expressed in our dataset reflected emotional support (e.g., expressing support for those who have experienced incidence of sexual harassment and violence), discussions centered on politics, and informational support (e.g., sharing news stories that highlight the pervasiveness of sexual harassment and violence). Other studies have highlighted that #MeToo focused on accusations of sexual misconduct by high-profile figures [8] and we observe supporting evidence of this. #MeToo peaked during the initial week and then again during the accusations of sexual violence perpetrated by Judge Kavanaugh. During the latter, new hashtag activism movements were born including #WhyIDidntReport and #BelieveSurvivors.

Consistent with Hosterman et al. (2018) [49], we also observe differential support strategies across gender: tweets authored by female users were more likely to express emotional support whereas tweets authored by male users were more likely to express informational support. These differential support strategies were also observed in the operationalization of hashtags. For example, tweets authored by female users were more likely to reflect emotional support and disclosure (#BelieveSurvivors, #WhyIDidntReport) whereas tweets authored by men were more likely to reflect contemporary politics using #MAGA and referring to #Trump both critically and supportively,

---

[11]Except #WCW, which was generally used in long strings of hashtags in tweets that did not consist of any non-hashtag text.

similar to to the findings of Vigil-Hayes et al. (2019) [101]. For example, tweets may subversively use #MAGA to express allyship or to dismiss #MeToo by perpetuating rape myths, similar to [40]'s findings. We also observed that tweets authored by black users were more likely to express emotional support than other racial/ethnic groups and distinctively use hashtags to reflect this (#EmbracingMyself, #Friendship).

Unlike Modrek and Chakalove (2019) [74] who found that tweets authored by white people and women had a higher frequency of disclosure, our dataset reflects that disclosure is a more distinctive topic among black and Hispanic women. The observation that women in general are more likely to disclose personal information on Twitter is consistent with Walton and Rice (2013) [103] and previous works finding that women were approximately twice as likely as men to disclose sexual violence online [7, 62].

Finally, there was distinctive co-opting of #MeToo among intersectional identities. Tweets authored by white female and male users were both more likely to include contemporary social and political issues. For example, tweets discussing #ClimateChange were more distinctive among white female users and tweets discussing religion (#Pray, #Jesus, #God) were more distinctive among white male users. In comparison, tweets authored by black female and male users were both more likely to include discussions on race-based discriminatory behaviors and differential treatment in the justice system. This is also reflected in the hashtags (#Brutality, #Humanity) that are distinctive among black women users. Similarly, tweets authored by Hispanic female and male users distinctively discussed police brutality. Both of these are likely a factor of historical and ongoing discriminatory practices and dissatisfaction with the police and the justice system because of inequitable treatment [1–4, 36, 37, 42, 80]. Less is known about how differing races and ethnicities employ Twitter for activism; however, a number of studies have examined how "Black Twitter" is used as a digital counterpublic by black people to comment about contemporary events to build collective framing [29, 43, 47], and our results exemplify how minority races and ethnicities co-opt the #MeToo movement to build on a historical and ongoing counter-narrative.

5.1 Implications for Digital Activism Research and Design

First, we discuss a practical implication of our work to the research community. Our study employed topic modeling to identify emergent topics in our dataset. At the surface-level, it appears that intersectional identities discuss similar topics. For example, there appears to be a convergence of the topics across tweets authored by males and females; however, the hashtags show a divergence. Through a qualitative review of the topics and associated tweets, we find that although users may employ the same language (hence appearing to discuss the same topic in the same way), the context of their tweet indicates differential framing of the issue at-hand. For example, tweets by female users tended to draw on their personal experiences or share others' personal experiences, thus highlighting the pervasiveness of sexual harassment and violence. Meanwhile, tweets by male users tended to discuss sexual harassment and violence in the context of politics, or as a problem that the general public should tackle. This implies that consideration of semantics is important for the framing of digital activism issues and that automated methods that can explore semantics, such as those employed by Chancellor et al. (2018) [28], may be more appropriate to disambiguate framing across groups of activists.

Second, we discuss the implications of our work to the design community to potentially raise marginalized voices participating in digital activism. One of the tenets of digital activism is the desire to democratize narratives and shape collective action. The ability of social media to support a public sphere, defined as "a constellation of communicative spaces in society that permit the circulation of information, ideas, debates, ideally in an unfettered manner, and also the formation of political will" [33] is contested [30]. "Black Twitter" possesses characteristics of a digital counterpublic

(i.e., a place where subordinated social groups can discuss alternative publics [38]) that has led to a number of influential hashtag activism movements [43] and enables new forms of activism [47]. However, previous research has shown a 'glass ceiling effect' whereby certain demographics receive more visibility. For example, in general, white males have more followers contributing to more visibility and more interaction elevating their tweets in the feed [69]. In contrast, Asian and black females have the lowest visibility. Trending topics often suffer from under-representation, persistently from middle-aged black females [27]. Ensuring that topics across demographic groups are seen is critical to forming a collective framing for digital activism.

In our study, we find that sheer volume of tweets authored by white users are over-represented and the topics that they discuss potentially obscure topics important to the framing of #MeToo, such as racial inequities. This may prohibit the objective of collective framing. Moreover, according to the Social Identity Model of Collective Action, collective efficacy and social identity are important for predicting collective action [100]. Designers could support those with marginalized voices by developing tools that allow users to leverage their social networks to promote visibility of their tweets. For example, Bisafar (2018) outlines a computing tool that provides informed views of the social capital (i.e., those with influence) of a user's network and allows users to have a birds-eye view of their impact from doing so (i.e., viewing aggregate interactions) [52]. In turn, leveraging social capital to increase influence may help users coalesce a collective identity and framing, ultimately translating to focused collective action [12, 56].

5.2 Limitations and Future Work

Our goal was to explore representation and storytelling across identities in the #MeToo movement. Our analysis has several limitations, which suggest areas for future work.

First, our study focuses on approximating large-scale representation of binary gender (female, male) and the most populous races and ethnicities in the U.S. (Asian, black, Hispanic, white). We acknowledge that is an oversimplification of these identities that overlooks non-binary genders, multiracial and multiethnic identities, and less populous racial and ethnic identities. For example, our study does not include non-binary or transgender identities, which is important given that people who hold these identities are subjected to sexual harassment and violence at disproportionately higher rates [44, 54], and given the criticism that the #MeToo movement included exclusionary language targeting cis-women [23, 94].

Second, numerous identities beyond gender, race, and ethnicity play a role in how people elect to use collaborative technologies [59] and even participate in digital activism. We do not consider other identities, like sexual orientation or socioeconomic and immigration status, for example. Socioeconomic and immigration statuses play a critical role in one's sense of agency and ability to report incidences of harassment in the workplace [80]. Most #MeToo analyses are English-centric (including ours) or U.S.-centric [64], although it is important to consider the varying cultural contexts under which sexual harassment and assault occurs [34].

Third, topic models only locate wide-scale patterns in themes of social media content; therefore, this method may not recover marginalized conversations that may happen among a subset of the communities our dataset. Additionally, as shown in our results, although some demographic groups discussed the same topic, further qualitative examination showed that despite using similar language, the expressions of the topics were semantically different. We suggest future work should focus on methods that can capture this variation [28], especially when comparing discourse across different identities.

The above limitations suggest that other participatory research methods like participatory design or in-depth interviews may be more apt for studying how these identities participate(d) in digital activism and the barriers that prevent them from safely doing so.

Finally, we limit our analysis to the first year of the #MeToo movement. Although our study encapsulates a significantly longer study period than previous works, future research should nonetheless examine how hashtags are revived or co-opted (or not) over even longer time periods.

6 CONCLUSIONS

The design of collaborative communication technologies emphasizes the ability of storytelling to aid our understanding of users' perspectives, especially those outside of one's identity [76]. Some argue that hashtag activism cannot effect real change while others argue that measuring change is difficult to impossible; however, one can infer metrics that indicate growing support [68]. With respect to the #MeToo movement, it is clear in the frequency, content, and participation of postings in this study that the movement was able to grow tremendous social support online. In our study, we find both supportive and critical voices of #MeToo, with most voicing support. However, we also find that under-represented intersectional groups are using #MeToo to also call attention to other interconnected historical and ongoing movements. Future work in social-media-based movements could observe these under-represented narratives and find ways to ensure that they can contribute democratically to digital activism.

ACKNOWLEDGMENTS

REFERENCES

[1] [n.d.]. Reasonable Woman. https://law.jrank.org/pages/9668/Reasonable-Woman.html
[2] 1988. *Lipsett v. Univ. of Puerto Rico, 864 F.2d. 881 (1st Cir.).* https://law.justia.com/cases/federal/appellate-courts/F2/864/881/239855/
[3] 1991. *Ellison v. Brady, 924 F.2d 872 (9th Cir.).* https://law.justia.com/cases/federal/appellate-courts/F2/924/872/224242/
[4] Kathryn Abrams. 1989. Gender discrimination and the transformation of workplace norms. *Vand. L. Rev.* 42 (1989), 1183.
[5] Ramona Alaggia and Susan Wang. 2020. "I never told anyone until the #MeToo movement": What can we learn from sexual abuse and sexual assault disclosures made through social media? *Child Abuse & Neglect* 103 (2020), 104312. https://doi.org/10.1016/j.chiabu.2019.104312
[6] Silvio Amir, Mark Dredze, and John W. Ayers. 2019. Population Level Mental Health Surveillance over Social Media with Digital Cohorts. In *Workshop on Computational Linguistics and Clinical Psychology (CLPsych)*.
[7] Nazanin Andalibi, Oliver L. Haimson, Munmun De Choudhury, and Andrea Forte. 2016. Understanding Social Media Disclosures of Sexual Abuse Through the Lenses of Support Seeking and Anonymity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3906–3918. https://doi.org/10.1145/2858036.2858096
[8] Monica Anderson and Skye Toor. 2018. How social media users have discussed sexual harassment since #MeToo went viral. *Pew Research Center* (October 2018). https://www.pewresearch.org/fact-tank/2018/10/11/how-social-media-users-have-discussed-sexual-harassment-since-metoo-went-viral/
[9] Mariam Asad and Christopher A. Le Dantec. 2015. Illegitimate Civic Participation: Supporting Community Activists on the Ground. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1694–1703. https://doi.org/10.1145/2675133.2675156
[10] John W Ayers, Benjamin M Althouse, and Mark Dredze. 2014. Could behavioral medicine lead the web data revolution? *Journal of the American Medical Association (JAMA)* 311, 14 (2014), 1399–1400.
[11] John W Ayers, Theodore L Caputi, Camille Nebeker, and Mark Dredze. 2018. Don't quote me: reverse identification of research participants in social media studies. *NPJ Digital Medicine* 1, 30 (2018). https://doi.org/10.1038/s41746-018-0036-2

[12] Robert D Benford and David A Snow. 2000. Framing processes and social movements: An overview and assessment. *Annual Review of Sociology* 26, 1 (2000), 611–639.

[13] Adrian Benton, Glen A Coppersmith, and Mark Dredze. 2017. Ethical Research Protocols for Social Media Health Research. In *Workshop on Ethics in NLP*. 94–102.

[14] Adrian Benton, Braden Hancock, Glen Coppersmith, John W. Ayers, and Mark Dredze. 2016. After Sandy Hook Elementary: A Year in the Gun Control Debate on Twitter. *CoRR* abs/1610.02060 (2016). arXiv:1610.02060 http://arxiv.org/abs/1610.02060

[15] Adrian Benton, Michael J. Paul, Braden Hancock, and Mark Dredze. 2016. Collective Supervision of Topic Models for Predicting Surveys with Social Media. In *Association for the Advancement of Artificial Intelligence (AAAI)* (Phoenix, Arizona). 2892–2898. http://dl.acm.org/citation.cfm?id=3016100.3016307

[16] Steven Bird, Edward Loper, and Ewan Klein. 2009. *Natural Language Processing with Python*. USA.

[17] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3 (March 2003), 993–1022. http://dl.acm.org/citation.cfm?id=944919.944937

[18] Su Lin Blodgett and Brendan O'Connor. 2017. Racial disparity in natural language processing: A case study of social media african-american english. *CoRR* abs/1707.00061 (2017). arXiv:1707.00061 http://arxiv.org/abs/1707.00061

[19] Jamillah Bowman Williams. 2020. Maximizing #MeToo: Intersectionality and the movement. (2020).

[20] Axel Bruns, Tim Highfield, and Jean Burgess. 2013. The Arab Spring and Social Media Audiences: English and Arabic Twitter Users and Their Networks. *American Behavioral Scientist* 57, 7 (2013), 871–898. https://doi.org/10.1177/0002764213479374 arXiv:https://doi.org/10.1177/0002764213479374

[21] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on Fairness, Accountability and Transparency*. 77–91.

[22] John D Burger, John Henderson, George Kim, and Guido Zarrella. 2011. Discriminating gender on Twitter. In *Empirical Methods in Natural Language Processing (EMNLP)*.

[23] Tarana Burke. 2017. #MeToo was started for black and brown women and girls. They're still being ignored. *The Washington Post* (November 2017). https://www.washingtonpost.com/news/post-nation%E2%80%8C/wp%E2%80%8C/2017/11/09/the-waitress-who-works-in-the-diner-needs-to-know-that-the-issue-of

[24] Theodore L Caputi, Alicia L Nobles, and John W Ayers. 2019. Internet Searches for Sexual Harassment and Assault, Reporting, and Training Since the #MeToo Movement. *JAMA Internal Medicine* 179, 2 (02 2019), 258–259. https://doi.org/10.1001/jamainternmed.2018.5094

[25] Jordan Carpenter, Daniel Preotiuc-Pietro, Lucie Flekova, Salvatore Giorgi, Courtney Hagan, Margaret L Kern, Anneke EK Buffone, Lyle Ungar, and Martin EP Seligman. 2017. Real men don't say "cute": Using automatic language analysis to isolate inaccurate aspects of stereotypes. *Social Psychological and Personality Science* 8, 3 (2017), 310–322.

[26] Nina Cesare, Christan Grant, and Elaine O Nsoesie. 2017. Detection of user demographics on social media: A review of methods and recommendations for best practices. *CoRR* abs/1702.01807 (2017).

[27] Abhijnan Chakraborty, Johnnatan Messias, Fabricio Benevenuto, Saptarshi Ghosh, Niloy Ganguly, and Krishna P Gummadi. 2017. Who makes trends? Understanding demographic biases in crowdsourced recommendations. *CoRR* abs/1704.00139 (2017). arXiv:1704.00139 http://arxiv.org/abs/1704.00139

[28] Stevie Chancellor, Andrea Hu, and Munmun De Choudhury. 2018. Norms matter: Contrasting social support around behavior change in online weight loss communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174240

[29] Meredith Clark. 2014. To tweet our own cause: A mixed-methods study of the online phenomenon "Black Twitter". (2014).

[30] Elanor Colleoni, Alessandro Rozza, and Adam Arvidsson. 2014. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication* 64, 2 (2014), 317–332.

[31] Kimberle Crenshaw. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *The University of Chicago Legal Forum* 140 (1989), 139–167.

[32] Aron Culotta, Nirmal Kumar Ravi, and Jennifer Cutler. 2016. Predicting Twitter user demographics using distant supervision from website traffic data. *Journal of Artificial Intelligence Research* 55, 1 (Jan. 2016), 389–408. http://dl.acm.org/citation.cfm?id=3013558.3013570

[33] Peter Dahlgren. 2005. The Internet, public spheres, and political communication: Dispersion and deliberation. *Political Communication* 22, 2 (2005), 147–162.

[34] Elizabeth Dartnall and Rachel Jewkes. 2013. Sexual violence against women: The scope of the problem. *Best Practice & Research: Clinical Obstetrics & Gynaecology* 27, 1 (2013), 3–13.

[35] Jill P. Dimond, Michaelanne Dye, Daphne Larose, and Amy S. Bruckman. 2013. Hollaback! The role of storytelling online in a social movement organization. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 477–490. https://doi.org/10.1145/2441776.2441831

[36] Kenneth Dowler and Raymond Sparks. 2008. Victimization, contact with police, and neighborhood conditions: Reconsidering African American and Hispanic attitudes toward the police. *Police Practice and Research* 9, 5 (2008), 395–415. https://doi.org/10.1080/15614260801980760 arXiv:https://doi.org/10.1080/15614260801980760

[37] Nancy S Ehrenreich. 1989. Pluralist myths and powerless men: The ideology of reasonableness in sexual harassment law. *The Yale Law Journal Forum* 99 (1989), 1177.

[38] Nancy Fraser. 1990. Rethinking the public sphere: A contribution to the critique of actually existing democracy. *Social Text* 25/26 (1990), 56–80.

[39] Deen Freelon, Charlton D. McIlwain, and Meredith D. Clark. 2016. *Beyond the hashtags: #Ferguson, #Blacklivesmatter, and the online struggle for offline justice*. Technical Report. Center for Media and Social Impact. https://cmsimpact.org/resource/beyond-hashtags-ferguson-blacklivesmatter-online-struggle-offline-justice

[40] Ryan J Gallagher, Elizabeth Stowell, Andrea G Parker, and Brooke Foucault Welles. 2019. Reclaiming stigmatized narratives: The networked disclosure landscape of #MeToo. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW, Article 96 (Nov. 2019), 30 pages. https://doi.org/10.1145/3359198

[41] Benjamin Gleason. 2013. #Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter. *American Behavioral Scientist* 57, 7 (2013), 966–982. https://doi.org/10.1177/0002764213479372 arXiv:https://doi.org/10.1177/0002764213479372

[42] Jennifer M Gómez and Robyn L Gobin. 2020. Black women and girls & #MeToo: Rape, cultural betrayal, & healing. *Sex Roles* 82, 1-2 (2020), 1–12.

[43] Roderick Graham and 'Shawn Smith. 2016. The Content of Our #Characters: Black Twitter as Counterpublic. *Sociology of Race and Ethnicity* 2, 4 (2016), 433–449. https://doi.org/10.1177/2332649216639067 arXiv:https://doi.org/10.1177/2332649216639067

[44] Jaime M Grant, Lisa A Mottet, Justin Tanis, Jack Harrison, Jody L Herman, and Mara Keisling. 2011. *Injustice at Every Turn: A Report of the National Transgender Discrimination Survey*. Technical Report. National Center for Transgender Equality, National Gay and Lesbian Task Force. https://transequality.org/sites/default/files/docs/resources/NTDS_Report.pdf

[45] Cristela Guerra. 2017. Where'd the #MeToo initiative really come from? Activist Tarana Burke, long before hashtags. *Boston Globe* (October 2017). https://www.bostonglobe.com/lifestyle/2017/10/17/alyssa-milano-credits-activist-tarana-burke-with-founding-metoo-movement-years-ago/o2Jv29v6ljObkKPTPB9KGP/story.html

[46] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. 2018. Gender recognition or gender reductionism? The social implications of embedded gender recognition systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173582

[47] Marc Lamont Hill. 2018. "Thank You, Black Twitter": State violence, digital counterpublics, and pedagogies of resistance. *Urban Education* 53, 2 (2018), 286–302. https://doi.org/10.1177/0042085917747124 arXiv:https://doi.org/10.1177/0042085917747124

[48] Liangjie Hong and Brian D Davison. 2010. Empirical study of topic modeling in Twitter. In *SOMA Workshop on social media analytics*.

[49] Alec R Hosterman, Naomi R Johnson, Ryan Stouffer, and Steven Herring. 2018. Twitter, social support messages, and the #metoo movement. *The Journal of Social Media in Society* 7, 2 (2018), 69–91.

[50] Xiaolei Huang, Michael C Smith, Amelia M Jamison, David A Broniatowski, Mark Dredze, Sandra Crouse Quinn, Justin Cai, and Michael J Paul. 2019. Can online self-reports assist in real-time identification of influenza vaccination uptake? A cross-sectional study of influenza vaccine-related tweets in the USA, 2013–2017. *BMJ open* 9, 1 (2019), e024018.

[51] Xiaolei Huang, Michael C Smith, Michael J Paul, Dmytro Ryzhkov, Sandra C Quinn, David A Broniatowski, and Mark Dredze. 2017. Examining patterns of influenza vaccination in social media. In *Workshop on Health Intelligence*.

[52] Farnaz Irannejad Bisafar. 2018. Catalyzing youth civic engagement through innovations in social computing. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Jersey City, NJ, USA) *(CSCW '18)*. Association for Computing Machinery, New York, NY, USA, 89–92. https://doi.org/10.1145/3272973.3272981

[53] Sarah J. Jackson, Moya Bailey, and Brooke Foucault Welles. 2020. *#HashtagActivism: Networks of Race and Gender Justice*. The MIT Press. https://doi.org/10.7551/mitpress/10858.001.0001

[54] Sandy E James, Jody L Herman, Susan Rankin, Mara Keisling, Lisa Mottet, and Ma'ayan Anafi. 2016. *The Report of the 2015 U.S. Transgender Survey*. Technical Report. National Center for Transgender Equality. https://www.

transequality.org/sites/default/files/docs/USTS-Full-Report-FINAL.PDF

[55] Holly Kearl. 2018. *The Facts Behind the #MeToo Movement: A National Study on Sexual Harassment and Assault*. Technical Report. Stop Street Harassment, GfK, Raliance, University of California San Diego Center on Gender Equity and Health. http://www.stopstreetharassment.org/wp-content/uploads/2018/01/Full-Report-2018-National-Study-on-Sexual-Harassment-and-Assault.pdf

[56] R Kelly Garrett. 2006. Protest in an information society: A review of literature on social movements and new ICTs. *Information, Communication & Society* 9, 02 (2006), 202–224.

[57] Christopher Kilmartin and Alan D. Berkowitz. 2001. *Sexual assault in context: Teaching college men about gender*. Mahwah, NJ: Psychology Press, Taylor and Francis Group.

[58] Rebecca Knowles, Josh Carroll, and Mark Dredze. 2016. Demographer: Extremely Simple Name Demographics. In *Natural Language Processing and Computational Social Science (NLP+CSS)* (Austin, Texas). http://aclweb.org/anthology/W16-5614

[59] Neha Kumar, Daniel A. Epstein, Catherine D'Ignazio, Amanda Lazar, Andrea Parker, Muge Haseki, and Anupriya Tuli. 2019. Women's Health, Wellbeing, & Empowerment. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (Austin, TX, USA) *(CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 116–121. https://doi.org/10.1145/3311957.3358606

[60] Jey Han Lau, Nigel Collier, and Timothy Baldwin. 2012. On-line trend analysis with topic models:#twitter trends detection topic model online. *COLING* (2012), 1519–1534.

[61] Rebecca Leung and Robert Williams. 2019. #MeToo and intersectionality: An examination of the #MeToo movement through the R. Kelly scandal. *Journal of Communication Inquiry* 43, 4 (2019), 349–371.

[62] Tetyana Lokot. 2018. #IAmNotAfraidToSayIt: Stories of sexual violence as everyday political speech on Facebook. *Information, Communication & Society* 21, 6 (2018), 802–817. https://doi.org/10.1080/1369118X.2018.1430161 arXiv:https://doi.org/10.1080/1369118X.2018.1430161

[63] Kimberly A. Lonsway and Joanne Archambault. 2012. The "justice gap" for sexual assault cases: Future directions for research and reform. *Violence Against Women* 18, 2 (2012), 145–168. https://doi.org/10.1177/1077801212440017 arXiv:https://doi.org/10.1177/1077801212440017 PMID: 22433226.

[64] Ritty Lukose. 2018. Decolonizing feminism in the #MeToo era. *The Cambridge Journal of Anthropology* 36, 2 (2018), 34–52.

[65] Erno Makinen and Roope Raisamo. 2008. Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 3 (2008), 541–547.

[66] Benjamin Mandel, Aron Culotta, John Boulahanis, Danielle Stark, Bonnie Lewis, and Jeremy Rodrigue. 2012. A demographic analysis of online sentiment during hurricane irene. In *Workshop on Language in Social Media*. 27–36.

[67] Lydia Manikonda, Ghazaleh Beigi, Huan Liu, and Subbarao Kambhampati. 2018. Twitter for sparking a movement, Reddit for sharing the moment: #MeToo through the lens of social media. *CoRR* abs/1803.08022 (2018). arXiv:1803.08022 http://arxiv.org/abs/1803.08022

[68] Dennis McCafferty. 2011. Activism vs. Slacktivism. *Communications of the Association for Computing Machinery* 54, 12 (Dec. 2011), 17–19. https://doi.org/10.1145/2043174.2043182

[69] Johnnatan Messias, Pantelis Vikatos, and Fabrício Benevenuto. 2017. White, man, and highly followed: Gender and race inequalities in Twitter. In *Proceedings of the International Conference on Web Intelligence* (Leipzig, Germany) *(WI '17)*. Association for Computing Machinery, New York, NY, USA, 266–274. https://doi.org/10.1145/3106426.3106472

[70] Marian Meyers. 1996. *News Coverage of Violence against Women: Engendering Blame*. Sage Publications, Inc.

[71] Ranit Mishori, Lisa Singh, Kenneth W. Lin, and Yifang Wei. 2019. #Diversity: Conversations on Twitter about Women and Black Men in Medicine. *American Board of Family Med* 32, 1 (2019), 28–36. https://doi.org/10.3122/jabfm.2019.01.180175

[72] Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and J Niels Rosenquist. 2011. Understanding the demographics of Twitter users. In *ICWSM*.

[73] Lewis Mitchell, Morgan R Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M Danforth. 2013. The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLOS One* 8, 5 (2013), e64417.

[74] Sepideh Modrek and Bozhidar Chakalov. 2019. The #MeToo movement in the United States: Text analysis of early Twitter conversations. *Journal of Medical Internet Research* 21, 9 (03 Sep 2019), e13837. https://doi.org/10.2196/13837

[75] Choon-Boon Ng, Yong-Haur Tay, and Bok-Min Goi. 2015. A review of facial gender recognition. *Pattern Analysis and Applications* 18, 4 (2015), 739–755.

[76] Ihudiya Finda Ogbonnaya-Ogburu, Angela D.R. Smith, Alexandra To, and Kentaro Toyama. 2020. Critical race theory for HCI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376392

[77] Abby Ohlheiser. 2017. #MeToo made the scale of sexual abuse go viral. But is it asking too much of survivors? *The Washington Post* (October 2017). https://www.washingtonpost.com/news/the-intersect/wp/2017/10/16/metoo-made-the-scale-of-sexual-abuse-go-viral-but-is-it-asking-too-much-of-survivors/

[78] Abby Ohlheiser. 2017. The woman behind 'Me Too' knew the power of the phrase when she created it – 10 years ago. *The Washington Post* (October 2017). https://www.washingtonpost.com/news/the-intersect/wp/2017/10/19/the-woman-behind-me-too-knew-the-power-of-the-phrase-when-she-created-it-10-years-ago/

[79] Alexandra Olteanu, Ingmar Weber, and Daniel Gatica-Perez. 2016. Characterizing the demographics behind the #BlackLivesMatter movement. In *Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposium Series.*

[80] Angela Onwuachi-Willig. 2018. What about #UsToo: The invisibility of race in the #MeToo movement. *The Yale Law Journal Forum* 128 (2018), 105–120. https://www.yalelawjournal.org/forum/what-about-ustoo

[81] Amy Dellinger Page. 2010. True colors: Police officers and rape myth acceptance. *Feminist Criminology* 5, 4 (2010), 315–334. https://doi.org/10.1177/1557085110384108 arXiv:https://doi.org/10.1177/1557085110384108

[82] Michael J Paul and Mark Dredze. 2011. You are what you tweet: Analyzing twitter for public health. In *Fifth International AAAI Conference on Weblogs and Social Media.*

[83] Michael J Paul and Mark Dredze. 2014. Discovering health topics in social media using topic models. *PLOS One* 9, 8 (2014), e103408.

[84] Michael J Paul and Mark Dredze. 2015. SPRITE: Generalizing topic models with structured priors. *Transactions of the Association for Computational Linguistics (TACL)* 3 (2015), 43–57. http://aclweb.org/anthology/Q15-1004

[85] Michael J Paul and Mark Dredze. 2017. Social monitoring for public health. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 9, 5 (2017), 1–183.

[86] Morgan E PettyJohn, Finneran K Muzzey, Megan K Maas, and Heather L McCauley. 2019. #HowIWillChange: Engaging men and boys in the #MeToo movement. *Psychology of Men & Masculinities* 20, 4 (2019), 612.

[87] Daniel Preoţiuc-Pietro, Svitlana Volkova, Vasileios Lampos, Yoram Bachrach, and Nikolaos Aletras. 2015. Studying user income through language, behaviour and affect in social media. *PLOS One* 10, 9 (2015), e0138717.

[88] Daniel Ramage, Susan Dumais, and Dan Liebling. 2010. Characterizing microblogs with topic models. In *ICWSM.*

[89] Delip Rao and David Yarowsky. 2010. Detecting latent user properties in social media. In *In Proceedings of the Neural Information Processings Systems (NeurIPS) MLSN Workshop.* Citeseer.

[90] Ellen Simpson. 2018. Integrated & alone: The use of hashtags in Twitter social activism. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Jersey City, NJ, USA) *(CSCW '18)*. Association for Computing Machinery, New York, NY, USA, 237–240. https://doi.org/10.1145/3272973.3274064

[91] Lauren Sinnenberg, Alison M Buttenheim, Kevin Padrez, Christina Mancheno, Lyle Ungar, and Raina M Merchant. 2017. Twitter as a tool for health research: A systematic review. *American Journal of Public Health* 107, 1 (2017), e1–e8.

[92] Aaron Smith and Monica Anderson. 2018. *Social Media Use in 2018.* Technical Report. Pew Research Center. http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/

[93] Shaked Spier. 2017. *Collective Action 2.0: The Impact of Social Media on Collective Action.* Cambridge, MA: Chandos Publishing.

[94] Verity Trott. 2020. Networked feminism: Counterpublics and the intersectional issues of #MeToo. *Feminist Media Studies* (2020), 1–18.

[95] Deborah Tuerkheimer. 1999. Teaching the retrenchment generation: When Sapphire meets Socrates at the intersection of race, gender, and authority. *William & Mary Journal of Race, Gender, and Social Justice* 53 (1999), 105–120.

[96] Deborah Tuerkheimer. 2017. Incredible women: Sexual violence and the credibility discount. *University of Pennsylvania Law Review* 166, 1 (2017), 105–120.

[97] John W. Tukey. 1949. Comparing individual means in the analysis of variance. *Biometrics* 5, 2 (1949), 99–114. http://www.jstor.org/stable/3001913

[98] Marlon Twyman, Brian C. Keegan, and Aaron Shaw. 2017. Black Lives Matter in Wikipedia: Collective memory and collaboration around online social movements. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1400–1412. https://doi.org/10.1145/2998181.2998232

[99] U.S. Census Bureau. 2011. Age and Sex Composition: 2010. U.S. Department of Commerce. https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf

[100] Martijn Van Zomeren, Tom Postmes, and Russell Spears. 2008. Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin* 134, 4 (2008), 504.

[101] Morgan Vigil-Hayes, Nicholet Deschine Parkhurst, and Marisa Duarte. 2019. Complex, contemporary, and unconventional: Characterizing the tweets of the #NativeVote movement and Native American candidates through the 2018

U.S. midterm elections. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW, Article 103 (Nov. 2019), 27 pages. https://doi.org/10.1145/3359205

[102] Svitlana Volkova, Theresa Wilson, and David Yarowsky. 2013. Exploring Demographic Language Variations to Improve Multilingual Sentiment Analysis in Social Media. In *Empirical Methods in Natural Language Processing (EMNLP)*. http://aclweb.org/anthology/D13-1187

[103] S Courtney Walton and Ronald E Rice. 2013. Mediated disclosure on Twitter: The roles of gender and identity in boundary impermeability, valence, disclosure, and stage. *Computers in Human Behavior* 29, 4 (2013), 1465–1474.

[104] Sherri Williams. 2015. Digital Defense: Black Feminists Resist Violence With Hashtag Activism. *Feminist Media Studies* 15, 2 (2015), 341–344. https://doi.org/10.1080/14680777.2015.1008744 arXiv:https://doi.org/10.1080/14680777.2015.1008744

[105] Zach Wood-Doughty, Nicholas Andrews, Rebecca Marvin, and Mark Dredze. 2018. Predicting Twitter User Demographics from Names Alone. In *Workshop on PEOPLES*. https://doi.org/10.18653/v1/W18-1114

[106] Zach Wood-Doughty, Paiheng Xu, Xiao Liu, and Mark Dredze. 2020. Using noisy self-reports to predict Twitter user demographics. *CoRR* abs/2005.00635 (2020). arXiv:2005.00635 http://arxiv.org/abs/2005.00635

[107] Shuang-Hong Yang, Alek Kolcz, Andy Schlaikjer, and Pankaj Gupta. 2014. Large-scale high-precision topic modeling on Twitter. In *KDD*. 1907–1916.