



# Mind games: A temporal sentiment analysis of the political messages of the Internet Research Agency on Facebook and Twitter

new media & society  
2023, Vol. 25(3) 463–484  
© The Author(s) 2021



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/14614448211014355  
journals.sagepub.com/home/nms



**Soyoung Park, Sharon Strover<sup>ID</sup>,  
Jaewon Choi<sup>ID</sup> and MacKenzie Schnell**

University of Texas at Austin, USA

## Abstract

This study examines the temporal dynamics of emotional appeals in Russian campaign messages used in the 2016 election. Communications on two giant social media platforms, Facebook and Twitter, are analyzed to assess emotion in message content and targeting that may have contributed to influencing people. The current study conducts both computational and qualitative investigations of the Internet Research Agency's (IRA) emotion-based strategies across three different dimensions of message propagation: the platforms themselves, partisan identity as targeted by the source, and social identity in politics, using African American identity as a case. We examine (1) the emotional flows along the campaign timeline, (2) emotion-based strategies of the Russian trolls that masked left- and right-leaning identities, and (3) emotion in messages projecting to or about African American identity and representation. Our findings show sentiment strategies that differ between Facebook and Twitter, with strong evidence of negative emotion targeting Black identity.

## Keywords

Affective polarization, emotional politics, Facebook, propaganda, sentiment, Twitter

## Introduction

As the United States moved through the 2020 presidential election cycle, it has yet to leave behind the last 2016 presidential election. Investigations of attempted foreign interference are still contentious in several ways. The Russian Internet Research Agency (IRA), a

---

### Corresponding author:

Sharon Strover, Moody College of Communication, University of Texas at Austin, Austin, TX 78712, USA.  
Email: sharon.strover@austin.utexas.edu

private “troll farm” owned and run by President Putin’s wealthy friend Yevginy Prigozhin (also responsible for a mercenary army), aimed to destabilize the 2016 election, according to the Office of the Director of National Intelligence and Special Counsel Robert Mueller’s investigation (Mueller, 2019). The IRA had begun to infiltrate social media accounts at least as early as 2012, using platforms including YouTube, Facebook, Twitter, Tumblr, and Instagram, and the same actions recurred in 2020, and probably will continue (Frenkel and Barnes, 2020). Many studies have sought to unravel the role and possible impacts of the Russian IRA’s social media operation during the election campaign, and while some questions regarding the provenance and aims of the Facebook messages, ads, videos, and tweets placed by the IRA have been resolved, others remain.<sup>1</sup> How the IRA’s messages in social media might have appealed to Americans is one of them.

From a communication research perspective, scholarly attention has attended to the elevated significance of social media that prompted people to be actively involved in disseminating or recirculating disinformation (Bail et al., 2020; Jamieson, 2018; Kim et al., 2018). Researchers identified incendiary and divisive social issues such as immigration and police brutality in evocative Russian-paced political ads (Jamieson, 2018; Ribeiro et al., 2018), and some note the characteristics and behaviors of “fake” or sock-puppet accounts using individual US personas or automated bots within the IRA interference operations (Freelon and Lokot, 2020; Linvill and Warren, 2020; Stukal et al., 2017).

Alongside the IRA’s rhetorical and emotional appeals, the technological affordances of social media platforms make them helpful propaganda vehicles, enabling microtargeting and rapid information dissemination. Apparatuses like Facebook ads that enable specific targeting, provide analytics, and allow customers to place financial budgets behind their messaging may be particularly useful (Dommert and Power, 2019; Kim et al., 2018). Twitter joins Facebook as two known sites where Americans encountered political information, also allowing the IRA chances to pose as faux sources and to target certain constituencies with emotional and partisan messages.

In the 2016 election context, while much attention focused on the false information such messaging promotes, little empirical work examines the strategic uses of emotion in social media over time or compares their temporal patterns across the two major social media platforms, Facebook and Twitter. The broader use of sentiment in messaging transcends issues of factuality and can be seminal in prompting people’s engagement with a message. The current research attempts to fill this gap by investigating the Russian IRA’s emotion-based strategies on Twitter and in Facebook advertising. Based on affective theories and assessments of platform characteristics, this study considers emotional sentiment in social media messages and the manipulation of partisan and racial identities.

Previous scholarship demonstrates the growing role of affective polarization in politics (Garrett et al., 2014; Iyengar et al., 2019; Mason, 2015), and some researchers examining the IRA propaganda identify partisan appeals in its messaging (Howard et al., 2018; Jamieson, 2018; Ribeiro et al., 2018). We extend the theorization of political messaging’s affective dimension to an analysis of IRA propaganda and demonstrate how Facebook strategies appear to differ from Twitter strategies, and how social identity, particularly racial representation, interacts with message affect. Broader understanding of propaganda appeals in domains marked by algorithmic circulation and affective polarization may explain how social media dynamics figure into contemporary politics.

## Literature review

Affect, emotion, and feeling in media content have been discussed in many different disciplines (Bargetz, 2015) and have a historical footing in early propaganda research. Because widespread use of social media has transformed the current decade's information environment, we focus on social media platforms and messages in the political sphere, particularly their affective dimensions. Emotional content in social media reshapes and complicates the idealized notion of a rational public sphere and political expression in this century.

The following review first examines affective politics, exploring ways that affect interacts with social identity in catalyzing polarization. To the extent that affective components in the social media-based IRA campaign messages are related to political affiliation and racial identity, two salient aspects of social identity, emotional expressions may be important vehicles to arouse divisive partisanship. Second, social media platforms like Facebook and Twitter exhibit different platform features that may contribute to strategic political applications and help to explain why the IRA used those sites as vehicles for propaganda. Third, we investigate some of the research to date that explores the IRA campaign, highlighting the need for a better understanding of its affective dimensions.

### *Affect, politics, and polarization*

The impact of emotional appeals in political messages has been widely recognized and applied in contemporary political ad campaigns. The type of appeal, the timing of its appearance, the medium in which it appears, and its user targeting all can be important.

Recognizable emotional expressions such as sadness or fear or happiness have long been incorporated into campaigns of all sorts. For example, using a large-scale content analysis of emotions in more than 1400 political ads, Brader (2006) found seven major types of emotions, consisting of fear, enthusiasm, anger, pride, sadness, amusement, and compassion. Each also can be expressed mildly or powerfully. Both affective *valence* (positive vs negative) and *arousal* (high vs low) can be important in producing different impacts, as Marcus' (2000) work demonstrates. In political psychology, researchers have found that negative political information has a more powerful effect on attitudes than positive information. It is easier to recall and then becomes a useful element of heuristic cognitive processing for simpler tasks (Baumeister et al., 2001; Utych, 2018). From the standpoint of information dissemination, however, scholars differ on whether valence or strength of arousal exerts greater influence (Brady et al., 2017; Stieglitz and Dang-Xuan, 2013a). For example, looking at Twitter posts on gun control, same-sex marriage, and climate change—all politically charged topics—Brady et al. (2017) found that people are 20% likelier to share or interact with emotionally evocative social media information, suggesting that arousal was the most important element. Arousal or emotional intensity is obvious in many of the IRA campaign messages.

As scholarly attention investigated the context in which emotional appeals can spread, the question of timing in a campaign context also emerges. For example, in the case of voting, negativity in ads ends up increasing turnout probability among voters who have not chosen their candidates while decreasing the turnout among those who have made

their choice (Krupnikov, 2014). However, the effect of exposure timing with emotional appeals is difficult to generalize largely because campaigners must consider several complex and sometimes local factors to isolate the impact of emotion from other variables (Druckman et al., 2010; Fowler et al., 2016). Nevertheless, political campaigns obviously have critical dates—election days, for example—that invite scrutiny in terms of when emotional appeals may be most useful. Earlier work from Krupnikov (2014) predicted fairly benign appeals in earlier stages of the campaign that would function to consolidate identity appeals, yielding to a more negative tone as the election date gets closer. Similarly, with respect to the Russian interference tactics, Jamieson (2018) identifies the “Russian Courtship,” referring to a conventional process in which IRA-constructed social media messages displayed more positive emotional appeals to left-leaning and African American handles during the initial campaign phases, presumably with the intent to generate trust and likability early to encourage continued engagement in later months.

Affective politics in the 21st century are capable of incorporating broader online social dynamics into their processes. Bennett and Iyengar (2008) presaged a growing role for the affective components in individual information processing within their broader conclusion that research in political communication must be more cognizant of changing socio-technical circumstances, particularly the new roles of various media in people’s lives. Their notion of “stratementation” signaled a significant change associated with how isolated individuals in American society interact with the media: “. . . what we find today, particularly among younger audience demographics, are shifting and far more flexible identity formations that require considerable self-reflexivity and identity management . . .” (p. 716). Departing from a conceptualization of individuals formerly engaged with mass media (and politics) through civic groups (such as the Kiwanis) or as members of formal institutions such as labor unions, they cite fan cultures, friend network sites, and games as new affinity sites, hinting at the flexible and multiple identity formations within new media spaces. In the past decade, audiences—and the electorate—have indeed fragmented; users and subcultures engage civic culture in new ways (Jenkins et al., 2015; Kim, 2018). Identity formation in media “high choice” spaces can interact with affect to create both affinities and polarization (Iyengar et al., 2019; Mason, 2016).

Going beyond a conceptualization of the atomized individual simply responding directly to media messages, researchers examining affect explain that online media spaces have proved useful to cultivate group-based emotional attachments. Papacharissi (2015), for example, explores the “soft structures of feeling” in online spaces (p. 116). She extends the idea of political affect by suggesting that as people contribute narratives to social media by developing stories, they become part of a larger story and feel connected to it. Her case studies on Occupy Wall Street, the #egypt hashtag during protests in Egypt, and political expression on Twitter argue for paying more research attention to what she calls “affective attunement” in those communication forums. In the case of the Russian IRA, this may be a crucial clue to the sorts of bonds that slip beyond the strictly partisan. Similarly, Boler and Davis (2018) conceptualize emotion and networked publics interacting in an “affective feedback loop,” sustaining relationships between highly charged messages and their constituencies in the digital media realm.

Extending a similar argument around what they call “empathic media,” Bakir and McStay (2018) sound an alarm about technologically based methods that “make emotions machine-readable,” warning of problems associated with fake news and misinformation’s enhanced capabilities to manipulate public sentiment (p. 155). Affective polarization, the linking of emotion to partisan identity (distinguished from other forms of polarization by its emotional roots), appears to be deepening political divisions in the United States (Garrett et al., 2014; Mason, 2015). To the extent that emotional messages can trigger or interact with core partisan identity, they also can be aligned with other salient social identity factors like race, gender, and religion. The strengths of social media platforms come into play as they create spaces for like-minded or homophilous individuals linked through some aspect of social identity.

Opportunities to manipulate affect in politics create distinct advantages to propagandists, including the IRA, especially in the context of emerging social media environments. How might the IRA have tailored emotion in their election messages, especially in social media, to elicit certain responses? Did their strategies vary over time in terms of emotional sentiment?

### *Social media platforms and politics*

Political uses of social media platforms have blown up in the past decade. The Pew Research Center’s Internet & American Life Project found that 66% of social media users used social network sites (SNSs) like Facebook or Twitter at least once for civic or political activities as early as 2012 (Rainie et al., 2012); more recent research illustrates that 18% of Americans “turn most” to social media for political information (Mitchell et al., 2020). The majority of Americans, about 70%, believe that these SNSs are very or somewhat important for accomplishing political goals (Anderson et al., 2018). Social media have been used in various domestic as well as international social movements (Bennett and Livingston, 2018; Howard and Hussain, 2013; Papacharissi, 2015). The average Twitter user as of 2019 is younger and somewhat better educated than the average Facebook user, and some research reports longer time periods of engagement for Twitter users as opposed to Facebook users (Sehl, 2020).

Social media do more than transmit information. In a political environment, their bridging and bonding capital-building capacities may amplify political unities, and also set up people for emotional antagonisms. The “sharing” or “liking” or commenting features of social media can strengthen or jeopardize social bonds and group solidarity among networked users (Pi et al., 2013; Placencia and Lower, 2013), building on known demographic and behavioral characteristics of the platforms’ users in terms of age and political engagement, for example (Jaidka et al., 2018; Mitchell et al., 2020; Stieglitz and Dang-Xuan, 2013b). In the context of political campaigns, the opportunity to use social media platform’s characteristics to learn about people’s media habits and their social media groups presents enhanced targeting opportunities.

Because they are heavily used and facilitate social relationships, social media hold the potential to foster hospitable environments for propaganda/ disinformation activities. Such activities can encourage citizens to be motivated by emotions, substituting arousal for deliberation and rationality (Stanley, 2015). Describing the “economics of emotions,”

Bakir and McStay (2018) argue that emotional arousal translates directly into profits for a social media company by encouraging attention and more use of the platform. The dynamics of arousal amplify engagement. Compared to text-heavy traditional print media, the form, speed and images of digital content on social media are more likely to trigger “fast thinking,” by which intuitive and emotional biases can be exacerbated (Chaudhuri and Buck, 1994; Kahneman, 2011). On top of this, the virality of social media messages can be driven by psychological arousal (Kramer et al., 2014; Stieglitz and Dang-Xuan, 2013a) so that circulating emotionally charged messages within the unique socio-technical infrastructures of sites such as Twitter and Facebook creates opportunities for a strategist to foment polarization.

Platform affordances affect the way that messages are constructed, distributed, and received (Bossetta, 2018), but Facebook and Twitter do not operate identically. On Facebook, organizations or people, once “liked,” can be invited into one’s broader circle of Facebook friends and become diffusion points for more messages from the source. It is a “social networking site” that is broadly used across all ages. Twitter is a “microblogging service” (Stieglitz and Dang-Xuan, 2013b), popular among media-savvy users, the press, and politicians, alongside the more general population. It has an advantage in news delivery and in reaching politically involved audiences, while Facebook elicits more self-disclosure behaviors (Jaidka et al., 2018). In addition, Twitter, unlike Facebook, has message limits of 280 characters. The ability to generate longer posts on Facebook can facilitate more self-disclosure and bonding behavior.

Both platforms build identity and reputation through handles, but differ in homophily characteristics: Twitter users connect primarily around common topical interests—often transitory—and/or specific content expressed in hashtags. In contrast, Facebook ties reflect many different offline social contexts and may have more longevity and even face-to-face correspondences (Bakshy et al., 2015). With features that accentuate higher levels of proximity and reciprocity and unlimited space for messages, Facebook is more popular than Twitter among extreme and opposition parties (Ernst et al., 2017). Results from Boczkowski et al. (2018) also suggest that people would rather depend on acquaintances’ judgment with respect to what news they should attend to as opposed to professional journalists’ recommendations, underscoring the efficacy of a platform like Facebook. It thus should come as no surprise that Hughes et al. (2019) characterize Facebook as a low-involvement platform that rewards hedonic posts with more engagement, echoing Bakir and McStay’s (2018) points.

Advertising strategies in the two platforms differ, but both take advantage of platform-specific capabilities. Both Facebook and Twitter have improved microtargeting capabilities that enable narrowly defined, individual-level targeting based on the accumulation of user data, making both sites popular with selective political campaigns that only expose messages to specific, highly targeted audiences (Kim et al., 2018). While microtargeting options have become imperative to enhancing advertising performance on many social media platforms, granular strategies for optimizing audience reach or engagement may have been platform-specific during the IRA campaign. A notable feature can be found in the way the IRA utilized native advertising, referring to ads that are deliberately designed to look like user-generated organic content. On Facebook, the IRA used sponsored news feeds or generic and benign group names (e.g. Stand For Freedom,



Native American United), most of which have a “sponsored content” disclaimer label (Kim et al., 2018), whereas most of the IRA-sponsored tweets appeared to have been posted by ordinary individuals without any “promoted” labels.<sup>2</sup> These strategies, intertwined with technological affordances, network features, and content or source features, can be another trigger to amplify or mitigate emotional responses in penetrating target audiences.

Given these platforms’ popularity and their affordances in terms of targeting and reach, they became obvious IRA vehicles for propaganda (Kim et al., 2018). However, the contrasting qualities of Twitter and Facebook and the range of research results reported in single platform studies raise questions about how the IRA might have constructed affective strategies to exploit the respective strengths of the two dominant platforms. To address the emotional sentiment in the IRA campaigns and the matter of temporal changes across the campaign timeline, our first research question explores platform comparisons:

*RQ1.* How does the emotion (valence and arousal) in IRA-created media compare between Facebook ads and Twitter? If present, do the emotional patterns change over time?

Furthermore, political contexts and constituents are important areas to understand any political campaigns, including emotionally based ones. The next section dives into the question of how aspects of social identity, namely, political and racial representation, were targeted or reflected through emotional messaging in the IRA campaign.

### *IRA strategies in political campaigns*

Research on possible IRA strategies during the 2016 election has explored the role of social media platforms and some of the attendant message features. Attesting to the utility of social media, the Mueller (2019) Report notes that “to reach larger U.S. audiences, the IRA purchased advertisements from Facebook that promoted the IRA groups on the newsfeeds of U.S. audience members” (p. 25). Social identity and partisan targeting were used extensively. Both Twitter and Facebook supported faux identities and faux groups that were aligned to certain partisan identity groups, particularly targeting Black users and “Liberals” and “Conservatives.” Mueller’s (2019) investigation and research. Research by Howard et al. (2018) conclude that the IRA sought to mobilize potential or current Trump supporters and right-wing voters while demobilizing/dividing left-leaning counterparts. Kim’s (2018) work arrives at similar conclusions, drawing attention as well to the microtargeting deployed by the IRA that linked specific topics to target groups. Golovchenko et al. (2020) found that the IRA coordinated across social media platforms broadly, with conservative trolls (targeting people who identify as conservative in various ways) being more active than liberal trolls. Linvill and Warren (2020) also explored left and right “troll” messages in IRA tweets, finding evidence of partisan targeting. The IRA’s attention to targeting specific social identities by using certain divisive subjects such as immigration and police brutality raises questions, however, regarding the weight and sentiment of the appeals, answers not explored in earlier research.

Social media features complicate the task of distinguishing between advertising and non-promotional or “genuine” content from a user’s perspective. While organic content may sometimes be more effective than ad-sponsored content (Lukito, 2020), in Facebook’s case it may have been difficult to tell the two apart. Ribeiro et al. (2018) found that the “click-through” rate on the Russian Facebook ads was 10 times that of typical Facebook ads, and other research suggests that the emotional content may have confused users’ abilities or interest in detecting source credibility (Arif et al., 2018).

Some scholars have begun to unravel the connection between platforms, partisanship, social identity, and emotional valence in political campaigns, including the IRA social media activities. For example, Farkas, Schou, and Neumayer (2018) demonstrate how “platformed antagonism” operates in their study of discursive practices across 6 months of 11 faked Facebook pages designed to create strife around Muslims living in Europe. Exploring the role of social identity, Freelon and Lokot (2020) identified liberals, conservatives, and Blacks as targets in the IRA 2016 election disinformation campaign using Twitter, noting that in each case the goal was to provoke anger toward oppositional groups. Freelon et al. (2020) also examine digital Blackface and ideological asymmetries in IRA political advertising, commenting that “race is a key vulnerability ripe for exploitation” by disinformation (p. 2). Ribeiro et al.’s (2018) content analysis of Russian Facebook ads concluded that “IRA ads reached audiences that are very biased towards African-Americans and Liberals” (p. 10). Arif et al. (2018) come to similar conclusions. However, are the strategies in Facebook parallel to those in Twitter? Where might emotional tenor enter the endeavor to foment polarization?

When Arif et al. (2018) analyzed #BlackLivesMatter messaging, they use the verb “inspire” to describe messages that go beyond simple misinformation; emotional content may be the core of the inspiration to which they refer (p. 20). The IRA imitated target groups, especially Black citizens, by astroturfing (the term used when groups pose as legitimate “sources,” feigning credibility and affiliation) to foster antagonism and other negative feelings (Arif et al., 2018; Freelon et al., 2020; US Senate, 2019). As seen in IRA-constructed Black Matters accounts (Bastos and Farkas, 2019) or IRA-constructed persona Jenna Abrams (Xia et al., 2019), feigned credibility and emotional appeals can be maintained across several platforms, new iterations of long-standing propaganda techniques that recall Becker’s (1949) early dissection of World War II disinformation.

To expand our understanding of the link between social media and emotional messages during the election, we examine the sentiment intensity and direction in emotional appeals in ads and messages linked to the IRA, comparing the two major platforms, Twitter and Facebook. We explore platform differences and how targeting racial representation as well as partisan identity (left and right trolls) was used.

Against the backdrop of specific platform characteristics, temporal dynamics across platforms, and political contexts and constituencies, three research questions address the emotional sentiment in the IRA campaigns and the aspects of social identity, namely, political and racial representation. We select two identities heavily targeted during the election—being a “Liberal” and being Black or African American, as highlighted in the literature (Arif et al., 2018; Freelon and Lokot, 2020; Freelon et al., 2020; Ribeiro et al., 2018). It is probable that emotional appeals were present in ads or messages addressed to them.



*RQ2.* How did the IRA's emotional valence and arousal on social media vary between right- and left-leaning handles? That is, does sentiment vary depending on whether the targets are liberal or conservative?

Building on Freelon's (2020) and Arif et al.'s (2018) discussions regarding negative emotion contributing to antagonism with messages targeting Black representation, we hypothesize the following:

*H2.1.* Left-leaning handles are associated with using more negative valence of emotions overall compared with right-leaning handles.

The IRA's emotion-based strategies may have followed a different trajectory within their construction of left- or right-leaning messages and representation of Black identities. In terms of valence, Russians may have wanted people connected with right-leaning accounts (conservative or pro-Trump) to be more integrated and supported with more positive emotion while disrupting those tied to left-leaning accounts ("Liberals," pro-Clinton or pro-Sanders) with dissonant or negative sentiment messages.

Similarly, the IRA may have treated traditionally liberal-leaning constituencies, such as African Americans differently, or represented them differently in left-leaning handles. Here, the issue of arousal intersects racial identity. Research questions specifically investigating political targeting and racial representation include the following:

*RQ3.* How did the IRA's emotional appeals on social media interact with Black or African American racial representation?

*H3.1.* Racialization of left- or right-leaning handles is associated with higher arousal of negative emotions compared with non-racialized left- or right-leaning handles.

*H3.2.* Racialization of left-leaning handles is associated with higher arousal of negative emotions compared with racialized right-leaning handles.

Furthermore, timing in the use of emotion or sentiment may be significant, as noted in earlier research (Krupnikov, 2014). We expect to see more attempts at positive emotional manipulation for messages linked to left-leaning and African American handles in earlier campaign stages:

*RQ4.* How does emotion in Facebook ads and Twitter enter into identities and representation over time?

*H4.1.* Left-leaning handles present Black identity with more positive emotions early in the election and become more negative close to the election date.

## Methods

This study applies a mixed-method design with both computational and qualitative analytic approaches. Drawing on definitions and methods applied to Twitter data that Linvill

and Warren (2020) developed to identify for right- and left-leaning trolls, we coded the congressionally released Facebook ads definitively associated with the IRA for right and left trolls as described below. We then analyzed textual data in Twitter and Facebook messages to explore emotional content and make comparisons. While Twitter and Facebook ads represent quite different datasets, examining emotional expression in each could illuminate IRA strategies that play to the strengths of these respective platforms.

**Sampling.** The two datasets used here include (1) 3519 Facebook ads provided by Facebook to Congress as identifiably placed by the IRA, which promoted 470 IRA-associated Facebook pages during the election months (roughly mid-2015 through mid-2017) (US House of Representatives, 2018), and (2) nearly 3 million tweets known to be associated with 3841 IRA Twitter handles (usernames) listed by the US House Intelligence Committee as placed by the Russian IRA. These were released publicly by Twitter and later coded by Linvill and Warren into left- and right-troll definitions (available at <https://russiatweets.com/>). The tweets spanned 19 June 2015 to 31 December 2017 in English language only. In the Twitter data, right trolls are associated with 663,740 tweets (617 handles,  $M = 1075.75$ ,  $SD = 2949.82$ ) and left trolls are associated with 405,549 tweets (230 handles,  $M = 1763.26$ ,  $SD = 2468.32$ ). This research acknowledges comparative limitations between Facebook ads and “organic” tweets (i.e. fake content designed to look like organic content) in terms of possible differences in platform or messaging characteristics; however, the dataset used here is the best public data available.

**Sentiment.** We analyzed the sentiment of Facebook ads and tweets using the RSentiment Package in the R development framework, which was developed by combining word-level sentiment analysis with natural language processing (NLP) techniques. By “sentiment,” this study refers to the valence or arousal of emotion automatically detected in a piece of text and whether (or how much) it is positive or negative, thus making an evaluative judgment of attitude or opinion in the posts expressed by a group of handles (e.g. right or left trolls). Using “Part of Speech tagging” to tag each word in a sentence, RSentiment evaluates the text units for positive or negative sentiment and assigns a numerical score to indicate the overall sentiment across sentences, which can be classified into five categories from very positive to very negative. The RSentiment algorithm calculates not only various degrees of emotions of verbs, adverbs, adjectives, and nouns, but also sequential occurrence of negators (e.g. “not,” “no”). Because the Facebook ad and tweet text generally occur in sentence units, we applied RSentiment which enables textual analysis at or near the sentence level of granularity. A comparative assessment of several sentiment analysis packages found that RSentiment was more effective than other popular variants for social media and SMS-length units of analysis (Bose et al., 2017).<sup>3</sup>

**Temporal patterns.** We identified the dates that tweets appeared in the Twitter data. In the Facebook data, we considered the ad creation date.

**Right- and left-troll coding.** Operationalized as accounts targeting left- or right-wing users, we coded left and right trolls as partisan identities. The troll position can be compared to partisan identity as it targets and represents like-minded users. For the Twitter

data, Linvill and Warren (2020) used publicly released IRA-associated tweets and coded five “handle categories” to assign tweets: right troll, left troll, hashtag gamer, news feed, and fearmonger.<sup>4</sup> Their right troll coding is sensitive to “nativist and right-leaning populist messages” and included specifications such as “supported Donald Trump’s candidacy.” The left troll definition included “tweets with socially liberal messages, with an overwhelming focus on cultural identity,” among other qualifications (pp. 7–8). We used their tweet definitions and identified the tweets associated with right and left trolls for our analyses.

Our research applied an analogous coding strategy to categorize Facebook ads that were identifiably right and left troll categories within the IRA-placed partisan world. To generate left- and right-troll codes, we evaluated the Facebook ad metadata containing fields helpful to the analysis. We used the “ad landing page” information as a primary classifier to identify left or right trolls, largely because it frequently contained information that reveals the targeted audience (e.g. <https://www.facebook.com/blackmattersus> or <https://www.facebook.com/patriototus>).<sup>5</sup> We examined the “ad content” and “ad targeting” information to support our coding, which includes targeted location and targeted interest, and excluded or included connections to help identify right versus left. This allowed us to properly code accounts such as Williams and Calvin, a well-known conservative Black duo typically promoting pro-Trump views. Three coders independently coded a sub-sample of 70 ads for reliability checks, yielding a high intercoder reliability on left- or right-troll identification (Krippendorff’s alpha of .92).

*Black or African American racial representation.* To infer the racial representation of African Americans in Facebook ads and tweets, this study developed a subset of data in a keyword-based fashion. First, we tokenized the message contents of Facebook and Twitter data into bigrams that were filtered based on the inclusion of words indicating relationship with African American communities (i.e. “Black,” “African,” words starting with the prefix “nigg-,” and “BLM”—an abbreviation of “Black Lives Matter,” the representative social movement of/for Black people). Next, the messages in the Twitter data were further parsed and filtered to exclude several keywords that are not connected to African American-related identity (e.g. Blackberry, Blackhawks, and Black Friday, among many others). The result yields messages that depict or present African American constituencies.<sup>6</sup> Given that messages mentioning Black people or communities contribute to represent Black identities in any way, we operationalize these as an expression of racial identity. While not directly parallel to identifying right and left trolls, racial representation can be a component of either right- or left-troll activity (Freelon et al. (2020), and the Senate investigation concluded that Twitter data present Black trolls as a primary component of left trolls).

*Analysis procedure.* We employed the RSentiment Package in the R development framework to explore the sentiment in the ad text and tweet text. To address the three main research questions, we provide sentiment score plots in addition to descriptive statistics, respectively. Using the mean daily sentiment scores and monthly scores, this study plots the temporal dynamics of positive and negative emotions that show patterned uses of emotions over time.

**Table 1.** Descriptive statistics of sentiment scores of the Internet Research Agency messages on Facebook and Twitter.

	<i>N</i>	<i>M</i>	Max	Min
Facebook	3425	0.55	21.00	-21.00
Twitter	2,025,211	-0.30	32.00	-15.00

**Results**

*Sentiment patterns (RQ1)*

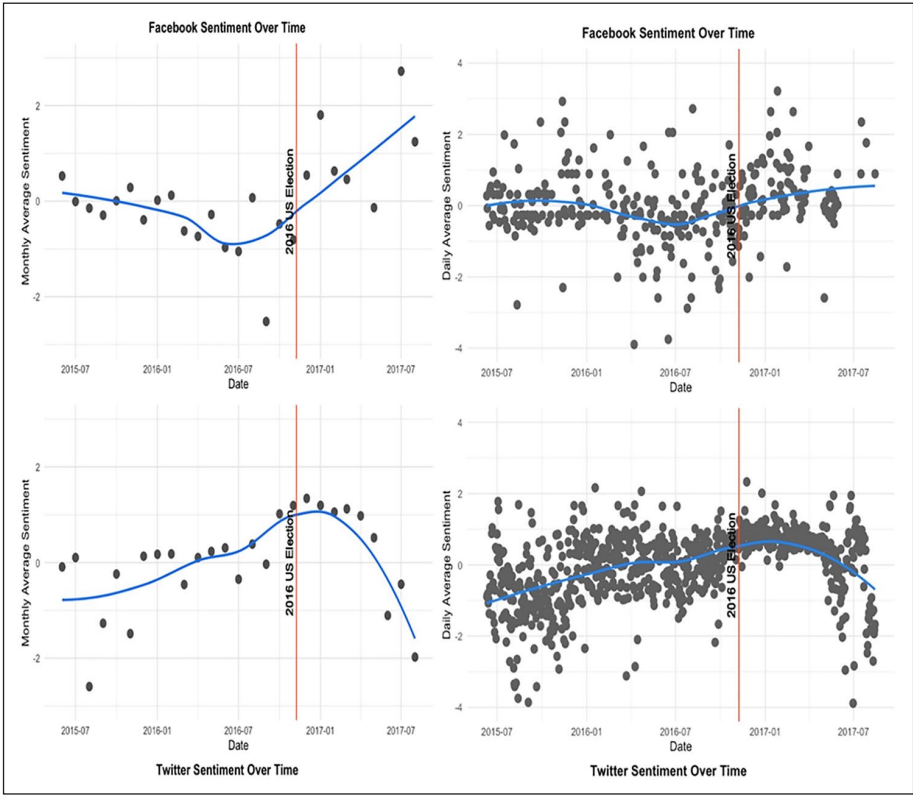
To explore the sentiment in the IRA messaging, we provide descriptive statistics from Facebook and Twitter messages and plot the monthly and daily mean sentiment scores in each platform over time. Using two different ways of measuring time captures greater variation in the patterns of emotion for monthly based mean sentiment as well as greater scale ranges for the daily-based mean sentiment. The mean sentiment scores of the Facebook ads and tweets show greater positive sentiment in Facebook messages and higher negative sentiment in tweets. The analysis also reveals slightly greater variance in emotions in tweets compared to Facebook messages (Table 1).

Over time, a relatively distinctive pattern of emotional expression on Twitter and Facebook can be seen in the following sentiment plots (Figure 1). On Facebook (first row), monthly and daily means of sentiment trend negative until about 5 months ahead of the election date (from July), when their messages began to display a slight positive swing. The emotional trend remained positive after the election.

In contrast, the temporal pattern of emotion on Twitter (bottom row) was always negative (below 0). It moved in a more positive direction before the election, but the monthly mean is always negative. Twitter’s emotional levels, in contrast to Facebook’s, become more negative after the election. Although daily mean sentiment (right side) reveals little dramatic change over time, monthly mean sentiment reveals evidence of the greater emotional variation in Facebook (its means ranging from -6.25 to +6) compared to Twitter (from -.92 to +.81). Throughout the entire period of analysis, which includes pre- and post-election timeframes, the pattern of emotion differs between Facebook and Twitter.

*Political leaning and sentiment (RQ2)*

To compare the IRA’s emotional appeals between right- and left-leaning troll handles on Facebook and Twitter, we compared the statistical difference of mean sentiment scores between right and left trolls in each platform, respectively, and plotted those scores over time. Left-troll messages were dominant in Facebook ads, about three times greater than right-troll messages, whereas on Twitter right-troll tweets numbered about twice as many as left-troll tweets (Table 2). In addition, an opposite pattern of sentiment was deployed for Facebook and Twitter: while the left-leaning handles exhibited statistically significantly less positive sentiment than the right-leaning handles on Facebook ( $M = 0.267$ ,  $SD = 2.97$  compared to  $M = 0.988$ ,  $SD = 2.37$ ), on Twitter the left handles expressed slightly more positive emotion ( $M = 0.024$ ,  $SD = 1.39$ ) than the right handles ( $M = -0.284$ ,  $SD = 1.38$ ).



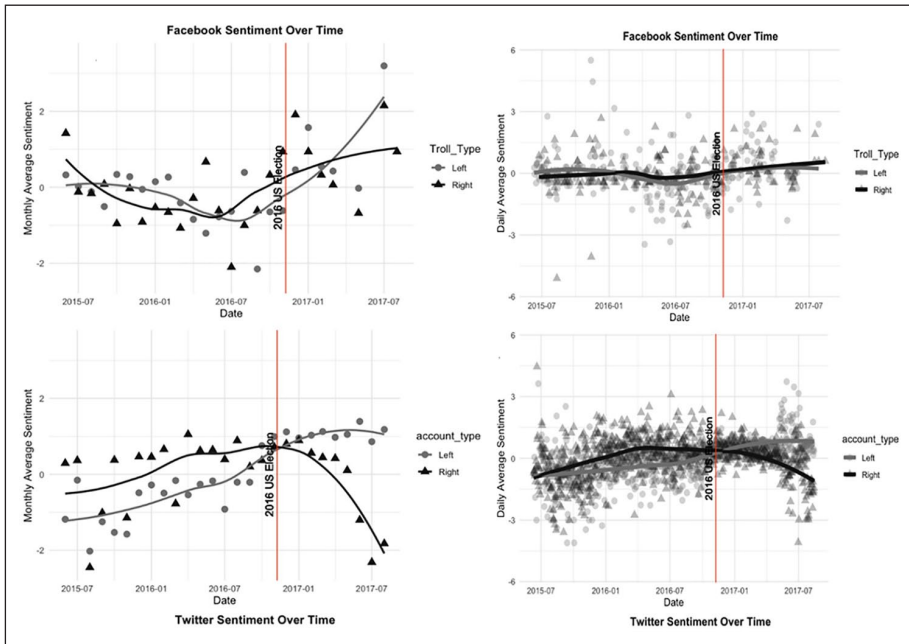
**Figure 1.** Mean sentiment of the IRA messages over time on Facebook ads and Twitter. Note. (Left) monthly mean sentiment; (right) daily mean sentiment.

**Table 2.** Descriptive statistics of sentiment scores of the IRA-placed left- and right-troll messages on Facebook ads and Twitter.

	N	M	SD	Min	Max
Facebook					
Left	2100	0.267	2.97	−18	21
Right	663	0.988	2.37	−21	15
Mean difference = −0.721, $t(2759) = -5.725^{***}$ (Hedge's $g = .254$ )					
Twitter					
Left	383,725	0.024	1.39	−12	11
Right	614,257	−0.284	1.38	−15	32
Mean difference = 0.312, $t(798,920) = 88.299^{***}$ (Hedge's $g = .379$ )					

SD: standard deviation; IRA: Internet Research Agency.  
\*\*\* $p < .001$ .

The temporal sentiment patterns of right and left trolls on Facebook and Twitter elaborate these patterns (Figure 2). Overall, before July 2016, left trolls on both platforms display relatively more negative emotions than right trolls, thus supporting H2.1. While



**Figure 2.** Sentiment means by time for left- and right-troll messages on Facebook and Twitter. Note. (Left) monthly mean sentiment; (right) daily mean sentiment.

Twitter's left trolls maintain positive increases in monthly/daily mean sentiment throughout the entire research period, even spiking near (from July 2016) and after the election, Facebook sentiment from left trolls exhibits more use of negative emotions early in the election, changing to the more positive sentiment as the election date gets closer. It is worth noting that the temporal patterns of positive and negative emotions between left and right trolls were reversed close to the election, especially on Twitter.

### *Projected racial identity (RQ3)*

Finally, we investigate how Black racial identity is reflected within left and right trolls by looking at sentiment patterns on Facebook and Twitter. Table 3 shows how trolls in both platforms used negative emotions in messages reflecting African American identities. First, it is striking that most messages invoking Black identity occur in left-leaning trolls in both Facebook and Twitter. The mean scores illustrate more use of negative emotions overall, exacerbated in the right-leaning handles compared to left-leaning handles.

The temporal sentiment plots in Figure 3, in general, show more negative appeals to or about people who referenced African American compared to "others" (non-Black). Looking primarily at the messages representing African Americans in Facebook and Twitter before the election, messages presenting the Black population (in blue and navy lines for left and right trolls, respectively) are always more negative than



**Table 3.** Descriptive statistics of sentiment scores of messages representing African Americans by left and right trolls on Facebook and Twitter.

		N	M	SD	Min	Max
Facebook						
Referencing African American identities	Left	807	−0.234	3.44	−18	21
	Right <sup>a</sup>	14	−1.79	5.71	−21	1
Twitter						
Referencing African American identities	Left	21,475	−0.602	1.63	−9	8
	Right	7005	−0.754	1.62	−10	6

SD: standard deviation.

<sup>a</sup>We refrain from interpretations here due to very few right trolls ( $N = 14$ ) referencing racial identity.



**Figure 3.** Mean sentiment of messages representing African Americans by left and right trolls on Facebook and Twitter.  
Note. (Left) monthly mean sentiment; (right) daily mean sentiment.

messages regarding non-Black identities (in orange and green lines). The above account supports H3.1 regarding emotional intensity among race-representing handles when these are aligned with political orientation. Overall, temporal sentiment was far

more negative throughout the entire period of the election, even including the post-election period, with respect to African American identity. However, because the projection of negative emotions was more intense among right than left trolls on Facebook and Twitter, we reject H3.2, which hypothesizes that racialized left-leaning handles would exhibit more use of negative emotional appeals over time than their respective counterparts on both platforms.

Finally, RQ4 examines questions regarding the sequence of messaging that could have reinforced and later threatened identity, using the case of African Americans. Our hypothesis has to do with identity disruption through left-leaning handles over time. First, only considering the period between July 2015 and July 2017, left trolls (blue lines) depicting African Americans on Facebook became slightly positive in late 2015 (or less negative) but gradually turn in a negative direction after January 2016. Other patterns are evident that our hypotheses cannot explain: for instance, a modest upswing of emotion in the months immediately before the election around July–November among left-leaning handles on both platforms. In addition, on Twitter, the hypothesized pattern was explicit among right trolls mentioning the Black population rather than the left-leaning counterpart. Despite that, it still seems probable that overall, the trolls representing African Americans used a different set of emotional manipulation tactics than those of non-Black trolls, given the sentiment divergence evident between these populations. Taken together, our result implies that the IRA projected more negative emotional feelings by mixing other salient social identities, like race, with partisan identities, a result that conforms to the conclusions around Twitter by Freelon et al. (2020).

## **Discussion and conclusion**

This study examined the Russian IRA's emotion-based campaign strategies on Twitter and Facebook during the 2016 US presidential election period. Drawing on three different dimensions of messages—the social media platforms, sentiment, and the abilities to facilitate targeting of partisan and racial identities, the current study investigates the IRA's emotional strategies that weave together different elements of political discourse. Emotion in messages functions to amplify impact and presumably activates the recipient to further pass along messages or perhaps to act in some way (to vote, refrain from voting, and so forth). In the current case, publicly released Facebook and Twitter data were analyzed qualitatively to identify political and racial identity targeting and then computationally analyzed to assess message sentiment over time. Our findings reveal patterns of emotional content that may have influenced public opinion or behavior during the election period.

First, we detected contrasting patterns of emotional appeals on Facebook and Twitter throughout the complete election period. Facebook messages were negative in the pre-election stages and then turn positive shortly before and after the election, whereas Twitter had the opposite pattern. The more positive sentiment in Facebook messages, using the targeting available in the ad framework, may have encouraged voting and more satisfaction with the political process and with one's reference group (perhaps with the older users using Facebook) compared to the contentious and generally negative sentiment on Twitter. Platform differences need more attention if we are to address and

predict the potential danger resident in the emotional content of information. It appears that the IRA treated Facebook ads and Twitter differently in terms of projecting positive and negative emotions, with potentially consequential behaviors for their different constituencies.

Second, the emotional patterns between left- and right-leaning handles support previous work examining IRA disinformation. We speculate that more negative appeals by left trolls as opposed to right trolls may illustrate dynamics that discourage left-leaning users from engaging normative political institutions, following the reasoning in the literature on the effect of negative campaigning that demobilizes voters and promotes political disaffection (Ansolabehere and Iyengar, 1995; Haselmayer, 2019). Partisanship maneuvering afforded by social media highlights the need to consider the broader social context and how affect interacts with social identity.

Third, the study found partial evidence for a possible IRA emotional strategy that introduces increasingly more negative emotions in messages reflecting African American identities. The left-leaning messages may have discouraged people from voting with their increasingly negative tone, and messages representing Black identity appear to be special targets for divisive messaging. This implies that race may have been targeted and subject to stronger forms of manipulation. As the United States continues to discuss regulating social media platform, this ability to target must be scrutinized lest democratic rights be imperiled.

Finally, this study documents the positive upswing in message emotion during the 5-month period before the election, around July 2016. At about that time, sentiments seem to intensify. Borrowing from Iyengar et al. (2019) regarding political polarization driven by negative appeals in political ads, this analysis possibly captures a “divide-and-conquer” strategy: messaging may have attempted first to destabilize and break up the larger public sphere and then later strengthen intra-group identities as the election date got closer.

These findings may be limited in several ways. First, as mentioned earlier, the possible inconsistency between Facebook ad data and organic tweet data requires cautious interpretation. Second, congruity in coding left and right trolls across Twitter and Facebook data is tricky. The Twitter data that were pre-coded by Linvill and Warren (2020) and their definitions and operationalizations may not align precisely with how we coded the Facebook trolls. Besides, there is a possibility that Black identity might still be represented in posts without referencing the four main keywords that we used in the current study. For example, some campaign hashtags (e.g. #KnowYourRights, #KnowYourHistory) could have separately been used to address African American communities. Furthermore, our analysis focused on interpreting trends in sentiment rather than conducting conventional statistical hypothesis testing. We do this to capture temporal changes in sentiment. Additional nonlinear models and time-series analysis may help to ascertain the significance of some sentiment changes. The computational analysis tool applied for this study is not designed to analyze visual content nor capture the actual content of the emotion, so more in-depth qualitative examination of the messages as well as visual approaches to the Facebook and Twitter data could expand our inferences, particularly with respect to the layers of emotion captured in message imagery.

Nonetheless, this study demonstrates platform-level differences in emotions within political discourses, or what passes for discourse. Since many analyses have investigated

the political use of social media by focusing on only a single network platform (Buccafurri et al., 2015) or on denotative meanings, the similarities or differences between two platforms emphasized here contribute new insights. Based on studies on cross-platform political campaign strategies (Golovchenko et al., 2020; Lukito, 2020), we call for multi-platform assessment efforts to better understand the provenance and consequences of online propaganda associated with different technological features, advertising policy, types of content (e.g. organic vs promoted), and the network characteristics. Moreover, looking at sentiment in the IRA-placed messages over time can uncover subtle transitions at discrete moments. The different patterns of emotional appeals on Facebook and Twitter suggest ways that disinformation strategies may influence truth perceptions or political tones (Zembylas, 2019). Investigating the political character of the left- and right-leaning trolls and the appeals to racial representation deepen our understanding of the IRA strategies. Their operational strategies appear attentive to political divisions and differences within the country.

Finally, our findings suggest counteracting disinformation requires special attention to certain populations because emotional strategies apparently target them more intensively. Echoing other authors, this territory is so highly contextualized, with messaging emulating one's in-group language so well (Arif et al. (2018) use the word "tribal" to describe that match), that new literacies must be developed to effectively discern genuine disinformation. This study provides evidence pertinent to the Black population. Policymakers might direct careful regulatory attention to the role of social media operations in dividing public groups and polarizing public opinion. Election interference worries prompted Twitter to announce in October 2019 that it would no longer allow sponsored political ads on its platform, and President Trump was banned outright from Twitter, Facebook, and YouTube after the Capitol invasion of January 2021. Examining emotional speech should be extended not just in the election period, but also during the pre- and post-election periods. Another useful response would be to educate the population about these manipulations and to help them critically evaluate and assess emotional appeals. Emotional content and intensity should figure into platform companies' efforts to identify and remediate manipulation. Reinvigorating other discursive practices that go beyond emotional manipulation would be an important addition to revitalizing democratic practices.

## Funding

The authors acknowledge the support of the Good Systems Grand Challenge Research effort at the University of Texas at Austin.

## ORCID iDs

Sharon Strover  <https://orcid.org/0000-0001-9385-2146>

Jaewon Choi  <https://orcid.org/0000-0003-1847-4709>

## Notes

1. Journalists and political activists were deceived, according to the Mueller (2019) report.
2. Twitter announced that in its comprehensive analysis of 6500 accounts that promoted election-related Tweets throughout 2016, "only nine accounts were Tweeting election-related content and linked to Russia" (Twitter, 2019, pp. 3–4).

3. To test reliability or validity of RSentiment, a slightly different sentiment calculation tool called Syuzhet was applied for comparison purposes. The consistency between these two methods was supported by a significant positive correlation of calculated sentiment scores and a substantial similarity between the visualized temporal patterns.
4. “Fearmongers” mainly include the accounts that spread news of a fabricated crisis event about poisoned turkeys near the 2015 Thanksgiving holiday (Linvill and Warren, 2020: 7). Several handles that tweeted briefly in the way of a right troll and changed completely to tweeting as a Fearmonger were coded into a Fearmonger.
5. Approximately 250 Facebook ads were checked for face validity, that is, to insure ad landing pages conformed to political identity in terms of message content. On that basis, ad landing pages were gauged to be appropriate coding sites and representative of message content.
6. We created a categorical variable that indicates whether a message on Facebook or Twitter data reflects African American communities. If a message contains bigrams related to African American, it is categorized as presenting African American communities. Otherwise, it was considered “other.”

## References

- Anderson M, Toor S, Rainie L, et al. (2018) Activism in the social media age. Washington, DC: Pew Internet & American Life Project.
- Ansolabehere S and Iyengar S (1995) *Going Negative: How Political Advertisements Shrink & Polarize the Electorate*. New York: Free Press.
- Arif A, Stewart LG and Starbird K (2018) Acting the part: examining information operations within #BlackLivesMatter discourse. *Proceedings of the ACM on Human-computer Interaction* 2: 20–27.
- Bail C, Guay B, Maloney E, et al. (2020) Assessing the Russian internet research agency’s impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proceedings of the National Academy of Science* 117: 243–250.
- Bakir V and McStay A (2018) Fake news and the economy of emotions: problems, causes, solutions. *Digital Journalism* 6(2): 154–175.
- Bakshy E, Messing S and Adamic L (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239): 1130–1132.
- Bargetz B (2015) The distribution of emotions: affective politics of emancipation. *Hypatia* 30(3): 580–596.
- Bastos M and Farkas J (2019) “Donald Trump Is My President!”: the internet research agency propaganda machine. *Social Media + Society* 5(3): 1–13.
- Baumeister R, Bratslavsky E, Finkenauer C, et al. (2001) Bad is stronger than good. *Review of General Psychology* 5: 323–370.
- Becker H (1949) The nature and consequences of black propaganda. *American Sociological Review* 14(2): 221–235.
- Bennett W and Iyengar S (2008) A new era of minimal effects? The changing foundations of political communication. *Journal of Communication* 58(4): 707–731.
- Bennett W and Livingston S (2018) The disinformation order: disruptive communication and the decline of democratic institutions. *European Journal of Communication* 33(2): 122–139.
- Boczkowski P, Mitchelstein E and Matassi M (2018) “News comes across when I’m in a moment of leisure”: understanding the practices of incidental news consumption on social media. *New Media & Society* 20(10): 3523–3539.
- Boler M and Davis E (2018) The Affective politics of the post-truth era: feeling rules and networked subjectivity. *Emotion, Space and Society* 27: 75–85.

- Bose S, Saha U, Kar D, et al. (2017) RSentiment: a tool to extract meaningful insights from textual reviews. In: Satapathy S, Bhateja V, Udgata S, et al. (eds) *Proceedings of the 5th international conference on frontiers in intelligent computing: theory and applications* (Advances in Intelligent Systems and Computing), vol. 516. Singapore: Springer, pp. 259–268.
- Bossetta M (2018) The digital architectures of social media: comparing political campaigning on Facebook, Twitter, Instagram and Snapchat in the 2016 election. *Journalism and Mass Communication Quarterly* 95 (2): 471–496. Available at: <https://doi-org.ezproxy.lib.utexas.edu/10.1177%2F1077699018763307>
- Brader T (2006) *Campaigning for Hearts and Minds: How Emotional Appeals in Political Ads Work*. Chicago, IL: University of Chicago Press.
- Brady W, Wills J, Jost J, et al. (2017) Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* 114(28): 7313–7318.
- Buccafurri F, Lax G, Nicolazzo S, et al. (2015) Comparing Twitter and Facebook user behavior: privacy and other aspects. *Computers in Human Behavior* 52: 87–95.
- Chaudhuri A and Buck R (1994) Media differences in rational and emotional responses to advertising. *Journal of Broadcasting & Electronic Media* 39(1): 109–125.
- Dommett K and Power S (2019) The political economy of Facebook advertising: election spending, regulation and targeting online. *The Political Quarterly* 90(2): 257–265.
- Druckman JN, Kifer MJ and Parkin M (2010) Timeless strategy meets new medium: going negative on congressional campaign websites, 2002–2006. *Political Communication* 27(1): 88–103.
- Ernst N, Engesser S, Büchel F, et al. (2017) Extreme parties and populism: an analysis of Facebook and Twitter across six countries. *Information, Communication & Society* 20(9): 1347–1364.
- Farkas J, Schou J and Neumayer C (2018) Platformed antagonism: racist discourses on fake Muslim Facebook pages. *Critical Discourse Studies* 15(5): 463–480.
- Fowler E, Ridout T and Franz M (2016) Political advertising in 2016: the presidential election as outlier? *In the Forum* 14(4): 445–469.
- Freelon D and Lokot T (2020) Russian Twitter disinformation campaigns reach across the American political spectrum. *The Harvard Kennedy School, Misinformation Review*. Available at: <https://doi.org/10.37016/mr-2020-003>
- Freelon D, Bossetta M, Wells C, et al. (2020) Black trolls matter: racial and ideological asymmetries in social media disinformation. *Social Science Computer Review*. Epub ahead of print 7 April. DOI: 10.1177/0894439320914853.
- Frenkel S and Barnes J (2020) Russians again targeting Americans with disinformation, Facebook and Twitter say. *New York Times*. Available at: <https://www.nytimes.com/2020/09/01/technology/facebook-russia-disinformation-election.html>
- Garrett K, Gvirsman S, Johnson B, et al. (2014) Implications of pro- and counterattitudinal information exposure for affective polarization. *Human Communication Research* 40: 309–332.
- Golovchenko Y, Buntain C, Eady G, et al. (2020) Cross-platform state propaganda: Russian trolls on Twitter and YouTube during the 2016 U.S. presidential election. *The International Journal of Press/politics* 25: 357–389.
- Haselmayer M (2019) Negative campaigning and its consequences: a review and a look ahead. *French Politics* 17: 355–372.
- Howard P and Hussain M (2013) *Democracy's Fourth Wave: Digital Media and the Arab Spring*. Oxford: Oxford University Press.
- Howard P, Ganesh B, Liotsiou D, et al. (2018) *The IRA, Social Media and Political Polarization in the United States, 2012–2018*. Working Paper. Oxford: Project on Computational Propaganda.
- Hughes C, Swaminathan V and Brooks G (2019) Driving brand engagement through online social influencers: an empirical investigation of sponsored blogging campaigns. *Journal of Marketing* 83(5): 78–96.



- Iyengar S, Lelkes Y, Levendusky M, et al. (2019) The origins and consequences of affective polarization in the United States. *Annual Review of Political Science* 22: 129–146.
- Jaidka K, Guntuku S, Buffone A, et al. (2018) Facebook vs. twitter: differences in self-disclosure and trait prediction. In: *Proceedings of the international AAAI conference on web and social media*, Palo Alto, CA, 25–28 June.
- Jamieson K (2018) *Cyberwar: How Russian Hackers and Trolls Helped Elect a President*. New York: Oxford University Press.
- Jenkins H, Ito M and boyd d (2015) *Participatory Culture in a Networked Era*. Malden, MA: Polity Press.
- Kahneman D (2011) *Thinking Fast and Slow*. London: Penguin Books.
- Kim M (2018) *Uncover: Strategies and Tactics of Russian Interference in US Elections*. Unpublished manuscript. Available at: [https://journalism.wisc.edu/wp-content/blogs.dir/41/files/2018/09/Uncover.Kim\\_v.5.0905181.pdf](https://journalism.wisc.edu/wp-content/blogs.dir/41/files/2018/09/Uncover.Kim_v.5.0905181.pdf)
- Kim Y, Hsu J, Neiman D, et al. (2018) The stealth media? Groups and targets behind divisive issue campaigns on Facebook. *Political Communication* 35(4): 515–541.
- Kramer A, Guillory J and Hancock J (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academic of Sciences USA* 111(29): 8788–8790.
- Krupnikov Y (2014) How negativity can increase and decrease voter turnout: the effect of timing. *Political Communication* 31(3): 446–466.
- Linville D and Warren P (2020) Troll factories: manufacturing specialized disinformation on Twitter. *Political Communication* 37: 447–467.
- Lukito J (2020) Coordinating a multi-platform disinformation campaign: internet research agency on three US social media platforms, 2015 to 2017. *Political Communication* 37: 238–255.
- Marcus G (2000) Emotions in politics. *Annual Review of Political Science* 3(1): 221–250.
- Mason L (2015) “I disrespectfully agree”: the differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science* 59(1): 128–145.
- Mason L (2016) A cross-cutting calm: how social sorting drives affective polarization. *Public Opinion Quarterly* 80: 351–377.
- Mitchell A, Jurkowitz M, Oliphant J, et al. (2020) July 30). Americans who mainly get their news on social media are less engaged, less knowledgeable. *Pew Research Center*. Available at: <https://www.journalism.org/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/>
- Mueller R (2019) *Report on the Investigation into Russian Interference in the 2016 Presidential Election*. U.S. Department of Justice. Available at: <https://www.justice.gov/storage/report.pdf>
- Papacharissi Z (2015) *Affective Politics: Sentiment, Technology and Politics*. New York: Oxford University Press.
- Pi S, Chou C and Liao H (2013) A study of Facebook Groups members’ knowledge sharing. *Computers in Human Behavior* 29(5): 1971–1979.
- Placencia M and Lower A (2013) Your kids are so stinkin’cute!(-): complimenting behavior on Facebook among family and friends. *Intercultural Pragmatics* 10(4): 617–646.
- Rainie L, Smith A, Schlozman K, et al. (2012) Social media and political engagement. *Pew Internet & American Life Project* 19: 2–13.
- Ribeiro F, Ufmg U, Henrique L, et al. (2018) On microtargeting socially divisive ads: a case study of Russia-linked ad campaigns on Facebook. In: *Presented at FAT conference on fairness, accountability and transparency FAT’19*, Atlanta, GA, 29–31 January. New York: ACM. Available at: <https://arxiv.org/pdf/1808.09218.pdf>. <https://doi.org/10.1145/3287560.3287580>
- Sehl K (2020) Top Twitter demographics that matter to social media marketers. *Hootsuite Blog*. Available at: <https://blog.hootsuite.com/twitter-demographics/>

- Stanley J (2015) *How Propaganda Works*. Princeton, NJ: Princeton University Press.
- Stieglitz S and Dang-Xuan L (2013a) Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems* 29(4): 217–248.
- Stieglitz S and Dang-Xuan L (2013b) Social media and political communication: a social media analytics framework. *Social Network Analysis and Mining* 3(4): 1277–1291.
- Stukal D, Sanovich S, Bonneau R, et al. (2017) Detecting bots on Russian political Twitter. *Big Data* 5(4): 310–324.
- Twitter (2019) *Retrospective review: Twitter, Inc. and the 2018 midterm elections in the United States*. Available at: [https://blog.twitter.com/content/dam/blog-twitter/official/en\\_us/company/2019/2018-retrospective-review.pdf](https://blog.twitter.com/content/dam/blog-twitter/official/en_us/company/2019/2018-retrospective-review.pdf)
- US House of Representatives (2018) Permanent select committee on intelligence: social media advertisements. Available at: <https://intelligence.house.gov/social-media-content/social-media-advertisements.htm>
- US Senate (2019) Russia's use of social media with additional views. Volume 2. Report of the Senate Intelligence Committee. Available at: [https://www.intelligence.senate.gov/sites/default/files/documents/Report\\_Volume2.pdf](https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf)
- Utych SM (2018) Negative affective language in politics. *American Politics Research* 46(1): 77–102.
- Xia J, Lukito J, Chang Y, et al. (2019) Disinformation, performed: self-presentation of a Russian IRA account on Twitter. *Information, Community & Society* 22(11): 1646–1664.
- Zembylas M (2019) The affective grounding of post-truth: pedagogical risks and transformative possibilities in countering post-truth claims. *Pedagogy, Culture & Society* 28: 77–92.

## Author biographies

Sharon Stover is a Professor in the School of Journalism and Media School at the University of Texas at Austin, where she also directs the Technology & Information Policy Institute. Her research interests include disinformation, ethical AI, the impact of technologies, and public policy regarding communication systems.

Soyoung Park is a Ph.D. candidate in Media Studies at the Moody College of Communication and a graduate researcher at the Technology and Information Policy Institute at UT Austin. Her research addresses the social and political implications of emerging technology, focusing on socio-technical systems, data, and digital literacy.

Jaewon Royce Choi is a PhD candidate in the Department of Radio-TV-Film, Moody College of Communication at the University of Texas at Austin. His academic interests include new media technology adoption and its implications on industry, policy, and society; digital inclusion; digital skills; IoT; and privacy.

Mackenzie Schnell is pursuing her masters in Organizational Communication & Technology and her MBA at the University of Texas at Austin. Her research focuses on sustainability and social change.