

SINGLE-STUDY PAPER

Using Twitter Data to Explore Public Discourse to Antiracism Movements

Rachel L. Walker and Linda K. Kaye

Department of Psychology, Edge Hill University



We used public Twitter data to explore Black Lives Matter (BLM)-related hashtags to explore collectivism and sentiment categories. Tweets containing BLM hashtags were collected at two time points; during Black History Month (BHM) and a non-BHM. At each time point, two data extractions were performed; one searching tweets containing BHM hashtags and another without. A 2×2 design was used to assess BHM hashtag and time point on emotional tone and personal pronoun use. Findings showed main effects of hashtag on all variables. Hashtags promoted greater use of collective pronouns, lesser use of singular pronouns, greater positivity, and lesser negative tone. Time point had no effect on plural pronoun use, and impacted differentially on emotional tone depending on hashtag use. Positive associations were found between plural pronoun use and positive tone but only when hashtags were used. Overall, our findings highlight the importance of online discourse to understand collectivism and sentiment in respect of antiracism movements.

Keywords: BlackLivesMatter, Black History Month, Twitter, hashtags, social identity theory

Supplemental materials: <https://doi.org/10.1037/tmb0000070.supp>

Despite legislative efforts designed to target discrimination, such as racism, Black, Asian, and Minority Ethnic (BAME) people still face an abundance of racism and discrimination (Alang et al., 2017; Ehrenfeld & Harris, 2020). Within the United States of America (USA), high-profile cases have illuminated the devastating consequences of racism, such as police brutality (Bullard, 1998; Chaney & Robertson, 2013; Ellison et al., 2008; Schaffer, 2014; Schwartz, 2020). Additionally, in the United Kingdom (U.K.), many BAME people are observed to be prone to be targets of Tasers by the police authorities (Dymond, 2020; Joseph–Salisbury et al., 2020).

The distress brought on by cases such as the above examples have motivated collective movements, such as Black Lives Matter (BLM). This originated in July 2013 by being presented as a hashtag (#BlackLivesMatter) on social media and was designed as a resistance against racially motivated violence against Black people. This has since

become a globally recognized movement, designed to promote social and political change surrounding antiracism. Similarly, another social movement that signals support for antiracism is Black History Month (also known as African American History Month), which originated in the U.S., but is now recognized in Canada, Ireland, the Netherlands, and the U.K. This is observed in the month of October, and in the year 2020, was specifically focused on raising awareness of Black experiences (Dennis, 2005; Farr, 2020; Joseph–Salisbury et al., 2020).

Given that social movements such as Black Lives Matter (BLM) and Black History Month (BHM) are designed to promote awareness, collectivity, and favorable attitudes surrounding antiracism, it would be expected that these sentiments are observable in public discourse. Indeed, findings have revealed that the popularity of BLM has changed over time (Pew Research Center, 2020). Specifically, in July 2018, figures show that only 38% of registered U.S.

Action Editor: C. Shawn Green was the action editor for this article.

ORCID iDs: Linda K. Kaye  <https://orcid.org/0000-0001-7687-5071>.

Disclosures: There are no potential conflicts of interests associated with this research.

Data Availability: A copy of our anonymized data can be found at https://osf.io/khevd/?view_only=cd060ea22ec64f67ad6eff379a8632e0

Data Use: The data reported have not been subjected to any prior uses by the authors in other publications or associated work.

Open Science Disclosures:

 The data are available at https://osf.io/khevd/?view_only=cd060ea22ec64f67ad6eff379a8632e0

Open Access License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC-BY-NC-ND). This license permits copying and redistributing the work in any medium or format for noncommercial use provided the original authors and source are credited and a link to the license is included in attribution. No derivative works are permitted under this license.

Contact Information: Correspondence concerning this article should be addressed to Linda K. Kaye, Department of Psychology, Edge Hill University, St Helens Road, Ormskirk L39 4QP, United Kingdom. Email: Linda.kaye@edgehill.ac.uk

voters were in support of BLM, which increased to 53% support in June 2020 following the infamous George Floyd case.¹ Since then, this has reduced to 48% support in January 2021 (CIVIQS, 2021).

However, much of what we know about public discourses surrounding racism and antiracist movements particularly around BLM and BHM are based on surveys such as from Government agencies who monitor national statistics. This can be problematic for a number of reasons. First, asking people to report on their racial attitudes may be prone to social desirability (Chan, 2009). Second, social change can occur at a fast pace, and so relying on national data can be time intensive. Alternative approaches are needed to address these limitations. One such approach is unobtrusively collected public online data relating to content on BLM and BHM. Indeed, this can be achieved from scraping data from platforms such as Twitter. The role of social media in movements such as BLM has been discussed as being rather important (Carney, 2016).

The advantages of using online data in research are becoming more apparent to researchers. Not only can this extend the reach of researchers, but some forms of online data can also provide objective indicators of human thought and behavior. Researchers have noted the benefits of Twitter data, for example, to gain understanding of public attitudes toward movements such as #WorldEnvironmentDay (Reyes-Menendez et al., 2018). When studying the specifics of online data, online language can reveal human perceptions and emotion (Yassine & Hajj, 2010), digital traces from smartphones can reveal individual differences and personality (Shaw et al., 2016), and digital data garnered through wearables and fitness apps can help us explore health-related behaviors (Piwek et al., 2016) and more widely help reveal insights into human movement patterns (Hinds et al., 2021).

Specifically, in relation to BLM and BHM, public Twitter data can reveal a snapshot of public attitudes and sentiment surrounding these movements, in which extensive data can be collected within just a few minutes. Recent work, for example, has indicated the role of social media in political communication (Matos et al., 2020). Not only does this address some of the aforementioned issues about social desirability in the research process, but also is much less time intensive for researchers (which, in turn, is beneficial for ensuring recency in how collected data represent current attitudes). This rectifies issues of data collection, but there are also considerations in respect of data analysis and how the information available in such data can be useful to draw valid inferences about public attitudes and sentiment. Here we can draw on the extensive literature in psycholinguistics.

There are two broad approaches which can be taken with regards to analyzing Twitter data; the small (netnographic, discourse analysis, etc.) and the big (e.g., frequency of language categories, net sentiment). The focus of the present article is on the latter approach. Specifically, we focused on two linguistic categories to draw out inferences of public sentiment and collectivity surrounding BLM and BHM. These were emotional tone (positive and negative) and first-person personal pronoun use (singular and plural), respectively. The psychological relevance of these will be discussed in the following sections, followed by a review of how these have been operationalized in previous research.

Collectivity is not a new phenomenon in social psychology. Indeed, social identity theory (SIT) has been extensively applied to explain how one's sense of self is dependent upon one's group membership (Tajfel, 1978, 1979; Tajfel & Turner, 1979). That is, a

personal self is merged with a social self, and the strength of one's social identity impacts upon one's self-regard (Abrams & Hogg, 1988; Ellemers et al., 2003). Formation of social identity is said to occur through three interrelated processes (Tajfel, 1978). First, "social categorisation" is when individuals see themselves and others as categories rather than as individuals. Second, "social identification" is when an individual's identity is formulated by their experiences within a social group or situation. Finally, "social comparison" is characterized by individuals assessing the worth of groups by comparing their relative features. Within this, "in-groups" and "out-groups" are established, and higher regard is afforded to those identified as part of the in-group. Although individual (e.g., "I") and social (e.g., "others and I") identity are well-established in the social psychology literature, it is important to note some further distinction of "collective" identity ("we") here (Priante et al., 2018). That is, unlike social identity, which may derive from identification to a social group or role, collective identity highlights "we-ness" and an emotional investment in a sense of commonality with others (Melucci, 1995; Priante et al., 2018). As such, this may be most relevant in relation to exploring social movements, such as BLM, as collectivity toward these may be equally as evident for those who do not themselves identify as BAME. Adopting the perspective of collective identity therefore, it may be expected that "we-ness" is a relevant construct of study in respect of BLM movements. Furthermore, given the fact that collective identity can be characterized by emotional investment, exploring sentiment associated with indicators of collectivity could be considered a fruitful endeavor.

However, the previous literature, which has applied SIT to racial issues, has more typically focused on formation and processes of racial identity (Matro et al., 2008; Pack-Brown, 1999) rather than collectivity more widely, as may be more relevant to social movements such as BLM and BHM. Further, there is a paucity of research which explores how antiracism collectivity may be represented in people's naturally occurring (online) language. Arguably, identity is malleable based on people's social contexts and interactions with others (Turner, 1981), and that people's language expression and behavioral choices may be a source of social or collective identity (Bucholtz & Hall, 2008). One may expect, for example, that such social movements may promote collectivity and be represented by linguistic categories such as personal pronoun use. Indeed, personal plural pronoun use (we, us, our) may signal collectivity to a given social group/movement (Best et al., 2018; Davis et al., 2019). This may especially be the case on platforms such as Twitter where functions such as hashtags are used as a mechanism for shared discourse on a given topic or issue. Similarly, it is also the case that features such as hashtags may encourage users to more openly express sentiment surrounding issues such as BLM (Bruns & Burgess, 2015; Harlow & Benbrook, 2019).

Twitter data could be entirely useful for exploring public discourse around antiracist social movements such as BLM and BHM. There is currently no published research which has made use of

¹ George Floyd was an African-American U.S. citizen. On the day of his death, he had been reported to the police following an allegation that he was using a counterfeited \$20 note in a grocery store. The authorities were alerted and they arrested Mr Floyd. However, footage obtained from the scene of the arrest showed that the leading officer (Mr Chauvin), while restraining Mr Floyd, used what was deemed excessive and prolonged force to hold him to the ground, by placing his knee on Mr Floyd neck. This was attributed to result in Mr Floyd's death due to asphyxiation.

Twitter data to make inferences about collectivity and sentiment on these specific antiracism movements. Among the research that does exist in relation to this issue, the findings are currently limited to general analysis of tweets surrounding previous police brutality incidents to demonstrate how group identity is fundamental for the BLM movement (Bonilla & Tillery, 2020; Harlow & Benbrook, 2019; Ince et al., 2017; Ray et al., 2017). However, there are wider areas of research in respect of using Twitter data to understand racism and antiracism. For example, research has collected racist language from Twitter to highlight the efficacy of these sorts of data collection methods for tracking racism online (Chaudhry, 2015). Other research explored race-related Tweets immediately following killings of Black people by law enforcement officers as a way of gauging public sentiment toward Black people (Nguyen et al., 2021). Additionally, other work has explored Twitter users' reports of Twitter in respect of content relating to race and racism (Criss et al., 2021). Finally, other work using corpus analysis has identified the use of terms such as "whitewashed" as a marker of racial identity (Nguyen, 2016). However, these do not specifically refer to BLM or BHM, nor do they focus on linguistic markers, which signal collectivity in relation to these. Further, as noted previously, given that identity is malleable based on contexts (Turner, 1981), we note there may be variations in people's discourses based on race-related events such as BHM. As such, it would be important to explore whether psychological constructs, which may be evident in people's Tweets, may vary during events such as BHM relative to other time points.

Despite the paucity of research which has used online data to apply the principles of SIT to collective identity in respect of BLM, there are a number of studies which explore social and/or collective identity from social media data (Alalwan et al., 2017; Fujita et al., 2018; Hardaker & Mcglashan, 2016), as well as exploring how collective identity is constructed on social media (e.g., DeCook, 2018; Gerbaudo, 2015). In respect of the former issue, previous research on different sociopolitical movements such as #MeToo have made effective use of Twitter to explore how hashtag use in this context may help identify identities behind this movement (Reyes-Menendez, Saura, & Thomas, 2020). Using a combination of corpus linguistics and discourse analysis, Reyes-Menendez, Saura, and Filipe (2020) noted that the #MeToo movement has two types of social identity; destructive negative and constructive positive. Other work outlines the principles of "cloud protesting" and how algorithmically mediated environments such as social media may serve collective activism (Milan, 2015). Further, other research indicates the role of interaction networks from social media data as a way of exploring collective identities (Monterde et al., 2015). However, these studies do not specifically establish how naturally occurring language on Twitter signals *collective* identity. As noted previously, collective identity is said to be distinctly different from *social* identity (Priante et al., 2018). As such, we sought to explore the linguistic markers which may signal "we-ness" as a relevant construct to antiracism movements, especially in light of the fact that "we-ness" may still exist even if an individual themselves may not occupy BAME social identity. Specifically, we focused on plural personal pronoun use based on people's Twitter discourses around BLM. We also were interested in the sentiment surrounding this movement and so explored the degree to which positive versus negative sentiment may be evident in BLM-related Twitter content.

To assess sentiment and collectivity from Twitter data, we used the linguistic categories available in the software Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2007, 2015). LIWC is a software which can process any written form of text. The application relies on an internal dictionary that defines words which should be counted based on the text which is input and the domain of the text (e.g., professional correspondence, personal writing). From this, it calculates the rate at which different word categories are used in a piece of text. These categories include summary language variables (analytical thinking, clout, authenticity, emotional tone), general descriptor categories (e.g., words per sentence), linguistic dimensions (e.g., percentage of words in the text that are pronouns, articles), word categories which relate to psychological constructs (e.g., affect, cognition, social processes), personal concern categories (e.g., work, home, leisure activities), informal language markers (swear words, netspeak), and punctuation categories (commas, etc.).

LIWC has been used extensively within research and previously been found to be a useful way of understanding a number of important psychological and social constructs (Tausczik & Pennebaker, 2010). For example, the nature of social relationships can be understood through analyzing the proportion of pronouns and emotional words (ibid). Similarly, social status can be revealed, in which those of lower status tend to use more tentative words and first-person singular pronouns. Of specific interest to the present research, better group cohesion/collectivity is associated with using more first-person plural pronouns. Indeed, previous research using discourse from online platforms has noted the value of linguistic categories such as plural personal pronouns for establishing collective identity (Best et al., 2018; Davis et al., 2019). For example, research on addiction recovery groups on Facebook highlights how the use of words such as "we" and their correspondence to affect words are relevant linguistic categories to explore collective identity for recovery (Best et al., 2018). Further, other work identifies plural personal pronouns as useful markers of collective identity to veganism (Davis et al., 2019). Therefore, the present research was particularly interested in categories such as linguistic dimensions (first-person singular and plural personal pronouns), given these would be expected to be most revealing of social processes relevant to collectivity. Furthermore, it was also interested in the emotional tone expressed in public discourse surrounding BLM and BHM, given that this may signal emotional investment surrounding collectivity. Specifically, we explored the extent to which this was more positive or negative in sentiment during a time point of BHM and in cases where BHM hashtags were used, and whether sentiment was related to level of collectivity.

In summary, the present research aimed to investigate collectivity and sentiment in public discourse in respect of BLM-related tweets and BHM. Specifically, it addressed the following research questions (RQs):

RQ1: How does the prevalence of collective identity (as evident from use of first-person plural pronouns) and positive sentiment toward BLM vary based on BHM hashtag use?

RQ2: How does the prevalence of collective identity (as evident from use of first-person plural pronouns) and positive sentiment toward BLM vary between BHM and other time points?

RQ3: To what extent does collectivity expressed in BLM-related tweets relate to sentiment, and does this vary based on BHM hashtag use and BHM time point?

Method

A 2×2 design was used in which BHM hashtag (BHM hashtag vs. no BHM hashtag) and time point (BHM time point vs. non-BHM time point) were studied in respect of their impact on four dependent variables: positive emotional tone, negative emotional tone, first-person singular pronouns, and first-person plural pronouns.

On the basis of ethical principles outlined by the British Psychological Society’s Internet-mediated Research Guidelines (BPS, 2021), and Twitter’s own privacy policy (Twitter, 2021), only public Twitter data were collected and used. This ensured that the data could be collected unobtrusively and did not require a participant consent process. As well as being important in the collection stage, ethical guidelines are also relevant in the dissemination stage. Namely, that tweet content has not been made available in the data sharing process, given that this can be reversed searched to reveal the original source of the tweet, therefore comprising anonymity. However, the anonymised raw data by condition containing the LIWC categories of interest can be found here: https://osf.io/khevd/?view_only=cd060ea22ec64f67ad6eff379a8632e0.

BLM hashtag related data were collected at two time points: during Black History Month (October 2020) and during a non-Black History Month (November 2020). Within each time point, public Tweets containing the following hashtags were extracted: #BlackLivesMatter and #BLM. However, at each time point, data were scraped twice; once with the hashtag #BlackHistoryMonth and once without, whereby this created our BHM hashtag and no BHM hashtag conditions.

Phantom Buster software was used to extract public Twitter data by inputting the hashtags: #BLM, #BlackLivesMatter in all conditions and #BlackHistoryMonth in the respective BHM hashtag conditions.² At the Black History Month time point, data were scraped on 26th October, 2020, with two back-to-back data scrapes in which the data for the two hashtag conditions were collected within 1 hr. The collections were set to the maximum number of tweets which could be collected in one run (5,000 tweets). This procedure was repeated for the non-Black History Month time point (November 30, 2020).

The extracted data from Phantom Buster generated an Excel file which was subjected to data cleaning. The data were cleaned based on the inclusion criterion that tweets should be in English, and on the exclusion criteria of tweets containing retweets, URLs, multimedia, threads, and replies. The exclusion criteria were chosen to eliminate any repetition of the same tweets, to reduce tweets with names and unnecessary data that could not be analyzed such as multimedia. Table 1 below displays the total number of tweets at the collection point and the final set following data cleaning.

The final data set was then uploaded into LIWC software, in which the four linguistic categories were selected to be processed for analysis.³ First, “I” and “we” which were from the listed items under the broader “total pronouns” category. “I” refers to any singular personal pronouns (I, me, my) and “we” to any plural personal pronouns (we, us, our). Second, positive emotional tone and negative emotional tone which were taken from the broader category of “affective or emotional processes.” For all variables, the scores generated represented percentage of total words. Previous work

Table 1

Total Tweets Per Condition Before and After Data Cleaning

Time point	Hashtag	Extracted data (n)	Cleaned data (n)
Black History Month	Hashtag	4,143	345
	No Hashtag	4,337	790
Non-Black History Month	Hashtag	3,804	170
	No Hashtag	4,006	461

indicates that these categories of the LIWC 2015 dictionary have adequate internal consistency (Pennebaker et al., 2015).

Although no materials are available to share given the unobtrusive data collection method, the anonymised raw data by condition containing the LIWC categories of interest can be found here: https://osf.io/khevd/?view_only=cd060ea22ec64f67ad6eff379a8632e0.

Results

Descriptive analyses of percentage of total words were performed on the study variables by hashtag and time point condition. See Table 2. It is worth noting that the typical average percentage of words from Twitter except for these categories has been found to be as follows: singular personal pronouns ($M = 4.75$), plural personal pronouns ($M = .74$), positive emotion ($M = 5.48$), and negative emotion ($M = 2.14$; Pennebaker et al., 2015).

To give some context to these values, Table 3 provides some examples of the types of Tweet content per condition. To adhere to BPS (2021) principles regarding anonymity of data, these are artificial examples which linguistically mimic the original tweets for the respective conditions.

Personal Pronoun Use

To ascertain how collectivity in language use varied by conditions, a series of two 2×2 factorial ANOVAs were undertaken. The first 2×2 factorial ANOVA was conducted to assess the impact of hashtag condition and event time point on personal singular pronoun use (variable labeled here as “i”). Significant main effects were found for hashtag condition, $F(1, 1,762) = 233.07$, $MSE = 31.48$, $p < .001$, $\eta_p^2 = .117$, and time point, $F(1, 1,762) = 124.67$, $MSE = 31.48$, $p < .001$, $\eta_p^2 = .066$. There was also a significant interaction effect of hashtag condition \times time point, $F(1, 1,762) = 4120.93$, $MSE = 31.48$, $p < .001$, $\eta_p^2 = .069$. Simple main effects analysis

² We used Phantom Buster for a number of pragmatic reasons. First, many other tools are designed to extract analytics from one’s own Twitter account rather than a more general scrape of all public data. Second, we were able to utilize a free version of Phantom Buster, which was preferable for the purposes of this small scale project (with no funding). Finally, it was more straightforward than other options with regards to the exporting capabilities while others we explored did not export separate metadata to individual Excel columns, which would have made the data cleaning ready for analysis very time consuming.

³ A further confidentiality assurance we took here was that all other metadata from the Excel files were removed before uploading into LIWC. This was to ensure that Twitter account names or other potentially identifying information was not added to this third party software. Therefore only the tweet content itself was contained in the respective file per condition.

Table 2
Descriptive Data by Hashtag and Time Point Condition for the Study Variables

Linguistic category	BHM				Non-BHM			
	BHM Hashtag		No BHM Hashtag		BHM Hashtag		No BHM Hashtag	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Singular personal pronoun	.78	.30	9.06	.20	.86	.43	2.05	.26
Plural personal pronoun	1.78	.18	1.52	.12	2.17	.25	1.48	.15
Positive emotion	3.50	.20	1.31	.13	3.13	.28	3.13	.17
Negative emotion	1.59	.21	2.82	.14	.73	.29	3.16	.18

Note. BHM = Black History Month.

suggested that the effect of hashtag on personal singular pronoun use was significant both in BHM, $F(1, 1,762) = 523.10, p < .001$, and non-BHM time points, $F(1, 1,762) = 5.55, p < .05$, whereby significantly fewer singular pronouns were used when hashtags were included on Tweets rather than not used. Further, the effect of time point on personal singular pronoun use was significant but only when BHM hashtags were not used, $F(1, 1,762) = 454.37, p < .001$, whereby more singular pronouns were used during BHM ($M = 9.06, SE = .20$) than non-BHM ($M = 2.05, SE = .26$). See Figure 1.

The second 2×2 factorial ANOVA was conducted to assess the impact of hashtag condition and event time point on personal plural pronoun use (variable labeled here as “we”). A significant main effect was found for the hashtag condition, $F(1, 1,762) = 6.84, MSE = 10.52, p < .01, \eta_p^2 = .004$, but not for time point, $F(1, 1,762) = 10.15, MSE = 10.52, p = .326, \eta_p^2 = .001$. Specifically, plural pronoun use was significantly higher in the BHM hashtag condition relative to the no BHM hashtag condition. There was not a significant interaction effect of hashtag condition \times time point, $F(1, 1,762) = 1.41, MSE = 10.52, p = .235, \eta_p^2 = .001$. See Figure 2.

Emotional Tone

To explore how sentiment varies by condition, two 2×2 factorial ANOVAs were undertaken. The first was conducted to assess the impact of hashtag condition and event time point on positive emotional tone (variable labeled here as “posemo”). Significant

main effects were found for hashtag condition, $F(1, 1,762) = 30.07, MSE = 13.15, p < .001, \eta_p^2 = .017$, and time point, $F(1, 1,762) = 13.12, MSE = 13.15, p < .001, \eta_p^2 = .007$. There was also a significant interaction effect of hashtag condition \times time point, $F(1, 1,762) = 29.71, MSE = 13.15, p < .001, \eta_p^2 = .017$. Simple main effects analysis suggested that the effect of hashtag on positive emotional language use was significant only during BHM, $F(1, 1,762) = 87.69, p < .001$, whereby more positive tone was evident when hashtags were used ($M = 3.49, SE = .20$) rather than not used ($M = 1.31, SE = .17$). Further, the effect of time point on positive emotional tone was significant but only when BHM hashtags were not used, $F(1, 1,762) = 73.19, p < .001$, whereby there was higher positive tone during a non-BHM ($M = 3.13, SE = .17$) than during BHM ($M = 1.31, SE = .13$). See Figure 3.

The final 2×2 factorial ANOVA was conducted to assess the impact of hashtag condition and event time point on negative emotional tone (variable labeled here as “negemo”). A significant main effect was found for the hashtag condition, $F(1, 1,762) = 75.65, MSE = 14.46, p < .001, \eta_p^2 = .041$, but not for time point, $F(1, 1,762) = 1.47, MSE = 14.46, p = .225, \eta_p^2 = .001$. Specifically, negative tone was significantly higher when there was no BHM hashtag compared to when one was present. There was also a significant interaction effect of hashtag condition \times time point, $F(1, 1,762) = 8.09, MSE = 14.46, p < .01, \eta_p^2 = .005$. Simple main effects analysis suggested that the effect of hashtag on negative emotional tone was significant both in BHM, $F(1, 1,762) = 25.13, p < .001$, and non-BHM time points, $F(1, 1,762) = 50.52, p < .001$, whereby not including hashtags elicited more negative tone than including them. Further, effect of time point on negative emotional tone was significant but only when BHM hashtags were used, $F(1, 1,762) = 5.73, p < .05$. Specifically, more negative tone was evident during BHM ($M = 1.59, SE = .21$), than non-BHM ($M = .73, SE = .29$). See Figure 4.

Proportion of Singular Versus Plural Personal Pronouns

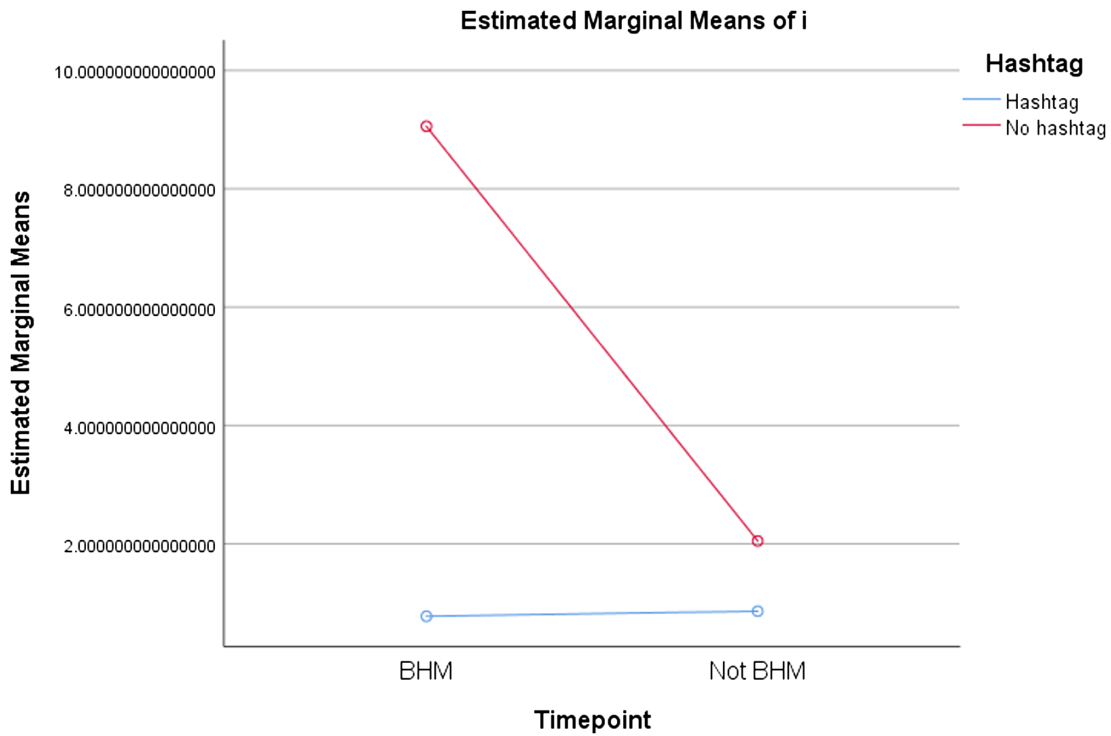
The aforementioned analysis identifies there to be some differences in use of pronouns by condition. However, to more firmly establish whether this was actually a reflection of collective identity, within-participant analyses were undertaken to compare the use of plural versus singular pronouns based on time point and hashtag.

Table 3
Indicative Tweet Content Per Condition

Condition	Example tweets
BHM Hashtag/BHM	“We celebrate Black History today in Politics!” “Here are tips on how students can hold universities to account in relation to racism”
No BHM Hashtag/BHM	“I will defend the black community until the end. They have been good allies in my community. I want to replay them. Black lives matter, today, tomorrow, and always. I will not stand for the oppression of my persons of colour siblings” “I want to live where people are compassionate, sane and kind”
BHM Hashtag/Non-BHM	“#BlackLivesMatter #BlackHistoryMonth #BlackWomen #BLM More facts and statistics to silence the deniers of systemic racism”
No BHM Hashtag/Non-BHM	“We are very proud of our Year 12 student Alice, for her original anti racist art. This painting is so wonderful and pertinent” “Let’s work together to do good in the world! Help us meet our target of raising \$20,000 for #BlackLivesMatter. An anonymous donor will match it! You can donate online” “With all this racism occurring in the world, I needed to raise my voice with the BLM movement we have been tackling this stuff for too long and we are fed up”

Note. BHM = Black History Month; BLM = Black Lives Matter.

Figure 1
Percentage of Singular Personal Pronouns by Condition



Note. BHM = Black History Month.

Specifically, if use of plural pronouns is indeed reflective of collective identity, one should expect more plural than singular pronouns when the hashtag was used, and when it was salient (BHM time point) compared to at a nonsalient time point and without a hashtag.

A mixed design ANOVA was conducted in which singular and plural pronouns were the within-condition variables and time point and hashtag were between-condition variables. A main effect was found, $F(1, 1,762) = 87.50$, $MSE = 1374.80$, $p < .001$, $\eta_p^2 = .047$. Namely, overall, a greater number of singular pronouns ($M = 4.82$, $SD = 6.80$) compared to plural pronouns ($M = 1.62$, $SD = 3.25$) were used. Exploring this by condition, however, revealed some interesting effects. That is, there was an interaction effect of both hashtag, $F(1, 1,762) = 282.06$, $MSE = 4431.64$, $p < .001$, $\eta_p^2 = .138$, and time point, $F(1, 1,762) = 137.93$, $MSE = 2167.11$, $p < .001$, $\eta_p^2 = .073$. When hashtags were used, significantly more plural ($M = 1.90$, $SD = 2.99$) compared to singular pronouns ($M = .80$, $SD = 2.40$) were used. However, when no hashtags were used, there were significantly more singular ($M = 6.47$, $SD = 7.31$) than plural pronouns ($M = 1.51$, $SD = 3.34$). In respect of time point, however, singular pronouns were used significantly more than plural pronouns, both during BHM and non-BHM. Namely, during BHM, singular pronouns ($M = 6.54$, $SD = 7.56$) and plural pronouns ($M = 1.60$, $SD = 3.33$), and non-BHM, singular pronouns ($M = 1.73$, $SD = 3.39$) and plural pronouns ($M = 1.67$, $SD = 3.10$).

Relationship Between Collectivism and Sentiment

Pearson correlations were conducted to explore the relationship between personal singular pronouns, personal plural pronouns,

positive emotional tone, and negative emotional tone. Data were split by hashtag condition, whereby the first correlation explored these associations under conditions in which BHM hashtags were used and the second where this hashtag was not used. Table 4 presents these analyses.

Use of singular pronouns was related to positive emotional tone in both the hashtag ($r = .26$, $p < .01$) and no hashtag conditions ($r = .06$, $p < .05$), but not to negative tone in either (both $p > .05$). Plural pronoun use was positively related to positive emotional tone in the hashtag condition ($r = .14$, $p < .01$), but interestingly, negatively related in the no hashtag conditions ($r = -.08$, $p < .01$). In relation to negative emotional tone, plural pronouns correlated negatively in the hashtag condition ($r = -.09$, $p < .05$), but were not significantly related at all in the no hashtag condition.

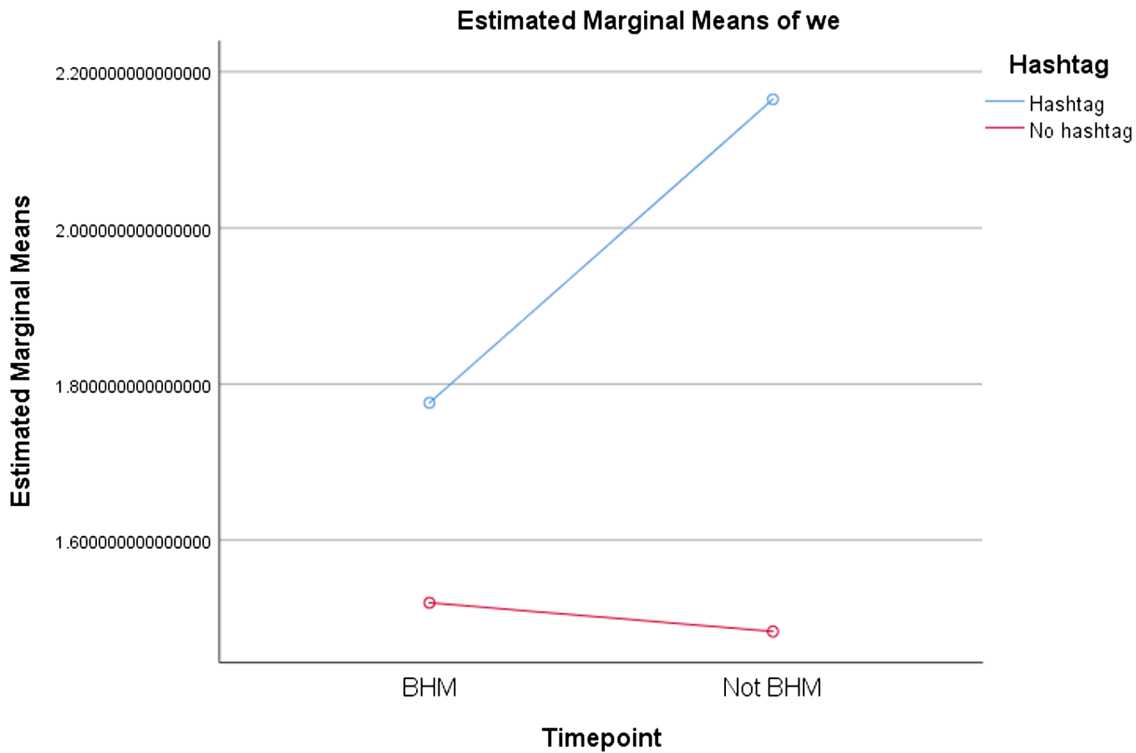
Finally, correlational analysis was undertaken as previously mentioned but split by time point, whereby the first correlation of variables was for BHM and the second for non-BHM. See Table 5.

Use of singular pronouns was related to negative emotional tone during BHM ($r = .11$, $p < .001$) and to positive emotional tone during a non-BHM ($r = .01$, $p < .01$). Plural pronoun use was not related to positive or negative emotional tone at either time point (all $p > .05$).

Discussion

Exploring public discourse around antiracism movements presents a key opportunity for researchers to assess attitudes and sentiments to these sociopolitical movements. Using public Twitter data is a highly fruitful approach in which to achieve this.

Figure 2
Percentage of Plural Personal Pronouns by Condition



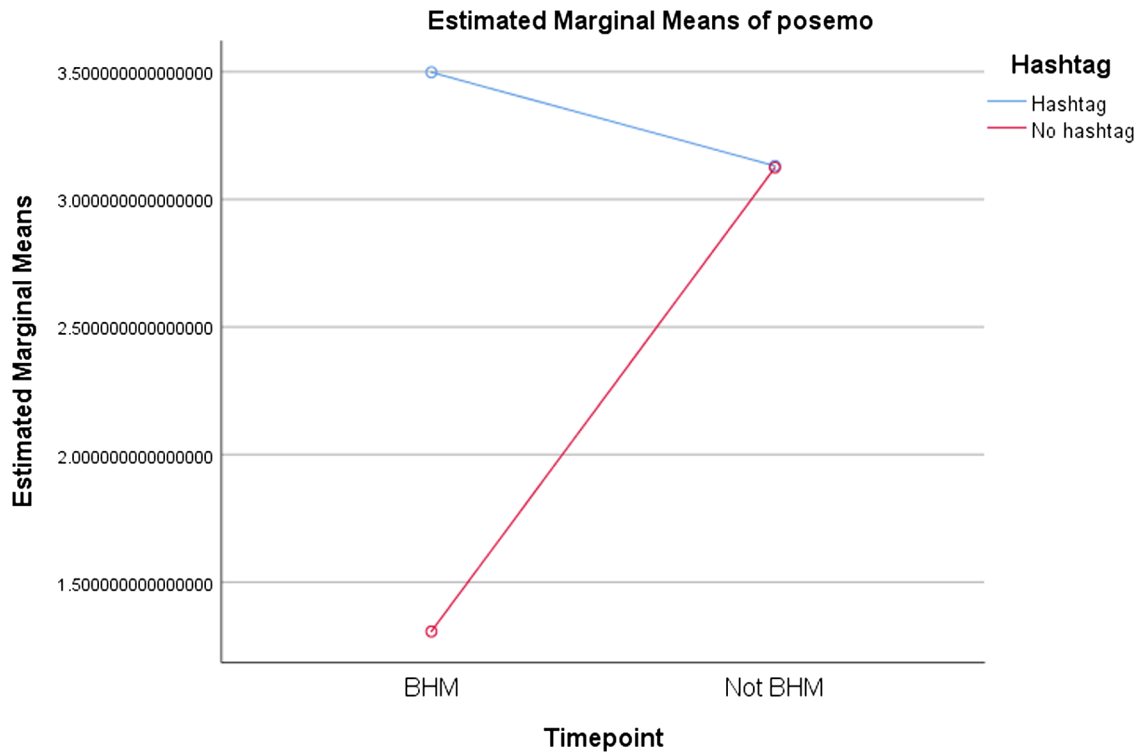
Note. BHM = Black History Month.

The present research utilized Twitter data for BLM-related hashtags to explore collectivity and sentiment in public discourse in respect of the antiracism movements of BLM and BHM. Specifically, we sought to understand the extent to which (a) BHM hashtags and (b) Black History Month as a celebration month may promote greater collectivism and positive sentiment surrounding the BLM movement. The key findings and implications will be discussed in the following sections.

In line with RQ1, in respect of BHM hashtags (relative to conditions without these), these had a number of significant effects. Tweets including BHM hashtags had greater use of collective pronouns and lesser use of singular pronouns, as well as more positivity in emotional tone when used during BHM, and lower negative tone. Consequently, these findings of positive sentiment toward the movement BLM when a hashtag is present reveal that the inclusion of #BlackHistoryMonth encourages a more collective, positive attitude toward BLM, as Twitter behaviors exhibit a communal acknowledgment and understanding of Black History Month. To provide additional strength to the assertion that plural pronouns were reflective of collective identity, we compared use of plural versus singular pronouns to establish whether greater use of plural pronouns (relative to singular pronouns) was more evident in conditions where one may expect collectivity to be salient (BHM and using BHM hashtags). Our findings provided some support for this, in that significantly more plural pronouns were used in tweets that used the BHM hashtag and more singular pronouns were used without the hashtag. These findings suggest something rather important about the use of hashtags in building collectivism and

positivity in public expression of antiracism movements. The implications of this should not be underestimated and may not only represent a snapshot of how we monitor public sentiment, but importantly, how these may influence subsequent discourse around these events. Indeed, further research which measures any direct effects of these discourses on future collective action or attitudes would be most enlightening. Hashtags encourage people to alter their behavior to suit the meaning of that hashtag as they are used to encourage conversations around similar topics (Cunha et al., 2011; Wang et al., 2016). There are many examples in which “real world” collective action has ensued from online activism, such as the #MeToo movement, whereby the use of hashtags has found an encouragement to act and promote these specific events, and perhaps a more infamous recent example of how a former U.S. President’s tweets were attributed to Far-Right protesters subsequently terrorizing the Capitol Building in Washington, DC (Suitner et al., 2013). Hashtags and other online cues, which can promote collective identity should not be considered insignificant in social and political discourse and action (see Priante et al., 2018, for a review of collective action via computer-mediated communication). Indeed, research suggests that sociopolitical hashtags such as #MeToo can elucidate different types of social identity, which may vary from being destructive negative to constructive positive (Reyes-Menendez, Saura, & Filipe, 2020). Although we did not seek to explore the nuances of this, the current findings contribute to the SIT literature in suggesting that such hashtags may be one mechanism by which individuals are able to categorize themselves into the in-group supporting the movement

Figure 3
Percentage of Positive Emotional Tone by Condition



Note. BHM = Black History Month.

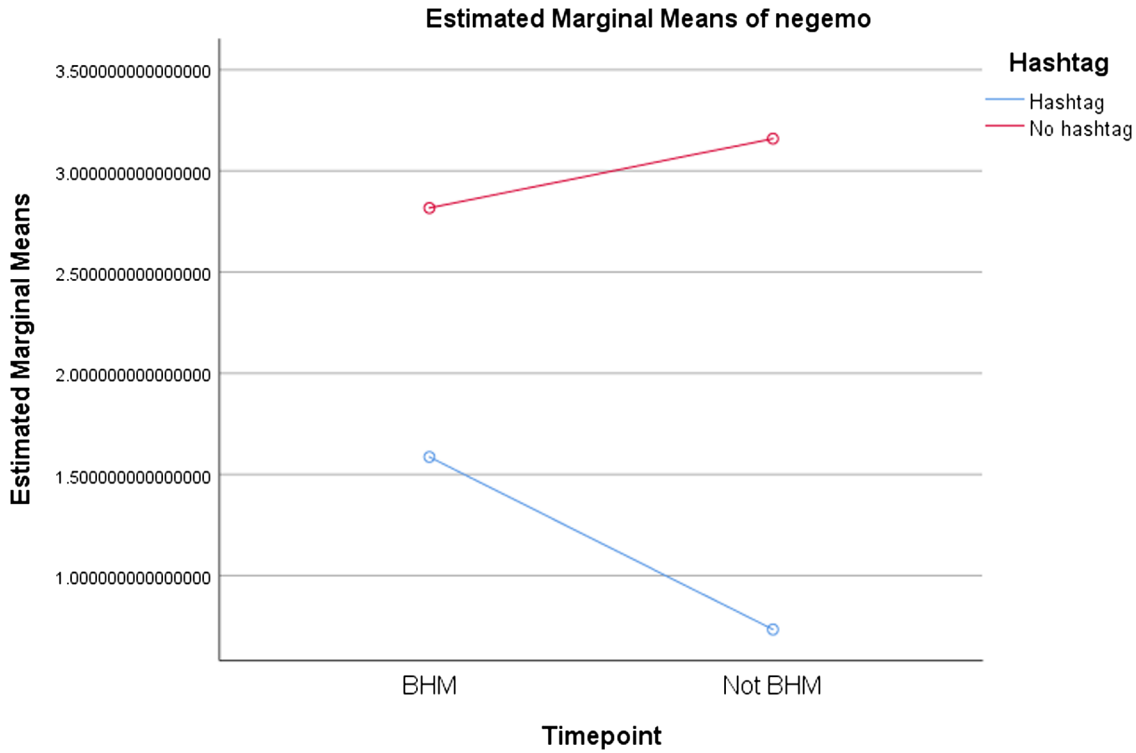
(Blascovich et al., 1997; Tajfel, 1978; Turner, 1981), and highlight how behavioral cues such as hashtags are entirely relevant to understand how collective identity is expressed in online sociopolitical discourse.

BHM event time point, however (RQ2), appeared to be less consistent in its effects. Namely, tweets during BHM (relative to a non-BHM) had greater use of singular pronouns but only when the BHM hashtags were not used. There were no significant effects of time point on use of plural pronouns. In relation to emotional tone, this presented some intriguing findings. Overall, time point had a significant effect on positive tone but not on negative tone, in which positive tone was significantly lower during BHM than non-BHM. Looking more closely at univariate effects, this effect only transpired when BHM hashtags were not used. Correspondingly, when focusing on when hashtags were used during BHM, a significantly higher positive tone was evident than when they were not used. In relation to negative emotional tone, this varied between time points in which it was more apparent during BHM than non-BHM, but only when hashtags were used. These findings are interesting and suggest sentiments surrounding BHM (irrespective of whether this is positive or negative) are expressed more strongly when relevant hashtags accompany the discourse. Although BHM time point in itself did not appear to be particularly influential on the types of language expressed on Twitter, this did provide a context in which BHM hashtags could be used and thus facilitate expression of sentiment. This raises an important issue about time and context for research on antiracism attitudes and the value of being able to extract public online data within a highly specified time frame (e.g., a few

minutes). Therefore, the impact of time/context is perhaps much better controlled than when collecting attitudes via surveys or other methods, which may be more time intensive.

It appears possible that celebratory events like BHM, on their own, are not enough to elicit attitudes and views toward BLM without directly using the hashtag #BlackHistoryMonth (Cunha et al., 2011; Givens, 2019). Therefore, these findings may help in understanding that hashtags are considerably more useful in exploring attitudes toward BLM, specifically to distinguish emotional sentiment, as well as collectivism. To corroborate this further for RQ3, the correlational findings highlight the association between the use of plural pronouns and positive emotional tone when BHM hashtags are used, which was not found when these hashtags were not used (indeed the converse effect was observed). These findings broadly support the work of other studies in this area where social media is an effective form of expression for groups about social justice issues to promote collective attitudes. For example, it supports previous work showing how hashtags create communities and discussions surrounding social movements to help promote collective attitudes (Cunha et al., 2011; DeLuca et al., 2012; Reyes-Menendez, Saura, & Filipe, 2020; Wang et al., 2016). Thus, these findings suggest that hashtags of events or movements are more powerful in promoting collective identity toward BLM than specific celebratory events like Black History Month itself. It would be interesting to see any subsequent effects of this, however, and how public attitudes and sentiments transmitted over Twitter may promote contagion and a range of associated “real-world” behaviors. This would further deepen our understanding of Twitter discourse

Figure 4
Percentage of Negative Emotional Tone by Condition



Note. BHM = Black History Month.

and how specific hashtags resonate and influence Twitter users (DeLuca et al., 2012; Papacharissi, 2016).

The present study focused on data from Twitter as the platform of choice. Like the majority of previous literature using these sorts of approaches, this was chosen for pragmatic reasons. Namely, this platform relative to other social media sites has the largest proportion of data which is publicly available and accessible. As such, this allowed a large population of data to be extractable for analysis. A further advantage of this platform and its largely open nature is that it can foster much larger and diverse networks of people to join together, and hashtags are an operational tool that can support this. Other social media sites are less successful in this regard as many users have profiles on these that are closed/protected, which makes consolidating these collective behaviors more difficult. It is important to note that there may be generalizability issues of a Twitter

sample as others have illustrated there may be distinct personality characteristics of Twitter users, which can be translated into the types of disclosures that are made online (Marshall et al., 2020). As such, it may be the case that the current findings are representative only of a subsample of the general population.

In relation to the findings about collectivity from pronoun use, a limitation of the present research is that Twitter users' race could not be established from the data. It would be interesting for additional research to ascertain how people's race or ethnicity may interact with linguistic expressions of collectivity. That is, previous research suggests that identity exploration during BHM is especially important to African American boys (Landa, 2012), and so user race seems to be something further to explore in this regard. It is entirely likely that the current data included people of many racial and ethnic backgrounds, and in this case, it is promising to note that collectivity can still be represented in these discourses. Alongside this, it is important to explore further how the use of plural pronouns is fully representative of collective identity and not perhaps as a result of a group's perception of status over another. That is, collective words may often be used by those in high status conditions compared to low status conditions (Dino et al., 2009; Kacewicz et al., 2014), and so it is important to ascertain how linguistic categories such as plural pronouns are actually being used in this regard.

Another limitation is that the tweets scraped may also have included a range of other semantically relevant hashtags such as #AllLivesMatter or #BlueLivesMatter, which largely oppose the views of BLM. As a result, these potentially disparate attitudes may explain findings relating to negative sentiment, rather than the

Table 4
Correlation Analyses of the Study Variables by Hashtag Condition

Linguistic category	1	2	3	4
1. Personal singular pronouns		-.14**	.26**	-.06
2. Personal plural pronouns	.23**		.14**	-.09*
3. Positive emotional tone	.06*	-.08**		-.20**
4. Negative emotional tone	.02	.03	-.03	

Note. Values on the top-right of the diagonal correspond to the Black History Month (BHM) hashtag condition and those on the bottom-left are the non-BHM hashtag condition.

* $p < .05$. ** $p < .001$.

Table 5
Correlation Analyses of the Study Variables by Time Point

Linguistic category	1	2	3	4
1. Personal singular pronouns		.34**	.06	.11**
2. Personal plural pronouns	-.13**		-.01	.01
3. Positive emotional tone	.09*	-.02		-.14**
4. Negative emotional tone	.01	-.05	-.06	

Note. Values on the top-right of the diagonal correspond to the Black History Month (BHM) time point and those on the bottom-left are the non-BHM time point.

* $p < .05$. ** $p < .001$.

negative sentiment being due to Twitter users discussing the pejorative side of BLM and how Black people are treated.

Finally, there have been recent discussions about the efficacy of dictionary-based sentiment analysis tools, such as those similar to LIWC. Although this specific platform was not included in the recent paper by van Atteveldt et al. (2021), their findings suggest that the agreement level of whether certain words correspond to positive or negative sentiment is often close to chance and there may be a heightened risk of type 2 errors in the classification process. As such, scholars recommend alternative approaches such as trained coders or machine learning to achieve better performance.

Irrespective of these limitations, the current findings hold some important implications. Namely, there are some key practical implications for communication campaigns via social media, which may surround antiracism and racial equality. This may also intersect with organizations efforts to promote socially inclusive content via social media and the social mobilization of these issues (Reyes-Menendez, Saura, & Filipe, 2020), which may generate further collective identity and collective action.

Conclusion

We sought to understand the extent to which (a) BHM hashtags and (b) Black History Month may promote greater collectivity and sentiment surrounding the BLM movement. Specifically, we explored linguistic dimensions (first-person singular and plural personal pronouns), given that we expected these to be most revealing of social processes relevant to collectivity. Furthermore, we also explored emotional tone and the extent to which this was more positive or negative in sentiment during a time point of BHM and in cases where BHM hashtags were used. On the basis of our analyses of BLM-related hashtag data from public Twitter posts, we highlight the prominent role of hashtags in facilitating expression toward sociopolitical movements such as BLM. This study is the first of its kind to use this type of data and linguistic categories to explore collective attitudes and sentiment toward the BLM movement during Black History Month. Our findings therefore contribute new understanding to these important societal issues, but specifically illuminate that these may elicit enhanced sentiment, which may both be positive and negative in nature. Although plural pronouns as a proxy of collectivism, appear to relate favorably to expression of positive emotional sentiment, it should also be noted that such collectivism may equally promote shared negative discourse. However, this demonstrates the mechanisms by which shared sentiment may influence subsequent collective action. Although the behavioral correlates of these discourses cannot be

determined from the current data, the findings do showcase how hashtags may have the power to shape conversations of race.

References

- Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and inter-group discrimination. *European Journal of Social Psychology, 18*(4), 317–334. <https://doi.org/10.1002/ejsp.2420180403>
- Alalwan, A. A., Rana, N. P., Dwivedi, Y. K., & Algharabat, R. (2017). Social media in marketing: A review and analysis of the existing literature. *Telematics and Informatics, 34*(7), 1177–1190. <https://doi.org/10.1016/j.tele.2017.05.008>
- Alang, S., McAlpine, D., McCreedy, E., & Hardeman, R. (2017). Police brutality and black health: Setting the agenda for public health scholars. *American Journal of Public Health, 107*(5), 662–665. <https://doi.org/10.2105/AJPH.2017.303691>
- Best, D., Bliuc, A.-M., Iqbal, M., Upton, K., & Hodgkins, S. (2018). Mapping social identity change in online networks of addiction recovery. *Addiction Research and Theory, 26*(3), 163–173. <https://doi.org/10.1080/16066359.2017.1347258>
- Blascovich, J., Wyer, N. A., Swart, L. A., & Kibler, J. L. (1997). Racism and racial categorisation. *Journal of Personality and Social Psychology, 72*(6), 1364–1372. <https://doi.org/10.1037/0022-3514.72.6.1364>
- Bonilla, T., & Tillery, A. B., Jr. (2020). Which identity frames boost support for and mobilization in the # BlackLivesMatter movement? An experimental test. *The American Political Science Review, 114*(4), 947–962. <https://doi.org/10.1017/S0003055420000544>
- British Psychological Society. (2021). *Ethics guidelines for internet-mediated research*. <https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research.pdf>
- Bruns, A., & Burgess, J. E. (2015). Twitter hashtags: from ad hoc to calculated publics. In N. Rambukkana (Ed.), *Hashtag publics: The power and politics of discursive networks* (pp. 12–28). Peter Lang.
- Bucholtz, M., & Hall, K. (2008). All of the above: New coalitions in sociocultural linguistics. *Journal of Sociolinguistics, 12*(4), 401–431. <https://doi.org/10.1111/j.1467-9841.2008.00382.x>
- Bullard, S. (Ed.). (1998). *The ku klux klan: A history of racism & violence*. Diane Publishing.
- Carney, N. (2016). All lives matter, but so does race: Black lives matter and the evolving role of social media. *Humanity & Society, 40*(2), 180–199. <https://doi.org/10.1177/0160597616643868>
- Chan, D. (2009). *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*. Taylor & Francis.
- Chaney, C., & Robertson, R. V. (2013). Racism and police brutality in America. *Journal of African American Studies, 17*(4), 480–505. <https://doi.org/10.1007/s12111-013-9246-5>
- Chaudhry, I. (2015). #Hashtagging hate: Using Twitter to track racism online. *First Monday, 20*(2). Advance online publication. <https://doi.org/10.5210/fm.v20i2.5450>
- CIVIQS. (2021). *Black Lives Matter*. Retrieved January 18, 2021, from https://civiqs.com/results/black_lives_matter?uncertainty=true&annotations=true&zoomIn=true&trendline=true
- Criss, S., Michaels, E. K., Solomon, K., Allen, A. M., & Nguyen, T. T. (2021). Twitter fingers and echo chambers: Exploring expressions and experiences of online racism using twitter. *Journal of Racial and Ethnic Health Disparities, 8*(5), 1322–1331. <https://doi.org/10.1007/s40615-020-00894-5>
- Cunha, E., Magno, G., Comarela, G., Almeida, V., Gonçalves, M. A., & Benevenuto, F. (2011, June). Analysing the dynamic evolution of hashtags on twitter: A language-based approach. In *proceedings of the workshop on language in social media (LSM 2011)* (pp. 58–65). Association for Computational Linguistics.

- Davis, J. L., Love, T. P., & Fares, P. (2019). Collective social identity: Synthesizing identity theory and social identity theory using digital data. *Social Psychology Quarterly*, 82(3), 254–273. <https://doi.org/10.1177/0190272519851025>
- DeCook, J. R. (2018). Memes and symbolic violence: #proudboys and the use of memes for propaganda and the construction of collective identity. *Learning, Media and Technology*, 43(4), 485–504. <https://doi.org/10.1080/17439884.2018.1544149>
- DeLuca, K. M., Lawson, S., & Sun, Y. (2012). Occupy wall street on the public screens of social media: The many framings of the birth of a protest movement. *Communication, Culture & Critique*, 5(4), 483–509. <https://doi.org/10.1111/j.1753-9137.2012.01141.x>
- Dennis, S. (2005). Cultural celebration: Black History Month is for everyone. *Nursing Standard*, 20(4), 34–35. <https://doi.org/10.7748/ns.20.4.34.s36>
- Dino, A., Reysen, S., & Branscombe, N. R. (2009). Online interactions between group members who differ in status. *Journal of Language and Social Psychology*, 28(1), 85–93. <https://doi.org/10.1177/0261927X08325916>
- Dymond, A. (2020). “Taser, Taser”! Exploring factors associated with police use of Taser in England and Wales. *Policing and Society*, 30(4), 396–411. <https://doi.org/10.1080/10439463.2018.1551392>
- Ehrenfeld, J. M., & Harris, P. A. (2020). *Police brutality must stop*. Retrieved January 18, 2021, from <https://www.ama-assn.org/about/leadership/police-brutality-must-stop>
- Ellemers, N., Haslam, S. A., Platow, M. J., & Knippenberg, D. (2003). Social identity at work: developments, debates and directions. In S. A. Haslam, D. V. Knippenberg, M. J. Platow, & N. Ellemers (Eds.), *Social identity at work: Developing theory for organisational practice* (pp. 3–26). Taylor and Francis Group.
- Ellison, C. G., Musick, M. A., & Henderson, A. K. (2008). Balm in Gilead: Racism, religious involvement, and psychological distress among African American adults. *Journal for the Scientific Study of Religion*, 47(2), 291–309. <https://doi.org/10.1111/j.1468-5906.2008.00408.x>
- Farr, L. T. (2020). Listen and Change. *Journal of the Academy of Nutrition and Dietetics*, 120(9), 1449–1451. <https://doi.org/10.1016/j.jand.2020.07.001>
- Fujita, M., Harrigan, P., & Soutar, G. N. (2018). Capturing and co-creating student experiences in social media: A social identity theory perspective. *Journal of Marketing Theory and Practice*, 26(1–2), 55–71. <https://doi.org/10.1080/10696679.2017.1389245>
- Gerbaudo, P. (2015). Protest avatars as memetic signifiers: Political profile pictures and the construction of collective identity on social media in the 2011 protest wave. *Information Communication and Society*, 18(8), 916–929. <https://doi.org/10.1080/1369118X.2015.1043316>
- Givens, J. R. (2019). “There would be nNo lynching if it did not start in the schoolroom”: Carter G. Woodson and the occasion of Negro History Week, 1926–1950. *American Educational Research Journal*, 56(4), 1457–1494. <https://doi.org/10.3102/0002831218818454>
- Hardaker, C., & Mcglashan, M. (2016). “Real men don’t hate women”: Twitter rape threats and group identity. *Journal of Pragmatics*, 91, 80–93. <https://doi.org/10.1016/j.pragma.2015.11.005>
- Harlow, S., & Benbrook, A. (2019). How #BlackLivesMatter: Exploring the role of hip-hop celebrities in constructing racial identity on Black Twitter. *Information Communication and Society*, 22(3), 352–368. <https://doi.org/10.1080/1369118X.2017.1386705>
- Hinds, J., Brown, O., Smith, L. G. E., Piwek, L., Ellis, D. A., & Joinson, A. N. (2021). Integrating insights about human movement patterns from digital data into psychological science. *Current Directions in Psychological Science*. Advance online publication. <https://doi.org/10.1177/096372142111042324>
- Ince, J., Rojas, F., & Davis, C. A. (2017). The social media response to Black Lives Matter: How Twitter users interact with Black Lives Matter through hashtag use. *Ethnic and Racial Studies*, 40(11), 1814–1830. <https://doi.org/10.1080/01419870.2017.1334931>
- Joseph-Salisbury, R., Connelly, L., & Wangari-Jones, P. (2020). The UK is not innocent”: Black Lives Matter, policing and abolition in the UK. *Equality, Diversity and Inclusion*, 40(1), 21–28. <https://doi.org/10.1108/EDI-06-2020-0170>
- Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33(2), 125–143. <https://doi.org/10.1177/0261927X13502654>
- Kaye, L. K. (2021). *Using Twitter data to explore public discourse to anti-racism movements*. Center for Open Science. <https://doi.org/10.17605/OSF.IO/KHEVD>
- Landa, M. H. (2012). Deconstructing black history month: Three African American boys’ exploration of identity. *Multicultural Perspectives*, 14(1), 11–17. <https://doi.org/10.1080/15210960.2012.646638>
- Marshall, T. C., Ferenczi, N., Lefringhausen, K., Hill, S., & Deng, J. (2020). Intellectual, narcissistic, or Machiavellian? How Twitter users differ from Facebook-only users, why they use Twitter, and what they tweet about. *Psychology of Popular Media*, 9(1), 14–30. <https://doi.org/10.1037/ppm0000209>
- Matos, N. D., Correia, M. B., Saura, J. R., Reyes-Menendez, A., & Baptista, N. (2020). Marketing in the Public Sector—Benefits and Barriers: A Bibliometric Study from 1931 to 2020. *Social Sciences*, 9(10), Article 168. <https://doi.org/10.3390/socsci9100168>
- Matro, D. E., Behm-Morawitz, E., & Kopacz, M. A. (2008). Exposure to television portrayals of Latinos: The implications of aversive racism and social identity theory. *Human Communication Research*, 34(1), 1–27. <https://doi.org/10.1111/j.1468-2958.2007.00311.x>
- Melucci, A. (1995). *The process of collective identity*. Temple University Press.
- Milan, S. (2015). When algorithms shape collective action: Social media and the dynamics of cloud protesting. *Social Media + Society*, 1(2). Advance online publication. <https://doi.org/10.1177/2056305115622481>
- Monterde, A., Calleja-López, A., Aguilera, M., Barandiaran, X. E., & Postill, J. (2015). Multitudinous identities: A qualitative and network analysis of the 15M collective identity. *Information Communication and Society*, 18(8), 930–950. <https://doi.org/10.1080/1369118X.2015.1043315>
- Nguyen, N. M. (2016). I tweet like a white person tbh! #whitewashed: Examining the language of internalized racism and the policing of ethnic identity on Twitter. *Social Semiotics*, 26(5), 505–523. <https://doi.org/10.1080/10350330.2015.1126046>
- Nguyen, T. T., Criss, S., Michaels, E. K., Cross, R. I., Michaels, J. S., Dwivedi, P., Huang, D., Hsu, E., Mukhija, K., Nguyen, L. H., Yardi, I., Allen, A. M., Nguyen, Q. C., & Gee, G. C. (2021). Progress and push-back: How the killings of Ahmaud Arbery, Breonna Taylor, and George Floyd impacted public discourse on race and racism on Twitter. *SSM - Population Health*, 15, Article 100922. <https://doi.org/10.1016/j.ssmph.2021.100922>
- Pack-Brown, S. P. (1999). Racism and White counselor training: Influence of White racial identity theory and research. *Journal of Counseling and Development*, 77(1), 87–92. <https://doi.org/10.1002/j.1556-6676.1999.tb02425.x>
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information Communication and Society*, 19(3), 307–324. <https://doi.org/10.1080/1369118X.2015.1109697>
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic inquiry and word count. LIWC 2001*. Manual Erlbaum Publishers.
- Pennebaker, J. W., Boyd, R. L., Jordan, K. N., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. University of Texas at Austin.
- Pew Research Center. (2020). *Amid protests, majorities across racial and ethnic groups express support for the Black Lives Matter movement*. Pew Research Center’s Social & Demographic Trends Project. Retrieved January 16, 2021, from <https://www.pewsocialtrends.org/2020/06/12/amid-protests-majorities-across-racial-and-ethnic-groups-express-support-for-the-black-lives-matter-movement/>

- Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The rise of consumer health wearables: Promises and barriers. *PLoS Medicine*, *13*(2), Article e1001953. <https://doi.org/10.1371/journal.pmed.1001953>
- Priante, A., Ehrenhard, M. L., van den Broek, T., & Need, A. (2018). Identity and collective action via computer-mediated communication: A review and agenda for future research. *New Media & Society*, *20*(7), 2647–2669. <https://doi.org/10.1177/1461444817744783>
- Ray, R., Brown, M., Fraistat, N., & Summers, E. (2017). Ferguson and the death of Michael Brown on Twitter: #BlackLivesMatter, #TCOT, and the evolution of collective identities. *Ethnic and Racial Studies*, *40*(11), 1797–1813. <https://doi.org/10.1080/01419870.2017.1335422>
- Reyes-Menendez, A., Saura, J. R., & Alvarez-Alonso, C. (2018). Understanding users' opinions of #World Environment Day on Twitter: A topic-based sentiment analysis approach. *International Journal of Environmental Research and Public Health*, *15*(11), Article 2537. <https://doi.org/10.3390/ijerph15112537>
- Reyes-Menendez, A., Saura, J. R., & Filipe, F. (2020). Marketing challenges in the #MeToo era: Gaining business insights using an exploratory sentiment analysis. *Heliyon*, *6*(3), Article e03626. <https://doi.org/10.1016/j.heliyon.2020.e03626>
- Reyes-Menendez, A., Saura, J. R., & Thomas, S. B. (2020). Exploring key indicators of social identity in the #MeToo era: Using discourse analysis in UGC. *International Journal of Information Management*, *54*, Article 102129. <https://doi.org/10.1016/j.ijinfomgt.2020.102129>
- Schaffer, G. (2014). Legislating against hatred: Meaning and motive in section six of the Race Relations Act of 1965. *Twentieth Century British History*, *25*(2), 251–275. <https://doi.org/10.1093/tcbh/hws053>
- Schwartz, S. A. (2020). Police brutality and racism in America. *EXPLORE: The Journal of Science and Healing*, *16*(5), 280–282. <https://doi.org/10.1016/j.explore.2020.06.010>
- Shaw, H., Ellis, D. A., Kendrick, L. R., Ziegler, F., & Wiseman, R. (2016). Predicting smartphone operating system from personality and individual differences. *Cyberpsychology, Behavior and Social Networking*, *19*(12), 727–732. <https://doi.org/10.1089/cyber.2016.0324>
- Suitner, C., Badia, L., Clementel, D., Iacovissi, L., Migliorini, M., Casara, B. G. S., Solimini, D., Formanowicz, M., & Erseghe, T. (2013). *The rise of# climateaction in the time of the FridaysForFuture movement: A semantic network analysis*. Retrieved June 10, 2021, from https://elea.rnring.unipd.it/scienzeumane/pluginfile.php/756888/mod_resource/content/0/The_evolution_of_online_activism_on_climate_change_after_fridays4future%20%281%29.pdf
- Tajfel, H. (1978). *Differentiation between social groups*. Academic Press.
- Tajfel, H. (1979). Individuals and groups in social psychology. *The British Journal of Social and Clinical Psychology*, *18*(2), 183–190. <https://doi.org/10.1111/j.2044-8260.1979.tb00324.x>
- Tajfel, H., & Turner, J. (1979). An integrative theory of inter-group conflict. In J. A. Williams & S. Worchel (Eds.), *The social psychology of inter-group relations* (pp. 33–47). Wadsworth.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. <https://doi.org/10.1177/0261927X09351676>
- Turner, J. C. (1981). Towards a cognitive redefinition of the social group. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *1*(2), 93–118.
- Twitter. (2021). *Privacy policy: Public information*. Retrieved January 18, 2021, from <https://twitter.com/en/privacy>
- van Atteveldt, W., van der Velden, M. A. C. G., & Boukes, M. (2021). The validity of sentiment analysis: Comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms. *Communication Methods and Measures*, *15*(2), 121–140. <https://doi.org/10.1080/19312458.2020.1869198>
- Wang, R., Liu, W., & Gao, S. (2016). Hashtags and information virality in networked social movement: Examining hashtag co-occurrence patterns. *Online Information Review*, *40*(7), 850–866. <https://doi.org/10.1108/OIR-12-2015-0378>
- Yassine, M., & Hajj, H. (2010). A framework for emotion mining from text in online social networks. In *proceedings of IEEE international conference on data mining workshops*. NSW.

Received August 26, 2021

Revision received February 7, 2022

Accepted February 9, 2022 ■