Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

The affiliative use of emoji and hashtags in the Black Lives Matter movement: A Twitter case study

Mark Alfano (mark.alfano@gmail.com)
Macquarie University https://orcid.org/0000-0001-5879-8033

Ritsaart Reimann
Macquarie University https://orcid.org/0000-0003-4742-2887

Ignacio Quintana
Australian National University

Marc Cheong
University of Melbourne https://orcid.org/0000-0002-0637-3436

Colin Klein
Australian National University

Article

Keywords: social media, protest, emoji, hashtags, social movements, Black Lives Matter

Posted Date: July 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-741674/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

The affiliative use of emoji and hashtags in the Black Lives Matter movement: A Twitter case study

Mark Alfano[1,*], **Ritsaart Reimann**[1], **Ignacio Ojea Quintana**[2], **Marc Cheong**[3], **and Colin Klein**[2]

[1]Department of Philosophy, Macquarie University, 25 Wally's Walk, Macquarie Park NSW 2109, AUSTRALIA
[2]School of Philosophy, Australian National University, Canberra ACT 2600, AUSTRALIA
[3]Centre for AI and Digital Ethics, Faculty of Engineering and IT, University of Melbourne, Parkville VIC 3010, AUSTRALIA
[*]corresponding author(s): Mark Alfano (mark.alfano@gmail.com)
[1,2,3]All authors contributed equally to this work

ABSTRACT

Protests and counter-protests seek to draw and direct attention and concern with confronting images and slogans[1–3]. In recent years, as protests and counter-protests have partially migrated to the digital space, such images and slogans have also gone online[4–6]. Two main ways in which these images and slogans are translated to the online space is through the use of emoji and hashtags. Despite sustained academic interest in online protests[7–9], hashtag activism[10–12] and the use of emoji across social media platforms[13–15], little is known about the specific functional role that emoji and hashtags play in online social movements. In an effort to fill this gap, the current paper studies both hashtags and emoji in the context of the Twitter discourse around the Black Lives Matter movement.

Introduction

Protests and counter-protests have long made effective use of images and slogans[1–3]. As protests and counter-protests have partially migrated to the digital space, such images and slogans have also gone online[4–6]. Two important ways in which these images and slogans appear is through the use of emoji and hashtags.

Some emoji have readily identifiable offline counterparts—such as the raised fist, which was first deployed as a standalone image in protests in the San Francisco Bay Area and Harvard University in 1968-9[1], and which now has its own emoji. Similarly, some hashtags (like #BlackLivesMatter) reflect well-known offline slogans. Indeed, on Twitter since 2016, this hashtag is automatically enhanced with a small emoji-like sticker featuring a trio of raised black, brown, and white fists. The use of other emoji and hashtags, however, can be more obscure. To shed some light on their functions, we here study emoji and hashtags embedded in tweets associated with the Black Lives Matter protests in 2020, including the right-wing backlash to those protests. We analyze a dataset covering both the lead-up to and the aftermath of the 25 May 2020 murder of George Floyd by Officer Derek Chauvin in Minneapolis. The nine-minute video of Floyd's murder set off a firestorm of activity both in the streets and online (see also *Supplementary Material*).

While the use of hashtags as an organizing mechanism in online activism has been studied[10], the role of emoji in social movements has, to the best of our knowledge, received no academic attention. At the same time, as emoji have become an increasingly popular form of communication, a growing body of work that tracks the various types and uses of emoji has emerged[13–15]. Extrapolating from this literature, we present and test four hypotheses regarding the use of emoji in online activism. First, emoji might be used for their straightforwardly semantic content, functioning as compact logograms that efficiently convey meaning within the tight character constraints of Twitter (**H1**). Second, emoji and hashtags might be employed to disambiguate tone in the context of highly-charged discursive exchanges (**H2**). This follows from the observation that emoji and hashtags enable us to track important linguistic subtleties—such as sarcasm and humour—that are otherwise hard to detect in computer-mediated communication[16–18]. Third, emoji might operate on par with ostensive interlocutory gestures that aim at drawing and directing attention to the content of a given message[16,19–21] (**H3**). Finally, emoji and hashtags might function as primarily affiliative gestures, drawing attention to the author of the tweet and demonstrating their *bona fides* within their group (**H4**). This fourth function is especially relevant with respect to the use of skin-tone modifiers, which have been associated with enhanced self- and group-identification[22]. Given that hashtags can be understood as organizing mechanisms that

connect people with shared interests[10] and systematically codify their shared interests under a common descriptor[23], people who employ the same hashtags may also do so to signal that they are members of the same community.

In addition to testing these four hypotheses (**H1-H4**), we are also interested in the broader question of whether there are discernible and meaningful differences in the ways that the various groups of participants involved in the Black Lives Matter discourse use emoji and hashtags. Accordingly, we employ social network analyses, classification algorithms, natural language processing techniques, conditional probability modelling, and regression models to answer the following two questions:

- **RQ1**. Are there informative and meaningful differences in the way that the various communities involved in the Black Lives Matter discourse employ emoji and hashtags?
- **RQ2**. Assuming there are differences, what is their functional significance?

Our work suggests that communities use emoji and hashtags in distinctive and meaningful ways. Further, it shows that emoji and hashtags are something of a mixed bag: they tend to decrease engagement with tweets, but increase engagement with the other tweets of authors who use them. This suggests that emoji and hashtags might play a primarily affiliative role in the communities we studied.

Methods

Data collection and cleaning

We queried the Twitter Streaming API with a series of Black Lives Matter (BLM)-related keywords, hashtags, and short expressions in a window between January and July 2020. We used a sliding window to take into account that between 80%-90% of retweets occur within 5-7 days, with diminishing returns beyond[24]. The dataset comprised ∼4.6M original tweets between January 13th and July 18th and ∼94.5M retweets from January 18th to July 23rd; these tweets were produced by ∼2.0M distinct authors. After the murder of George Floyd (May 25th 2020), the number of daily tweets increased by several orders of magnitude (from ∼255k to ∼4.35M).

Social Network Construction

We generated a *retweet network*[25], a weighted directed network where nodes are authors and the weight of an edge from node u to node v represents the number of times that user v retweeted user u. Self-retweeting was disregarded. Given this definition, users who retweeted but who did not author any tweets could not be nodes in the network. Having built the retweet network, we took the largest connected component (∼689k nodes, ∼13M edges) for further analysis. (See *Supplementary Material* for technical details).

Community Clustering and retweet statistics

To find clusters, we used *igraph*[26] and the Python *leidenalg* package which implements the Leiden community detection algorithm[27]. We found first-level clusters using Modularity Vertex Partitioning, preserving clusters with more than 10% of the original nodes. This gave 4 clusters, covering 83% of the graph. Next, we manually inspected the 100 most-influential nodes within each group. Based on this, we characterize the four communities as follows.

Activists: this cluster represents the core of the movement and reflects the grass-roots nature of Back Lives Matter. It features a heterogeneous collection of individual activists, many of whom explicitly endorse the movement by placing #BlackLivesMatter in their profile bio.

Progressives: this cluster contains a range of high-profile individuals and organizations that are generally supportive of the Black Lives Matter movement. In addition to prominent Democratic politicians (e.g., Kamala Harris, Bernie Sanders) and liberal media outlets (e.g., ABC, NBC, CBS), there are various non-profit organizations and legal aid foundations (e.g., ACLU).

Reactionaries: this cluster features conservative politicians and public figures (e.g., Donald Trump, James Woods), as well as right-leaning media outlets (e.g., Breitbart News, The Gateway Pundit), anti-BLM activists, supporters of the police, and a large number of conspiratorially-minded and openly racist individuals. Additionally, this community features an exceptionally large number of suspended accounts. Though we cannot interpret these accounts, it stands to reason that they violated Twitter's community guidelines regarding false (conspiratorial) and offensive (xenophobic) content.

Boosters: this cluster comprises a diverse collection of individuals whose primary involvement with the Black Lives Matter movement seems to consist of link-sharing and fundraising. With respect to fundraising, we identify a large contingent of fans of the Korean pop phenomenon, commonly called K-pop. While the link between K-pop and Black Lives Matter might seem tenuous at first, the fact that one such band—Bangtan Sonyeondan (BTS)—donated

a million dollars to the Black Lives Matter foundation, and encouraged its followers to match that donation, explains this group's engagement[28].

Based on these initial impressions, and in effort to test whether these different communities use emoji and hashtags in differential ways, we next constructed a classifier.

Classifier Construction

The prospect of using emoji to train classifiers has received considerable attention and produced impressive results[13,29]. Classifiers are able to disambiguate the tone and intention of a given statement by differentiating between positive and negative valences of the same emoji[30–32]. Extending this line of research, we investigated **RQ1** by constructing a classifier of our own, with the goal of using emoji and hashtags for community detection. The goal was to provide a measurement of the linguistic cogency of the communities we identified through modularity analysis, and to provide a measure of entropy that shows how much additional information about community membership is contained within each type of communication.

We constructed a standard data-science workflow in Python for automated text classification, using *Tensorflow*[33] for neural network approaches and *Scikit-Learn*[34] for classical classification techniques.

Excluding members of the Boosters cluster because there were too few of them for meaningful analysis, we generated a document containing the plain text of all tweets by each user in the network, together with the emoji and hashtags they used. Tweet text was pre-processed using a typical text processing workflow (removing non-alphanumeric, non-hash identifier (#), and non-emoji characters, standardizing case, etc.), and emoji were encoded using Python's *emoji* package[35]. Further details on the data sampling, train/test splits, as well as evaluation, are available in the *Supplementary Material*.

Distinctiveness

To examine the distinctiveness of each of the communities, we first determined the 15 most popular emoji for each community c and then for each emoji e. From that set we calculated $Pr(c|e)$. This gives a rough measure of how much information emoji and hashtag use carries about community membership, as well as which specific emoji and hashtags are most distinctive within each community.

Emoji and Hashtag Co-occurrence Network Construction

Though informative in its own right, we noted that a single tweet may well contain multiple emoji, multiple hashtags, and various combinations thereof. On this score, previous research has found that multiple hashtags are often combined to draw attention to interrelated issues[23,36]. Likewise, various emoji are frequently used together to refine a user's stance, attitude, or sentiment[13–15](see also *Supplementary Material)*. Accordingly, it is informative to consider whether and in what ways the various communities combine emoji and hashtags. To this end, we conducted a co-occurrence analysis.

For each of the four communities identified by the social network analysis, we identified the fifty most-commonly-used emoji and the fifty most-commonly-used hashtags. We then constructed a co-occurrence network of these emoji and hashtags for each community, where the nodes are either emoji or hashtags and the edges between them represent co-occurrence in the same tweet. These networks enable us to answer the question, "What do people (in this community) talk about, and how do they talk about it, when talking about X?" We then visualized these networks using Gephi[37] and the Image Preview plugin[38] to provide a snapshot of the imagery and slogans most distinctive of each community.

Results

Community Structure

As Figure 1 shows, the retweet network is bipolar, with Activists, Progressives, and Boosters on one side and Reactionaries on the other. This outcome is consistent with numerous findings from political science suggesting substantial polarisation in the political landscape[39,40].

As Figure 2 shows, the murder of George Floyd triggered an outpouring of tweets – first among Activists, then among Progressives and Boosters, and finally among Reactionaries. The decay in the volume of tweets among these groups is also worth considering, as Reactionaries decay much less quickly than the three other communities, suggesting a self-sustaining dynamic within that community.

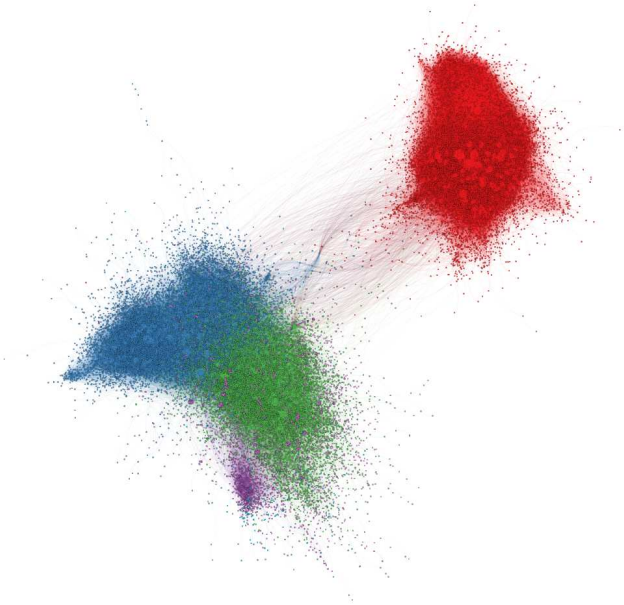


Figure 1. Community rendering. Green: Activists, Blue: Progressives, Red: Reactionaries, Purple: Boosters. *ForceAtlas2* used for layout.

Classification Task

The results for the classification task are shown in Table 1 and warrant the following observations.

To begin, all classifiers with all data types greatly outperformed random classification, which for this task had an expected accuracy of ∼0.3333. Even the worst performing classifier—Linear Stochastic Gradient Descent trained on emoji—obtained an accuracy greater than 0.5. More generally, we note that deep learning techniques marginally outperformed traditional approaches. More specifically, we find that the best-performing classifiers were GRU and LSTM neural architectures which take ordering into account, suggesting that the order in which emoji are presented in a given tweet makes a difference, perhaps because they occur in decreasing order of priority for the user.

With respect to **RQ1**, these results confirm that the various communities involved in the Black Lives Matter discourse use emoji and hashtags in distinctive ways. What is more, emergent patterns of emoji and hashtag use correspond with patterns of retweet behaviour (Figure 1), indicating that retweet engagement is associated with the use of emoji and hashtags. In addition to this, we find that hashtags are a particularly powerful marker of

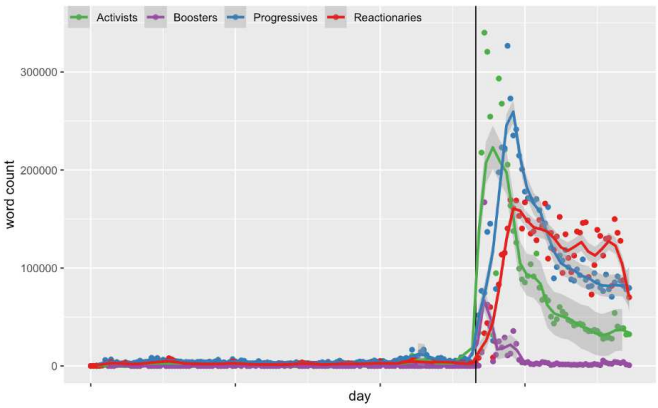
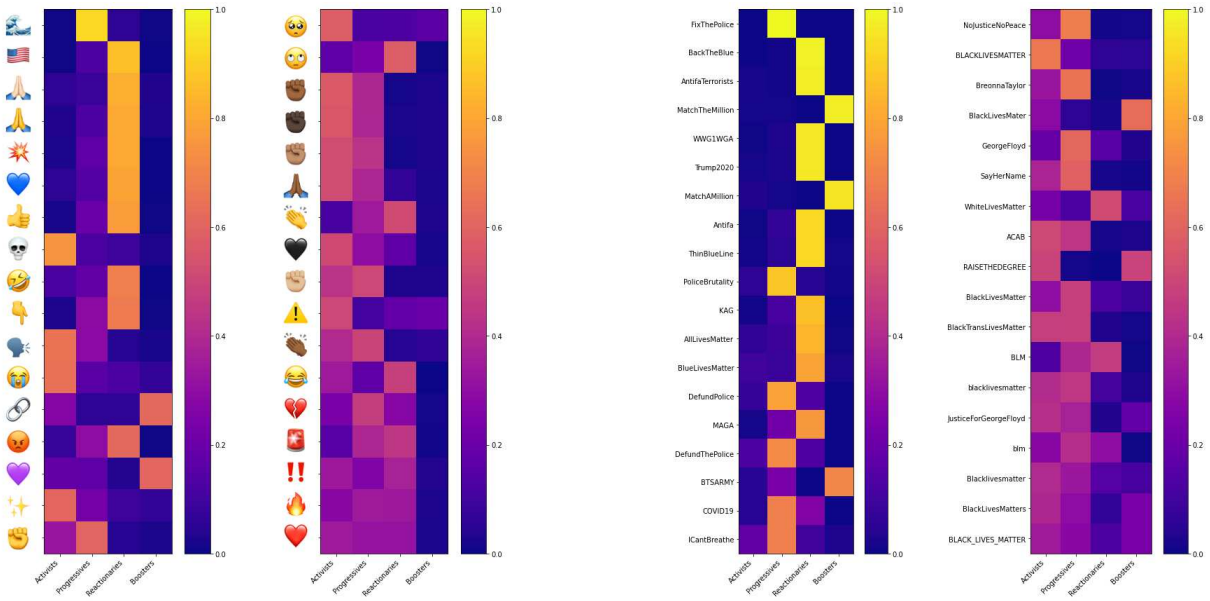


Figure 2. Daily word-count sums of tweets associated with different communities. The vertical line indicates 25 May 2020, the day on which George Floyd was murdered. Date range: 17 January 2020 to July 23 2020.

Table 1. Classification task results

Data Type	Log Reg.	Rand. For.	Linear SGD	DNN	RNN	LSTM
emoji (E)	0.62	0.61	0.56	0.58	0.64	0.64
hashtags (H)	0.72	0.70	0.70	0.71	0.73	0.73
$E+H$	0.72	0.70	0.70	0.71	0.73	0.73
text	0.74	0.69	0.73	0.72	0.74	0.73
all data	0.76	0.73	0.76	0.76	0.77	0.76



(a) emoji **(b)** hashtags

Figure 3. Conditional probability of community membership given emoji (left) and hashtags (right) for top 15 most popular in each community.

community membership and that, taken together, emoji and hashtags are roughly as informative as text in this regard.

Distinctiveness

Figure 3 shows the conditional probability of group membership given both emoji (3a) and hashtag (3b) usage. For each community, there are both emoji and hashtags that are extremely diagnostic of cluster membership. For example, the probability of belonging to the Reactionary group conditional on using a US Flag emoji (🇺🇸) is 87%. Angry 'pouting' (😡) and contemptuous 'rolling eyes' (🙄) emoji are likewise significantly more likely to be used by members of the Reactionary cluster. Moreover, many of the hashtags used by Reactionaries are virtually pathognomic, particularly those associated with the QAnon conspiracy theory. Though the distinction between Progressives and Activists is less sharp, there are discernible differences. #DefundThePolice and the blue wave emoji (🌊) are both strongly associated with Progressives, while the sparkle emoji (✨) is more common among Activists. We also note that the use of skin-tone modifiers is coupled to conditional probabilities of group membership, with Activists using darker modifiers than Progressives, and Progressives using darker modifiers than Reactionaries. In fact, Reactionaries frequently use the 'light' (as opposed to 'medium-light') skin-tone, which is rare in other communities. Therefore, despite only slight probabilistic variance for a number of emoji and hashtags, there are systematic differences that are large enough that, on aggregate, patterns of emoji and hashtag usage are good markers of group affiliation.

Emoji and Hashtag Co-Occurrence

As Figure 4 shows, the emoji and hashtags favored by these communities differ in meaningful ways and cluster together even at the level of individual tweets.

Activists (Figure 4a) were strongly associated with the the use of raised-fist emoji (✊,✊🏻,✊🏼,✊🏽,✊🏾,✊🏿), and frequently used darker skin-tone modifiers when using other emoji. There is also a notable focus on LGBTQ issues via the use of the rainbow of hearts (a series of heart emoji with different colours) and the rainbow emoji itself (🌈), along with hashtags such as #blacktranslivesmatter and #pride2020. (See *Supplementary Material* for context).

Like the Activists, Progressives (Figure 4b) favored raised-fist emoji (✊,✊🏻,✊🏼,✊🏽,✊🏾,✊🏿), hearts of different colors, and warning signals such as exclamation points (‼️). However, they tended to use lighter skin-tones and the default cartoon-yellow skin-tone more than their Activist counterparts. In addition, Progressives used the down-pointing finger (👇,👇🏻,👇🏼,👇🏽,👇🏾,👇🏿) more than Activists, suggesting that they were seeking less to demand recognition for themselves and more to redirect attention and concern for the demands of recognition being made by their Activist allies. Progressives also used emoji associated with electoral politics, especially the blue heart (💙) and the blue wave (🌊), both of which are associated with support for Democratic political candidates[41].

In contrast to both Activists and Progressives, Reactionaries (Figure 4c) did not appear to center their attention on any single topic. Instead, different elements of this group pushed back against the Black Lives Matter movement in different ways. For instance, we see a large contingent drawing attention to the police via both emoji (blue heart 💙, police officer 👮, police cruiser 🚔) and hashtags (e.g., #backtheblue, #bluelivesmatter), while other tweets seem to focus more on electoral politics, either by identifying with the Trump reelection campaign (e.g., #kag2020, #trump2020) or by derogating enemies (e.g., #democratsaredestroyingamerica, #liberalismisamentaldisorder). In addition, we see the centrality of the QAnon conspiracy theory to this community in its use of hashtags such as #qanon, #wwg1wga ("where we go one, we go all," a popular slogan in the QAnon movement), and #thegreatawakening. The Reactionary community does not seem to unite around a single cause or message; instead, they are primarily defined in terms of what they oppose. It appears that this reactionary movement in part reflects an attempt to hijack the Black Lives Matter discourse in order to bootstrap its own political agenda[42].

It is also worth remarking that Progressives and Reactionaries used distinctive hashtags to collate tweets about the protests in Seattle, Washington: Progressives employed #seattleprotest and #seattleprotests, whereas Reactionaries used #chop and #chaz (referring to the anarchist zone that protesters set up on 8 June 2020, and which the Seattle Police Department cleared on 1 July 2021 after an unlawful assembly was declared). This suggests a potential filter bubble effect, in which users who followed one set of hashtags about the Seattle protests would encounter the corresponding set of information and sentiment about it, while users who followed the other hashtags would encounter a radically different set of information and sentiment about the same topic.

Finally, the Booster community (Figure 4d) exhibits many similarities with Activists and Progressives. For instance, they use raised fists (✊,✊🏻,✊🏼,✊🏽,✊🏾,✊🏿) , down-pointing fingers (👇,👇🏻,👇🏼,👇🏽,👇🏾,👇🏿) and a range of exclamation points (‼️, ❗) to draw and redirect attention and concern. In contradistinction to Activists and Progressives, however, Boosters make more use of the link emoji (🔗, associated with links to websites for petitions and donations). As outlined earlier, a substantial chunk of these users are associated with the K-Pop band BTS, which activated its followers by making a sizeable donation to the Black Lives Matter foundation. In addition to the link emoji, this affiliation is reflected through the purple heart emoji (💜, a symbol of BTS) and a range of hashtags referencing BTS's million-dollar donation and the fans' efforts to match this donation.

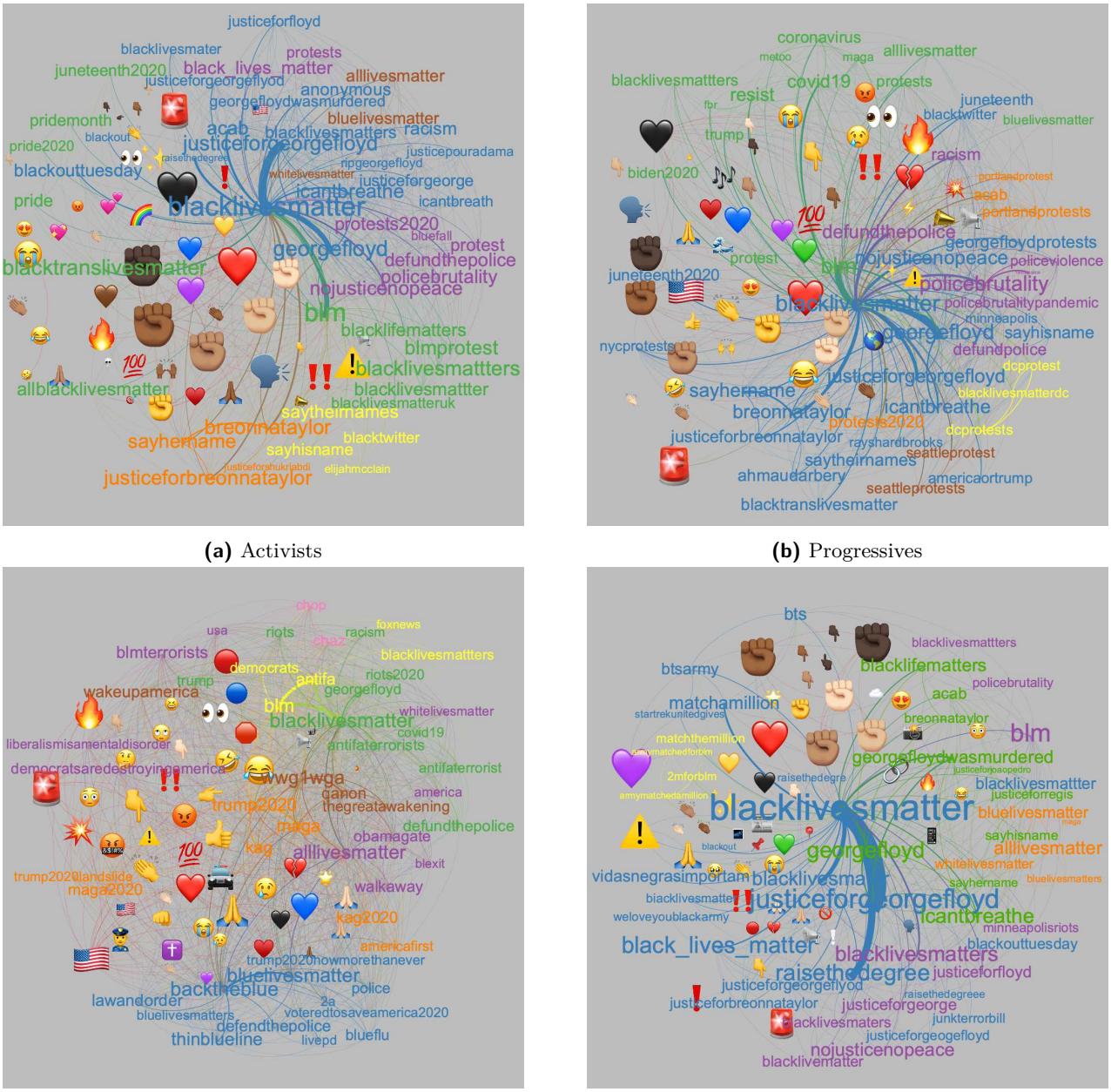
Usage Statistics

Descriptive statistics of emoji/hashtag (EH) use are given in Table 2. Across communities, about 25% of users have at least one emoji in one of their tweets. Hashtag usage is much higher, averaging about 57% but ranging as high as 90% in the Booster group. For all communities, both emoji- and hashtag-users have a higher PageRank, suggesting that they are more embedded in the network as a whole.

Retweet statistics show several curious patterns for both emoji (Table 3) and hashtag (Table 4) usage. For ease of interpretation, we divide these into three patterns:

Type I patterns: when both the following conditions are met:

- EH-tweets by EH-users have fewer retweets than either non-EH tweets by EH users or by non-EH users (e.g., Activists who sometimes use emoji have their emoji-using tweets retweeted on average 55.39 times, compared to 74.76 times for non-emoji using Activists);
- where EH-users' *non*-EH tweets are more frequently retweeted than either (e.g., the non-emoji tweets by Activists who used emoji in other tweets are retweeted far more than either; by 105.6 times).



(a) Activists

(b) Progressives

(c) Reactionaries

(d) Boosters

Figure 4. Co-occurrence network of the emoji and hashtags used by by each cluster. Node size = degree. Text and edge color = modularity class.

Table 2. Descriptive statistics for emoji and hashtag use. $\%E$ and $\%H$ give percentage of users in each group who use emoji and hashtags, respectively. E-PR and ¬E-PR are mean pagerank of emoji and non-emoji users, each $\times 10^{-6}$; H-PR and¬H-PR give the same for hashtags.

Community	$\%E$	$\%H$	E-PR	¬E-PR	H-PR	¬H-PR
Overall	26	57	3.69	1.92	2.86	1.72
Activists	26	53	3.75	2.18	2.97	1.70
Progressives	26	67	4.25	1.96	2.52	1.22
Reactionaries	25	51	3.06	1.48	2.65	1.42
Boosters	25	90	3.76	2.11	2.86	1.72

Table 3. Retweet statistics for emoji and non-emoji users. E-w and E-w/o: mean retweets of tweets by emoji-users with and without emoji, respectively. ¬E stands for mean retweets for users who never use emoji. See main text for definition of type.

Community	E-w/	E-w/o	¬E	Type
Overall	34.96	47.57	37.12	I
Activists	55.51	105.79	74.69	I
Progressives	22.24	30.74	17.6	II
Reactionaries	27.07	26.52	25.29	III
Boosters	61.86	44.71	46.77	III

Table 4. Mean retweets for hashtag and non-hashtag users. Labels as per table 3.

Community	H-w/	H-w/o	¬H	Type
Overall	31.58	51.21	45.07	I
Activists	63.19	96.33	91.92	I
Progressives	19.1	33.04	18.44	II
Reactionaries	24.65	34.81	16.22	II
Boosters	39.45	79.65	68.73	I

Type II patterns: like Type I, save that EH-tweets by EH-users are roughly equivalent or have a slight advantage over non-EH users' tweets.

Type III patterns: anything not in Types I and II.

Note that, with two exceptions, both the overall and community-level patterns fall into either Type I or Type II. That is, in general, EH-usage does not provide a substantial advantage in mean retweets, and often gives a substantial disadvantage, compared to tweets by people who never use them. However, EH-users' *other* tweets are retweeted on average much more than non EH-users.

The two Type-III exceptions are found in Reactionaries' use of emoji (which appears to make no particular difference to retweets) and Boosters' use of emoji (which overall confers substantial advantages and may reflect the multilingual nature of this community and the fact that emoji can serve as a sort of *lingua franca*).

Discussion

With respect to **RQ1**, Table 1, Figure 3, and Figure 4 provide compelling evidence that there are informative and meaningful differences in the way that the various communities involved in the Black Lives Matter discourse employ emoji and hashtags. The results of our classification task confirm that there are distinct patterns in emoji and hashtag usage across the various communities and that these differences are informative with respect to detecting community membership. Furthermore, we find that hashtags are as informative as textual tweets in inferring communities, while emojis are almost as informative. This is significant for two reasons.

First, in terms of informational entropy, emoji and hashtags are more compact than text. Hence, they provide a computationally-efficient way of determining community membership. This is useful when dealing with large data-sets like the one analyzed here. Additionally, and in contrast to text, emoji and hashtags are freestanding expressions that can be interpreted even before considering syntactic complexities like word order or sentence structure. Based on nothing more than conditional probability analyses (Figure 3) and a co-occurrence matrix of emoji and hashtags (Figure 4), we were able to learn a great deal about the various communities involved in the Black Lives Matter discourse.

In addition to capturing the most obvious slogans (e.g., #BlackLivesMatter) and counter-slogans (e.g., #AllLivesMatter), our analysis reveals that the perspectives of queer and trans people of colour are conveyed through the use of emoji and hashtags. This result is significant in that it reflects the movement's emphasis on giving voice to historically-marginalised victims of oppression[43–45]. Consonant with previous research, we are also able to confirm that the right-wing backlash to the Black Lives Matter movement is spearheaded by loosely-related, racist, and conspiratorially-minded conservative partisans who come together in support of the police and former president

Donald Trump[42,46]. More surprisingly, Figure 4 accurately captured the link-sharing and fundraising efforts of a small but loud Booster contingent, as well as the political ambitions of the Activists' Progressive allies.

Next, consider the use of interlocutory gestures and skin-tone modifiers. With respect to interlocutory gestures, we note the differential usage of 'raised-fist' (✊,✊🏻,✊🏼,✊🏽,✊🏾,✊🏿) and 'pointing-finger' emoji (👇,👇🏻,👇🏼,👇🏽,👇🏾,👇🏿). Compared to Activists' focus on the attention-grabbing raised-fist, Progressives and Boosters use the point-finger much more frequently. On the assumption that pointing-fingers direct rather than draw attention, we infer that one of the contributions of Progressives and Boosters to the Black Lives Matter movement online is redirecting attention to Activists. Interestingly, these same interlocutory emoji reveal a great deal with respect to skin-tone modification. The fact that the Activist cluster uses darker modifiers than the other communities lends support to Robertson et. al's[22] observation that skin-tone modifiers enhance self- and group-identification. At the same time, we note the conspicuous use of non-modified, yellow, 'default' emoji within the Reactionary cluster. In light of this community's racist attitudes, it is worth considering why its members rarely employ the skin-tone modification function to signal their whiteness. Here, we flag the possibility that non-modified, yellow emoji might be a manifestation of the ideology of colorblindness. In much the same way that the ideology of colorblindness masks racism by rejecting it[47,48], non-modified emoji may maintain a pretence of neutrality by ignoring any alternative. Furthermore, it may be that widespread usage of this 'default' setting among white people implicitly equates whiteness with normality.

In terms of **RQ2**, we started by proposing four hypotheses about the functional roles of emoji and hashtags. In light of our results, we come to the following conclusions. First, **H1** proposed that emoji and hashtags are principally used for their semantic properties. There is some evidence for this, e.g., Progressives' use of the blue wave (🌊) to predict and encourage the 'Blue Wave' election of Democratic politicians. As for **H2**, which proposed that emoji are used to disambiguate tone, our results suggest that this is not widespread. Indeed, if this were the case, we would expect to see more emoji such as sarcasm (🙃) and disdain (😕) to modify the tone of a retweet. Looking at Figure 4, this expectation is not borne out. **H3** posited that emoji might operate on par with ostensive interlocutory gestures that draw and direct attention. While this hypothesis holds to a certain extent in each of the four communities, the prevalence of Type I and II patterns (Table 3) suggests that this form of emoji use is not a sustainable strategy. To the contrary, we find that emoji use is by and large negatively correlated with retweet count and in effect diminishes attention via engagement.

Finally, **H4** proposed that emoji and hashtags are primarily affiliative gestures that call attention to individuals as *bona fide* members of a given group. In light of the evidence, this strikes us as the most plausible response to **RQ2**. For instance, Table 1, Figure 3, and Figure 4 all suggest that emoji are reliable markers of group affiliation. Hence, it stands to reason that they are also used as such. Further evidence in support of this conclusion can be inferred from the results of our retweet analysis (see Table 3 and Table 4). Recall here that, even though using emoji and hashtags in a given tweet generally decreases the amount of attention awarded to that tweet, doing so simultaneously *increases* the prospect of receiving more attention for future tweets.

Accordingly, it stands to reason that emoji and hashtags can be interpreted as affiliative gestures that impose an initial cost on signaling one's commitment to the group. Indeed, as as Figure SI1 shows, there is a substantial overall retweet penalty for both emoji and hashtags, with a correlation of -0.73 and -0.66, respectively between number of items and mean retweets. However, signaling does so with the payoff of increasing one's standing and following within that group. If this is right, our results suggest an interesting tension between **H3** and **H4**. On the one hand, using emoji and hashtags signals commitment to one's group and increases the prospect of receiving more attention from that group at some future moment. At the same time, it decreases the amount of attention awarded to one's present tweets.

In sum, we started by noting that hashtags can be understood as indexing mechanisms that systematically codify certain topics under a common descriptor[23] and thus potentially connect people who are interested in those topics[10]. Hence, we expected, and have now established, that hashtags are a strong marker of group affiliation. This is especially true for hashtags used by the conspiratorial wing of the Reactionary community (e.g., #qanon, #wwg1wga) and the K-Pop wing of the Booster community (e.g., #matchthemillion, #matchamillion). We note that the use of hashtags to signal affiliation is not entirely foolproof. For instance, our analyses confirm that the Boosters at one point briefly dragooned various Reactionary hashtags such as #bluelivesmatter and #whitelivesmatter. However, trolling is not likely to be a sustainable strategy in the long term, and we would expect that the attention economy would quickly discard such efforts (as in fact happened in this case).

At the same time, we are surprised to see that this indexing function does not drive engagement. Contrary to previous research[49], we conclude that hashtags are negatively correlated with retweet count and serve a primarily affiliative function—at least as they are used in connection with the Black Lives Matter movement.

With respect to emoji, our expectations are again only partially confirmed. On the one hand, emoji such as the

raised fist (✊,✊🏻,✊🏼,✊🏽,✊🏾,✊🏿) signal in-group membership and community alliances among Activists, Progressives, and Boosters. Likewise, the American flag (🇺🇸) and various police-related emoji (👮‍♀️) are all clear makers of belonging to the Reactionary community. Conversely, we find that interlocutory gestures like pointing fingers (👇) and exclamation marks (‼️) do not succeed at drawing and directing attention to specific tweets. Hence, emoji too can be understood as principally performing a kind of affiliative function. Together, these results suggest that emoji and hashtags play a complex role in the attention economy, operating at the level of both individual tweets and their authors.

Future research could, amongst other things, examine the ongoing activities of the communities studied in this paper. Examples include the November 2020 US Presidential Election; the 6 January 2021 insurrection by Reactionaries; and Twitter's suspension of Donald Trump and purge of QAnon accounts.

Additionally, the use of emoji and hashtags in bios—as opposed to tweets—remains under-studied. One hypothesis prompted by the current research is that here too, emoji and hashtags would play an affiliative role. It would also be illuminating to test whether the affiliative use of emoji and hashtags generalizes across other topics (or is constrained to the Black Lives Matter movement), and to examine discourse on other platforms to test whether this use is confined to Twitter.

These directions for future research reflect some of the limitations of the current study, which does not cover the full multi-year history of the Black Lives Matter movement on Twitter, let alone on all platforms.

As protests and counter-protests continue to migrate to the digital space, the need to understand the use of emoji and hashtags in online activism becomes increasingly important. The current paper responds to this need and provides novel insights into the use of emoji and hashtags in online activism that we hope will be useful for further research.

Code availability

For all studies using custom code in the generation or processing of datasets, a statement must be included under the heading "Code availability", indicating whether and how the code can be accessed, including any restrictions to access. This section should also include information on the versions of any software used, if relevant, and any specific variables or parameters used to generate, test, or process the current dataset.

References

1. Rippon, J. *The Art of Protest* (Palazzo, 2019).
2. Young, R. *Make Art Not War: Political Protest Posters from the Twentieth Century* (NYU Press, 2016).
3. Siegler, B. *Signs of Resistance: A Visual History of Protest in America* (Artisan, 2018).
4. Tufekci, Z. *Twitter and Tear Gas: The Power and Fragility of Networked Protest* (Yale University Press, 2017).
5. Kübler, J. A review of zeynep tufekci - twitter and tear gas: The power and fragility of networked protest. *Int. J. Polit. Cult. Soc.* **32**, 357–63, https://doi.org/10.1007/s10767-019-9316-3 (2019).
6. Zeynep, T. A response to johanne kübler's a review of zeynep tufekci - twitter and tear gas: The power and fragility of networked protest. *Int. J. Polit. Cult. Soc.* **32**, 365–69, https://doi.org/10.1007/s10767-019-9317-2 (2019).
7. Kidd, D. & McIntosh, K. Social media and social movements. *Sociol. Compass* **10**, 785–794 (2016).
8. Freelon, D., McIlwain, C. & Clark, M. Quantifying the power and consequences of social media protest. *New Media & Soc.* **20**, 990–1011, https://doi.org/10.1371/journal.pone.0195644 (2016).
9. Lewis, K., Gray, K. & Meierhenrich, J. The structure of online activism. *Sociol. Sci.* **1**, 1–9, 10.15195/v1.a1 (2014).
10. Jackson, S., Bailey, M. & Foucault, B. *#HashtagActivism* (MIT Press, 2020).
11. Dadas, C. Hashtag activism: The promise and risk of "attention". In *Social Writing/Social Media: Publics, Presentations, Pedagogies*, 17–36 (Fort Collins, Colorado: The WAC Clearinghouse and University Press of Colorado, 2017).
12. Yang, G. Narrative agency in hashtag activism: The case of #blacklivesmatter. *Media Commun.* **4**, 13–17, 10.17645/mac.v4i4.692 (2016).
13. Bai, Q., Dan, Q., Mu, Z. & Yang, M. A systematic review of emoji: Current research and future perspectives. *Front. Psychol.* **10**, 2221–2237 (2019).

14. Danesi, M. *The Semiotics of Emoji* (Bloomsbury, 2017).

15. Miller, H. *et al.* "blissfully happy" or "ready tofight": Varying interpretations of emoji. In *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 10 (2016).

16. Dresner, E. & Herring, S. Functions of the nonverbal in cmc: Emoticons and illocutionary force. *Commun. Theory* **20**, 249–268 (2010).

17. Felbo, B., Mislove, A., Søgaard, A., Rahwan, I. & Lehmann, S. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (2017).

18. Prasad, A., Sanjana, S., Bhat, S. & Harish, B. Sentiment analysis for sarcasm detection on streaming short text data. In *2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA)*, 1–5, 10.1109/ICKEA.2017.8169892 (2017).

19. Herring, S. & Dainas, A. Receiver interpretations of emoji functions: A gender perspective. In *Proceedings of the 1st International Workshop on Emoji Understanding and Applications in Social Media (Emoji2018)* (2018).

20. Gwane, L. & McCulloch, G. Emoji as digital gestures. *Language@Internet* **17** (2019).

21. Marsili, N. Retweeting: its linguistic and epistemic value (2020).

22. Robertson, A., Magdy, W. & Goldwater, S. Self-representation on twitter using emoji skin color modifiers. In *ICWSM 2018*, 680–683 (2018).

23. Cheong, M. Inferring social behavior and interaction on Twitter by combining metadata about users & messages (2013). PhD thesis. Monash University.

24. Kwak, H., Lee, C., Park, H. & Moon, S. What is Twitter, a social network or a news media? In *Proc. WWW 2010* (2010).

25. Sullivan, E. *et al.* Vulnerability in social epistemic networks. *Int. J. Philos. Stud.* **28** (2020).

26. Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal* **Complex Systems**, 1695 (2005).

27. Traag, V., Waltman, L. & van Eck, N. J. From louvain to leiden: guaranteeing well-connected communities. *Sci. Reports* **9**, 5233, 10.1038/s41598-019-41695-z (2019).

28. Cha, S. K-pop boyband bts' fans match group's 1 million donation to black lives matter (2020). Retrieved from <https://www.reuters.com/article/us-southkorea-bts-minneapolis-police-idUSKBN23E08T>.

29. Hallsmar, F. & Palm, J. Multi-class sentiment classification on twitter using an emoji training heuristic. *Ind. Manag. & Data Syst.* **118**, 1804–1820, 10.1108/IMDS-12-2017-0582 (2016).

30. Li, M., Ch'ng, E., Chong, A. & See, S. Multi-class twitter sentiment classification with emojis. *Ind. Manag. Data Syst.* **118**, 1804–1820 (2018).

31. Chen, Y., Quanzeng, Y., Yuan, J. & Luo, J. Twitter sentiment analysis via bi-sense emoji embedding and attention-based lstm. In *2018 ACM Multimedia Conference*, 1–9, 10.1145/3240508.3240533 (2018).

32. Wang, Z. *et al.* A hybrid model of sentimental entity recognition on mobile social media. *J. on Wirel. Commun. Netw.* **253**, 10.1186/s13638-016-0745-7 (2016).

33. Abadi, M. *et al.* TensorFlow: Large-scale machine learning on heterogeneous systems (2015). Software available from tensorflow.org.

34. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

35. Kim, T. & Wurster, K. emoji 1.2.0 (2021).

36. Ray Chowdhury, J., Caragea, C. & Caragea, D. On identifying hashtags in disaster twitter data. *Proc. Conf. AAAI Artif. Intell.* **34**, 498–506 (2020).

37. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. In *Third international AAAI conference on weblogs and social media* (2009).

38. Lockerman, Y. Image preview (2020). Retrieved from <https://gephi.org/plugins/#/plugin/imagepreview>.

39. Bawn, K. *et al.* A theory of political parties: Groups, policy demands and nominations in american politics. *Perspectives on Polit.* **10**, 571–97 (2019).

40. Grossmann, M. & Hopkins, D. Ideological republicans and group interest democrats: The asymmetry of american party politics. *Perspectives on Polit.* **13**, 119–39, 10.1017/S1537592714003168 (2015).

41. Hagen, L. *et al.* Emoji use in twitter white nationalism communication. In *Proceedings of the Computer Supported Cooperative Work and Social Computing Conference*, https://doi.org/10.1145/3311957.3359495 (2019).

42. Gallagher, R., Reagan, A., Danforth, C. & Dodds, P. Divergent discourse between protests and counter-protests: #blacklivesmatter and #alllivesmatter. *PlosOne* **13**, https://doi.org/10.1371/journal.pone.0195644 (2018).

43. Collins, P. H. On violence, intersectionality and transversal politics. *Ethn. Racial Stud.* **40**, 1460–1473 (2017).

44. Clark, A., Dantzler, P. & Nickles, A. Black lives matter: (re)framing the next wave of black liberation. *Res. Soc. Movements, Conflicts Chang.* **42**, 145–172 (2018).

45. Crenshaw, K., Ritchie, A., Anspach, R., Gilmer, R. & Harris, L. Say her name: Resisting police brutality against black women. Tech. Rep., Centre for Victim Research (2015).

46. Stewart, L. G., Arif, A., Nied, A. C., Spiro, E. S. & Starbird, K. Drawing the lines of contention: Networked frame contests within #blacklivesmatter discourse. *Proc. CSCW* **1**, 10.1145/3134920 (2017).

47. Tarca, K. Colorblind in control: The risks of resisting difference amid demographic change. *Educ. Stud.* **38**, 99–120, https://doi.org/10.1207/s15326993es3802_3 (2005).

48. Atkins, A. Black lives matter or all lives matter? color-blindness and epistemic injustice. *Soc. Epistemology* **33**, 1–22 (2019).

49. Pervin, N., Phan, T., Datta, A., Takeda, H. & Toriumi, F. Hashtag popularity on twitter: Analyzing co-occurrence of multiple hashtags. In *International Conference on Social Computing and Social Media*, 169–182 (2015).

Acknowledgements

(not compulsory) This publication was made possible through the support of grants from the John Templeton Foundation (61378) and the Australian Research Council (DP190101507). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation or the Australian Research Council. The authors would like to acknowledge XXX, YYY, and ZZZ for their helpful comments in improving this publication.

Author contributions statement

M.A., R.R., and C.K. drafted the main body of the paper and constructed and visualized the emoji and hashtag networks. M.A., R.R., I.O.Q., M.C., and C.K. edited the main body of the paper. I.O.Q. cleaned the data and trained the classifiers. R.R. conducted the literature review. M.C. helped with all aspects of coding/preprocessing in Python. C.K. collected the data and coded the network graph and clustering.

Competing interests

(mandatory statement)

The corresponding author is responsible for providing a competing interests statement on behalf of all authors of the paper. This statement must be included in the submitted article file.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BLMTwitterEmojiandHashtags25NatureSupplemental.pdf
- BLMTwitterEmojiandHashtags25NatureSupplemental.pdf