The psychology of volunteer moderators: Tradeoffs between participation, belonging, and norms in online community governance

Beril Bulat (iD)
University of California, Davis, USA

Hannah Wang
University of California, Davis, USA

Stephen Fujimoto
University of California, Davis, USA

Seth Frey
University of California, Davis, USA

Abstract

Online communities rely on effective governance for success, and volunteer moderators are crucial for ensuring such governance. Despite their significance, much remains to be explored in understanding the relationship between community governance processes and moderators' psychological experiences. To bridge this gap, we conducted an online survey with over 600 moderators from Reddit communities, exploring the link between different governance strategies and moderators' needs and motivations. Our investigation reveals a contrast to conventional views on democratic governance within online communities. While participatory processes are associated with higher levels of perceived fairness, they are also linked with reduced feelings of community

Corresponding author:
Beril Bulat, Department of Communication, University of California, Davis, 376 Kerr Hall, Davis, CA 95616, USA.
Email: bbulat@ucdavis.edu

belonging and lower levels of institutional acceptance among moderators. Our findings challenge the assumption that greater democratic involvement unequivocally leads to positive community outcomes, suggesting instead that more centralized governance approaches can also positively affect moderators' psychological well-being and, by extension, community cohesion and effectiveness.

Introduction

Social media platforms have fundamentally changed the dynamics of human interaction, ushering in new forms of engagement and community development. Among these platforms, Reddit has emerged as a prominent hub for diverse communities centered around a wide spectrum of topics. Volunteer moderation plays a critical role within the Reddit ecosystem, overseeing communities that span from small groups of individuals to millions of subscribers. Moderators shape their community's ethos, uphold rules, facilitate discussions, and ensure quality interactions. Consequently, the success of Reddit's online communities depends largely on the effectiveness of these voluntary moderators.

As both the founders and enforcers of authority within their communities, moderators are key stakeholders in governance outcomes. They not only establish and implement policies, but also directly experience their effects. Consequently, moderators have a vested interest in the design and implementation of effective governance systems that promote community well-being, though it is not clear what constitutes an effective governance system within online communities. Conventional approaches to governance assume that the advantages of democratic processes are meant to be shared by all (Ostrom, 1998), yet the rewards of online community governance are primarily and most directly realized by those who govern. And online platforms, by allowing anyone to create and manage their own community, extend governance experiences to an unprecedented number of users. Despite their integral function, however, our current knowledge about the interplay between different community governance styles and moderators' psychological experiences remains limited. This article seeks to bridge this gap by exploring the extent of member involvement in rule-making processes and its relationship with moderators' psychological needs. The significance of our research lies in challenging the assumptions that online communities inherently foster horizontal social structures that uniformly benefit all users. Our findings present a nuanced perspective, suggesting that more centralized approaches to governance can also have positive implications for online communities and their volunteer moderators.

To frame our study, we draw on the humanistic rational choice theory (HRCT; DeCaro, 2019), which integrates Ostrom's (1998) theory of collective action with self-determination theory (SDT; Deci and Ryan, 2012) to produce an interdisciplinary behavioral theory of cooperation. The premise of this framework is that governance

structures, when aligned with fundamental human needs, such as security and belonging, foster legitimacy and group cohesion, thereby facilitating enhanced collaborative efforts. Applying this framework to Reddit's online communities, our study examines the interaction between different governance procedures and moderators' psychological needs and motivations.

Our work delves into various governance styles, or rule-making processes in subreddits, ranging from democratic to autocratic governance, and their relationship with elements critical for collaborative governance. These factors include moderators' sense of justice, self-determination, belonging, security, competence, institutional acceptance, and internalization. Drawing insights from an extensive survey of over 600 Reddit moderators, our findings reveal that democratic decision-making, or in other words, participatory rule processes do not necessarily correlate with greater feelings of community belonging and institutional acceptance. Instead, our data point to a surprising finding: more centralized forms of governance appear to be more effective in cultivating group cohesion and institutional acceptance within the unique ecosystem of Reddit communities.

The findings of this study have important implications for understanding the complex dynamics of online communities and tradeoffs related to different governance structures. The divergence between democratic processes and feelings of community belonging and institutional acceptance underscores the significance of factoring in governance structures when devising and implementing them. In addition, our findings highlight the need for platforms like Reddit to provide tools and resources that could enhance participatory decision-making processes, to better support moderators and overall community success.

Literature review

Volunteer community moderation

The sustainability of an online community is related to its ability to provide its members with high-quality content and resources (Hercheui, 2011), which requires some form of moderation (Lazar and Preece, 2002). Since the early days of the Internet, community governance in online spaces has undergone significant evolution. Initially, research in this domain was rooted in ethnographic studies from the late 1970s (Hiltz and Turoff, 1978), intensifying in the coming decades with a focus on digital spaces such as electronic bulletin board systems (Chesebro, 1985), electronic information exchange systems (Kerr, 1986), Internet Relay Chat communities (Reid, 1991), and public list servers (Butler, 1999). These communities mainly facilitated minimal text-based interactions, and were characterized by their technical simplicity and decentralized nature, compared with today's sophisticated online platforms. Despite being unmoderated, they still required some level of technical and social maintenance work, in the form of infrastructure administration, and at times management of interpersonal dynamics (Butler et al., 2007; Kollock and Smith, 1996). Moving into the 2000s, the research focus has shifted toward peer production groups, such as Wikipedia (Burke and Kraut, 2008; Butler et al., 2008; Giles, 2005; Kiesler et al., 2011; Ransbotham and Kane, 2011), or open source software (OSS) communities (Crowston and Howison, 2006; O'Mahony and Ferraro, 2007; Weber, 2005), exploring new paradigms of collaboration and governance in online

spaces. This period saw the emergence of distributed, crowd-sourced moderation, marked with fluid, informally emergent leadership structures (Faraj et al., 2015; Larson and DeChurch, 2020). Their self-regulation was reflective of the early Internet culture, which emphasized freedom of expression and a communal approach to information sharing and governance.

As the Internet evolved, so too did the nature and scale of online communities. The advent of Web 2.0 introduced platforms that supported user-generated content and social networking at an unprecedented scale, leading to a more dynamic and interactive online experience. As online communities grow in size and complexity, their leaders encounter a range of challenges that extend beyond the technical and social maintenance into establishing new communities and dealing with newcomers, encouraging quality contributions, ensuring long-lasting commitment, and regulating user behavior (Butler et al., 2007; Kiesler et al., 2011; Lazar and Preece, 2002). Consequently, community governance became increasingly multifaceted, necessitating more structured moderation approaches to manage the growing volume of content and the diverse range of user interactions.

Grimmelmann (2015) defines moderation as "the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse" (p. 52). While some platforms adopt a commercial strategy, involving moderation algorithms or professional moderators, or a combination of both (Gillespie, 2020; Gorwa et al., 2020), others like Reddit depend on the voluntary efforts of their users for community governance. The nature of moderation work varies significantly across platforms. For example, on Mastodon, moderation not only involves establishing policies for each node, or community, but also requires server administration (Jhaver et al., 2023). Moderating a live stream channel on Twitch demands real-time decision-making to manage dynamic, interactive content (Cai et al., 2021). Similarly, on Discord, server moderators are responsible for regulating voice-based, ephemeral communications among users, a unique challenge described by Jiang et al. (2019). In the context of Reddit, users not only create and manage their communities but also develop distinct governance structures and community-specific guidelines. Across all these platforms, volunteer moderators deeply engage with their communities, establishing rules and maintaining standards for quality contributions, often with minimal guidance or intervention from the platforms they inhabit.

Reddit's structure of subreddits, each with its own set of rules and established governance mechanisms, presents a stark contrast to the early Internet communities. A Reddit moderator's main role usually involves setting and enforcing rules and norms, debating and negotiating with other moderators, conducting regular content reviews, and effectively addressing posts or users that transgress the boundaries of community guidelines (Matias, 2019; Seering et al., 2019; Wohn, 2019). Subreddit moderation is a costly task, which requires not only emotional labor (Wohn, 2019), but also "civic" labor (Matias, 2019). Despite this civic quality of the task, however, moderators commonly do not consult with community members when they are establishing or enforcing rules (Seering, 2020; Zhang et al., 2020). Oftentimes, this strictly hierarchical form of governance causes disagreements between moderators and users (Jhaver et al., 2019; Matias, 2019; Seering, 2020; Seering et al., 2019, 2022; Srinivasan et al., 2019).

There seems to be a discrepancy between community members and moderators that is also apparent in their ideal form of governance. As with many other social media platforms, Reddit lacks a comprehensive mechanism to systematically aggregate input from their members and integrate it into the decision-making processes of its subreddits (Juneja et al., 2020). Consequently, this responsibility rests on the shoulders of volunteer moderators. Yet, a study conducted with Reddit users shows that while the majority of them expressed a desire for more democracy, more than half the moderators who participated in the survey were not in favor of democratic governance (Weld et al., 2022). Considering how moderation decisions could shape and limit not only the community discourse and participation (Gibson, 2019), but also the general public discourse and civic engagement (Dosono and Semaan, 2018; Schneider, 2022), moderators' ambivalence about democratic norms becomes an important issue that warrants further exploration.

In "Moderation as Civic Participation," Matias (2019) argues that moderation work in online communities can be compared with civic leadership roles in offline settings, such as neighborhood watch groups, where the responsibilities and boundaries of the role are constantly evolving. But online communities are also different from civic organizations. For one, online communities lack a primary purpose of serving the public. Moreover, volunteer moderators assume leadership roles within these communities, departing from formal processes electing experienced leaders in civic organizations. Online communities can be much larger, and interactions can happen much faster, often anonymously, posing challenges in managing undesirable conduct. Despite these differences, online communities have the potential to serve as important avenues for civic engagement and democratic participation, particularly among younger generations (Ekström and Östman, 2015). Though, they can also have negative effects on prosocial attitudes and behaviors, such as in cases of online trolling (Bacovsky, 2021). Considering that socialization within online communities has the potential to mold individual behaviors and shape perceptions of social norms and values (Strimling and Frey, 2020), moderators wield a significant influence over the outcomes of online experiences for individuals. Consequently, examining moderators and their self-governance strategies is essential for comprehending the dynamics of online communities, and their impact on social behavior.

HRCT

A growing community of psychologists and collective action scholars is attending to the psychological foundations of effective collective action and community governance. Building on this work, we explore the psychosocial processes behind cooperative self-governance in online communities (DeCaro and Stokes, 2008; Drury and Reicher, 2005; Freeman et al., 2020; Gavrilets, 2015; Kelly and Breinlinger, 1995; Van Vugt, 2009). While there is a long history of social–psychological or personality testing of online community participants (Malinen, 2015), these inventories often have a rather indirect connection to self-governance. What is needed is an approach to individual qualities that is couched within a framework for the psychological prerequisites of collaborative self-governance. To build this bridge, we rely on HRCT (DeCaro, 2019).

Combining and expanding on Elinor Ostrom's (1998) theory of collective action and SDT (Deci and Ryan, 2000), HRCT provides the psychological prerequisites of

effective self-rule in resource systems such as online communities. The core argument of this theory is that cooperative governance arises from three core psychosocial processes: fundamental need satisfaction, institutional acceptance, and in-group cohesion. Governance systems that satisfy these basic needs of individuals are perceived as legitimate; this legitimacy then provides a ground for institutional acceptance to develop among individuals, which facilitates group cohesion and further collaboration within the group (DeCaro et al., 2021).

Social–psychological needs are shown to be universal across different cultures (Chen et al., 2015), and influential across a variety of domains (Lian et al., 2012; Park and Rainey, 2008; Rosen et al., 2014; Sheldon et al., 2013). Within the domain of cooperative governance, De Caro has conducted several lab experiments that replicated ecological social dilemmas, to observe how the organizational aspects of governance structures influence a set of five social–psychological needs: procedural justice, self-determination, security, belonging, and competence (DeCaro et al., 2021). The HRCT suggests that satisfaction of these needs facilitates cooperation among individuals through development of legitimacy, in the form of internalization and institutional acceptance (DeCaro et al., 2021). Once a governing mechanism is perceived as legitimate by individuals, they are more likely to comply with and cooperate in achieving common goals. When norms or rules are internalized, norm-related behaviors become self-sustaining, and maintained without external enforcement. Institutional arrangements that fulfill an individual's fundamental needs for procedural justice and self-determination, such as democratic decision-making, tend to promote greater internalization of the established norms and decisions (Kiesler et al., 2011). Individuals become more likely to comply when their needs for autonomy are met, which results in increased stability and cooperation.

For Reddit moderators, attaining legitimacy seems to present a formidable challenge, marked by multifaceted considerations and intricate dynamics. Sarah Gilbert's (2020) analysis of r/AskHistorians provides insights into how the legitimacy of a community's governing structure is intertwined with the visibility of their efforts. While some of the work moderators perform, like content removal, is conspicuous, many other crucial tasks that are vital for the community's functioning remain hidden from the audiences. The lack of transparency in establishing and upholding community rules and norms causes skepticism and hostility toward moderators (Kiesler et al., 2011; Lazar and Preece, 2002), likely diminishing their feelings of security and autonomy. At times, disagreements arise between moderators and members due to the hierarchical nature of community governance structures (Jhaver et al., 2019; Matias, 2019; Seering, 2020; Seering et al., 2019, 2022; Srinivasan et al., 2019). Conversely, Dosono and Semaan's (2018) work on Reddit communities where Asia and Pacific Islanders engage in identity work highlights how moderators can enforce collectivist norms, such as inclusivity, civility, and rationality, to increase cooperation and build solidarity among users. These studies exemplify the highly delicate balance moderators need to uphold, between asserting authority and cultivating a sense of community (Lazar and Preece, 2002). In this context, the HRCT framework emerges as a promising approach, providing a conceptual lens to explore the intricate interplay between different community governance styles and the social–psychological needs of moderators.

The theoretical framework of HRCT has previously been tested in two lab experiments of a common-pool resource dilemma by DeCaro et al. (2015, 2021), providing evidence that different forms of governance differentially affect the social–psychological needs and internal motivations of individuals. However, the governance modes considered in the study—voting versus no voting, enforcement versus no enforcement in Study 1, compared with an index of decision-making that ranges from no decision to majority decision in Study 2—weakly represent the continuum of governance structures that are available. Furthermore, the very limited sample size of only four individuals per condition/group is a narrow representation of the complexities and variations in social groups and interactions. Reddit, for instance, encompasses a diversity of communities across a wide range of topics, and includes millions of individuals from various countries, cultures, and backgrounds.

While there has been increased attention to voluntary moderation, a large part of the previous work has mostly relied on either content analysis (Dosono and Semaan, 2018; Gibson, 2019; Juneja et al., 2020; Srinivasan et al., 2019; Topinka, 2018; Trujillo and Cresci, 2022) or interviews with a limited number of participants (Bozarth et al., 2023; Dosono and Semaan, 2019; Gilbert, 2020; Hwang and Foote, 2021; Jhaver et al., 2019; Seering et al., 2019, 2022; Squirrell, 2019), with few exceptions (Seering and Kairam, 2023; Wang et al., 2022; Weld et al., 2022). Diverging from this approach, Seering and Kairam (2023) surveyed moderators of live streaming communities on Twitch, focusing on their motivations, tasks, and roles, but this study did not consider the variations in rule-making processes and their potential impact on suggested relationships. Surveying a large group of Reddit users, Weld et al. (2022) examined community values such as diversity, inclusion and trustworthiness, and found that moderators and users differed in terms of their attitudes toward democracy. While this study did not consider the different governance processes employed within communities, it suggests that democratic participation may not be perceived by moderators as an efficient governance approach. Lately, a partial application of the HRCT framework on Reddit moderators provided survey evidence for an inverse relationship between democratic governance and moderator's attitudes of procedural justice and self-determination (Wang et al., 2022). Accordingly, the more control moderators had over their communities, the higher perceived procedural justice and self-determination they reported. However, this study considered only two of the specific fundamental needs proposed in the HRCT framework and further research is warranted to draw more robust conclusions about the relationship between different governance strategies and Reddit moderator's needs and motivations.

Current study

This study aims to address the gap in understanding the relationship between the level of community involvement in the rule development processes, and moderators' underlying psychological needs and motivations. As direct stakeholders in governance outcomes, we focus our analysis on community moderators, recognizing their role in shaping community dynamics and their unique perspective over governance processes. Our sampling choice is grounded in the premise that the primary beneficiaries of an online community governance approach are the individuals who voluntarily invest the effort to govern and manage the

community itself. Moderators occupy a distinct position within online communities, serving as both the enforcers and implementers of their own policies, shaping the community culture in which they also take part as users. It is therefore important to understand how different governance structures affect moderator experiences and community outcomes. Reddit provides the ideal grounds for this purpose, for it allows users to create and rule their own communities through established governance mechanisms, offering a unique opportunity to study moderator attitudes toward their governance strategies.

We build upon previous work on cooperative self-governance, but also recognize that subreddits and social–ecological dilemmas differ in significant ways. While both scenarios involve shared, limited resources (e.g. water or forest resources vs user time, energy, and attention; Butler, 2001; Kiesler et al., 2011; Kollock and Smith, 1996), subreddit moderators operate within a distinct context. Anyone can start new subreddits and exercise control without interference (Juneja et al., 2020). However, the anonymity and vast volume of interactions present both opportunities and challenges for fostering engagement, a critical factor in community success (Butler, 2001; Lazar and Preece, 2002; Lin and Lee, 2006). Moderators play a key role in promoting engagement by setting norms, enforcing rules, and removing disruptive content or users (Gilbert, 2020; Lazar and Preece, 2002; Seering et al., 2022). Nonetheless, their efforts can often lead to conflicts with other users and moderators, causing stress and burnout (Schöpke-Gonzalez et al., 2022). Although moderators highly value their communities and get fulfillment out of what they voluntarily do, moderation seems to also involve a lot of invisible work (Gilbert, 2020; Thach et al., 2022), accompanied by tensions. Considering the limited governance tools available to moderators due to the platform's reliance on their voluntary labor (Schneider, 2022; Thach et al., 2022), it may not be surprising that moderators seem to favor hierarchical governance structures (Weld et al., 2022).

Democratic decision-making, or in other words, participatory rule processes in online communities can be inclusive and empowering, allowing members to actively participate in shaping the direction and policies of the community (Dosono and Semaan, 2018; Seering, 2020; Thach et al., 2022). This would provide legitimacy to the governing authority, helping other members comply with moderator decisions (Kiesler et al., 2011; Lazar and Preece, 2002). But participatory governance may also lead to a large number of competing viewpoints, uninformed opinions, or even symptoms of mob rule, resulting in difficulties for moderators when reaching a consensus, enforcing a rule, or defending a decision (Gilbert, 2020; Juneja et al., 2020; Schöpke-Gonzalez et al., 2022; Squirrell, 2019). This may cause a perceived lack of procedural justice, as moderators can feel that their opinions are not adequately considered, or respected, which leaves very little room for anything but asserting more control over the community. Since challenging user behaviors can become serious obstacles to effective management, they could prompt moderators to give precedence to their own understanding of procedural fairness rather than concerns related to democratic involvement. Therefore we hypothesize as follows:

H1. Participatory rule processes will negatively correlate with greater feelings of procedural justice.

Early work on communities suggest that when members are given a chance to argue their viewpoints, their commitment to the social group and willingness to comply with group

decisions greatly enhance (Colquitt, 2001; Kiesler et al., 2011). But previous work also indicates that online communities that prioritize free speech frequently become breeding grounds for the dissemination of discriminatory, racist, and sexist content (Gilbert, 2020; Massanari, 2017; Topinka, 2018). These communities, while promoting open dialogue and deliberation, can inadvertently create an environment where harmful and offensive viewpoints flourish unchecked. This suggests that in communities with hierarchical governance structures, where power is more centralized, moderators would have greater control over decision-making processes and user behaviors (Dosono and Semaan, 2018; Juneja et al., 2020; Weld et al., 2022), potentially resulting in an increased sense of autonomy for them. Therefore, we predict as follows:

H2. Participatory rule processes will negatively correlate with greater feelings of self-determination.

Most community leaders look for the recognition and acknowledgment of their voluntary labor, which enhances their commitment and feelings of belonging toward the community (Burke and Kraut, 2008; Kim, 2006). But the inherent inclusiveness of participatory governance within online communities can sometimes give rise to conflicts among community members (Lazar and Preece, 2002), which in turn has the potential to diminish the overall sense of cohesion and camaraderie within the community (Gilbert, 2020; Juneja et al., 2020). This intricate dynamic has been explored in prior studies, where the management of recurring conflicts was found to contribute to feelings of frustration and burnout among community moderators (Dosono and Semaan, 2019; Schöpke-Gonzalez et al., 2022; Squirrell, 2019). It is plausible that engagement with these communities might inadvertently lead moderators to perceive a lack of recognition and appreciation for their knowledge and experience from fellow community members (Juneja et al., 2020). Consequently, this perception could erode their sense of belonging to the community (Tausczik et al., 2014), and result in a decline in their perceived competence and effectiveness as moderators as follows:

H3. Participatory rule processes will negatively correlate with greater feelings of sense of belonging.

H4. Participatory rule processes will negatively correlate with greater feelings of competence.

Security is an important issue for moderators, as it enhances collaborative efforts, and ensures healthy and free communication among community members, which are essential to any community's success (Larson and DeChurch, 2020; Lazar and Preece, 2002; Preece, 2000). Prior studies underscore that moderators often grapple with toxic dynamics within their communities, confronting racism, sexism, and even death threats, ultimately leading to psychological distress (Dosono and Semaan, 2019; Gilbert, 2020). The implementation of participatory decision-making processes may be exposing moderators to a wider array of intricate challenges (Juneja et al., 2020), potentially resulting in a diminished sense of security. By contrast, communities featuring a more hierarchical

decision-making structure, where moderators exercise greater assertiveness without accountability, might be experiencing fewer security concerns. Therefore, we hypothesize as follows:

> *H5.* Participatory rule processes will negatively correlate with greater feelings of security.

In a participatory decision-making structure, moderator opinions may not always align with the majority viewpoints or the decisions reached by the community (Lazar and Preece, 2002). A hierarchical governance mechanism, on the contrary, can ensure an organized, stable, decision-making structure, promoting a sense of safety and order in the community (Dosono and Semaan, 2018; Gibson, 2019), which may potentially lead moderators to feel a higher sense of authority and fairness. Therefore, we predict as follows:

> *H6.* Participatory rule processes will negatively correlate with greater feelings of institutional acceptance.

In hierarchical structures, the concentration of authority within a limited group of individuals often facilitates moderators' internalization of the established rules and norms. Conversely, with the distribution of power in participatory systems, conflicting viewpoints might lead moderators to feel as if they do not have enough influence on the direction the community is following, making it more difficult for them to internalize the community standards and values. Therefore, we hypothesize as follows:

> *H7.* Participatory rule processes will negatively correlate with greater feelings of internalization.

Method

Participants and procedure

To understand the relationship between different community governance modes on Reddit, and moderators' psychological needs and motivations, we conducted a survey with a total of 2269 respondents on Qualtrics. However, we had to exclude a large part of the responses from further analysis due to either incomplete questionnaires or failure to pass the two attention checks we included in the study. These attention checks were strategically integrated in the survey at two different time points and asked participants to select a specific response from a list of options, with clear instructions, to ensure data validity. Ultimately, our analysis is based on a sample of 605 respondents who completed the questionnaire and passed the two attention checks. To recruit participants we utilized both public and private messages sent through Modmail, Reddit's internal messaging system for moderators. Participants were informed that the study was designed to explore the interplay between Reddit use, leadership, and community participation, and were required to provide their consent before taking the survey. The compensation for completing the survey was the option to enter a raffle, and an Amazon gift card valued at US$50 was offered as the raffle prize.

Table 1. Examples of sample questions, mean, standard deviation, and Cronbach's alpha.

Variable	Sample item	M	SD	α
Participatory process	*Please select the option that best describes the process you use to make rules in your community*	2.6	1.13	
Procedural justice	*The procedures used to create rules make me feel as if I am able to exercise my views and desires*	2.11	1.01	.84
Self-determination	*The procedures used to create rules make me feel as if I have a genuine sense of choice*	2.86	1.19	.91
Perceived belonging	*I feel like I'm seen as a valuable person in this community*	3.09	1.15	.8
Perceived competence	*I feel as if I am meeting the challenge of making valuable contributions to my community*	2.87	1.16	.85
Security	*I feel that the rules we have make the community more structured and predictable*	2.24	1.35	.85
Institutional acceptance	*I support the procedures used to create rules in the community.*	2.96	1.08	.93
Internalization	*I feel that the community rules match with my own desires and values*	2.39	1.14	.89

Participants in our sample had a mean age of 31 years, with 69.2% of participants between the ages 18 and 34 years (n=438). The majority of the sample was identified as male (70.9%, n=450), with some or more with a college education (74.6%, n=472). Most moderators in our sample were experienced, with 37% having 3 years of experience or more (n=234), and 54.4% of the sample reported spending more than 5 hours moderating their community on a weekly basis (n=344).

Measures

Respondents were asked to respond to each item using a 7-point Likert-type scale that ranged from "strongly agree" to "strongly disagree." Sample survey items are displayed in Table 1 alongside the mean, standard deviation, and Cronbach's alpha score for each variable (Supplemental Appendix A).

To identify the governance structure employed in each community, we asked the respondents to select the option that best described the process they used to make community rules (Seering et al., 2019). Items of this construct were arranged from most hierarchical (e.g. *Moderators enact new rules on their own without consulting members*) to most democratic (e.g. *Moderators work with the entire group of community members to enact new rules*) ($M=2.60$, $SD=1.13$).

Procedural justice refers to the perceived fairness or "justice of the processes that lead to decision outcomes" (Colquitt, 2001: 386). We assessed perceived fairness of rule-making procedures with four items (DeCaro et al., 2015; Van Prooijen, 2009) (e.g. *The procedures used to create rules make me feel as if I am able to exercise my views and desires;* $\alpha=.84$, $M=2.11$, $SD=1.01$).

Self-determination refers to the extent to which individuals feel they have self-agency and act in accordance with their personal values and goals (Ryan and Deci, 2006). Perceptions of self-determination were assessed using four items, where two measured the dimension of self-concordance (Sheldon and Houser-Marko, 2001) (e.g. *The procedures used to create rules make me feel as if I am free to do things that express, or exercise, my "true self.*) and two measured sense of choice (DeCaro et al., 2015) (e.g. *The procedures used to create rules make me feel as if I have a genuine sense of choice;* α=.91, *M*=2.86, *SD*=1.19).

Belonging is described as perceived acceptance from a social group or community (Baumeister and Leary, 1995). The sense of belonging was measured with six items that targeted two dimensions (DeCaro et al., 2015). Both perceived social standing (Leary et al., 2013) (e.g. *I feel like I'm seen as a valuable person*), and social connectedness was measured with three items (Sheldon and Houser-Marko, 2001) (e.g. *I feel close and connected with other users who are important to me;* α=.80, *M*=3.09, *SD*=1.15).

Competence denotes the perceived ability to exercise control and effect change in one's environment (Deci and Ryan, 2000). Perceived competence of moderators was measured with four items (DeCaro et al., 2015), where two measured the intelligence dimension (Ryan, 1982) (e.g. *I feel intelligent*) and another two measured the mastery (Sheldon and Houser-Marko, 2001) (e.g. *I feel as if I am meeting the challenge of making valuable contributions to my community;* α=.85, *M*=2.87, *SD*=1.16).

Security denotes the perceptions of structure and safety that stem from predictability (Sheldon et al., 2001). This construct was measured with two items (DeCaro et al., 2015) (e.g. *I feel that the rules we have make the community more structured and predictable;* α=.85 *M*=2.24, SD=1.35).

Institutional acceptance is described as "the extent to which individuals endorse a set of rights, rules and decision-making procedures" (DeCaro and Stokes, 2013: 6). Moderator's institutional acceptance was measured using three items (Colquitt, 2001; DeCaro et al., 2015) (e.g. *I support the procedures used to create rules in the community;* α=.93, *M*=2.06, *SD*=1.08).

Internalization refers to the process by which individuals adopt the rules and norms that govern their behaviors as part of their own values and belief systems (DeCaro and Stokes, 2013). Moderators' level of internalization was measured with two items (DeCaro et al., 2015) (e.g. *I feel that the rules match with my own desires and values;* α=.89, *M*=2.39, *SD*=1.14).

Results

We conducted a multiple regression analysis to examine the relationship between rule-making processes and Reddit moderators' perceptions of procedural justice, self-determination, belonging, competence, security, institutional acceptance, and internalization. All predictor variables were mean-centered prior to the regression analysis. We checked for multicollinearity by assessing zero-order correlations among variables and their variance inflation factor (VIF) statistics (Table 2). While the zero-order correlation matrix revealed that some of the independent variables were moderately correlated, the accompanying VIF statistics were all found to be within an acceptable range (VIF < 5), indicating limited collinearity (Akinwande et al., 2015).

Table 2. Zero-order correlations with tolerance and variance inflation factor statistics.

	1	2	3	4	5	6	7	8	Tolerance	VIF
Rule process	**1.00**									
Procedural justice	0.19***	**1.00**							0.41	2.44
Self-determination	−0.05	0.40***	1.00						0.51	1.96
Community belonging	−0.23***	0.43***	0.60***	1.00					0.50	2.02
Competence	0.02	0.47***	0.56***	0.53***	1.00				0.51	1.96
Security	0.08	0.44***	0.33***	0.28***	0.36***	1.00			0.66	1.51
Institutional Acceptance	−0.10**	0.53***	0.36***	0.37***	0.39***	0.38***	1.00		0.45	2.23
Internalization	0.04	0.54***	0.50***	0.44***	0.46***	0.53***	0.43***	1.00	0.53	1.89

Table 3. Multinomial regression results with participatory process as the dependent variable.

	β	SE
(Intercept)	2.49(***)	0.04
Procedural justice	.32(***)	0.07
Self-determination	−.02	0.05
Community belonging	−.30(***)	0.05
Competence	.08	0.05
Security	.04	0.04
Institutional acceptance	−.14*	0.06
Internalization	.02	0.05

$*p < .05; **p < .01; ***p < .001.$

Overall, the model was significant $F(7, 597) = 8.86, p < .001, R^2 = .093$. Our findings suggest that moderators who utilize participatory decision-making procedures for community governance perceive greater levels of procedural justice toward the governance processes. However, participatory decision-making also seems to have its own drawbacks, as moderators also reported lower levels of community belonging and institutional acceptance when members are involved with community governance (Table 3).

We found a significant positive correlation between participatory governance and procedural justice ($β = .32, p < .001$). Moderators who reported higher feelings of procedural justice were more likely to use democratic processes in their communities. Hypothesis 1 was not supported. However, the correlation between sense of belonging and participatory governance was significant in the negative direction ($β = −.30, p < .001$), supporting Hypothesis 3. This suggests that fostering an inclusive decision-making environment might inadvertently reduce feelings of belonging, making it difficult to maintain a strong sense of community cohesion. Similarly, the relationship between participatory rule processes and institutional acceptance was found to be negatively correlated with significance ($β = −.14, p < .01$). Hypothesis 6 was supported: moderators who governed their communities with more participatory processes reported lower levels of acceptance toward their governance structures.

We found a negative association between feelings of self-determination and participatory rule processes, however, this relationship lacked significance ($β = −.02, p = .659$). Hypothesis 2 was not supported. We also found nonsignificant positive associations between participatory governance and feelings of competence ($β = .08, p = .115$) and feelings of security ($β = .04, p = .245$). Hypotheses 4 and 5 were therefore not supported. The participatory rule processes and internalization were positively correlated, but lacked significance ($β = .022, p = .677$). Hypothesis 7 was not supported.

In addition, we conducted an analysis to explore the potential interaction effects among procedural justice, institutional acceptance, belonging, and rule processes. However, our findings revealed that these interactions did not yield statistically significant effects in their respective domains. The interaction between procedural justice and institutional acceptance did not have a significant effect on participatory processes

($\beta = -.061$, $p = .104$). The interaction between community belonging and institutional acceptance lacked a significant effect on rule processes ($\beta = -.015$, $p = .684$). Similarly the interaction between procedural justice and belonging was found to be nonsignificant ($\beta = -.041$, $p = .193$). We tested the effects of gender by adding a covariate to our regression model, and found the effect to be positive but insignificant, with little impact on other model coefficients ($\beta = .019$, $p = .846$).

Discussion

This study explored the relationship between community governance processes and the psychological needs and motivations of community moderators through the lens of HRCT. We focus on moderators rather than users out of a perspective that their substantial investment in building, leading, and taking responsibility for their community may concentrate in them the civic skills and psychological benefits that online engagement can bring, regardless of the form of government they impose. Traditionally, in physical communities, increased democratic participation is often correlated with higher perceptions of fairness, a greater sense of belonging, and increased institutional acceptance, all considered to be crucial elements for effective cooperation (DeCaro et al., 2021; Ostrom, 1998). However, our findings suggest a contrasting pattern for online communities. While moderators who employ democratic rule-making procedures report higher levels of procedural justice, supporting previous research (Kiesler et al., 2011), participatory approaches to online governance do not necessarily translate into a stronger sense of community or acceptance. Contrary to traditional perspectives, we found that more centralized forms of governance may be facilitating group cohesion and institutional acceptance in the context of Reddit communities. The study findings shed light on the unique challenges and the inherently conflicting nature of volunteer moderation on Reddit, and the difficult task of balancing between democratic participation, community cohesion, limited volunteer energy, and effective governance.

At first glance, it seems counterintuitive that a more hierarchical, centralized governance on Reddit is associated with greater feelings of community belonging and acceptance among moderators. This is because democratic processes are often associated with inclusivity and equality (Kiesler et al., 2011; Ren et al., 2011), while centralized authority can at times feel oppressive and divisive (Lazar and Preece, 2002). However, it is important to recognize that online communities are strikingly different from their offline counterparts, that is, civic organizations, as the former provide greater anonymity for members, lack physical presence, and therefore have lower barriers to entry and participation than the latter. Moreover, online communities, like Reddit, provide opportunities for members to take on leadership roles and contribute to decision-making processes. Anyone can start their own community and shape it to their liking without much intervention, which makes it difficult to ensure that moderators have the necessary skills to effectively govern their community while promoting an inclusive, participatory environment for all their members.

On one hand, participatory governance processes in online communities meet the moderator's needs for procedural justice; a critical component for cooperation in organizational behavior and psychology (Colquitt et al., 2001; Lind and Tyler, 1988; Tyler and

Degoey, 1995). The importance of perceived fairness in rule compliance (Kiesler et al., 2011; Walters and Bolger, 2019), cooperation (Tyler and Degoey, 1995), and individual wellbeing (Moliner et al., 2013) has been established in previous work. On the other hand, participatory governance seems to be at odds with moderators' feelings of community cohesion and institutional acceptance, which are crucial for maintaining a successful community (Blanchard and Markus, 2004; Frey and Sumner, 2019; Kiesler et al., 2011; Lin and Lee, 2006).

Previous research has shown that a higher sense of belonging positively correlates with increased time and effort spent in online community participation and greater intentions to continue participating in the future (Roberts, 1998; Teo et al., 2003; Zhao et al., 2012). This highlights the importance of establishing a sense of togetherness and reciprocity in communities, particularly for moderators. Our results showed that shared decision-making is inversely related to the moderator's feelings of community belonging. Those who adopt more centralized governance structures in their communities reported greater feelings of community belonging. We also found evidence of an inverse relationship between participatory governance and institutional acceptance. Moderators who share their decision-making powers reported lower levels of approval for the procedures they employ. This suggests that community moderators might be feeling greater institutional acceptance when they have more control over their communities, and are able to make decisions without having to consult with a larger group of people.

Participatory governance on Reddit seems to be presenting a multitude of challenges for moderators, despite all its potential benefits. While involving community members in decision-making can increase perceptions of procedural justice, it can also be a demanding and time-consuming task for moderators who already provide their services voluntarily. Democratic participation promotes open debate and discussion in decision-making, but it can also result in dissenting views and conflict that may hinder group cohesion and engagement quality (Gilbert, 2020). Reddit is known to be a fertile ground for harassment (Dosono and Semaan, 2019; Schöpke-Gonzalez et al., 2022), misogyny, and racism (Topinka, 2018), adding to the difficulty of managing and dealing with unwarranted behaviors. These factors are potentially contributing to the reasons why moderators are hesitant to implement democratic governance on Reddit (Weld et al., 2022), as it is likely to increase the existing challenges and issues.

It is important to note that even though volunteer moderators hold a unique position in the Reddit ecosystem, their powers are not absolute, and their ability to influence governance is shaped by the power dynamics at play. While moderators play a crucial role in maintaining the quality of the content and fostering user engagement, their authority is constrained by the broader platform structure and guidelines set by the platform administrators (Gilbert, 2020). Although Reddit administrators aim to create a consistent and cohesive platform experience, instances of platform-wide modifications that disregard the specific needs of individual subreddits can create friction between administrators and moderators (Schneider, 2022). As a result, power struggles emerge, highlighting the discord between top–down decision-making and bottom–up community-driven governance.

The effectiveness of volunteer moderators is also impeded by the scarcity of available tools for efficient governance, making democratic participation perhaps too challenging to implement. This deficiency places an undue burden on moderators, hindering their

ability to effectively uphold community standards. Without adequate resources to monitor and moderate content and interactions, moderators struggle to ensure the visibility of their decision-making processes, and at times opt for non-disclosure to mitigate potential conflicts (Gilbert, 2020; Juneja et al., 2020; Seering et al., 2019; Weld et al., 2022). This lack of transparency can fuel suspicions among community members, as the mechanisms behind content removals, rule enforcement, and other governance decisions remain obscure (Jhaver et al., 2019; Juneja et al., 2020; Kiesler et al., 2011; Lazar and Preece, 2002; Squirrell, 2019). The absence of clear visibility into these processes undermines the credibility of moderators and creates tensions between them and the community they voluntarily serve.

To address these challenges, it is imperative for platforms to invest in developing user-friendly tools that empower moderators to create vibrant, safe, and well-governed online communities. For instance, platforms can develop and provide better content filtering options, and features that allow for easier communication and collaboration among moderators. Clearer guidelines on handling specific situations and access to support resources can also enhance moderators' decision-making processes. New features that increase transparency for users can be introduced, which may include options for users to appeal decisions and provide more context on how specific rules should be applied. A meta-moderation system could also be useful, for users to systematically evaluate and provide feedback on moderator decisions and actions, such as the one established by the Slashdot (Kiesler et al., 2011). Platforms can also actively support moderators in fostering diverse and inclusive communities, which may involve offering resources on promoting respectful dialogue, addressing hate speech, and managing sensitive topics. By ensuring a more balanced power distribution within the platform, it is possible to create an environment where administrators, moderators and community members work cohesively to build and maintain a thriving digital ecosystem.

Limitations

There are several limitations to this work. First, the study relied on self-reported, cross-sectional data from Reddit moderators, which may be subject to social desirability bias. Second, the study focused on community moderators as the direct stakeholders of governance outcomes. However, considering the perspective of community members may provide a more comprehensive understanding of community governance structures and their influence. Future work could examine how different governance mechanisms affect community belonging among members. Third, this study was limited in its ability to comprehensively explore the gender-related nuances in moderator behaviors. While our analysis did not yield statistically significant effects for gender on the psychological constructs we examined, we acknowledge the potential influence of gender on community governance structures. Future work could address these limitations by actively seeking a more diverse and representative participant sample, which would enable a more granular analysis of gender-related effects on community governance and psychological needs of community moderators. Fourth, the study only examined a specific set of psychological constructs that are deemed necessary for effective collaborative governance in previous work on physical communities, and did not consider other potential factors

that may influence community governance online. As such, there might be potentially important variables that have been left out of our model. Future work could consider replicating the findings of this study in other online communities and further investigate the psychological mechanisms through which governance structures may affect moderators' needs and motivations.

Conclusion

Overall, this study sheds light on the complex interplay between online community governance processes, and moderators' social–psychological needs and motivations. It directly acknowledges the reality that, despite the democratic ideals of online community research, most volunteer moderators on Reddit do not implement democracies. Rather, they seem to develop valuable civic literacies and perspectives despite their vulnerability to the implicit feudalism of the platform. By highlighting the unique challenges and tradeoffs inherent in volunteer moderation work on Reddit, this study contributes to a deeper understanding of community dynamics and the factors that facilitate effective governance in online communities. We found that moderators who employ participatory processes to govern their communities perceive higher levels of procedural justice at a cost, as these moderators also reported lower levels of community belonging and acceptance. Our results show that centralized power structures can foster a sense of community coherence and institutional acceptance among moderators, which carries the potential to positively influence engagement quality and community success. However, it is also important to recognize the resulting power imbalances, which require further efforts to mitigate the risks through transparent policies and inclusive procedures. By understanding the relationship between community governance styles and moderator perceptions, it is possible to design and implement governance structures that can promote both community cohesion and effective cooperation in online communities.

ORCID iD

Beril Bulat [ID] https://orcid.org/0000-0002-0966-129X

Supplemental material

Supplemental material for this article is available online.

References

Akinwande MO, Dikko HG and Samson A (2015) Variance inflation factor: as a condition for the inclusion of suppressor variables in regression analysis. *Open Journal of Statistics* 5(7): 754–767.

Bacovsky P (2021) Gaming alone: videogaming and sociopolitical attitudes. *New Media & Society* 23(5): 1133–1156.

Baumeister RF and Leary MR (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin* 117(3): 497–529.

Blanchard AL and Markus ML (2004) The experienced "sense" of a virtual community: characteristics and processes. *SIGMIS Database* 35(1): 64–79.

Bozarth L, Im J, Quarles C, et al. (2023) Wisdom of two crowds: misinformation moderation on Reddit and how to improve this process—a case study of COVID-19. *Proceedings of the ACM on Human-Computer Interaction* 7(CSCW1): 155.

Burke M and Kraut R (2008) Taking up the mop: identifying future Wikipedia administrators. In: *CHI EA '08: CHI '08 extended abstracts on human factors in computing systems*, Florence, 5–10 April, pp. 3441–3446. New York: ACM.

Butler B, Joyce E and Pike J (2008) Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in Wikipedia. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, Florence, 5–10 April, pp. 1101–1110. New York: ACM.

Butler B, Sproull L, Kiesler S, et al. (2007) Community effort in online groups: who does the work and why? In: Weisband SP (ed.) *Leadership at a Distance*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 187–210.

Butler BS (1999) *The Dynamics of Cyberspace: Examining and Modelling Online Social Structure* (ed. K Carley and R Kraut). Ann Arbor, MI: Carnegie Mellon University.

Butler BS (2001) Membership size, communication activity, and sustainability: a resource-based model of online social structures. *Information Systems Research* 12(4): 346–362.

Cai J, Wohn DY and Almoqbel M (2021) Moderation visibility: mapping the strategies of volunteer moderators in live streaming micro communities. In: *IMX '21: Proceedings of the 2021 ACM international conference on interactive media experiences*, Virtual Event, 21–23 June, pp. 61–72. New York: ACM.

Chen B, Vansteenkiste M, Beyers W, et al. (2015) Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion* 39(2): 216–236.

Chesebro JW (1985) Computer-mediated interpersonal communication. *Information and Behavior* 1: 202–222.

Colquitt JA (2001) On the dimensionality of organizational justice: a construct validation of a measure. *Journal of Applied Psychology* 86(3): 386–400.

Colquitt JA, Conlon DE, Wesson MJ, et al. (2001) Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology* 86(3): 425–445.

Crowston K and Howison J (2006) Hierarchy and centralization in free and open source software team communications. *Knowledge, Technology & Policy* 18(4): 65–85.

DeCaro DA (2019) Humanistic rational choice understanding the fundamental motivations that drive self-organization and cooperation in commons dilemmas. In: Hudson B, Cole D and Rosenbloom J (eds) *Routledge Handbook of the Study of the Commons*. New York: Taylor & Francis, pp. 117–131.

DeCaro DA and Stokes M (2008) Social-psychological principles of community-based conservation and conservancy motivation: attaining goals within an autonomy-supportive environment. *Conservation Biology* 22(6): 1443–1451.

DeCaro DA and Stokes MK (2013) A social–psychological perspective. *Ecology and Society* 18(4): 40.

DeCaro DA, Janssen MA and Lee A (2015) Synergistic effects of voting and enforcement on internalized motivation to cooperate in a resource dilemma. *Judgment and Decision Making* 10(6): 511–537.

DeCaro DA, Janssen MA and Lee A (2021) Motivational foundations of communication, voluntary cooperation, and self-governance in a common-pool resource dilemma. *Current Research in Ecological and Social Psychology* 2: 100016.

Deci EL and Ryan RM (2000) The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry* 11(4): 227–268.

Deci EL and Ryan RM (2012) Self-determination theory. In: Higgins E (ed.) *Handbook of Theories of Social Psychology*, Vol. 1. Thousand Oaks, CA: Sage, pp. 416–436.

Dosono B and Semaan B (2018) Identity work as deliberation: AAPI political discourse in the 2016 US presidential election. In: *CHI '18; Proceedings of the 2018 CHI conference on human factors in computing systems*, Montreal, QC, Canada, 21–26 April, pp. 1–12. New York: ACM.

Dosono B and Semaan B (2019) Moderation practices as emotional labor in sustaining online communities: the case of AAPI identity work on Reddit. In: *CHI '19: Proceedings of the 2019 CHI conference on human factors in computing systems*, Glasgow, 4–9 May, pp. 1–13. New York: ACM.

Drury J and Reicher S (2005) Explaining enduring empowerment: a comparative study of collective action and psychological outcomes. *European Journal of Social Psychology* 35(1): 35–58.

Ekström M and Östman J (2015) Information, interaction, and creative production: the effects of three forms of internet use on youth democratic engagement. *Communication Research* 42(6): 796–818.

Faraj S, Kudaravalli S and Wasko M (2015) Leading collaboration in online communities. *Mississippi Quarterly* 39(2): 393–412.

Freeman J, Baggio JA and Coyle TR (2020) Social and general intelligence improves collective action in a common pool resource system. *Proceedings of the National Academy of Sciences* 117(14): 7712–7718.

Frey S and Sumner RW (2019) Emergence of integrated institutions in a large population of self-governing communities. *PLoS ONE* 14(7): e0216335.

Gavrilets S (2015) Collective action and the collaborative brain. *Journal of the Royal Society Interface* 12(102): 1067.

Gibson A (2019) Free speech and safe spaces: how moderation policies shape online discussion spaces. *Social Media + Society* 5(1): 2056305119832588.

Gilbert SA (2020) 'I run the world's largest historical outreach project and it's on a cesspool of a website.' Moderating a public scholarship site on Reddit: a case study of r/AskHistorians. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW1): 1–27.

Giles J (2005) Internet encyclopedias go head to head. *Nature* 438: 900–901.

Gillespie T (2020) Content moderation, AI, and the question of scale. *Big Data & Society* 7(2): 2053951720943234.

Gorwa R, Binns R and Katzenbach C (2020) Algorithmic content moderation: technical and political challenges in the automation of platform governance. *Big Data & Society* 7(1): 2053951719897945.

Grimmelmann J (2015) The virtues of moderation. *Yale Journal of Law & Technology* 17: 42.

Hercheui MD (2011) A literature review of virtual communities. *Information, Communication & Society* 14(1): 1–23.

Hiltz SR and Turoff M (1978) *The Network Nation: Human Communication Via Computer*. Boston, MA: Addison-Wesley Publishing Company.

Hwang S and Foote JD (2021) Why do people participate in small online communities? *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW2): 1–25.

Jhaver S, Bruckman A and Gilbert E (2019) Does transparency in moderation really matter? User behavior after content removal explanations on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1–27.

Jhaver S, Frey S and Zhang AX (2023) Decentralizing platform power: a design space of multi-level governance in online social platforms. *Social Media + Society* 9(4): 20563051231207857.

Jiang JA, Kiene C, Middler S, et al. (2019) Moderation challenges in voice-based online communities on discord. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1–23.

Juneja P, Rama Subramanian D and Mitra T (2020) Through the looking glass: study of transparency in Reddit's moderation practices. *Proceedings of the ACM on Human-Computer Interaction* 4: 1–35.

Kelly C and Breinlinger S (1995) Identity and injustice: exploring women's participation in collective action. *Journal of Community & Applied Social Psychology* 5(1): 41–57.

Kerr EB (1986) IEEE transactions on professional communication. *IEEE Transactions on Professional Communication* 29(1): 12–18.

Kiesler S, Kraut R, Resnick P, et al. (2011) Regulating behavior in online communities. In: Kraut R and Resnick P (eds) *Building Successful Online Communities: Evidence-Based Social Design*. Cambridge, MA: MIT Press, pp. 125–177.

Kim AJ (2006) *Community Building on the Web: Secret Strategies for Successful Online Communities*. Berkeley, CA: Peachpit Press.

Kollock P and Smith M (1996) Managing the virtual commons: cooperation and conflict in computer communities. In: Herring SC (ed.) *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*. Philadelphia, PA: John Benjamins, pp. 109–128.

Larson L and DeChurch L (2020) Leading teams in the digital age: four perspectives on technology and what they mean for leading teams. *The Leadership Quarterly* 31(1): 101377.

Lazar J and Preece J (2002) Social considerations in online communities: usability, sociability, and success factors. In: Van Oostendorp H (ed.) *Cognition in the Digital World*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 1–46.

Leary MR, Kelly KM, Cottrell CA, et al. (2013) Construct validity of the need to belong scale: mapping the nomological network. *Journal of Personality Assessment* 95(6): 610–624.

Lian H, Lance Ferris D and Brown DJ (2012) Does taking the good with the bad make things worse? How abusive supervision and leader–member exchange interact to impact need satisfaction and organizational deviance. *Organizational Behavior and Human Decision Processes* 117(1): 41–52.

Lin H-F and Lee G-G (2006) Determinants of success for online communities: an empirical study. *Behaviour & Information Technology* 25(6): 479–488.

Lind EA and Tyler TR (1988) *The Social Psychology of Procedural Justice*. New York: Plenum.

Locke EA and Schweiger DM (1979) Participation in decision-making: one more look. In: Staw BM (ed.) *Research in Organizational Behavior*. London: JAI Press, pp. 265–339.

Malinen S (2015) Understanding user participation in online communities: a systematic literature review of empirical studies. *Computers in Human Behavior* 46: 228–238.

Massanari A (2017) #Gamergate and the fappening: how Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society* 19(3): 329–346.

Matias JN (2019) The civic labor of volunteer moderators online. *Social Media + Society* 5(2): 2056305119836778.

Moliner C, Martínez-Tur V, Peiró JM, et al. (2013) Perceived reciprocity and well-being at work in non-professional employees: fairness or self-interest? *Stress and Health* 29(1): 31–39.

O'Mahony S and Ferraro F (2007) The emergence of governance in an open source community. *Academy of Management Journal* 50(5): 1079–1106.

Ostrom E (1998) A behavioral approach to the rational choice theory of collective action: presidential address, American political science association, 1997. *American Political Science Review* 92(1): 1–22.

Park SM and Rainey HG (2008) Leadership and public service motivation in U.S. federal agencies. *International Public Management Journal* 11(1): 109–142.

Preece J (2000) *Online Communities: Designing Usability and Supporting Sociability*. 1st ed. Hoboken, NJ: John Wiley & Sons.

Ransbotham S and Kane GC (2011) Membership turnover and collaboration success in online communities: explaining rises and falls from grace in Wikipedia. *Mississippi Quarterly* 35(3): 613–627.

Reid E (1991) *Electropolis: electropolis: communication and community on internet relay chat*. PhD Thesis, The University of Melbourne, Melbourne, VIC, Australia. Available at: http://www.aluluei.com/electropolis.htm (accessed 19 November 2023).

Ren Y, Kraut R, Kiesler S, et al. (2011) Encouraging commitment in online communities. In: Kraut R and Resnick P (eds) *Building Successful Online Communities: Evidence-Based Social Design*. Cambridge, MA: MIT Press, pp. 100–124.

Roberts TL (1998) Are newsgroups virtual communities? In: *CHI '98: Proceedings of the SIGCHI conference on human factors in computing Systems*, Palo Alto, 18–23 April, pp. 360–367. New York: ACM.

Rosen CC, Ferris DL, Brown DJ, et al. (2014) Perceptions of organizational politics: a need satisfaction paradigm. *Organization Science* 25(4): 1026–1055.

Ryan RM (1982) Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *Journal of Personality and Social Psychology* 43(3): 450–461.

Ryan RM and Deci EL (2006) Self-regulation and the problem of human autonomy: does psychology need choice, self-determination, and will? *Journal of Personality* 74(6): 1557–1585.

Schneider N (2022) Admins, mods, and benevolent dictators for life: the implicit feudalism of online communities. *New Media & Society* 24(9): 1965–1985.

Schöpke-Gonzalez AM, Atreja S, Shin HN, et al. (2022) Why do volunteer content moderators quit? Burnout, conflict, and harmful behaviors. *New Media & Society*. Epub ahead of print 4 December. DOI: 10.1177/14614448221138529.

Seering J (2020) Reconsidering self-moderation: the role of research in supporting community-based models for online content moderation. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW2): 1–28.

Seering J and Kairam SR (2023) Who moderates on twitch and what do they do? Quantifying practices in community moderation on twitch. *Proceedings of the ACM on Human-Computer Interaction* 7(CSCW): 1–18.

Seering J, Kaufman G and Chancellor S (2022) Metaphors in moderation. *New Media & Society* 24(3): 621–640.

Seering J, Wang T, Yoon J, et al. (2019) Moderator engagement and community development in the age of algorithms. *New Media & Society* 21(7): 1417–1443.

Sheldon KM and Houser-Marko L (2001) Self-concordance, goal attainment, and the pursuit of happiness: can there be an upward spiral? *Journal of Personality and Social Psychology* 80(1): 152–165.

Sheldon KM, Elliot AJ, Kim Y, et al. (2001) What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology* 80(2): 325–339.

Sheldon KM, Zhaoyang R and Williams MJ (2013) Psychological need-satisfaction, and basketball performance. *Psychology of Sport and Exercise* 14(5): 675–681.

Squirrell T (2019) Platform dialectics: the relationships between volunteer moderators and end users on Reddit. *New Media & Society* 21(9): 1910–1927.

Srinivasan KB, Danescu-Niculescu-Mizil C, Lee L, et al. (2019) Content removal as a moderation strategy: compliance and other outcomes in the changemyview community. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1–21.

Strimling P and Frey S (2020) Emergent cultural differences in online communities' norms of fairness. *Games and Culture* 15(4): 394–410.

Tausczik YR, Dabbish LA and Kraut RE (2014) Building loyalty to online communities through bond and identity-based attachment to sub-groups. In: *CSCW '14: Proceedings of the 17th*

ACM conference on computer supported cooperative work & social computing, Baltimore, MD, 15–19 February, pp. 146–157. New York: ACM.

Teo H-H, Chan H-C, Wei K-K, et al. (2003) Evaluating information accessibility and community adaptivity features for sustaining virtual learning communities. *International Journal of Human-Computer Studies* 59(5): 671–697.

Thach H, Mayworm S, Delmonaco D, et al. (2022) (In)visible moderation: a digital ethnography of marginalized users and content moderation on Twitch and Reddit. *New Media & Society*. Epub ahead of print 18 July. DOI: 10.1177/14614448221109804.

Topinka RJ (2018) Politically incorrect participatory media: racist nationalism on r/ImGoingToHellForThis. *New Media & Society* 20(5): 2050–2069.

Trujillo A and Cresci S (2022) Make Reddit great again: assessing community effects of moderation interventions on r/The_Donald. *Proceedings of the ACM on Human-Computer Interaction* 6(CSCW2): 1–28.

Tyler TR and Degoey P (1995) Collective restraint in social dilemmas: procedural justice and social identification effects on support for authorities. *Journal of Personality and Social Psychology* 69(3): 482–497.

Van Prooijen J-W (2009) Procedural justice as autonomy regulation. *Journal of Personality and Social Psychology* 96(6): 1166–1180.

Van Vugt M (2009) Averting the tragedy of the commons: using social psychological science to protect the environment. *Current Directions in Psychological Science* 18(3): 169–173.

Walters GD and Bolger PC (2019) Procedural justice perceptions, legitimacy beliefs, and compliance with the law: a meta-analysis. *Journal of Experimental Criminology* 15(3): 341–372.

Wang HM, Bulat B, Fujimoto S, et al. (2022) Governing for free: rule process effects on Reddit moderator motivations. In: Stephanidis C, Antona M, Ntoa S, et al. (eds) *HCI International 2022—Late Breaking Posters*. Cham: Springer, pp. 97–105.

Weber S (2005) *The Success of Open Source*. Cambridge, MA: Harvard University Press.

Weld G, Zhang AX and Althoff T (2022) What makes online communities 'better'? Measuring values, consensus, and conflict across thousands of subreddits. In: *Proceedings of the sixteenth international AAAI conference on web and social media*. Available at: https://cdn.aaai.org/ojs/19363/19363-28-23376-1-2-20220531.pdf

Wohn DY (2019) Volunteer moderators in twitch micro communities: how they get involved, the roles they play, and the emotional labor they experience. In: *CHI '19: Proceedings of the 2019 CHI conference on human factors in computing systems*, Glasgow, 4–9 May, pp. 1–13. New York: ACM.

Zhang AX, Hugh G and Bernstein MS (2020) PolicyKit: building governance in online communities. In: *UIST '20: Proceedings of the 33rd annual ACM symposium on user interface software and technology*, Virtual Event, 20–23 October, pp. 365–378. New York: ACM.

Zhao L, Lu Y, Wang B, et al. (2012) Cultivating the sense of belonging and motivating user participation in virtual communities: a social capital perspective. *International Journal of Information Management* 32(6): 574–588.

Author biographies

Beril Bulat is a PhD candidate in the Department of Communication and a member of the Computational Communication Lab at the University of California, Davis. She specializes in governance and control in online communities, alongside the impact of computational propaganda on online political discourse.

Hannah Wang received her degree in Communication from UC Davis, where she conducted research focusing on online community moderator motivations. Following this, she obtained her

MS in Marketing Intelligence from the University of San Francisco and now works as a market research consultant specializing in healthcare.

Stephen Fujimoto is a graduate of University of California, Davis with a degree in Cognitive Science and Statistics.

Seth Frey is a Professor in Communication at the University of California, Davis and an affiliate of the Ostrom Workshop at Indiana University. His expertise is in computational approaches to self-governance and the cognitive science of strategic behaviour.