The Gender Dimensions of Foreign Influence Operations

SAMANTHA BRADSHAW[1]
Stanford University, USA

AMÉLIE HENLE
Oxford University, UK

Drawing on a qualitative analysis of 7,506 tweets by state-sponsored accounts from Russia's GRU and the Internet Research Agency (IRA), Iran, and Venezuela, this article examines the gender dimensions of foreign influence operations. By examining the political communication of feminism and women's rights, we find, first, that foreign state actors co-opted intersectional critiques and countermovement narratives about feminism and female empowerment to demobilize civil society activists, spread progovernment propaganda, and generate virality around divisive political topics. Second, 10 amplifier accounts—particularly from the Russian IRA and GRU—drove more than one-third of the Twitter conversations about feminism and women's rights. Third, high-profile feminist politicians, activists, celebrities, and journalists were targeted with character attacks by the Russian GRU. These attacks happened indirectly, reinforcing a culture of hate rather than attempting to stifle or suppress the expression of rights through threats or harassment. This comparative look at the online political communication of women's rights by foreign state actors highlights distinct blueprints for foreign influence operations while enriching the literature about the unique challenges women face online.

Keywords: foreign influence operations, feminism, gender, disinformation, social media, social movements

Following the outcome of the 2016 U.S. presidential election, foreign influence has been a growing topic of academic inquiry (Bradshaw & Howard, 2017; Lin & Kerr, 2019; Martin & Shapiro, 2019; McGregor, Barrett, & Kreiss, 2021; Moore, 2018; Rid, 2020; Singer & Brooking, 2019; Starbird, Arif, & Wilson, 2019).

Since then, a wave of research has examined how disinformation—the purposeful spread of false or misleading information—was used by foreign states, among others, to manipulate public opinion via social media (Allcott & Gentzkow, 2017; Benkler, Faris, & Roberts, 2018; Bennett & Livingston, 2018; Grinberg, Joseph, Friedland, Swire-Thompson, & Lazer, 2019). Researchers have documented how these campaigns have homed in on national issues and local communities, spreading disinformation to inflame debates about race and religion and deepen polarization across the political spectrum (DiResta et al., 2018; François, Nimmo, & Eib, 2019; Freelon et al., 2020; Friedberg & Donovan, 2019; Howard, Ganesh, Liotsiou, Kelly, & Francois, 2018; Woolley & Joseff, 2019). However, the gender dimension of these campaigns has received less attention.

Social media is used by millions of women around the world to express their freedom of speech, contribute to the global digital economy, and participate in public life. Platforms like Facebook and Twitter provide space for women and girls to connect with others, develop a sense of collective identity, and participate in online activism (Crossley, 2015; Locke, Lawthom, & Lyons, 2018; Lopez, Muldoon, & McKeown, 2019). But for many women—especially women of color or with diverse gender identities—social media is where they experience sexism, harassment, and threats (Brown, Ray, Summers, & Fraistat, 2017; Cole, 2015; Eckert, 2018; Hardaker & McGlashan, 2016; Jankowicz et al., 2021; Mantilla, 2015; Mendes, Ringrose, & Keller, 2018; Sobieraj, 2020; Phillips, 2015). The vocabulary of gender-based attacks has also been adopted by state actors to suppress women's rights and civic participation through accusations of collusion, sexualized tropes, and threats of rape and violence (Bradshaw, 2019a; Jankowicz, 2017; Judson, Atay, Krasodomski-Jones, Lasko-Skinner, & Smith, 2020; Monaco, Woolley, & Nyst, 2018). Qualitative accounts have highlighted how these kinds of targeted attacks against women can have measurable impacts on the online behaviors of victims (Amnesty International, 2017; Zeiter, Pepera, & Middlehurst, 2020).

Although women experience disinformation and harassment differently, there have been no systematic studies that examine the gender dimensions of foreign influence operations. Some theories about contemporary influence operations describe how state actors do not galvanize support for a particular person or idea but instead exploit ideological differences (Freelon & Lokot, 2020; Freelon et al., 2020). Indeed, academic, journalistic, and independent investigations into the U.S. 2016 presidential elections highlighted concentrated efforts targeting conservative and Black American voters with polarizing disinformation (DiResta et al., 2018; Howard et al., 2018; Mueller, 2019; Nimmo, 2016). However, no studies have systematically examined how foreign state actors instrumentalize narratives about gender—and narratives about women—in contemporary influence operations.

Our research fills this gap by exploring how foreign state actors engage in covert influence operations targeting feminist activists and politicians. Drawing on a qualitative analysis of 7,506 tweets from Twitter's information operations transparency reporting, we examine the narratives and strategies foreign state actors use to target feminist movements and their advocates. Our research not only provides a unique opportunity to comparatively assess how different states approach foreign influence operations but also provides insight into how these campaigns interact with gender and feminism. While critical research about foreign influence operations has focused largely on racial and political asymmetries that emerge between conservatives and liberals, or between Black and White Americans, we explore how gender identity is instrumentalized.

Contemporary Foreign Influence Operations

Social media platforms have become vectors for foreign influence operations—clandestine operations by foreign state actors that seek to undermine information systems and manipulate civic discourse (Martin & Shapiro, 2019; Waltzman, 2017). Sometimes considered a tool of "asymmetric," "nonlinear," or "hybrid" warfare, social media are used as alternatives for or complements to traditional kinetic warfare (Lin & Kerr, 2019; Rid, 2020; Starbird et al., 2019). Although foreign influence operations are not new, the affordances of social media platforms—including algorithms, automation, and data—can enhance the scope, scale, and precision of these campaigns (Bradshaw, 2019b; Bradshaw & Howard, 2018; Hwang & Rosen, 2017). The ubiquity of social networking technologies, combined with the low cost to produce and disseminate content online, presents a qualitatively new landscape for persuasion, manipulation, and hybrid forms of warfare (Lin & Kerr, 2019; Moore, 2018; Rid, 2020).

The study of foreign influence operations regained prominence following the 2016 U.S. presidential election, when major social media companies confirmed their platforms had been co-opted by both foreign and domestic actors (Sanders, 2016). This motivated a dedicated corps of researchers, technology integrity teams, governments, and journalists to detect, study, and combat influence operations on social media. Academic studies and journalistic investigations have looked at the variety of actors, narratives, and technological tools used to influence publics during critical political moments and around contentious political topics such as immigration, climate change, and the novel coronavirus (Benkler et al., 2018; Freelon et al., 2020; Marwick & Lewis, 2017; Nguyen & Catalan, 2020; Starbird et al., 2019; Woolley & Howard, 2018). One aspect of these operations that has received less attention is the gender dimension of these campaigns.

State actors have used social media to exclude women from politics and public life through intimidation tactics, harassment, and gendered disinformation (MacKinnon, 2012; Monaco et al., 2018). Anecdotal evidence about these state-sponsored gender attacks highlights how government-run fake accounts target women with threats of rape and violence, accusations of collusion, and disinformation to undermine their legitimacy in public spaces. By relying on gender stereotypes and sexualized tropes, state-sponsored gender attacks suppress the participation of women in public life while advancing other political objectives (Fichman & McClelland, 2021; Jankowicz et al., 2021; Judson et al., 2020). Qualitative accounts highlight how targeted attacks against women have measurable impacts on the behaviors, professions, and economic security of victims (Amnesty International, 2017; Zeiter et al., 2020).

Although women experience disinformation differently, most academic research on foreign influence operations has focused on political and, to a lesser extent, racial identities. Contemporary theories about foreign influence operations highlight how race and political identity are used to fracture social cohesion (Freelon & Lokot, 2020; Freelon et al., 2020; Friedberg & Donovan, 2019; Tucker et al., 2018). When it comes to political identity, empirical studies have found that foreign state actors generate more content targeting conservative users, and conservative users tend to share disinformation more than their liberal counterparts (Fichman & McClelland, 2021; Guess, Nagler, & Tucker, 2019; Osmundsen, Bor, Vahlstrup, Bechmann, & Petersen, 2020). There have also been extensive influence operations targeting Black American communities, where fake accounts masquerading as Black Americans spread disinformation to demobilize activist communities and suppress voter turnout (DiResta et al., 2018; Freelon et al., 2020;

Howard et al., 2018). Although research shows that certain groups of users experience disinformation disproportionately, no studies have systematically examined how foreign state actors target high-profile women or feminist movements as part of their influence operations.

While some traditional theories of propaganda focus on ways that actors foster support for a particular person or idea, contemporary theories suggest that influence and persuasion are functions of political and group identity, and that covert operations exploit digital affordances to share, amplify, and target content that provokes resentment against oppositional groups (Freelon & Lokot, 2020; Freelon et al., 2020; Friedberg & Donovan, 2019; Tucker et al., 2018). Scholars have enriched our understanding of these processes by examining the role of race in these "ideological asymmetries," which can be highly divisive and engaging narratives (Freelon et al., 2020, p. 2). We build on this critical work by asking whether gender leads to a qualitatively different experience or understanding of contemporary influence operations. In addition to exploring how harassment is perpetuated through foreign influence operations, we also explore how state actors co-opt narratives around feminism as part of their campaigns.

Feminism, Online Social Movements, and Collective Identity Formation

To properly conceptualize how feminist movements are co-opted for foreign influence operations, it is important to discuss contemporary feminist movements and the relationship between collective identity formation and mobilization. Understanding how and why people come together and mobilize clarifies how these processes can be instrumentalized for demobilization. Theories about collective action mobilization suggest that the development of collective identity is a key factor for mobilizing action (Buchan et al., 2011; Jenkins, 1983; Van Zomeren, Spears, Fischer, & Leach, 2004). Before mobilization occurs, individual members must identify and share common concerns, experiences, and feelings (Klandermans, 1997; Polletta & Jasper, 2001; Simon et al., 1998). This sense of collective identity is not static; rather, it is constantly defined, redefined, and renegotiated through everyday interactions among group members (Barassi, 2018; Hopkins & Blackwood, 2011; Khazraee & Novak, 2018). Empirical studies suggest that sharing a sense of collective identity can facilitate feelings of togetherness and commonality, which can increase trust and support and improve a group's ability to collectively mobilize. When individual group members attach their sense of self to their group membership, the pursuit of group interests becomes interchangeable with pursuing one's own interests (Buchan et al., 2011; Van Zomeren et al., 2004).

Feminist movements are no exception to these findings, and social media provides a new environment for women and girls to learn about feminist activism, connect with others around the globe, and develop a sense of collective identity (Crossley, 2015; Locke et al., 2018). Scholars have documented the rise of various digital feminist campaigns that speak to broad issues of violence, rape, injustice, and inequality facing women in both online and offline spaces (Fileborn & Loney-Howes, 2019; Loney-Howes, 2018; Sobieraj, 2018). By leveraging the unique affordances of digital platforms, feminist movements have used platforms like Twitter to organize, structure, and make accessible the movements to an ever-growing audience, facilitating a sense of collective identity, support, and empathy (Turley & Fisher, 2018). But the high visibility of feminist movements opens participants to a number of vulnerabilities, such as the further proliferation of misogyny and harassment (Banet-Weiser & Miltner, 2016; Boynton, 2012; Cole, 2015; Hardaker & McGlashan, 2016; Mantilla, 2015). Global, visible, and performative movements also have the

potential to exclude people based on specific elements of their identities, and many scholars and activists have critiqued feminist movements for maintaining White and Western-centric perspectives (Daniels, 2015; Liska, 2015; Phipps, 2019). These intersectional critiques and countermovement narratives are spaces where collective identity and shared norms are challenged.

As with other movements, there has been some evidence of foreign influence operations co-opting the language of feminism and its countermovements and critiques (Bradshaw, 2019a). For example, pro-Kremlin outlets have spread content around notions of totalitarian feminism: the idea that women supporting feminist movements want to bully men and discourage "men's interest in women" by "turning them homosexual" (EU vs. Disinfo, 2019, para. 4). Similarly, other narratives have described feminism as incompatible with Islam, stating that in Islamic communities women are "forced to be prostitutes," "wear hijabs," or "undergo genital mutilation" (EU vs. Disinfo, 2019, para. 1). Despite the use of these frames, there have been no systematic analyses of how feminism is co-opted as part of foreign influence operations, and our study contributes to filling this gap.

In sum, our data allow us to both look at how high-profile women are targeted and harassed, and explore how foreign agents engage in rhetoric and narratives of women's rights and empowerment as part of contemporary influence operations. Thus, our core research questions ask:

RQ1:   *What are the key narratives foreign state actors use in influence operations that discuss women's rights?*

RQ2:   *What tactics do foreign state actors use when targeting high-profile feminist politicians, journalists, and activists?*

Methodology

Overview of Data: Twitter Data

We use data released by Twitter to examine the narratives and tactics foreign state actors use in covert influence operations targeting women and feminists. Although these operations occur across and on multiple platforms, Twitter has made available the most open and comprehensive data set about foreign influence operations to date as part of its transparency initiative. Twitter provides researchers with publicly available data sets about accounts and related content associated with potential influence operations that have taken place on the platform (Twitter, 2018). At the time of data collection and analysis, Twitter uploaded six data sets about influence operations from three different countries. The data used for this article are summarized in Table 1.

Table 1. Twitter's Transparency Reporting Data Set Summary.

Twitter Attribution	Date of Takedown	Number of Accounts
Russia (IRA)	October 2018	3,841 Accounts
Iran	October 2018	770 Accounts
Russia (GRU)	January 2019	418 Accounts
Iran	January 2019	2,320 Accounts
Venezuela	January 2019	1,196 Accounts
Venezuela	January 2019	764 Accounts

Note: Data were collected from Twitter's transparency reporting. Iran and Venezuela each had separate takedowns, resulting in separate data sets.

Twitter Data Limitations

While Twitter's data are currently the largest publicly available data set on state-sponsored influence operations, there are some limitations to working with these data. First, although conversations about politics often take place on Twitter, the platform has unique affordances that make social media manipulation campaigns different from those taking place on other platforms, such as character limits, anonymity through nonpermanent pseudonyms, and ease of creating new accounts. Thus, some of the findings may not easily transfer to other social media platforms. Second, Twitter does not provide any methodological information for detecting state-sponsored accounts. While offering researchers information on how these accounts were identified would help remove false positives, it could simultaneously improve the ability of these accounts to evade future detection. Thus, we can only assume the attribution conducted by Twitter is sufficient for this study. Third, Twitter did not provide any information about how many real people followed the accounts or clicked through on content. While retweet information is available, the data do not distinguish between real accounts and those run by state actors. Thus, it is not possible to measure the impact these campaigns had on real individuals. Finally, since the influence operations in this data set come from diverse geographic locations, the accounts published content in various languages. Twitter did not provide tweet language information for all the data in public repositories. For those where language information was available, the five most-used languages were Russian, American English, French, Hungarian, and Arabic. For our analysis, only tweets in American English (3,940,094 tweets) were analyzed. Thus, while influence operations target various countries and audiences, our findings are relevant only for English-speaking audiences, particularly in the United States, where most English-language tweets were labeled.

Sampling Strategy

State actors talked about various issues on Twitter. To focus on tweets specifically related to feminism, we applied topic-based sampling as a strategy to identify relevant data points. Topic-based sampling is a common methodology to study specific phenomena on Twitter, though representativeness is a major limitation when collecting tweets through Twitter's Streaming application programming interface (API) because it limits third-party queries to 1% of all Twitter traffic (boyd & Crawford, 2012). For our analysis, we assume that Twitter's information operations transparency data set is sufficiently representative for capturing the variety of conversations and activities carried out by foreign state actors on its platform.

Many Twitter studies begin the sampling process by using search terms or hashtags to identify relevant data (Kim et al., 2013). To build a clean and exhaustive sample of tweets about feminism, we developed a predetermined list of hashtags related to women's rights and activism and queried the data sets for tweets that contained these terms. The terms included #feminism, #feminist, #feminists, #womensrights, and #genderequality, leading to a sample of 5,149 tweets. We selected these hashtags to avoid under- and overestimating the volume of discussion about feminist movements.

We then used a snowball sampling strategy to identify any additional hashtags that could be present in the database. Building on our predefined list, we searched for collocated hashtags. To confirm new hashtags were predominantly concerned with issues to do with feminism, each collocated hashtag was independently reviewed, and those that were too general (such as #news or #local) were removed from the sample. Additionally, many hashtags were shared only one or two times. To capture the most relevant tweets, only collocated hashtags that were shared more than 20 times were included. This snowballing strategy added six new hashtags—#FeminismIsCancer, #FeminismIsAwful, #InternationalWomensDay, #Feminazi, #MeToo, and #WomensMarch—to the data set, and 2,357 new tweets. In total, 11 keywords and hashtags and 7,506 tweets in English were analyzed.

Overall, topic-based sampling can be accurate and representative, particularly because of the flexible nature that allows researchers to adapt and extend their lists of terms (Bartlett, Miller, Reffin, Weir, & Wibberley, 2014). Nevertheless, this approach introduces limitations concerning comprehensiveness (Bartlett et al., 2014; Cameletti et al., 2020; Hull & Grefenstette, 1996). By purposefully selecting specific terms and manually reviewing collocated hashtags, our sampling strategy tended to value precision over comprehensiveness (Tulkens, Hilte, Lodewyckx, Verhoeven, & Daelemans, 2016). Although this sample might not capture *all* of the gender-related conversations held by foreign state actors, best efforts were made to ensure the data in this sample were clean, representative, and relevant for the topic of study.

Typology Building and Coding

Tweets were coded using a grounded theory (Charmaz, 2006). We began with a line-by-line review of a small subset of data (20%, or 1,500 tweets) to identify the discourse types emergent from the data set. Then we manually grouped the discourse types into 11 categories to articulate common themes (e.g., tweets that celebrated a specific person for promoting female empowerment or women's rights in some way, tweets that distorted feminist principles or values, polemic tweets that attacked individuals or the women's rights movement, or informational tweets that shared resources relevant to those interested in feminism). We then applied focused coding to the rest of the data set (6,006 tweets; Charmaz, 2006). Reliability scores showed that there was substantial agreement among the coders in terms of how to categorize tweets relative to our discourse types (Krippendorff's α = .84). All disagreements were reviewed and collectively resolved. To identify additional relational patterns, we manually grouped the 11 discourse types into three sentiment categories: (1) positive tweets that supported, celebrated, or promoted women's rights, (2) negative tweets that attacked individuals or the movement more broadly, or (3) neutral tweets that were informational in nature, without a

positive or negative sentiment. Table 2 provides a summary of the discourse types and sentiments of the tweets. A full description of the typology can be found in the Appendix.[2]

Table 2. Discourse Type and Coding Decisions.

Discourse Type	Full Data Set	IRA	GRU	Iran	Venezuela
Positive tweets					
Positive—character	125 (2%)	73 (3%)	21 (2%)	8 (3%)	23 (1%)
Solidarity	562 (9%)	331 (15%)	110 (9%)	57 (21%)	64 (3%)
Calls to action	53 (1%)	32 (1%)	9 (1%)	7 (3%)	5 (0.2%)
Negative tweets					
Attacks—movement	1,092 (16%)	735 (28%)	206 (16%)	8 (3%)	143 (6%)
Attacks—individuals	283 (4%)	77 (3%)	94 (7%)	3 (1%)	109 (4%)
Attacks—Clinton	264 (4%)	74 (3%)	68 (5%)	4 (1%)	118 (5%)
General value attacks	228 (3%)	150 (6%)	56 (4%)	2 (1%)	20 (1%)
Islamophobic tweets	633 (10%)	150 (6%)	209 (17%)	12 (4%)	262 (10%)
Liberal feminism critiques	226 (3%)	150 (6%)	46 (4%)	10 (4%)	20 (1%)
Male rights juxtaposition	128 (2%)	81 (3%)	38 (3%)	0	9 (0.3%)
Female rights juxtaposition	346 (5%)	95 (4%)	33 (3%)	4 (2%)	214 (9%)
Willful ignorance	156 (2%)	69 (3%)	79 (6%)	3 (1%)	5 (0.2%)
Tabloid tweets	1,616 (24%)	121 (5%)	87 (7%)	37 (14%)	1,371 (55%)
Neutral tweets					
Information sharing	497 (7%)	242 (9%)	75 (6%)	26 (10%)	154 (6%)
NA or spam	163 (2%)	68 (3%)	21 (2%)	15 (5%)	59 (2%)
Hashtag spamming	336 (5%)	146 (5%)	106 (9%)	80 (29%)	4 (0.1%)

Note: Data are based on qualitative analysis of tweets. Percentages are calculated based on the proportion of each discourse type per state actor.

The Narratives About Feminism and Women's Rights

Polarization and Gender Asymmetries

Research about foreign influence operations describes how foreign state actors instrumentalize political and group identity to share, amplify, and target content that promotes in-group solidarity and provokes resentment against oppositional groups (Freelon & Lokot, 2020; Freelon et al., 2020; Guess et al., 2019; Osmundsen et al., 2020; Tucker et al., 2018). This is particularly relevant for racial identity, where empirical research has demonstrated a concentrated strategy to amplify racial divisions in the United States (DiResta et al., 2018; Freelon & Lokot, 2020; Howard et al., 2018). We found similar patterns in the

---

2 The online appendix can be found at: https://purl.stanford.edu/cq131gj2513.

narratives foreign state actors employed around feminist movements, but rather than promoting in-group solidarity, conversations about feminism were skewed toward amplifying negative sentiment about the movement and its activists more broadly.

First, foreign influence operations disproportionately shared content that attacked feminist movements over content that promoted in-group solidarity. The IRA was the most antagonistic, where more than 64% of the IRA tweets focused on spreading negative sentiment about feminism, compared with less than 20% of tweets that were coded as expressing positive sentiment. IRA accounts focused largely on misogynistic polemics that attacked the feminist movement and its activists. For example, one IRA tweet stated, "Feminists are more likely to rely on petty insults in spite of giving valid arguments #FeminismIsAwful" (Twitter Transparency, 2018).[3] Like other gender-based attacks, state-sponsored accounts discussing feminism on Twitter focused on the appearances of women and activists: "RT Handful of self-entitled fat ugly feminists trying to get arrested in desperate attempt to impress any man" or "RT #FeministsAreUgly is trending by feminists themselves just post pics amp get attention don't y'all have women to save?" (Twitter Transparency, 2018). In contrast to the IRA, Venezuelan accounts used clickbait, sensation, and fear to drive negative sentiment and polarization about feminism. For example, one tweet read, "VIDEO Feminists Now Embracing RADICAL ISLAM" (Twitter Transparency, 2019c). Overall, tweets by state actors and, to a greater extent the IRA, focused on amplifying negative sentiment about feminism rather than promoting solidarity.

Additionally, research has highlighted how racial and political identities are leveraged to provoke oppositional groups and exacerbate polarization. We wanted to explore the asymmetries between men and women in the context of foreign influence operations. IRA and GRU accounts were the only state actors to share content that discussed asymmetries about gender identity, juxtaposing male versus female rights. Here, tweets focused on distorting principles of equality—for example, by making arguments that feminism is about suppressing male rights as opposed to supporting equality for men and women. Narratives often presented feminists as "man-hating" oppressors, claiming that modern-day feminism leads to the suppression of male rights. One tweet, for example, stated, "#IHaveADreamThat feminism will stop being a man hating ideology and girls stop accusing us of rape if sex wasn't good" (Twitter Transparency, 2018). Content describing these divisions—between men and women—made up only a small proportion of the overall tweets (6% of the total tweets shared by both the IRA and GRU accounts). Thus, theories about contemporary influence operations that focus on amplifying divisions among identity groups do not fully fit with an analysis on gender.

Undermining Women's Collective Identity

Rather than focusing on divisions *between* men and women, foreign influence operations made use of narratives that undermined women's sense of collective identity. This was a strategy primarily adopted by the Russian-backed accounts, where Russian state actors co-opted intersectional critiques that have

---

3 In our analysis, all tweets listed as examples have been removed by Twitter. However, the data used in this article, including all the tweets referenced, can be requested for download at: https://transparency.twitter.com/en/reports/information-operations.html.

arisen from and exist within the broader feminist movement. Their strategies were not based on driving divisions between men and women, or between feminism and society more broadly, but rather were about fracturing the feminist movement and undermining women's ability to formulate a shared identity.

Tweets in our data discussed three main intersectional critiques: that feminism is (1) too White to represent Black women, (2) too liberal to represent conservative women, and (3) too wealthy to represent poor women. In all three cases, the IRA co-opted the narratives that feminist activists have themselves raised about contemporary feminist movements in both online and offline spaces (Daniels, 2015; Liska, 2015; Phipps, 2019). Intersectional critiques about racial or political identities were predominantly discussed by Russian accounts, where more than 86% of the tweets came from IRA (66%) or GRU (20%) accounts. These accounts used discourses about feminism being rooted in "White" and "liberal" values, and that Black women or conservative women are not represented by the movement. For example, one IRA tweet stated, "RT A quick search of white feminist & #BlackLivesMatter yielded ZERO results. Gee I wonder why?" (Twitter Transparency, 2018). Another example includes "RT Funny how feminists are all for the advancement of women unless that woman is a Conservative #TCOT #MAGA @FoxNews" (Twitter Transparency, 2018). In addition to race and political identity, IRA accounts also discussed class identity and the representation of poor or marginalized groups of women; one example reads, "RT Liberal feminists ignore poor women entirely in their activism That's why their activism is garbage" (Twitter Transparency, 2018). The co-option of these narratives about feminism is used to undermine women's ability to form the collective identity necessary for political mobilization.

Strategies for Targeting Feminists and Their Movements

Harassment and Character Attacks Against Women

Research about women and disinformation has consistently demonstrated that high-profile women face distinct patterns of hate, harassment, and gender-based attacks online (Jankowicz, 2017; Judson et al., 2020; Monaco et al., 2018). When these kinds of attacks are perpetrated by state actors, oppressive regimes can combine online harassment with threats of real-world violence or imprisonment to stifle fundamental freedoms (Deibert, 2013; MacKinnon, 2012). However, no systematic studies have examined how foreign state actors use gender-based attacks as part of their foreign influence operations.

Our data provide some insight into these questions. We found concentrated efforts to delegitimize and discredit female activists, celebrities, politicians, and, to a lesser extent, journalists and users, who discussed or participated in feminist movements on Twitter. The IRA, GRU, and Venezuelan accounts launched the most character attacks against high-profile feminists (see Figure 1), reflecting tactics of suppression and discrediting often seen in these countries' domestic contexts (MacKinnon, 2012; Monaco et al., 2018). However, across all three data sets in the sample, the accounts rarely used the @mention feature when launching character attacks. Instead, attacks against women often used their names in plaintext along with hashtags that were popular or trending. This could suggest that foreign influence operations make less use of *publicly* targeted attacks on prominent voices as part of their campaigns abroad. We have two explanations for this less-visible approach. The first is that attacking or harassing individuals domestically comes with better technical or legal measures for enforcing actions that would silence speech—such as

physical harm or arrests that are so often combined with state-sponsored attacks. Another explanation could be that targeted harassment happens more through direct messages as opposed to public-facing tweets, which are more likely to be flagged by users or by Twitter's automatic detection measures.
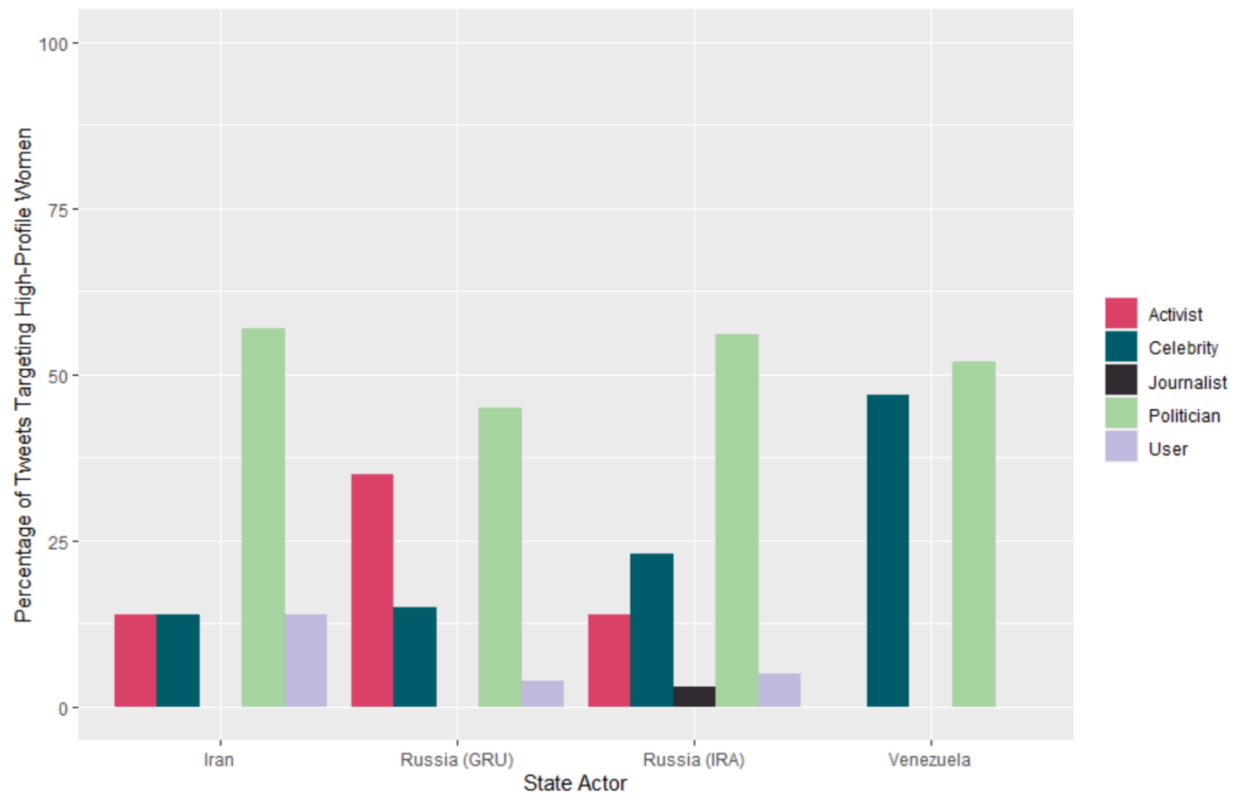

Figure 1. Targeted attacks against high-profile women.
Note. Data were collected from Twitter. Percentages were calculated based on the total number of tweets targeting high-profile women per state actor.

Nevertheless, we still identified a number of character attacks about specific women. Across all three data sets, politicians were discussed the most by state-sponsored accounts, with attacks against Hillary Clinton accounting for more than 90% of all attacks made against politicians. Users (everyday Twitter users who are not considered "high profile," i.e., not recognizable or famous) were rarely attacked, with less than 5% of the total character attack tweets targeting users. The GRU launched the most character attacks against women on Twitter (30% of all character attack tweets). In particular, they focused on creating and amplifying ad hominem attacks about Hillary Clinton and Linda Sarsour, the American political activist who organized the 2017 Women's March. GRU accounts tweeted content suggesting Sarsour wanted to implement sharia law in the United States or was secretly a jihadi terrorist infiltrating America (Twitter Transparency, 2019a). One explanation for the use of Islamophobic narratives could be that traditional propaganda campaigns have used narratives about Islam, and orientalism more broadly, to spread fear to justify military intervention (Khalid, 2011). While the dissemination of Islamophobic disinformation is

modern in its distribution methods, the topics and content are highly similar to Soviet operations of the past (Lukito, 2020; Tucker et al., 2018). Thus, the use of these narratives could simply be an extension of traditional military propaganda strategies into the digital realm.

While GRU accounts had more character attacks targeting high-profile activists with traditional propaganda frames, the IRA focused attacks on high-profile feminist celebrities or women supporting feminist causes. Madonna was the biggest celebrity target of the IRA accounts, with more than one-third (35%) of all character attacks coded in the data set mentioning her. Venezuelan accounts also made character attacks against many female and feminist celebrities, but these accounts relied on clickbait and sensation to discredit female figures. For example, one tweet stated, "CAT FIGHT Legendary Lesbian Feminist Declares Hillary Exploits Feminism" (Twitter Transparency, 2019c). Overall, character attacks against female politicians, activists, and celebrities by accounts from the IRA, the GRU, and Venezuela may have differing tactics, but they all work to discredit and delegitimize women involved in feminist discourse.

Finally, it is important to highlight one outlier: In contrast to the IRA, GRU, and Venezuela, Iranian accounts were positive in their discussions about women's rights and empowerment and celebrated feminist figureheads on Twitter. This was a surprising finding considering Iran as one of the worst countries in the world for women's rights (Freedom House, 2019; Human Rights Watch, 2019). Iranian accounts focused their engagement with English-speaking audiences on timely and high-profile events, generating positive tweets about various feminist movements. More than 20% of all the Iranian tweets in the sample positively tweeted out solidarity statements about the 2017 Women's March or the #MeToo movement. Theories about Iranian influence operations describe how Iran promotes its geopolitical interests abroad by employing a "distorted truth" strategy that exaggerates "Iran's moral authority while minimizing the repression of its citizens" (Brooking & Kianpour, 2020, p. 2). Supporting feminist movements could be an extension of this "distorted truth" strategy, especially considering that the #MeToo movement and the 2017 Women's March were highly critical of Donald Trump's presidency, which is in line with Iran's anti-Trump stance and geopolitical interests.

Amplification Through Automation and Hashtag Hijacking

The use of political bots to spread disinformation has become a growing area of academic research as investigations have uncovered networks of "bot" accounts used in foreign influence operations (Bessi & Ferrara, 2016; Borak, 2019; Bradshaw, Bailey, & Howard, 2021; Bradshaw & Howard, 2019; Elswah, Howard, & Narayanan, 2019; Murthy et al., 2016; Owen Jones, 2016). The detection and identification of bots is technically complex, with multiple definitions and measures for capturing this kind of activity (Gorwa & Guilbeault, 2020). One useful way for defining automated activity is what McKelvey and Dubois (2017) call "amplifier accounts": those that deliberately seek to increase the number of voices speaking about or the attention paid to certain messages (p. 9). Unlike traditional "bot" definitions, amplifier accounts include automated, semiautomated, and highly active human-curated accounts on social media (McKelvey & Dubois, 2017, p. 9).

State-sponsored amplifier accounts were present in the data set. In total, 10 accounts tweeted more than one-third of the content in the sample. Nine out of the 10 accounts were Russian-backed, with

four accounts operated by the GRU and five operated by the IRA. All nine Russian-operated accounts amplified negative content about feminism and women's rights. The most active GRU account amplified character attacks to suppress and delegitimize high-profile feminists. Similarly, the most active IRA account amplified demobilization messages that attacked the feminist movement more broadly. These amplifier accounts not only focused on issues to do with feminism and women's rights but also tweeted about many other issues in the data provided by Twitter. For issues covered in the sample, the GRU amplifier accounts mainly retweeted content. This fits in line with other analyses that suggest the GRU creates content for fake media websites, think tanks, or political commentators and amplifies this content with other fake accounts to generate a false sense of legitimacy and credibility (DiResta & Grossman, 2019). In contrast, the IRA amplifier accounts created original content, with the most active amplifier account tweeting 502 original tweets out of the IRA sample.

It is important to note that amplifier accounts might be overrepresented in the data set because detecting automated account activity is a common method Twitter uses for identifying coordinated inauthentic behavior. Additionally, amplified content does not need to be automated by highly active amplifier accounts. For example, many tweets in the Venezuelan data set appear to have been copied and pasted; one such tweet, "Donald Trump Responds To #WomensMarch Liberal Snowflakes Explode On Twitter," was posted by the Venezuelan accounts more than 70 times (Twitter Transparency, 2019c).

In addition to amplifier accounts, Iranian operations used conversations about feminism to amplify other state-sponsored propaganda unrelated to women's rights. Almost one-third (29.2%) of the Iranian tweets in the sample "hijacked" hashtags, where accounts simply used prominent or trending hashtags to spread unrelated messages by linking them to popular conversations. One example is the use of the #WomensMarch hashtag to share news about Saudi air strikes in Yemen (Twitter Transparency, 2019b). Iran also had the largest proportion of tweets coded as NA—tweets were simply spam or were not comprehensible to the coders. Hashtag hijacking and spam are reflective of trends identified by other investigations into Iranian influence operations, where accounts are much less sophisticated and make use of crude automation techniques to hijack conversations rather than engaging with issues that are important to the audiences discussing them (Leprince-Ringuet, 2018; Nimmo et al., 2020).

Conclusion

Foreign influence operations on social media are increasingly important areas of study, as state actors exploit the affordances of new technologies for geopolitical gain. Like other online social movements, conversations about women's rights and empowerment have been affected by such activities. This article sheds some light on how foreign actors use narratives around gender identity to influence, polarize, and disrupt online discussions of feminism. As they have done with other social movements, foreign state actors have distorted the values of feminism to polarize and alienate the movement from broader communities. But attacks did not focus only on divisions *between* men and women or between feminism and society more broadly; rather, foreign state actors made use of narratives to undermine women's sense of collective identity, instrumentalizing in-group tensions *in addition to* provoking resentment against oppositional groups.

By examining the narratives and strategies four state-backed actors used when discussing feminism, we also found distinct blueprints for contemporary influence operations. First, Russian accounts focused their efforts on the demobilization of feminist movements. IRA accounts hijacked legitimate criticisms of the activist communities to exploit fracture lines within the community itself, whereas GRU accounts perpetuated long-standing xenophobic narratives to spread disinformation and undermine the work of activists in this space. Second, Iranian accounts focused operations on the dissemination of information that supported their own political agendas. They perpetuated positive narratives that reinforced their political goals while "hijacking" popular hashtags about women's rights to spread propaganda about unrelated political issues. Finally, Venezuelan accounts relied mainly on clickbait, crime, and sensation to generate virality around divisive political topics. These blueprints fit in with other research about influence operations by Russian, Iranian, and Venezuelan accounts.

While direct attacks on high-profile individuals certainly occurred, contrary to what we expected, most targeted attacks against high-profile women were indirect in nature: foreign state actors relied on amplification to spread false claims and establish narratives to discredit prominent voices, as opposed to threatening or harassing activists who supported these causes. We suggest more research is needed to understand the mechanisms of online harassment by foreign state actors outside the context of Twitter. A broader cross-platform approach that looks at the different affordances of platforms—such as private messaging—would be an area of future inquiry. Additionally, while we examined gender through the lens of female experiences, further research should examine how other dimensions of gender, in particular LGBTQ+ communities, experience disinformation and foreign influence operations.

In conclusion, our research brings attention to the way civil rights movements in general, and feminism in particular, are hijacked to polarize and disrupt democratic action. The narratives, targeting strategies, and tactics we identified not only influence the ways women participate in public spaces and online movements but also influence how society more broadly views feminist activism. These campaigns can also influence the ways women themselves think about their shared sense of identity and online activism. Civil society is the heart of democratic societies—indeed, people's ability to collectively organize and push for change on issues that matter to individuals and to society encompasses the very definition of democracy. Online spaces, such as Twitter, have certainly created new spaces for civic action. But they have simultaneously opened opportunities for other actors to silence speech and discourage mobilization. Influence operations targeting women and feminist movements ultimately harm gender equality and stifle women's ability to advocate for social change.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives, 31*(2), 211–236. doi:10.1257/jep.31.2.211

Amnesty International. (2017). *Amnesty reveals alarming impact of online abuse against women*. Retrieved from https://www.amnesty.org/en/latest/news/2017/11/amnesty-reveals-alarming-impact-of-online-abuse-against-women/

Banet-Weiser, S., & Miltner, K. M. (2016). #MasculinitySoFragile: Culture, structure, and networked misogyny. *Feminist Media Studies, 16*(1), 171–174. doi:10.1080/14680777.2016.1120490

Barassi, V. (2018). Social media activism, self-representation and the construction of political biographies. In G. Meikle (Ed.), *The Routledge companion to media and activism* (pp. 142–151). London, UK: Routledge.

Bartlett, J., Miller, C., Reffin, J., Weir, D., & Wibberley, S. (2014). *Social media is transforming how to study society*. Retrieved from https://www.demos.co.uk/files/Vox_Digitas_-_web.pdf?1408832211

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford, UK: Oxford University Press.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication, 33*(2), 122–139. doi:10.1177/0267323118760317

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US Presidential election online discussion. *First Monday, 21*(11). doi:10.5210/fm.v21i11.7090

Borak, M. (2019, August 23). Bots or people? Pro-China disinformation campaigns make it hard to tell. *South China Morning Post*. Retrieved from https://www.scmp.com/abacus/culture/article/3029567/bots-or-people-pro-china-disinformation-campaigns-make-it-hard-tell

boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679. doi:10.1080/1369118X.2012.678878

Boynton, P. (2012). Getting the press we deserve: Opportunities and challenges for innovative media practice. *Feminism & Psychology, 22*(4), 536–540. doi:10.1177/0959353512463557

Bradshaw, S. (2019a). *The gender dimensions of foreign influence operations: Challenges and opportunities*. Ottawa, Canada: Global Affairs Canada.

Bradshaw, S. (2019b). Disinformation optimised: Gaming search engine algorithms to amplify junk news. *Internet Policy Review, 8*(4). doi:10.14763/2019.4.1442.

Bradshaw, S., Bailey, H., & Howard, P. (2021). *Industrialized disinformation: 2020 global inventory of organized social media manipulation.* Computational Propaganda Research Project. Retrieved from https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/02/CyberTroop-Report20-Draft9.pdf

Bradshaw, S., & Howard, P. (2017). The global organization of social media disinformation campaigns. *Journal of International Affairs, 71*(1.5). Retrieved from https://jia.sipa.columbia.edu/global-organization-social-media-disinformation-campaigns

Bradshaw, S., & Howard, P. (2018). *Why does junk news spread so quickly across social media? Algorithms, advertising and exposure in public life*. Knight Foundation. Retrieved from https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/142/original/Topos_KF_White-Paper_Howard_V1_ado.pdf

Bradshaw, S., & Howard, P. N. (2019). *Global disinformation disorder: 2019 inventory of social media manipulation*. Computational Propaganda Project. Retrieved from https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf

Brooking, E. T., & Kianpour, S. (2020, February 12). *Iranian digital influence efforts: Guerrilla broadcasting for the twenty-first century*. Atlantic Council. Retrieved from https://www.atlanticcouncil.org/in-depth-research-reports/report/iranian-digital-influence-efforts-guerrilla-broadcasting-for-the-twenty-first-century/

Brown, M., Ray, R., Summers, E., & Fraistat, N. (2017). #SayHerName: A case study of intersectional social media activism. *Ethnic and Racial Studies, 40*(11), 1831–1846. doi:10.1080/01419870.2017.1334934

Buchan, N. R., Brewer, M. B., Grimalda, G., Wilson, R. K., Fatas, E., & Foddy, M. (2011). Global social identity and global cooperation. *Psychological Science, 22*(6), 821–828. doi:10.1177/0956797611409590

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London, UK: SAGE Publications.

Cole, K. K. (2015). "It's like she's eager to be verbally abused": Twitter, trolls, and (en)gendering disciplinary rhetoric. *Feminist Media Studies, 15*(2), 356–358. doi:10.1080/14680777.2015.1008750

Crossley, A. D. (2015). Facebook feminism: Social media, blogs, and new technologies of contemporary US feminism. *Mobilization: An International Quarterly, 20*(2), 253–268. doi:10.17813/1086-671X-20-2-253

Daniels, J. (2015). The trouble with White feminism: Whiteness, digital feminism and the intersectional internet. In A. Brock, S. Noble, & B. Tynes (Eds.), *Intersectional Internet: Race, sex and culture online* (pp. 41–61)*.* New York, NY: Peter Lang.

Deibert, R. (2013). *Black code: Surveillance, privacy and the dark side of the internet*. Toronto, Canada: McClelland & Stewart.

DiResta, R., & Grossman, S. (2019). *Potemkin pages & personas: Assessing GRU online operations, 2014–2019*. Stanford Internet Observatory. Retrieved from https://fsi.stanford.edu/publication/potemkin-think-tanks

DiResta, R., Shaffer, D. K., Ruppel, B., Sullivan, D., Matney, R., Fox, R., . . . Johnson, B. (2018). *The tactics & tropes of the Internet Research Agency.* New Knowledge. Retrieved from https://digitalcommons.unl.edu/senatedocs/2/

Eckert, S. (2018). Fighting for recognition: Online abuse of women bloggers in Germany, Switzerland, the United Kingdom, and the United States. *New Media & Society, 20*(4), 1282–1302. doi:10.1177/1461444816688457

Elswah, M., Howard, P., & Narayanan, V. (2019). *Digital confrontations: The Iranian intervention in the Arab world*. Computational Propaganda Project. Retrieved from https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/04/Iran-Memo.pdf

EU vs. Disinfo. (2019). Faces of Western women. *EU vs DISINFORMATION*. Retrieved from https://euvsdisinfo.eu/faces-of-western-women/

Fichman, P., & McClelland, M. W. (2021). The impact of gender and political affiliation on trolling. *First Monday, 26*(1). doi:10.5210/fm.v26i1.11061

Fileborn, B., & Loney-Howes, R. (2019). *#MeToo and the politics of social change*. Cham, Switzerland: Palgrave, Macmillan.

François, C., Nimmo, B., & Eib, C. S. (2019). *The IRA copypasta campaign*. Graphika. Retrieved from https://graphika.com/reports/copypasta/

Freedom House. (2019). *Freedom on the Net | Iran*. Freedom House. Retrieved from https://freedomhouse.org/country/iran/freedom-net/2019

Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y., & Adams, K. (2020). Black trolls matter: Racial and ideological asymmetries in social media disinformation. *Social Science Computer Review*. doi:10.1177/0894439320914853

Freelon, D., & Lokot, T. (2020). Russian disinformation campaigns on Twitter target political communities across the spectrum. *Harvard Misinformation Review*. doi:10.37016/mr-2020-003

Friedberg, B., & Donovan, J. (2019). On the Internet, nobody knows you're a bot: Pseudoanonymous influence operations and networked social movements. *Journal of Design and Science,* (6). doi:10.21428/7808da6b.45957184

Gorwa, R., & Guilbeault, D. (2020). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet, 12*(2), 225–248. doi:10.1002/poi3.184

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science, 363*(6425), 374–378. doi:10.1126/science.aau2706

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances, 5*(1). doi:10.1126/sciadv.aau4586

Hardaker, C., & McGlashan, M. (2016). "Real men don't hate women": Twitter rape threats and group identity. *Journal of Pragmatics, 91*, 80–93. doi:10.1016/j.pragma.2015.11.005

Hopkins, N., & Blackwood, L. (2011). Everyday citizenship: Identity and recognition. *Journal of Community & Applied Social Psychology, 21*(3), 215–227. doi:10.1002/casp.1088

Howard, P., Ganesh, B., Liotsiou, D., Kelly, J., & Francois, C. (2018). *The IRA and political polarization in the United States*. The Computational Propaganda Project. Retrieved from https://comprop.oii.ox.ac.uk/research/ira-political-polarization/

Hull, D. A., & Grefenstette, G. (1996). Querying across languages: A dictionary-based approach to multilingual information retrieval. In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR 96)* (pp. 49–57). New York, NY: Association for Computing Machinery. doi:10.1145/243199.243212

Human Rights Watch. (2019). *World report 2019: Rights trends in Iran*. Human Rights Watch. Retrieved from https://www.hrw.org/world-report/2019/country-chapters/iran

Hwang, T., & Rosen, L. (2017). *Harder, better, faster, stronger: International law and the future of online psyops*. Computational Propaganda Project. Retrieved from http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/02/Comprop-Working-Paper-Hwang-and-Rosen.pdf

Jankowicz, N. (2017). How disinformation became a new threat to women. *World Policy*. Retrieved from https://worldpolicy.org/2017/12/20/how-disinformation-became-a-new-threat-to-women/

Jankowicz, N., Hunchak, J., Pavliuc, A., Davies, C., Pierson, S., & Kaufmann, Z. (2021). *Malign creativity: How gender, sex, and lies are weaponized against women online*. Wilson Center. Retrieved from https://www.wilsoncenter.org/publication/malign-creativity-how-gender-sex-and-lies-are-weaponized-against-women-online

Jenkins, J. C. (1983). Resource mobilization theory and the study of social movements. *Annual Review of Sociology, 9*(1), 527–553. doi:10.1146/annurev.so.09.080183.002523

Judson, E., Atay, A., Krasodomski-Jones, A., Lasko-Skinner, R., & Smith, J. (2020). *Engendering hate: The contours of state-aligned gendered disinformation online*. Demos. Retrieved from https://www.ndi.org/sites/default/files/Engendering-Hate.pdf

Khalid, M. (2011). Gender, orientalism and representations of the 'other' in the war on terror. *Global Change, Peace & Security, 23*(1), 15–29. doi:10.1080/14781158.2011.540092

Khazraee, E., & Novak, A. N. (2018). Digitally mediated protest: Social media affordances for collective identity construction. *Social Media + Society, 4*(1), 313–325. doi:10.1177/2056305118765740

Kim, A. E., Hansen, H. M., Murphy, J., Richards, A. K., Duke, J., & Allen, J. A. (2013). Methodological considerations in analyzing Twitter data. *JNCI Monographs, 2013*(47), 140–146. doi:10.1093/jncimonographs/lgt026

Klandermans, P. G. (1997). *The social psychology of protest*. New York, NY: Wiley.

Leprince-Ringuet, D. (2018, October 25). *Iran has its own fake news farms, but they're complete amateurs*. Wired UK. Retrieved from https://www.wired.co.uk/article/iran-fake-news

Lin, H., & Kerr, J. (2019). On cyber-enabled information warfare and information operations. In P. Cornish (Ed.), *Oxford handbook of cybersecurity*. Oxford, UK: Oxford University Press.

Liska, S. (2015). Talking back to White feminism: An intersectional review. *Liberated Arts: A Journal for Undergraduate Research, 1*(1), 1–7. Retrieved from https://ir.lib.uwo.ca/lajur/vol1/iss1/8/

Locke, A., Lawthom, R., & Lyons, A. (2018). *Social media platforms as complex and contradictory spaces for feminisms: Visibility, opportunity, power, resistance and activism*. London, UK: SAGE Publications.

Loney-Howes, R. (2018). Shifting the rape script: "Coming out" online as a rape victim. *Frontiers: A Journal of Women Studies, 39*(2), 26–57. Retrieved from muse.jhu.edu/article/698452

Lopez, K. J., Muldoon, M. L., & McKeown, J. K. L. (2019). One day of #feminism: Twitter as a complex digital arena for wielding, shielding, and trolling talk on feminism. *Leisure Sciences, 41*(3), 203–220. doi:10.1080/01490400.2018.1448022

Lukito, J. (2020). Coordinating a multi-platform disinformation campaign: Internet Research Agency activity on three US social media platforms, 2015 to 2017. *Political Communication, 37*(2), 238–255. doi:10.1080/10584609.2019.1661889

MacKinnon, R. (2012). *Consent of the networked*. New York, NY: Basic Books.

Mantilla, K. (2015). *Gendertrolling: How misogyny went viral*. Santa Barbara, CA: Prager.

Martin, D. A., & Shapiro, J. N. (2019). *Trends in online foreign influence efforts*. Princeton University. Retrieved from https://scholar.princeton.edu/sites/default/files/jns/files/trends_in_foreign_influence_efforts_2019jul08_0.pdf

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data & Society. Retrieved from https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

McGregor, S. C., Barrett, B., & Kreiss, D. (2021). Questionably legal: Digital politics and foreign propaganda. *Journal of Information Technology & Politics*. doi:10.1080/19331681.2021.1902894

McKelvey, F., & Dubois, E. (2017). *Computational propaganda in Canada: The use of political bots*. Computational Propaganda Project. Retrieved from http://geography.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Canada-FINAL.pdf

Mendes, K., Ringrose, J., & Keller, J. (2018). #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies, 25*(2), 236–246. doi:10.1177/1350506818765318

Monaco, N., Woolley, S. C., & Nyst, C. (2018). *State-sponsored trolling: How governments are deploying disinformation as part of broader digital harassment campaigns*. Institute for the Future. Retrieved from https://www.iftf.org/fileadmin/user_upload/images/DigIntel/IFTF_State_sponsored_trolling_report.pdf

Moore, M. (2018). *Democracy hacked: Political turmoil and information warfare in the digital age*. London, UK: Oneworld Publications.

Mueller, R. S. (2019). Report on the investigation into Russian interference in the 2016 presidential election. *US Department of Justice*. Retrieved from https://www.justice.gov/archives/sco/file/1373816/download

Murthy, D., Powell, A. B., Tinati, R., Anstead, N., Carr, L., Halford, S. J., & Weal, M. (2016). Bots and political influence: A sociotechnical investigation of social network capital. *International Journal of Communication, 10*, 4952–4971. Retrieved from https://ijoc.org/index.php/ijoc/article/view/6271

Nguyen, A., & Catalan, D. (2020). Digital mis/disinformation and public engagement with health and science controversies: Fresh perspectives from Covid-19. *Media and Communication, 8*(2), 323–328. doi:10.17645/mac.v8i2.3352

Nimmo, B. (2016, August 17). *Understanding the role of Russian propaganda in the US election*. Atlantic Council. Retrieved from https://www.atlanticcouncil.org/blogs/new-atlanticist/understanding-the-role-of-russian-propaganda-in-the-us-election/

Nimmo, B., Eib, C. S., Ronzaud, L., Ferreira, R., Lederer, T., & Smith, M. (2020). *Iran's broadcaster: Inauthentic behavior*. Graphika. Retrieved from https://graphika.com/reports/irans-broadcaster-inauthentic-behavior/

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2020). Partisan polarization is the primary psychological motivation behind political fake news sharing on Twitter. *American Political Science Review, 115*(3), 999–1015. doi:10.1017/S0003055421000290

Owen Jones, M. (2016, June 21). *The automation of sectarianism: Are twitter bots spreading sectarianism in the Gulf?* Marc Owen Jones. Retrieved from https://marcowenjones.wordpress.com/2016/06/21/the-automation-of-sectarianism/

Phillips, W. (2015). *This is why we can't have nice things*. Cambridge, MA: MIT Press.

Phipps, A. (2019). "Every woman knows a Weinstein": Political whiteness and white woundedness in #MeToo and public feminisms around sexual violence. *Feminist Formations, 31*(2), 1–25. doi:10.1353/ff.2019.0014

Polletta, F., & Jasper, J. M. (2001). Collective identity and social movements. *Annual Review of Sociology, 27*(1), 283–305. https://www.jstor.org/stable/2678623

Rid, T. (2020). *Active measures: The secret history of disinformation and political warfare*. New York, NY: Farrar, Straus and Giroux.

Sanders, S. (2016, November 8). *Did social media ruin election 2016?* NPR. Retrieved from https://www.npr.org/2016/11/08/500686320/did-social-media-ruin-election-2016

Simon, B., Loewy, M., Stürmer, S., Weber, U., Freytag, P., Habig, C., . . . Spahlinger, P. (1998). Collective identification and social movement participation. *Journal of Personality and Social Psychology, 74*(3), 646–658. doi:10.1037/0022-3514.74.3.646

Singer, P. W., & Brooking, E. T. (2019). *LikeWar: The weaponization of social media*. New York, NY: Houghton Mifflin Harcourt Publishing Company.

Sobieraj, S. (2018). Bitch, slut, skank, cunt: Patterned resistance to women's visibility in digital publics. *Information, Communication & Society, 21*(11), 1700–1714. doi:10.1080/1369118X.2017.1348535

Sobieraj, S. (2020). *Credible threat: Attacks against women online and the future of democracy*. Oxford, UK: Oxford University Press.

Starbird, K., Arif, A., & Wilson, T. (2019). Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1–26. doi:10.1145/3359229

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., . . . Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the scientific literature*. Hewlett Foundation. Retrieved from https://www.hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

Tulkens, S., Hilte, L., Lodewyckx, E., Verhoeven, B., & Daelemans, W. (2016). A dictionary-based approach to racism detection in Dutch social media. *ArXiv Preprint ArXiv:1608.08738.* Retrieved from https://arxiv.org/abs/1608.08738

Turley, E., & Fisher, J. (2018). Tweeting back while shouting back: Social media and feminist activism. *Feminism & Psychology, 28*(1), 128–132. doi:10.1177/0959353517715875

Twitter. (2018). Election Integrity. *Twitter Blog*. Retrieved from https://about.twitter.com/en_us/values/elections-integrity.html

Twitter Transparency. (2018). IRA 2018 dataset. Twitter Information Operations Data Archive*.* Retrieved from https://transparency.twitter.com/en/reports/information-operations.html

Twitter Transparency. (2019a). Russia GRU 2019 dataset. Twitter Information Operations Data Archive*.* Retrieved from https://transparency.twitter.com/en/reports/information-operations.html

Twitter Transparency. (2019b). Iran 2019 dataset. Twitter Information Operations Data Archive*.* Retrieved from https://transparency.twitter.com/en/reports/information-operations.html

Twitter Transparency. (2019c). Venezuela 2019 dataset. Twitter Information Operations Data Archive*.* Retrieved from https://transparency.twitter.com/en/reports/information-operations.html

Van Zomeren, M., Spears, R., Fischer, A. H., & Leach, C. W. (2004). Put your money where your mouth is! Explaining collective action tendencies through group-based anger and group efficacy. *Journal of Personality and Social Psychology, 87*(5), 649–664. doi:10.1037/0022-3514.87.5.649

Waltzman, R. (2017). *The weaponization of information and the need for cognitive security*. RAND
        Corporation. Retrieved from
        https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT473/RAND_CT473.pdf

Woolley, S., & Howard, P. (2018). *Computational propaganda: political parties, politicians, and political
        manipulation on social media*. Oxford, UK: Oxford University Press.

Woolley, S., & Joseff, K. (2019). *The human consequences of computational propaganda: Eight case
        studies from the 2018 US midterm elections*. Institute for the Future. Retrieved from
        https://www.iftf.org/fileadmin/user_upload/downloads/ourwork/IFTF_Executive_Summary_comp.
        prop_W_05.07.19.pdf

Zeiter, K., Pepera, S., & Middlehurst, M. (2020). Tweets that chill: Analyzing online violence against
        women in politics. National Democratic Institute. Retrieved from
        https://www.ndi.org/sites/default/files/NDI%20Tweets%20That%20Chill%20Report.pdf