

Article

## Raging Against the Machine: Network Gatekeeping and Collective Action on Social Media Platforms

Sarah Myers West

Annenberg School of Communication and Journalism, University of Southern California, Los Angeles, CA 90031, USA;  
E-Mail: sarahmye@usc.edu

Submitted: 7 April 2017 | Accepted: 25 August 2017 | Published: 22 September 2017

### Abstract

Social media platforms act as networked gatekeepers—by ranking, channeling, promoting, censoring, and deleting content they hold power to facilitate or hinder information flows. One of the mechanisms they use is content moderation, or the enforcement of which content is allowed or disallowed on the platform. Though content moderation relies on users' labor to identify content to delete, users have little capacity to influence content policies or enforcement. Despite this, some social media users are turning to collective action campaigns, redirecting information flows by subverting the activities of moderators, raising the visibility of otherwise hidden moderation practices, and organizing constituencies in opposition to content policies. Drawing on the example of the campaign to change Facebook's nudity policy, this paper examines the strategies and tactics of users turning to collective action, considering which factors are most influential in determining the success or failure of a campaign. It finds that network gatekeeping salience is a good model for assessing which collective action efforts are most likely to be effective in achieving individual user goals. This indicates that the users who are already most able to harness the attention economy of social media platforms are more likely to successfully navigate the content moderation process. The analysis concludes by attending to what users might learn from the dynamics of network gatekeeping as they seek to resist the asymmetrical power relations of platforms.

### Keywords

collective action; content moderation; network gatekeeping; platforms; social media

### Issue

This article is part of the issue "Acting on Media: Influencing, Shaping and (Re)Configuring the Fabric of Everyday Life", edited by Sigrid Kannengießner and Sebastian Kubitschko (University of Bremen, Germany).

© 2017 by the author; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

### 1. Introduction

Social media platforms act as networked gatekeepers in the contemporary information space. As users provide streams of posts, photos and videos, platforms rank, channel, promote, censor, and delete content, facilitating or hindering information flows (Barzilai-Nahon, 2008; Nahon & Hemsley, 2013). This power to shape content is a form of information control enacted at multiple levels and through differing mechanisms, including platform design, algorithmic curation, and active moderation of posted content.

Researchers have already begun to explore how power flows on and through platforms: at the level

of design, platform affordances, such as Twitter's 140-character length limit, shape which user behaviors are encouraged or discouraged on the platform (Nagy & Neff, 2015; Neff, Jordan, McVeigh-Schultz, & Gillespie, 2012). The workings of these affordances are often made most visible when they change: the elimination of the character limit for direct messages, for example, had an effect on the kinds of discourse—and therefore the broader culture—of the community of Twitter users.

The power to shape content is also enacted through algorithmic curation of which content is made visible or invisible (Gillespie, 2012, McKelvey, 2014). Facebook's newsfeed algorithms subtly shape which content users are exposed to, without explicitly making users aware

of how the algorithm defines the information they are seeing—a form of network gatekeeping that has come under considerable critique for its potential influence on political discussion (Tufekci, 2015) and discriminatory effects (Crawford, 2015; Noble, 2013).

This paper focuses primarily on a third, and perhaps most explicit, form through which social media companies act as networked gatekeepers: active moderation of the content users post on platforms. Most major social media companies have developed complex systems to moderate content at scale, which require immense human resources—from salaried and freelance moderators, as well as users—and largely operate in obscurity (Gillespie, 2017). These systems seek explicitly to accomplish the three main objectives of network gatekeeping: protecting norms within communities from unwanted entry from outside, “locking in” users to the gatekeeper’s network, and maintaining ongoing activities within the network without disturbances (Barzilai-Nahon, 2008). Like the other two forms of network gatekeeping on social media, the influence of companies over information flows through content moderation is largely asymmetric: though social media companies rely heavily on the labor of users to generate content, users have little recourse to petition the companies when their content is taken down, or to make demands of companies when they would like to see changes to the content policies set by the companies.

This paper examines the conflicted relationship between companies and their users when users contest the policies set by companies. I attend particularly to users’ adoption of techniques of collective action that explicitly protest, subvert, and raise the visibility of moderation practices, as well as companies’ response to these tactics. I aimed to understand more clearly two aspects in particular: first, how are users engaging in collective action efforts? Where do these interest networks come from, and what kinds of tactics do they adopt? Secondly, which factors are most influential in determining the success or failure of a collective action effort?

I adopted a qualitative case study approach (Stake, 2005), examining the series of campaigns enacted by Facebook users around its policies on female nudity under the hashtag #FreetheNipple. I find a striking range of strategies and tactics were adopted by Facebook users, resulting in a diverse coalition of different interest groups that converged around the goal of changing Facebook’s content policies. Though they ultimately achieved only partial success in changing the policy, I was able to find that certain factors did influence the likelihood of an individual member of a campaign finding redress when their content is removed. In particular, Barzilai-Nahon’s (2008) typology of networking gatekeeping salience is a good model for assessing which collective action efforts are most likely to be effective in achieving individual user goals. This indicates that the users who are already most

able to harness the attention economy of social media platforms are more likely to successfully navigate the content moderation process. I conclude my analysis by attending to what users might learn from the dynamics of network gatekeeping revealed in this case study as they seek to resist the asymmetrical power relations of platforms.

## 2. Methods

Given that there are relatively few successful examples of collective action efforts driving a change in social media content policies, I decided to adopt a qualitative case study approach (Stake, 2005) that examines one long-standing campaign in depth.

The #FreetheNipple case is illuminating in a number of respects: it is the longest ongoing example of a collective action campaign targeting a social media platform. It also joined together a coalition of users with different interests, and involved techniques that bridged online and offline practices. These factors combined make it a useful instrumental example: though it is not broadly generalizable, it is nevertheless suggestive of a number of underlying dynamics that can further our understanding of how user practices engage and respond to network gatekeeping power. Moreover, this case is important on its own merits, as a persistent example drawn upon in subsequent collective action campaigns by users. In his study of the campaign, Tarleton Gillespie notes that the disagreement “powerfully shaped not only Facebook’s policies, which did change in response, but also how Facebook came to understand the job of moderation, and how users slowly came to understand how Facebook moderates” (Gillespie, in press, p. 154).<sup>1</sup>

I selected a relatively well-known and well-documented case that enabled me to consider an assortment of broadly comparable collective action efforts by different groups. However, there are many others at smaller levels of scale that I excluded, a drawback to this study worth making note of given the salience of visibility in this example. As I find, groups that already have visibility through other platforms are those most likely to succeed in garnering a company response, even if the response of the company only relates to the circumstances of individuals. Those that lack visibility, do not have connections to existing advocacy organizations that can capture the company’s attention, or are unable to attain visibility through other media platforms, are less successful in their efforts. Understanding this dynamic in further detail would require additional research examining cases of collective action at smaller scale. Though this was beyond the scope of this article, it is worth future study.

To examine the case study, I systematically collected documents produced by the users engaged in the campaign, by searching for texts posted on Facebook by the groups and searching news databases for blog posts and

<sup>1</sup> Gillespie’s chronicle of this case, in the book *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*, provides an insightful extended examination of the campaign over breastfeeding photos on social media platforms.

commentary. I also collected and analyzed public statements issued by Facebook. Wherever I encountered a new campaign, I ran additional searches for texts produced by that campaign. I complemented my analysis with contextual information provided in media reports describing events related to the campaign, both to situate my findings as well as to triangulate interpretations of how the campaign was received by the company and broader public. In my analysis of the texts, I used thematic content analysis, identifying patterns within and among the various groups that made up the campaign, attending particularly to the types of tactics adopted by users and when and where new tactics were introduced. In addition, I identified when and where statements were made, repeated and circulated by the companies to understand better how the companies conceptualized their relationships to users. Finally, I looked for moments of policy change or inaction in order to understand at which points during the campaign collective action resulted in substantive policy change.

### 3. Companies Acting on Social Media: Content Moderation

Content moderation is a central part of the way social media companies exert their influence over information flows. As Tarleton Gillespie puts it, what unites US-based social media platforms “is their central offer: to host and organize user content for public circulation, without having produced or commissioned it. They don’t make the content, but they make important choices about that content: what they will distribute and to whom, how they will connect users and broker their interactions, and what they will refuse” (Gillespie, 2017). As a general rule, any intervention a company makes into what content is hosted on their platforms is executed at their own behest: most of the major global social networks are headquartered in the US and thus are not liable under US law for the content posted by users as they are protected by the intermediary liability provisions of Section 230 of the Communications Decency Act. However, most do actively monitor and take down some of the content posted by users. Even though they are not compelled to do so by law, they have economic incentives to take on this curatorial role by moderating material likely to make users feel uncomfortable, such as content that is obscene or violent (Klonick, 2017).

Many social media platforms outline a set of community guidelines that specify the types of content they prohibit on the platform in order to encourage users to police themselves. These community guidelines generally include, at a minimum, provisions against violent, graphic or threatening content, obscenity and nudity, content that violates trademarks or copyright, and fraudulent content or spam. Community guidelines tend to articulate these provisions to users at a relatively high level in order to allow a broad scope for interpretation by the company—particularly given differing cultural norms

and expectations of users around content, as well as differences in legal obligations from country to country. However, content moderators enforce the community guidelines using a much more detailed and concrete set of internal rules, which operationalize and make explicit exactly how much blood, skin, or obscene language constitutes a violation. These operationalized guidelines are not made public, though at times versions of them have been leaked to the public by anonymous moderators (Roberts, 2014).

Community guidelines are not fixed documents; they change and evolve over time as the company’s self-perception and the demands of users evolve. For example, while an early version of Facebook’s Community Standards emphasized the company’s protection of users’ expression, in 2015 the company announced a substantive redraft of the Standards foregrounding users’ safety and security—a move that was likely a response to growing criticism by users that social media platforms failed to protect them from harassment (Gillespie, 2015). Changes to the community guidelines thus can manifest the tensions a company is facing at any point in time over its content policies, and indicate how they navigate competing imperatives to keep as much content as possible online while removing offensive material.

The tension between these two imperatives is an ever-present reflection of social media companies’ network gatekeeping power. Global social media companies increasingly face a challenge of scale: for instance, Facebook’s content moderation system must now police the 300 million photos uploaded every day by Facebook’s 1.86 billion monthly active users (Zephoria, 2017). Though the legal and policy teams—those who set the policies but do not directly enforce them—of many major US social media companies are heavily influenced by First Amendment norms that favor free expression (Amori, 2014; Klonick, 2017), companies increasingly must seek out efficiencies in order to manage the flood of offensive and pornographic content posted by a fraction of its growing user base.

Historically, major platforms such as Facebook, Instagram, Twitter and YouTube have done so by relying heavily on user reporting of posts that violate content policies, rather than actively policing the content themselves—what James Grimmelmann (2015) has termed *reactive* as opposed to *proactive* moderation. Content that is flagged by a user is sent on to a team of content moderators, who are often freelancers working on contract with the company, who check the flagged material against a detailed set of internal guidelines designed to operationalize the broader content policies established by the company. Though information about content moderation processes is generally not made public, researchers such as Roberts (2014) point to a number of challenges with this approach: moderators are required to assess content quickly and without context, face burnout from watching the most graphic and violent of the Internet’s content, and may have differing cultural interpretations

of what content violates the guidelines. These factors combined are likely to introduce a high degree of error in to the system, but because companies do not include figures on content moderation in their transparency reports, this is hard to validate.

Through content moderation policies, companies act on the user-generated media submitted to them: by setting policies, hiring teams of moderators, and, increasingly by introducing new technologies to automatically filter content. Users are demonstrably absent in this system, other than as laborers flagging content they deem objectionable. They also have relatively little recourse to seek accountability from the company within existing channels. Many social media platforms do offer some form of appeal to users when their content is taken down, but users often report they are unaware of this or are unsuccessful in seeing the content restored after appealing (Onlinecensorship.org, 2016). Moreover, users have few venues to influence the policies themselves. In a rare exception, Facebook engaged in a short-term experiment in 2009 during which users could collectively “vote” on potential policy changes, but scrapped the initiative when it failed to achieve sufficient engagement for the vote to move ahead (Stein, 2013).

Thus, many users have turned to collective action as a means to push back on the network gatekeeping power wielded by companies. Taking these efforts by users into account, a more complex picture of platforms emerges: one in which users respond creatively to these power discrepancies, seeking to subvert and resist them in order to reshape information flows in the directions they find more desirable.

#### 4. Users Acting on Social Media: Collective Action

Collective action campaigns take as their starting point the notion that “groups of individuals with common interests usually attempt to further those common interests” (Olson, 1965), an idea that has been explored and challenged thoroughly by social scientists. Adopting a collective action approach to resist content moderation intuitively makes sense given the power dynamics described above. As individuals, social media users have relatively little capacity to push back on the content moderation policies and enforcement of companies. But they do share a common interest in doing so, both as a means to respond to the rate of error in enforcing policies and to the terms set by the policies themselves. Given companies’ reliance on advertising to their user base, it further seems likely that acting collectively would be more likely to result in success in influencing companies than acting alone.

In this section, I examine one such effort in detail: the extended campaign to encourage Facebook to change its gendered policies on images of female nudity. For nearly a decade, collectives of Facebook users have engaged in

various forms of demonstration around the company’s community guidelines on nudity, which allow male toplessness but not female toplessness. These campaigns have taken up a variety of tactics, including petitions, use of hashtags, humorous memes, virtual sit-ins and in-person protests at Facebook’s headquarters. It is characterized by a series of surges: phases in which content produced by the protesters went viral, spreading rapidly and joining together interested users, and phases in which the campaign loses attention. The most visible of these campaigns uses the hashtag #FreeTheNipple, which has become an umbrella term encapsulating a variety of efforts by different actors to encourage Facebook to change its policies. Though the Free the Nipple campaign was launched several years after the initial outcry by users, it became a powerful symbol under which many disparate groups of Facebook users united, and a means through which the collective history of the campaign can be preserved and accessed over time.

The first large-scale protest that I was able to identify began in 2007, and centered on the removal of images of mothers breastfeeding from their profile pages. This first initiative was led by Kelli Roman, who, after noticing a photo she posted to her profile of herself breastfeeding her new daughter had been deleted, wrote Facebook asking why the photo had been taken down. When she did not receive a response from the company, she started the Facebook group “Hey Facebook, Breastfeeding is Not Obscene”, which became a place for other mothers to congregate who had experienced the same issue (Belkin, 2008). The Facebook group became an important node in the emergent network around the issue, linking together interested users and mobilizing them to spread and share information about the removal of content.

As the group grew in size, an offshoot, calling itself the Mothers International Lactation Campaign, decided to hold an online “nurse-in” protest on December 27, 2008, which garnered attention from mainstream media including the “Parenting” section of the *New York Times*. In what could be qualified as the campaign’s first viral information event,<sup>2</sup> over 11,000 Facebook users changed their profile photos in protest to the image of a mother nursing a child. The success of the protest translated in to additional momentum for the campaign, including an online petition that over 82,000 users signed in support of (Sweney, 2008), as well as a real-life demonstration by a smaller group outside of Facebook’s headquarters, at which members of the group sang, chanted and breastfed (Noguchi, 2008).

Despite the substantial public attention the campaign received, Facebook refused to budge. In statements to the *Guardian*, *New York Times*, and other media outlets, Facebook spokesman Barry Schnitt gave a statement saying “Breastfeeding is a natural and beautiful act and we’re very glad to know that it is so important to some mothers to share this experience with oth-

<sup>2</sup> In Nahon & Hemsley’s (2013) description, a *viral information event* “creates a temporally bound, self-organized interest network in which membership is based on an interest in the information content or in belonging to the interest network of others” (p. 34).

ers on Facebook” (Belkin, 2008), but that “photos containing a fully exposed breast, as defined by showing the nipple or areola, do violate those terms (on obscene, pornographic or sexually explicit material) and may be removed. The photos we act upon are almost exclusively brought to our attention by other users who complain” (Sweney, 2008). Schnitt’s statement reflects both Facebook’s values: “breastfeeding is a natural and beautiful act”, and “we are glad to know it is important to some mothers to share it”, as well as its operationalized principles: “fully exposed breasts, defined by showing the nipple, violate the terms of service”. Moreover, he is drawing implicit boundaries around the “some” mothers who want to share their experiences on Facebook and the “other users” who complain about these images. By suggesting there is a tension between two constituencies, Facebook is placing itself in the position of a neutral arbiter, enforcing the operationalized rules that will not change. Thus, while the first round of the campaign was successful in achieving visibility for the issue, it did not result in a substantive change to policy.

The Facebook group steadily grew, as did offshoot groups oriented around short-term campaigns and expressions of solidarity with members whose content had been taken down. A few years later, it began to forge ties with other communities and their campaigns, such as one oriented around Facebook’s removal of images of young breast cancer survivors that prominently featured their mastectomy scars from the page of the breast cancer awareness group, SCAR Project. After the photos were taken down, cancer survivor Scorchy Barrington began a Change.org petition cataloguing the experiences of a number of survivors whose photos were taken down. Citing Facebook’s statement on breastfeeding, she asked “So, why is breast cancer considered a violation? Women fighting breast cancer are also beautiful, and I can’t think of a more important experience to share with others than one that raises awareness of the disease and helps other women who are facing treatment” (Barrington, 2013).

After receiving over 21,000 signatures, Facebook responded with a statement similar to the one on breastfeeding: “We agree that undergoing a mastectomy is a life-changing experience and that sharing photos can help raise awareness about breast cancer and support the men and women facing a diagnosis, undergoing treatment, or living with the scars of cancer. The vast majority of these kinds of photos are compliant with our policies” (Goldhill, 2013). Here, Facebook signaled rhetorically that there was not a problem with the *category* of images—in the “vast majority” of instances these photos do not violate the policies, there were problems with *individual photos*, which were exceptions to this general rule. They again emphasized the underlying value of sharing, while leaving the underlying policy intact.

During roughly the same period, a new group of protesters joined the campaign under the moniker #FreeTheNipple. Lina Esco first started the campaign in

2012 as part of an effort to combat public toplessness laws in New York City that enact different standards for men and women (Esco, 2013). Esco filmed a documentary in which she runs topless through Times Square, and posted a teaser trailer on Facebook. The social media platform suspended Esco’s profile in December 2013 for violating Facebook’s community guidelines, sparking Esco’s outrage—and activating her network.

Several well-known celebrities, including Miley Cyrus, Lena Dunham, Chelsea Handler, Rihanna, and Chrissy Teigen, rallied behind Esco’s cause, posting photos of themselves exposing their bare chests or wearing t-shirts in support of the documentary (Esco, 2014). Others posted humorous memes critical of the policy, such as Cyrus photoshopping her head on to the image of a naked Barbie doll (Tejada, 2014), or Handler posting a photo of herself parodying the famous image of Russian President Vladimir Putin riding horseback without a shirt on (Marcotte, 2014). The involvement of celebrities boosted attention to the cause, again leading to a surge in user protests and leading to coverage from media outlets, several of which published articles explicitly condemning Facebook’s policy.

The campaign also found an intersection with similar efforts by Women, Action and the Media, the Everyday Sexism Project, and author/activist Soraya Chemaly, who were already working on a campaign against gender-based violence on social media. As Chemaly put it later, “of equal importance to gender-based hate was the issue of the *context* in which content passes moderation. As a reflection of the world’s culture, Facebook continues to be a place in which depictions of women as sexually objectified...or debased is broadly allowable, but others, in which women represent their own bodies...is largely not” (Chemaly, 2014). They collected 60,000 tweets and 5,000 emails from users about the issue, during which 15 advertisers said they would leave the platform (Women, Action & the Media, 2013). Again, despite attaining greater visibility for the cause, the campaign was unsuccessful in forcing a change to the policy.

The issue remained in the public eye, though sporadically, in the following weeks and months, receiving media coverage from time to time when a celebrity or public figure had their image taken down for violating the policy. Typical of this phase was an incident caused when Facebook’s subsidiary company, Instagram, briefly disabled the singer Rihanna’s account several times for posting images that featured nudity from a cover shoot with the French magazine *Lui*. Given Rihanna’s popularity, there was an immediate outcry over the suspension and a representative from Instagram quickly responded, restoring the account and saying that its deletion was due to a technical glitch: “This account was mistakenly caught in one of our automated systems and very briefly disabled. We apologize for any inconvenience” (Smith, 2014). Rihanna responded by posting a fan drawing of her topless cover, mocking Instagram for the takedown (Muhammad, 2014).

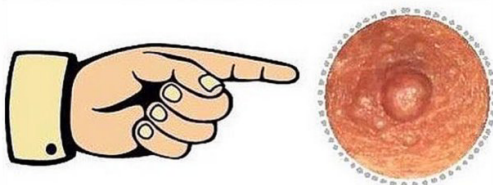


Shortly afterward, Heather Bays, a maternity photographer, had her Instagram account shut down after receiving a negative comment on a photo of her breastfeeding her daughter. Unlike Rihanna, Bays initially lacked the public platform to attract attention from the company to her case. Her account was only reinstated after she used other social media accounts to draw attention to the issue (Corregan, 2014). Scout Willis also had her account deleted over what Instagram called “instances of abuse”, for posting a photo of herself in a sheer top and a photo of a jacket that featured an image of two of her close friends topless. In protest, Willis walked topless in public through New York City and wrote a blog about her protest on the website XOJane (Willis, 2014).

In June 2014, Soraya Chemaly reported that Facebook had quietly made a change to its community guidelines, allowing exceptions for breastfeeding mothers (Chemaly, 2014), as well as for mastectomy scarring. In the new guidelines the company responded to the criticism, “our policies can sometimes be more blunt than we would like and restrict content shared for legitimate purposes. We are always working to get better at evaluating this content and enforcing our standards” (Facebook, 2014).

This only partially addressed the protesters’ grievances, leading to a new wave of creative responses by social media users that sought to push the boundaries of the policy. Electronic Frontier Foundation Director of International Freedom of Expression Jillian York called the policy “the new fig leaf, a new standard on the corporate Internet” (Pizzi, 2015), a theme users picked up on as they turned to humor and subversion of the process of content moderation in additional more traditional protest tactics. For example, after an image in which she appeared topless at a breast cancer fundraising art event was taken down from Instagram, the artist Micol Hebron circulated a template she designed for users to cover images of female nipples with a man’s (see Figure 1). The artists Our Lady J and La Sera shared her post, which

## THIS IS A MALE NIPPLE:



If you are going to post pictures of topless women, please use this acceptable male nipple template to cover over the unacceptable female nipples.

(Simply Cut, Resize and Paste)

THANK YOU FOR HELPING TO MAKE  
THE WORLD A SAFER PLACE.

**Figure 1.** Image created by the artist Micol Hebron to critique Instagram’s nudity policy. Source: MicolHebron.com

quickly went viral after it was shared by the comedian Sarah Silverman and celebrity writer Perez Hilton. Others put the template to creative use and shared it over Facebook and Instagram, subverting the moderation process by taking the parts of the image that would violate the policy out of their context (Ferrier, 2015).

As Hebron described it, “With the digital pasty, I was offering a satirical response to the double standards of Instagram’s sexist and senseless “community guidelines”. I was taking their guidelines VERY literally in an effort to point out how absurd their restrictions were (not to mention the fact that in censoring female nipples but not male nipples, Instagram was also inadvertently presuming to know people’s gender simply by looking at a picture of them—which is also offensive, and absurd). I advocate for all bodies to be treated equally, and for all people to have autonomy over their bodies, their gender, the image of their bodies, and how their bodies are treated in public space” (M. Hebron, personal communication, August 21, 2017).

Celebrities also played a role in subverting the policy: when a photo of the model Chrissy Teigen was taken down, she played humorously on the company’s exception for nudity in art, reposting a previously banned image of herself with a filter on it that made it appear like an oil painting, and again as a pencil sketch (Noman, 2015). Male actor Matt McGorry turned the protest on its head, photoshopping a topless image of himself with cut and pasted images of Miley Cyrus’ and Teigen’s nipples from the photos they had taken down (Plank, 2015). Despite continued attention, Facebook has not made any further changes to its rules on toplessness.

### 4.1. Discussion

The extended campaign over Facebook’s policies provides a number of lessons in how users engage in collective action on social media platforms. First, it demonstrates the remarkable creativity of online protesters. They adopted a wide range of tactics, from virtual sit-ins, petitions, letter-writing campaigns, protests in physical space, media coverage, hashtagging and satirical art to sustain attention to the cause. It illustrates how a diverse coalition of campaigners can form around a single issue: mothers, breast cancer survivors, celebrities, activists for gender equality and artists all joined together over the course of campaign in protest. They did so not because of prior institutional structures, but as a result of a groundswell of protests arising from multiple corners of the internet, which found common cause in several viral information events and channeled around the hashtag #FreeTheNipple. The broad-based coalition may have been one reason why the campaign has been able to sustain itself over such a long period of time, and is something that subsequent collective action campaigns such as the Nameless Coalition have sought to recreate.

The ultimate change to the policy came quietly after years of protest, making it difficult to ascertain which,

if any, of these tactics influenced the campaign's partial success. In general, Facebook tended to emphasize continuity over change in its discussion of its policies; emphasizing how the protestors' objections worked within the broad framework of its nudity policy rather than presenting them as a radically different approach. The creative approaches of protestors worked at cross-purposes with this, using virality to leverage the attention economy of social media platforms and rally newcomers behind the protest. These efforts were most successful when they leveraged humor and novelty, or when they successfully reached someone with broad influence, such as celebrities like Silverman and Cyrus or activists like Esco. Though viral information events would have a short-term impact on the visibility of the campaign, their long-term impact contributed to the persistence of the interest network around the issue.

Despite this, the change to the policy ultimately had narrow benefits for only two of the coalitions involved (breastfeeding mothers and breast cancer survivors), even though all the campaigners based their arguments on the issue of gender discrimination, which remained unaddressed in the new guidelines. Thus, while the campaign was successful in building a network, it was unsuccessful in translating its network capacity in to the kind of pressure needed to force a company policy change.

The case did, however, provide insight in to how Facebook responds to user complaints, suggesting that network gatekeeping salience (Barzilai-Nahon, 2008) is a relatively good model for assessing which kinds of collective action efforts are most likely to move the needle with the company in addressing individual complaints. The network gatekeeping salience typology suggests that if users have any of four key attributes (political power, information production ability, relationship with the gatekeeper, or alternative choices), they are more likely to have greater salience in the network.

In this example, salience appeared to take the form of visibility: users who were highly visible, such as Rihanna, held a greater amount of political power relative to the platform and were able to get their accounts restored (despite a clear violation) without even asking for it. Celebrities who were less visible, like Willis, were punished for violating the policy in an explicit act of protest. Some non-celebrity users were able to get their accounts restored by leveraging alternative attributes of network gatekeeping salience to increase their visibility, as did Bays when she engaged in a concentrated campaign of information production, leveraging her accounts on other social media platforms to raise attention to her cause. This demonstrates the value to users of leveraging the principles of virality in order to seek redress: even where the underlying policy has not changed, its application to users may be inconsistently applied depending on the amount of attention they are able to garner behind their cause.

This also suggests that collective action is particularly important to users who do not, on their own, have much visibility—the very same users who benefit most from the additional channel that social media platforms like Facebook provide them. So even if a campaign is unsuccessful in pressuring a company to change its content policies, there may be ancillary benefits to users who take part in coalitions, work to garner media attention and draw visibility to an issue.

## 5. Conclusion

The network gatekeeping power of companies is conventionally thought of as the ability to facilitate information flows and bridge networks. This is a power that companies like Facebook have substantively amassed through ranking, channeling, promoting content, and, at times, censoring and deleting it (Nahon & Hemsley, 2013). At the other side of Nahon and Hemsley's equation is the formation of user-driven interest networks that form around particular issues, which can drive attention to content from the bottom-up.

Often, we think of these networks using terms of *connectivity*—their capacity to make connections with one another and circulate content through networks. As Jose van Dijck (2013) notes, connectivity is a valuable resource to companies—thus situating the power of individual users in their ability to forge networks that can be monetized by companies risks staking the organization of social exchange on neoliberal economic principles. But for the users examined in this study, the operating principle for acting on social media companies' network gatekeeping power may have been *visibility*, not connectivity. The mere fact of writing letters, spreading images, signing petitions and engaging in protests did not result in a change to company policies. And in fact, some of the women involved in the protests were already part of communities and networks on related issues before taking part in the campaign against Facebook (Gillespie, in press). Instead, influencing the visibility of an issue was a key operating force that shaped how the company responded to the demands of users.

This dynamic is problematic in different ways, most notably in that it inherently benefits those who already have a platform. But it also is suggestive of the kinds of tactics users seeking to influence social media content policies might adopt: self-publishing, building coalitions, and working across media platforms to create viral information events as a means to raise the visibility of a cause. In so doing, they push back directly on the power of platforms to define the content we see.

## Acknowledgments

The author wishes to acknowledge Jessica Anderson, Kim Carlson, and Jillian York for their feedback on the manuscript.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- Ammori, M. (2014). The “new” New York Times: Free speech lawyering in the age of Google and Twitter. *Harvard Law Review*, 127(8), 2259–2295.
- Barrington, S. (2013). Facebook: Stop censoring photos of men and women who have undergone mastectomies. *Change.org*. Retrieved from <https://www.change.org/p/facebook-stop-censoring-photos-of-men-and-women-who-have-undergone-mastectomies>
- Barzilai-Nahon, K. (2008). Toward a theory of network gatekeeping: A framework for exploring information control. *Journal of the American Society for Information Science and Technology*, 59(9), 1493–1512.
- Belkin, L. (2008, December 19). Censoring breastfeeding on Facebook. *The New York Times*. Retrieved from <https://parenting.blogs.nytimes.com/2008/12/19/censoring-breastfeeding-on-facebook>
- Chemaly, S. (2014, June 9). #FreeTheNipple: Facebook changes breastfeeding mothers photo policy. *The Huffington Post*. Retrieved from [http://www.huffingtonpost.com/soraya-chemaly/freethenipple-facebook-changes\\_b\\_5473467.html](http://www.huffingtonpost.com/soraya-chemaly/freethenipple-facebook-changes_b_5473467.html)
- Corregan, S. (2014, May 23). Instagram breastfeeding fuss shows bias. *Times Colonist*. Retrieved from <http://www.timescolonist.com/opinion/columnists/shannon-corregan-instagram-breastfeeding-fuss-shows-bias-1.1071948>
- Crawford, K. (2015). Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology & Human Values*, 41(1), 77–92.
- Esco, L. (2013, December 9). Why I made a film called free the nipple and why I’m being censored in America. *The Huffington Post*. Retrieved from [http://www.huffingtonpost.com/lina-esco/free-the-nipple\\_b\\_4415859.html](http://www.huffingtonpost.com/lina-esco/free-the-nipple_b_4415859.html)
- Esco, L. (2014, January 7). Facebook wages war on the nipple. *The Huffington Post*. Retrieved from [http://www.huffingtonpost.com/lina-esco/facebook-war-on-nipples\\_b\\_4548832.html](http://www.huffingtonpost.com/lina-esco/facebook-war-on-nipples_b_4548832.html)
- Facebook. (2014). Encouraging respectful behavior. *Facebook*. Retrieved from <https://m.facebook.com/communitystandards/encouraging-respectful-behavior>
- Ferrier, M. (2015, July 7). How to beat Instagram and Facebook’s female nipple ban? Use male nipples instead. *The Guardian*. Retrieved from <https://www.theguardian.com/lifeandstyle/shortcuts/2015/jul/07/instagram-facebook-female-nipple-ban-use-male-nipples-instead>
- Gillespie, T. (2012). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–193). Cambridge, MA: MIT Press.
- Gillespie, T. (2015, March 18). Facebook’s improved “community standards” still can’t resolve the central paradox. *Culture Digitally*. Retrieved from <http://culturedigitally.org/2015/03/facebooks-improved-community-standards-still-cant-resolve-the-central-paradox>
- Gillespie, T. (2017). Governance of and by platforms. In J. Burgess, T. Poell, & A. Marwick (Eds.), *SAGE handbook of social media*. Thousand Oaks, CA: SAGE Publications. Book accepted for publication.
- Gillespie, T. (in press). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. New Haven, CT: Yale University Press.
- Goldhill, O. (2013, June 13). Facebook allows mastectomy photos after breast cancer patient’s petition. *The Telegraph*. Retrieved from <http://www.telegraph.co.uk/technology/facebook/10118016/Facebook-allows-mastectomy-photos-after-breast-cancer-patients-petition.html>
- Grimmelmann, J. (2015). The virtues of moderation: Online communities as semicommons. *Yale Journal of Law and Technology*, 17(42), 42–109.
- Klonick, K. (2017). The new governors: The people rules, and processes governing online speech. *Harvard Law Review*, 131. Retrieved from [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2937985](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2937985)
- Marcotte, A. (2014, October 31). Chelsea Handler stands up for freed nipples. *Slate*. Retrieved from [http://www.slate.com/blogs/xx\\_factor/2014/10/31/chelsea\\_handler\\_puts\\_a\\_topless\\_photo\\_on\\_instagram\\_and\\_protests\\_when\\_it\\_s.html](http://www.slate.com/blogs/xx_factor/2014/10/31/chelsea_handler_puts_a_topless_photo_on_instagram_and_protests_when_it_s.html)
- McKelvey, F. (2014). Algorithmic media need democratic methods: Why publics matter. *Canadian Journal of Communication*, 39(4), 597–613.
- Muhammad, L. (2014, May 6). Rihanna’s Instagram account ‘briefly disabled’. *BET*. Retrieved from <http://www.bet.com/news/music/2014/05/06/instagram-briefly-disabled-rihanna-s-account.html>
- Nagy, P., & Neff, G. (2015). Imagined affordance: Reconstructing a keyword for communication theory. *Social Media + Society*, 1(2), 1–9.
- Nahon, K., & Hemsley, J. (2013). *Going viral*. Cambridge: Polity Press.
- Neff, G., Jordan, T., McVeigh-Schultz, J., & Gillespie, T. (2012). Affordances, technical agency, and the politics of technologies of cultural production. *Journal of Broadcasting & Electronic Media*, 56(2), 299–313.
- Noble, S. (2013). Google search: Hyper-visibility as a means of rendering black women and girls invisible. *InVisible Culture*, 19. Retrieved from <http://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible>
- Noguchi, S. (2008, December 27). Protests mount over Facebook ban on breast-feeding photos; bigger turnout online than in Palo Alto. *The Mercury News*. Retrieved from <http://www.mercurynews.com/2008/12/27/protests-mount-over-facebook-ban-on>



- breast-feeding-photos-bigger-turnout-online-than-in-palo-alto
- Noman, N. (2015, June 30). Chrissy Teigen freed her nipple on Instagram—And the photos got taken down. *Mic*. Retrieved from <https://mic.com/articles/121551/chrissy-teigen-exposes-nipple-on-instagram#.jtf4k55Yv>
- Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Onlinecensorship.org. (2016). Censorship in context: Insights from crowdsourced data on social media censorship. *Onlinecensorship.org*. Retrieved from <https://onlinecensorship.org/news-and-analysis/onlinecensorship-org-launches-second-report-censorship-in-context-pdf>
- Pizzi, M. (2015, March 16). Facebook clarified, confuses with new content rules. *Al Jazeera*. Retrieved from <http://america.aljazeera.com/articles/2015/3/16/facebook-clarifies-confuses-with-new-content-rules.html>
- Plank, E. (2015, July 8). Actor Matt McGorry just freed the nipple in the best way possible. *Mic*. Retrieved from <https://mic.com/articles/121976/orange-is-the-new-black-matt-mcgorry-free-the-nipple-instagram#.DfvTagM8d>
- Roberts, S. (2014). *Behind the screen: The hidden digital labor of commercial content moderation* (Doctoral dissertation). University of Illinois at Urbana-Champaign, Champaign, IL.
- Smith, L. (2014, November 2). Bad Gal RiRi is back! Rihanna's Instagram has returned. *Glamour*. Retrieved from <http://www.glamourmagazine.co.uk/article/rihanna-instagram-deleted-after-topless-photos>
- Stake, R. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 443–466). Thousand Oaks, CA: SAGE Publications.
- Stein, L. (2013). Policy and participation on social media: The cases of YouTube, Facebook, and Wikipedia. *Communication, Culture & Critique*, 6(3), 353–371.
- Sweney, M. (2008, December 30). Mums furious as Facebook removes breastfeeding photos. *The Guardian*. Retrieved from <https://www.theguardian.com/media/2008/dec/30/facebook-breastfeeding-ban>
- Tejada, C. (2014, December 29). Miley Cyrus supports 'free the nipple' with topless photo. *The Huffington Post*. Retrieved from [http://www.huffingtonpost.ca/2014/12/29/miley-cyrus-topless\\_n\\_6390150.html](http://www.huffingtonpost.ca/2014/12/29/miley-cyrus-topless_n_6390150.html)
- Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal*, 13(2), 203–218.
- Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford: Oxford University Press.
- Willis, S. (2014, June 2). I am Scout Willis and this is the only thing I have to say about walking topless down the streets of New York last week. *XOJane*. Retrieved from <http://www.xojane.com/issues/scout-willis-topless-instagram-protest>
- Women, Action & the Media. (2013). Facebook agreement. *Women, Action & the Media*. Retrieved from <http://www.womenactionmedia.org/fb agreement>
- Zephoria. (2017). Top 15 valuable Facebook statistics. *Zephoria*. Retrieved from <https://zephoria.com/top-15-valuable-facebook-statistics>

### About the Author



**Sarah Myers West** is a doctoral candidate at the Annenberg School for Communication and Journalism, where her research centers on international policymaking and activism around privacy, security, and freedom of expression. She is an affiliate researcher at the Berkman Klein Center for Internet and Society and the managing editor for Global Voices Advocacy's Netizen Report. Her research is published in journals including *Policy & Internet* and *Business and Society*.