ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/335043873

Values versus Rules in Social Media Communities

Chapter · July 2019

DOI: 10.4324/9780429266140-3

CITATION
1

READS
63

3 authors, including:

Michael Trice
Massachusetts Institute of Technology
6 PUBLICATIONS **15** CITATIONS

SEE PROFILE

Liza Potts
Michigan State University
48 PUBLICATIONS **374** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Experience Architecture View project

All content following this page was uploaded by Michael Trice on 09 January 2020.

The user has requested enhancement of the downloaded file.

Values vs. Rules in Social Media Communities:

How Platforms Generate Amorality on reddit and Facebook

Michael Trice, Liza Potts, and Rebekah Small

In this chapter, we examine issues of ethics in online communities, presenting an historical account of an online community ethos juxtaposed against modern rules-based communities of harm. By communities of harm, we mean communities that place the individual's freedom to act aggressively above any common community value or long-term goal. Of particular interest in these communities of harm are the rules used to maintain their presence on platforms, like Facebook, Twitter, and reddit, by either subverting platform rules or following the rules while disengaging with the values behind those rules. This rule subversion might manifest as a reddit thread banning doxxing[1] and dogpiling[2] while still serving to aggregate targets for a campaign or harassment. It might also manifest itself as a closed Facebook group that tightly regulates how posts are reported so that concerns about content never reach Facebook itself.

For this critique, we examine historical online communities such as the WELL, an early online community, and the writings about these spaces written by Rheingold (1993) among others. The ethos and practices of these foundational communities are then juxtaposed with the GamerGate controversy as it played out on reddit's /kotakuinacton (KiA) subreddit, a news aggregator and forum for the GamerGate community, and the transformation of LeftBook, the network of left leaning Facebook groups that began as places for community discourse and slowly devolved into spaces for shaming and personal attacks. The KiA subreddit is widely

[1] Doxxing is the revealing of personal identification in an online community. This might include the name of pseudonymous accounts or contact information, like addresses, employment, and phone numbers.
[2] Dogpiling is signalling for a mass of users to respond to a thread or specific account, often to create a flood of overwhelming responses.

recognized as the primary reddit home for GamerGate with over 30,000 subscribers and its persistence as one of the last centralized hubs of GamerGate rests in its ability to follow the rules of reddit while promoting the hostile ethos of GamerGate. LeftBook is not as well known, but the activity of some of their "Shaming" groups--where members post images with the intent of the community shaming individuals for rings, weddings, nails, etc--has recently garnered the attention of news media for their dogpiling and criticism of those not involved in the community. LeftBook intentionally circumvents Facebook moderation to protect its own shaming ethos. These two spaces represent dynamic opposites with regards to the ways in which community ethos is built and content is moderated--GamerGate is quite public about following the rules of reddit, while shaming groups hide from Facebook as closed communities with tight internal moderation. Our findings suggest that, even in the absence of the aggressive ideology found in the KiA subreddit, groups that are predicated on rules rather than values fall to the same demise. When values are not communicated through positive messaging but rather through rules which steer the community in the direction of certain behaviors, those communities find ways to skirt positive values while still following the rules.. Ultimately, regardless of the intentions behind the creation of groups like KiA and LeftBook, their rules-based moderation has led them into similar patterns of messaging and behavior.

To explore this issue of rules versus values in full, we first examine the intersections between community governances, rules, and ethics. Here we take a look at both the history of ethical moderation within online communities as well as how the field of technical communication has wrestled with the issue of ethical governance historically: from codes of ethics to social justice to community *phronesis* as palpable outcome of UX design. We then apply this understanding to KotakuInAction and LeftBook to demonstrate the ethical limitations

of rules-based moderation and how the practice of rules-based moderation bridges ideological classifications in its ethical short-sightedness.

Governance Versus Ethos in Online Communities

In comparing ethos and rules, it is worth discussing a bit of the history of online governance. To start, it is worth recalling that ideological radicalization online has been a concern since the turn of the century. In fact, following the U.S. presidential elections in 2000, Jenkins and Thorburn (2004) felt it necessary to rebuff these concerns by explaining that while radicalism on the left and right were rampant online, the current audience of Internet media paled by comparison to cable and network television. For them, broadcast media would never allow such radicalization to expand due to the enormous audience advantage of broadcast media. Of course, much has since changed in the size of online audiences.

Less than 10 years later, Zittrain (2008) proposed the rise of a new kind of civil libertarian ideal online: a *netizenship* informed by the communal practice, success, and ethics of Wikipedia. Roughly concurrently, Coleman and Blumler (2009) questioned ceding public deliberation to Twitter and Facebook by suggesting that only a BBC-like government-funded (yet politically independent) platform could properly capture the values of civic deliberation. By 2013, Tufekci documented the rise of the activist celebrity as a newly christened master node in networked deliberation. Tufekci’s work in particular has proven prescient as it anticipated the importance of individual broadcast nodes as tastemakers on Twitter and Facebook. Quickly the debate moved from why we need governance online to what form of online governance would best serve the public good.

Haas and Eble (2018) recently drew technical communication directly into this debate by calling attention to how the tensions between globalization and local communities create

opportunities for oppression that we must be attuned to. In many ways, the communities of harm we examine are responding to global platforms by resisting the stated values of these platforms in order to indulge more freely in antisocial and amoral behavior. It is worth examining how they do that and why it matters when rules arise from values and not individualistic rebellion.

Code of Ethics and Accountability

At the close of World War II, Bush (1945) asked a pivotal question in the history of science and engineering: what next? In "As We May Think," Bush's proposal of the memex offers a less compelling moment than the ethical pivot the article personifies as a description for how disciplines should adapt to the post-war era. In hindsight, we can note that the decades that followed gave rise to technical achievements from ARPANET to the moon landing. That same period also led to a growing codification of the ethics within engineering. In 1946, the National Society of Professional Engineers released their initial Canon of Ethics for Engineers expanding upon the rules of ethics that had preceded it. The first canon (NSPE, 2018) expressed a particular commitment to public welfare as ethos: "Hold paramount the safety, health and welfare of the public in the performance of their professional duties."

While codes of ethics existed prior to the end of the war, many of those that came after the war had changed. In the late 19th and early 20th century, professional codes of ethics often focused upon issues of individual honor and honesty. For example, the 1912 code of ethics for the American Institute of Electrical Engineers opened with: "While the following principles express, generally, the engineer's relations to client, employer, the public, and the engineering fraternity, it is not presumed that they define all of the engineer's duties and obligations." Yet by in 1950, the American Institute of Electrical Engineers code opened:

> Honesty. justice, and courtesy form a moral philosophy which, associated with mutual interest among men, constitutes the foundation of ethics. The engineer should recognize such a standard, not in passive observance. but as a set of dynamic principles guiding his conduct and way of life. It is his duty to practice his profession according to these Canons of Ethics. (AIEE, 1950)

Additionally, by the second half of the 20th century, the codes of organizations like the National Society of Professional Engineers (NSPE) (1961) and American Society of Mechanical Engineers (ASME) (1976) shifted to prioritize canons that placed a commitment to the "welfare of the public" or "human welfare" as a key principle. Many trends within the engineering organizations likely fed such shifts. The Work Progress Administration led to numerous nationalized public works, World War II generated both a sense of national professional unity and opposition to the horrors of science and engineering run amok in Germany, and communication technologies began to challenge the local versus the national versus the global. Perhaps few codes epitomized all three movements quite so well as the World Medical Association, which was formed explicitly to address the concerns about global ethical issues in medicine after the atrocities of World War II (World Medical Association, 2018). This ethics-centered shift in professional community behavior from personal integrity to the ethics of public welfare is an important historical note to revisit when considering how governance has evolved online, especially given the technical communications' historic relationship with these codes.

Katz (1992) addressed part of what inspired this shift in codified engineering ethos in "The Ethic of Expediency: Classical Rhetoric, Technology, and the Holocaust." In his analysis of Nazi communication regarding the improvement of trucks for more efficient slaughter of human life, Katz identified an ethos that prioritized scientific and engineering efficiency over human

life. This efficiency ran deeper than the communication itself. Historians have noted that Germany adapted Soviet technology in its gas vans as it eased the minds of the executioners. Where soldiers often suffered ill effects from the shooting of unarmed women and children, the vans removed this emotional and personal connection. Importantly, the process as well as the communication of the process was designed to empower unethical, inhumane, monstrous behavior by distancing the actor from the action and the victim. The entire Nazi platform was predicated on creating an environment where the unconscionable became psychologically palatable. The resistance to such an atmosphere became the centering of humanistic values as central to governance across national, professional, and organizational communities. Canons centered on a commitment to the public welfare became the norm.

The move from individual accountability as ethos to an ethos of service to humanity offers a compelling comparison for those of us looking at social media platforms. As we discuss later with the WELL, personal accountability can be derived from a community ethos. In fact, Zittrain's (2008) evaluation of Wikipedia's ethos and the *netizenship* that rose from it suggests that it is this shared collaborative vision that promotes the governance necessary to maintain the community. These two examples of the WELL and Wikipedia contrast sharply with our later modern examples of community online, KotakuInAction and LeftBook's shaming groups. Importantly, these communities maintain strict community rulesets, but they do so around an ethos less shaped on positive, ethical contributions than amoral recreation and harassment. Essentially, they propose rules of accountability to the community, but not to a greater ethical consideration, ether the platform, society, or humanity at large. In this way, it is not simply a rejection of social justice and ethics, but a redefining of what it means to be social and a rejection of justice outside of the community's standards.

In many ways, the response of communities like KotakuInAction and Ring Shamers can be viewed as part of the populist response to growing globalization that Haas and Eble (2018) ask us to consider as technical communicators. Populist anger and a desire to lash out with that anger is a known nemesis of the social justice Haas and Eble promote. It is not unlike the one Bush (1945) responded to as part of his post-war call to something greater—the Internet has channeled anger into destructive communities that lack accountability beyond the community as well as lacking a community ethic that seeks positive change. The flexibility, anonymity, and generative power of online networks make them excellent realms for fighting oppression, but those features make these networks efficient at enacting oppression at the individual, organizational, national, and global level (Tufekci, 2017).

The connection between ethics and governance matters as it relates to both organizations and platforms. Part of the reason for exploring amoral groups on platforms arises from our wish to focus upon how platforms might offer a means to create ethos rather than simply codes of conduct. For us that means a similar transition as occurred after War War II, a move from personal accountability as rules to an affirmation of a community's positive contribution to the world around them and the sense of personal accountability that might arise from such a mission. In this way, we also hope to express how a commitment to positive applications of social justice might alleviate the oppressive populism that arisen in the face of tensions between globalization and the individual.

Early Online Communities

Social networks, fora, content management systems, wikis, mobile apps, and dedicated messaging services all offer a range of rhetorical purposes, possible audiences, and functional capabilities that make a universal online community difficult to reliably define. Before we

referred to these digital spaces as social media platforms, researchers and practitioners were examining what was primarily text-based technologies such as email, chat, and online bulletin board systems. Researchers noted that "cyberspace is essentially a reconceived public sphere for social, political, economic, and cultural interaction" (Fernback, 1997 p. 37), and as such, studied it as a phenomenon. Many of the early, celebrated work about online communities focused on the positive potential of the Internet, often framed around ideas connecting counterculture with technoculture (Johnson, 1997; Raymond, 1999; Rheingold, 1993). This work served to inspire a growing community of leaders and participants in these digital spaces. While academic scholarship questioned ethical issues of identity and representation in these communities (Baym, 2000; Ess, 1996; Herring, 1993), it was often drowned out in industry by cyberlibertarian voices in early incarnations of publications such as *Wired* magazine and others (Warnick, 2002).

The ethos of these early digital spaces were buoyed by the influential, often utopianistic writings of the major influencers of this time. These spaces were heralded as places where "the common man and the information worker—cowboy or infocraft—can search, manipulate, create or control information directly; he can be entertained or trained, seek solitude or company, win or lose power" (Benedikt, 1991, p. 123). Virtual communities thought leaders such as Howard Rheingold, open source leaders such as Richard Stallman, and popular culture shaped a lot of these ideas. In these spaces, there was a strong spirit of positivity with regards to creating community, volunteerism, and shared governance. While trolling and worse certainly existed (Turkle, 1995), there was a sense that the "virtual" Internet would be better than "real life." There were expectations for civil discussion. The kind of anonymity that existed was still tied to a name, or user "handle" of some sort, that would often voluntarily follow you from online board to online board.

While many of these spaces saw the early Internet as a space of utopian vision, others noted that “cyberspace is an arena of power” (Fernback, 1997, p. 37). These individualistic beliefs held in many early spaces can be summed up in the hacker ethic by Levy (1984):

1. Access to computers—and anything which might teach you something about the way the world works—should be unlimited and total.
2. All information should be free.
3. Mistrust authority—promote decentralization.
4. Hackers should be judged by their hacking, not bogus criteria such as degrees, age, race, or position.
5. You can create art and beauty on a computer.
6. Computers can change your life for the better. (pp. 27-33)

Not too surprisingly, even though this utopian vision carried a promise of something better, the examples explored later in this chapter took that spirit of individualism and connected it to power without any ethical safeguard. In fact, they frequently placed the right of the individual to misbehave as the core value of the community. Rather than individualism feeding art and beauty, it would feed hate, envy, and self-indulgence.

WELL

The WELL (Whole Earth ‘Lectronic Link) is an online community that began in 1985 and is still running today. On the website’s “About” section, the writers describe the site as “a cherished destination for conversation and discussion… widely known as the primordial ooze where the online community movement was born—where Howard Rheingold first coined the term “‘virtual community’” (2019). Built in Sausalito, a city near San Francisco and north of Silicon Valley, it was built as a place for writers, artists, thinkers, and tinkers populated by

people with community-building skills (Evans, p. 132). Its co-founder, Stewart Brand, an early counterculture entrepreneur, had built the Whole Earth Review from which the WELL's name sprung.

Today, the WELL describes itself as a space for "artisanal conversations" where members create a space "distinguished by the quality of our non-anonymous participants, and by uncommon policies." These policies include the axiom "real people, real names" (2019). It is this "real names" policy that immediately distinguishes it from platforms like reddit where anonymity is a core value. And the WELL's policies on creating a space where "mutual respect and cooperation" are paramount makes for a very different kind of community than the combative nature of many of subreddits.

To examine this ethos further, consider this statement on the WELL's website (2019):

> The WELL was launched in 1985 with minimal rules and a freewheeling spirit. The aphorism crafted to greet users at the time was "You Own Your Own Words." This now classic and carefully tended policy has been expanded into The WELL Member Agreement with detailed explanations of the implications of "YOYOW," as members refer to it.

This focus on YOYOW is symbolic of the ethos of responsibility devoid of government oversight. This catch-phrase seems jovial, but it demands that participants be responsible for their actions and work together to create a better digital society.

Participants on the WELL refer to themselves as "WELL-beings" who are engaged with their community for positive growth. Describing his entry into this online community, Rheingold (1993) states, "Finding the WELL was like discovering a cozy little world that had been flourishing without me, hidden with the walls of my house; an entire cast of characters welcomed

me to the troupe with great merriment as soon as I found the secret door" (p. xv). This secret door created a space that flourished in its community.[3]

reddit's KotakuInAction

Whereas the WELL offers a view of a community focused upon non-anonymous communities of high quality conversations, reddit's KotakuInAction serves as something of a direct counterexample for how chan culture has shifted away from the WELL's vision. Much like the wider chan culture, KotakuInAction strongly encourages anonymity over real names. It also aims for a populist and raw discourse more interested in cutting insults than quality of conversations. As discussed below, this attitude bleeds over to a rejection of most publishers and sources considered valid outside the community and an ethos closely tied to the harrassing nature of the GamerGate movement than founded KiA.

GamerGate began in mid-2014 from an angry rumor-filled break-up rant posted online. The details of that story have been recounted by other articles (Chess & Shaw, 2015; Mortensen, 2016) and by the initial target of GamerGate, Zoë Quinn (2017), but a quick primer will assist in contextualizing the key issues around Gamerate as it relates to current ethical issues in community governance. First, from Zoë Quinn to Seth Rich, online conspiracies often move from blog rumors to 4chan fodder to social media activism as anarchy, even as they remain the tragedy and drama of actual human beings. Secondly, they are first and foremost rumor mills.

---

[3] Much of the early design and discussion around online communities focused on the counterculture of the times. Noting that the demographic of the WELL was "thirty-year old white guys," Evans (2018) quotes an early WELL employee, Nancy Rhine, as saying "The WELL stood for Whole Earth 'Lectronic Link, but it did not represent the whole Earth" (p. 133). While men were the majority of participants in these online spaces, this did not mean that women were not also engaging. In a similar vein, examining the following modern examples suggests that while a gender divide persists ideologically online that divide has still led to similar failures in ethical governance, leading us to conclude that neither gender is absolved when ethics are developed with an ethos of chaotic neutrality and applied inconsistently across a platform.

They start as personal accusations that are made to appear as vetted claims via focused discussion on chan boards and YouTube. They develop into deep networks of actors and technical platforms to support and defend the rumors, they often build a culture and counterculture around the rumors, and they apply this sense of community to target both ideological opponents and longstanding institutions they see as threats (Chess & Shaw, 2015). As a matter of ethics, the method and purpose of these networks matter. Rather than guided by a principle of bettering the public welfare, they are attack engines. In addition, they are attack engines driven by rumor and drama over knowledge and information. With all of that in mind, a brief recap of GamerGate follows.

The originating post, full of sordid, personal, largely unsubstantiated claims about infidelity and betrayal quickly found its way to 4chan. As the target of the blog post was an independent game developer—and the accusations revolved around unsupported claims of sex with journalists for positive coverage of her games—channers used the blog post as a rallying cry against games journalists and feminists in game design and critique. From 4chan, conversations about the rumors spilled over into a network of YouTube channels dedicated to criticizing feminism and cultural criticism. Eventually the rumor mill moved to Twitter.

On Twitter, 4chan users attempted to trend several hashtags, including #quinnspiracy, #fiveguys, and #FiveGuysBurgersAndLies. The hashtags make clear that Quinn was the target and that the ridicule was focused upon the rumors of supposed infidelities in the original blog post. The hashtags failed to capture much attention until actor Adam Baldwin coined the tag #GamerGate while retweeting a YouTube video about the incident. Organizers quickly took advantage of Baldwin's interest and latched onto his hashtag in a behavior pattern that would become central to GamerGate success: involve conservative celebrities and pundits to amplify

the message. From there a number of articles covered Baldwin's involvement and the backlash to it. Pushback against these articles critical of GamerGate soon spiraled into a massive Twitter fray pitting culture critics against GamerGate and conservative pundits like Milo Yiannopoulos, Christina Hoff Sommers, Cathy Young, and Mike Cernovich. The inclusion of these pundits would do little to abate the original harassment and rumor-mongering even as a veneer of media criticism focused on challenging feminist culture critics was added to the discussion. GamerGate as a machine for rumor and accusation relished in flat structure and largely rejected anything resembling governance or accountability.

When GamerGate did generate rules, those rules tended to be technical, not ethical. There are no canons of GamerGate. For example, as shown in Figure 3.1, GamerGate frames its Twitter instructions as a how-to for new users. Yet buried in these instructions are a number of implicit values: anonymity, connected networks, in-group/out-group definitions, and prioritization of certain accounts. What GamerGate saw as an ethos was largely volume and power.

[Figure 3.1. GamerGate Instructions for Creating a Twitter Account, by KotakuInAction, n.d., copyright reddit.]

As GamerGate transitioned to reddit, this rule structure was modified (see Figure 3.2). The site's rules explicitly banned many of the tactics popular in the community. Actions such as linking to a target's tweet or article could result in the subreddit being banned. In fact, as numerous aggressive and harmful subreddits were removed, KiA survived due to an aggressive enforcement of its rules. Thus, KiA's effort to survive on reddit resulted in mitigating the acceptable behavior of the subreddit's users. What did not happen was a shift in values. The rules do not state an affirmative community purpose or a set of guiding principles. What they outline

are the behaviors that are unacceptable, thus leaving anything else open to the community. Thus, the community remained heavily focused upon its animosity towards targets. While this prevented direct organization and harassment of targets from the subreddit, it also maintained the ethos of hostility and dehumanizing behavior that fed the culture of the harassment.

[Insert Figure 3.2. Outline of KiA rules from the subreddit, by KotakuInAction, n.d., copyright reddit.]

The discourse of the KiA rules remained resolutely that of GamerGate's discourse, starting with the language of Rule 1, a reference to a rape joke from a popular online comic. The commitment to anonymity (Rule 2) also remained present. This anonymity is an interesting feature when compared to YOTOW of the WELL, which regulated a free-wheeling community via accountability for the type of speech a person used. This accountability is largely lacking on KiA. While certain behaviors were banned, KiA aggressively disavows any responsibility beyond these rules for what a user says. Another rule worth mentioning is Rule 5: "We are not your personal army." In the expanded clarifications for these rules, KiA moderators state:

> Don't post a call to action to downvote some submission on reddit you disagree with. In fact, all links to other subreddits' comment sections will be automatically removed by AutoModerator.
>
> Don't make posts like "let's give that idiot a piece of our mind!" if you come across something stupid someone said on the internet. (KotakuInAction, 2019)

These rules highlight a significant move away from the 4chan operation side of GamerGate that focused heavily upon directing mass movement to and away from certain publications and individuals. It is one clear case where the rules completely altered the standard behavior of

GamerGate on reddit. The discourse of the rules became a means to express the aggrieved purpose of the community even as the rules restrict the old behavior patterns.

While KiA became a muted form of GamerGate due to the need to meet the rules of the reddit platform, KiA's discourse and rules help demonstrate that the core values of aggrievement within the community remained largely unchanged even as the activities allowed by the community were limited. Since KiA strictly enforced rules against targeting and mob action, those elements exceptionally common to the Twitter and 4chan versions of GamerGate disappeared from the site (though not necessarily disappearing from incarnations on other platforms). KiA also maintained a focus on multiple targets (social justice, feminists, and journalists), though the method of linking shifted to images rather than live links. Thus the foes remained the same even as the engagement with those foes shifted.

The ability to rein in the GamerGate community so that it could survive on reddit demonstrates some intriguing possibilities about the role of community rules in other online spaces. First, it suggests that community values can be entwined into the functional representations of platform, similar to hashtags on Twitter and Facebook. Word choice and personalization matter in these rule systems. Yet a tension remains in that the values of reddit are not reflected in the KiA community. By following the reddit rules as a technical obstacle, KiA users managed to maintain values of aggrievement and harassment even as they took the actual behaviors offsite.

Facebook's LeftBook as a Space for Shaming Groups

In *The Presentation of Self in Everyday Life*, Goffman (1959) posits a scenario in which a woman observes a man at a party in order to perceive the man's impression of a third party-goer. This woman carefully notes the man's facial expressions, proximity, and body language, and in

doing so becomes the "unobserved observer," who watches but is not watched. As we examine our next group of users attempting to build community, in an age where platforms preach stoicism and neutrality instead of affirmative values, we witness the consequences of an unobserved observer becoming the observed. The network of individuals we examine next have created a system of their own to vet users and content. LeftBook is not official, unified, or transparent, but it includes a vast web of Facebook users who have joined groups of like-minded users to discuss almost everything, from tacky weddings to late stage capitalism. While part of the downfall of this network was the far-right groups opposed to their activity (upset at their inability to troll from outside of the closed groups), the real undoing came down to Facebook's inability to moderate its own users.

Named for the many groups with a far left political affiliation, LeftBook refers to the unknown number of Facebook groups that utilize rule-based moderation and are quick to ban users who abuse said rules. While like-minded users see this as a valuable effort to promote a specific environment within a group, others (whose ideals are not valued or welcome) perceive the insular nature of this as toxic and threatening. Regardless of personal politics, this network subverted Facebook's functionality, utilizing it for their own purposes and evading the Facebook moderators' attention as much as possible. And as a result, the groups that began as facets of a value-based community morphed over time into a strict rules-based system focused on subverting Facebook's platform and approaches to moderation (see Figure 3.3).

[Insert Figure 3.3. Screenshot of Rules for Ring Shaming Group, Rebekah Small, 2018.]

Three factors contribute to governance issues on LeftBook:

- The practice of official Facebook moderators' completely shutting down groups which generate too many blocks/reports.

- The authority placed in the hands of a select few users/moderators, who control the atmosphere of the group.
- Clashes with the public (on Facebook and other media) over the content, rules, or values of the groups.

Facebook's lack of nuanced moderation led to the authority being placed in the hands of a few individuals who seem to be as ill-suited to moderate as the 7,500 moderators employed by Facebook (Koebler and Cox, 2018). Because of this failure to effectively govern their platforms, Facebook moderation has become a feature of the website that is avoided at all costs in these groups. The rhetoric behind the antifacist statement "we don't call the police" as response to police brutality is echoed in LeftBook groups' adoption of the practice "we don't report users," the insinuation being that to flag a group or user would cause more problems than it would solve. One change led to another, and the vast majority of LeftBook addressed their lack of moderation by instituting their own rules. Facebook's failure to moderate their users directly led to the rules-based moderation we see throughout LeftBook today.

Because users do not trust Facebook to properly moderate their groups and pages, LeftBook groups are run by individual moderators or administrators, called "modmins" by some. These adminstrators are charged with making sure that the rules of their group are followed, admitting new members on a case-by-case basis, and dealing with interpersonal conflicts within the group swiftly. Instead of alerting the officials, users are told to tag a modmin who will quickly review the situation and do what needs to be done so that the group can continue to function. This rule is paramount especially considering that a vast majority of these groups are intended for humorous but anonymous criticism of any expression of personal taste, meaning that personal attacks are common, and they span across themes including cultural appropriation,

cosmetics, or even humorous examples of technology-illiterate generations using the platform "incorrectly." Because of the fine line holding the ethos of these groups together they simply cannot tolerate those who would ignore their rules. As such these groups are known for being "ban-happy", sometimes so much so that splinter groups form, who self-proclaim to not ban as many users (see Figure 3.4).

[Insert Figure 3.4. Screenshot of "That's it i'm wedding shaming" Facebook search, by Rebekah Small, 2018.]

It is hard to consider the amount of time these modmins invest in this endeavor without drawing comparisons to the early AOL days and the labor-intensive volunteer program used to moderate chats and update community bulletins (Postigo, 2009). But while AOL recognized (albeit with miniscule reimbursement) the labor it took to run their community, Facebook does not. In fact, that lack of consistency from Facebook when it comes to moderation is expressly what led to the authority now given to modmins. Facebook users, like many others on social media, have found disappointment when looking to large social media platforms for help handling abuse and harassment. Twitter and Facebook have both famously failed to suspend or ban bad actors. Due to their ad hoc "Community Standards," which rely on hard lines and "objective" workflows, Facebook claims they have not had the time to think ethically about their policy as they are already overwhelmed with time sensitive cases (Koebler & Cox, 2018). The problems that users, especially those in groups, have found with Facebook led the modmins of these groups to, like KiA, take measures to avoid appearing on the radar of official moderators and the public.

The effect of viral posts reported through media outlets like *Slate* and *The Today Show* had a profound effect on these groups as they found themselves in the public eye (Withers,

2018). While the content and tone of these viral posts were very different, the result was the same: splintering groups, anger expressed at those in the groups who shared the content outside of it, and even more dramatic efforts to hide from the platform and the public.

One attempt to hide from Facebook moderators, and the general public, was employed by the group titled "That's It, I'm Ring Shaming," which received viral attention several times regarding content that was cross posted to other social and news media sites. In order to hide from the media attention, as well as the deluge of new member applications that the attention attracted, they changed the name of the group to "gnimahS gniR m'I ,tI s'tahT" which is the name of the group written backwards so that it would not show up in Facebook's or Google's search engine.

In the end, as these groups garner attention and then lose it again, one thing is clear: the rules holding them together are not enough to shape their actions and attitudes. And while this case may seem like the confluence of several missteps the consistent differences between groups who follow rules (or do not with disastrous consequences as seen here and above in the case of /KotakuinAction) and those who base their community around affirmative, positive values suggests otherwise. It is also worth mentioning that within the network of LeftBook there are several groups that do focus on affirmative values. The flip side to all of the shaming groups are "That's it, I'm ___ Praising" groups, which have not garnered media attention due to the harmless, positive nature of their behavior. These groups employ the same set of rules, but do not have the same problems as the shaming groups. The center *can* hold if online communities mold themselves around productive, positive, fruitful goals and values.

Conclusions

The goal of this chapter is to update Rheingold's (1993) call "more than ever before, we need to ask the right questions today about what kind of people, what kind of societies might emerge from social cyberspaces tomorrow" (p. 323). Now that tomorrow has become today, the types of communities have diversified with disturbing trends. Rules that focused on personal accountability have gradually lost sight of ethical value. YOYOW now means skirting Facebook moderators or behaving just well enough on reddit to avoid subreddit shutdown. Accountability only goes so far as survival. Minus some greater call to public welfare, social justice, or even positive contribution, many digital communities have descended into an ethos of aggrievement and harm. As platforms enact universal rules of conduct to promote sitewide values and behaviors, KiA and LeftBook demonstrate how these values are often resisted. The result being groups only avoiding offenses that might invite outside human intervention within the community. Platforms must ask themselves whether these rules have ethical meaning, and if they do not, what purpose do the rules serve?

References

American Institute of Electrical Engineers. (1912). Code of Principles of Professional Conduct (1912). Retrieved from http://ethics.iit.edu/ecodes/node/5068.

American Institute of Electrical Engineers. (1950). Code of Principles of Professional Conduct (1950). Retrieved from http://ethics.iit.edu/ecodes/node/3232.

American Society of Mechanical Engineers. (1976). Codes of Ethics of Engineers (1976). Retrieved from http://ethics.iit.edu/ecodes/node/6297.

Baym, N. (2000). *Tune in, log on: Soaps, fandom, and online community*. Thousand Oaks, CA: Sage Publications.

Benedikt, M. (1991). Cyberspace: Some proposals. In M. Benedikt (Ed.), *Cyberspace: First Steps* (pp. 119-224). Cambridge, MA: MIT Press.

Bush, V. (1945). As we may think. *The Atlantic Monthly, 176*(1), 101-108.

Chess, S., & Shaw, A. (2015). A conspiracy of fishes, or, how we learned to stop worrying about #GamerGate and embrace hegemonic masculinity. *Journal of Broadcasting & Electronic Media, 59*(1), 208-220.

Coleman, S., & Blumler, J. G. (2009). *The Internet and democratic citizenship: Theory, practice and policy*. Cambridge, UK: Cambridge University Press.

Ess, C. (1996). *Philosophical Perspectives on Computer-Mediated Communication*. Albany, New York: SUNY Press.

Fernback, J. (1997). “The individual within the collective: Virtual ideology and the realization of collective principles.” In S. G. Jones (Ed.), *Virtual culture: Identity and communication in cybersociety* (pp. 36-54). Thousand Oaks, CA: Sage.

Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Random House.

Haas, A. M., & Eble, M. F. (Eds.). (2018). *Key theoretical frameworks: Teaching technical communication in the twenty-first century*. Logan, UT: Utah State University Press.

Herring, S. C. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication*. 3(2).

Jenkins, H., & Thorburn, D.. (2004). *Democracy and new media.* Cambridge, MA: MIT Press.

Johnson, S. (1997). *Interface culture: How new technology transforms the way we create and communicate*. New York, NY: Basic Books.

Katz, S. B. (1992). The ethic of expediency: Classical rhetoric, technology, and the Holocaust. *College English, 54*(3), 255-275.

Koebler, J, & Cox, J. (2008, August 23). "Here's how Facebook is trying to moderate its two billion users." *Motherboard*. Retrieved from motherboard.vice.com/en_us/article/xwk9zd/how-facebook-content-moderation-works

KotakuInAction. (2019). Rules. *CODE OF CONDUCT FOR KOTAKUINACTION*. Retrieved from https://www.reddit.com/r/KotakuInAction/wiki/rules#wiki_5._we_are_not_your_personal_army

Levy, S. (1984). *Hackers: Heroes of the computer revolution*. Garden City, NY: Anchor Press/Doubleday.

Massanari, A. (2017). # GamerGate and The Fappening: How reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society, 19*(3), 329-346.

Moorville, P. (2000). The ethics of information architecture. *Argus Center for Information Architecture* Retrieved from https://argus-acia.com/strange_connections/strange008.html

Mortensen, T. E. (2016). Anger, fear, and games: The long event of# GamerGate. *Games and Culture, 13*(8), 787-806.

National Society of Professional Engineers. (2018). NSPE code of ethics for engineers. *History of the Code of Ethics for Engineers.* Retrieved from https://www.nspe.org/resources/ethics/code-ethics/history-code-ethics-engineers

National Society of Professional Engineers. (1961). Code of Ethics for Engineers (1961). Retrieved from http://ethics.iit.edu/ecodes/node/3262

Postigo, H. (2009). America Online volunteers: Lessons from an early co-production community. *International Journal of Cultural Studies,* 12(5), 451-469.

Quinn, Z. (2017). *Crash override: How GamerGate (nearly) destroyed my life, and how we can win the fight against online hate*. New York, NY: Hachette.

Raymond, E. (1999). "The cathedral and the bazaar." *Knowledge, Technology, and Policy*. 12(3), 23–49.

Rheingold, H. (1993, 2000). *The virtual community: Homesteading on the electronic frontier*. (Revised Edition). First published by Addison Wesley in 1993. Then published by MIT Press in 2000.

The WELL (n.d.). *The WELL*. Retrieved from https://www.well.com/about-2/.

Tufekci, Z. (2013). "Not this one" social movements, the attention economy, and microcelebrity networked activism. *American Behavioral Scientist, 57*(7), 848-870.

Tufekci, Z. (2017). *Twitter and tear gas: The power and fragility of networked protest*. New Haven, CT: Yale University Press.

Turkle, S. (1995). *Life on the screen: Identity in the age of the internet*. New York, NY: Simon & Schuster.

Warnick, B. (2002). *Critical literacy in a digital age: Technology, rhetoric, and the public interest*. Mahwah, NJ: Lawrence Erlbaum Associates.

World Medical Association. (2018). *History*. Retrieved from https://www.wma.net/who-we-are/history/

Zittrain, J. (2008). *The future of the Internet—and how to stop it.* New Haven, CT: Yale University Press.

View publication stats