The Impact of Governance Bots on Sense of Virtual Community: Development and Validation of the GOV-BOTs Scale

C. ESTELLE SMITH, University of Colorado Boulder, USA
IRFANUL ALAM, University of Colorado Boulder, USA
CHENHAO TAN, University of Chicago, USA
BRIAN C. KEEGAN, University of Colorado Boulder, USA
ANITA L. BLANCHARD, University of North Carolina Charlotte, USA

Bots are increasingly being used for governance-related purposes in online communities, yet no instrumentation exists for measuring how users assess their beneficial or detrimental impacts. In order to support future human-centered and community-based research, we developed a new scale called GOVernance Bots in Online communiTies (GOV-BOTs) across two rounds of surveys on Reddit (N = 820). We applied rigorous psychometric criteria to demonstrate the validity of GOV-BOTs, which contains two subscales: *bot governance* (4 items) and *bot tensions* (3 items). Whereas humans have historically expected communities to be composed entirely of humans, the social participation of bots as *non*-human agents now raises fundamental questions about psychological, philosophical, and ethical implications. Addressing psychological impacts, our data show that perceptions of effective *bot governance* positively contribute to users' *sense of virtual community* (SOVC), whereas perceived *bot tensions* may only impact SOVC if users are more aware of bots. Finally, we show that users tend to experience the greatest SOVC across groups of subreddits, rather than individual subreddits, suggesting that future research should carefully re-consider uses and operationalizations of the term "community."

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Reddit, bots, online community, sense of virtual community, human-bot interaction, community-bot interaction, moderation, governance, psychometric scale, confirmatory factor analysis, structural equation modeling

ACM Reference Format:
C. Estelle Smith, Irfanul Alam, Chenhao Tan, Brian C. Keegan, and Anita L. Blanchard. 2022. The Impact of Governance Bots on Sense of Virtual Community: Development and Validation of the GOV-BOTs Scale. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 462 (November 2022), 30 pages. https://doi.org/10.1145/3555563

INTRODUCTION

"Bots" are automated software programs that take advantage of application program interfaces (APIs) to perform a wide variety of behaviors across online platforms. Bots are becoming increasingly common in platforms that host online communities such as Reddit [16, 42, 57], Wikipedia [30,

Authors' addresses: C. Estelle Smith, c.estelle.smith@colorado.edu, University of Colorado Boulder, Boulder, Colorado, USA, 80304, c.estelle.smith@colorado.edu; Irfanul Alam, University of Colorado Boulder, Boulder, Colorado, USA, 80304, irfanul.alam@colorado.edu; Chenhao Tan, University of Chicago, Chicago, Illinois, USA, 80304, chenhao@uchicago.edu; Brian C. Keegan, University of Colorado Boulder, Boulder, Colorado, USA, 80304, brian.keegan@colorado.edu; Anita L. Blanchard, University of North Carolina Charlotte, Charlotte, North Carolina, USA, 28223, anita.blanchard@uncc.edu.

2573-0142/2022/11-ART462
https://doi.org/10.1145/3555563

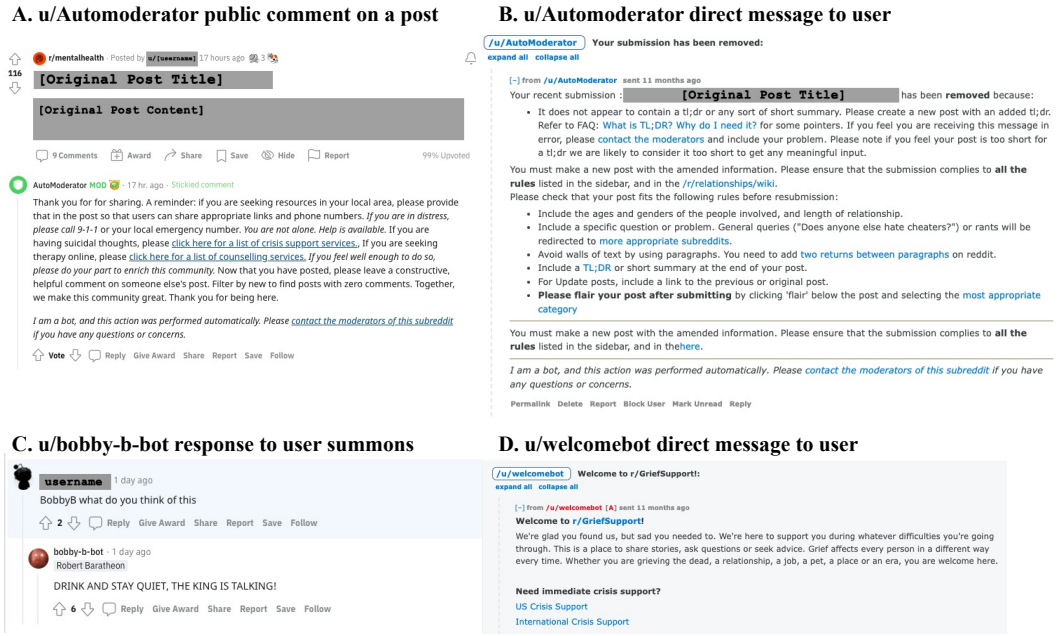


Fig. 1. Example screenshots of Reddit governance bot activities. u/Automoderator is a prolific bot available to all moderators. A. In r/mentalhealth, u/Automoderator leaves a public comment on every first-time post to provide resources. B. In r/relationships, u/Automoderator sends a direct message to inform a user why their post was removed. C. On any sub, users can playfully summon u/bobby-b-bot to leave a public comment containing a random quotation from Game of Thrones character, Robert Baratheon. D. In r/GriefSupport, u/welcomebot sends a direct message to support new users after joining the sub.

31, 97, 116], Facebook [70, 106], Twitter [62, 93, 104], and Twitch [91]. Whether through "hidden" commercial moderation strategies [82] or volunteer community self-moderation [65, 90], moderators of online communities are increasingly likely to adopt bots as semi- or fully-automated mechanisms for coping effectively with scale as online communities grow in size [46, 47, 97].

As Steve Jones wrote, *"The question is no longer whether bots can pass [as human], but how social interaction with them may be meaningful."* [44, p. 1] Including and beyond moderation, bots fill many social roles in communities, such as prompting play and games [64], improving content quality, driving interaction, and maintaining the community [57]. In this paper, we refer broadly to **governance bots** as bots that can assist moderators and users of online communities through activities such as posting, commenting, sending direct messages, removing content, or banning users. We selected Reddit as our study site due to the availability of public data on millions of unique online spaces called subreddits (or "subs"), the diversity of *types* of subs that exist, the increasing presence of governance bots present in subs, and growing research interest in Reddit [75] and human-bot interaction, moderation, and online community governance [37, 65, 88, 91, 97]. Figure 1 provides example screenshots of public and private bot activities across several arbitrary subs. The prevalence of such bots varies across subs: some subs configure one or more bots to perform specific functions there, whereas others do not use any sub-specific bots. Moreover, some bots are active *across* Reddit–thus, even if a given sub has not directly implemented bots, bots may nonetheless appear there from time to time.

Bots can impact sociological processes in important and meaningful ways [12]. Lebeuf suggests that bots should be classified in terms of: (1) environments in which they operate; (2) intrinsic properties (*i.e.,* technical features built by designers); and (3) interactions within the environment [54]. For the purposes of this paper, bots' environments are subreddits. Many papers, as well as ubiquitous website copy on Reddit.com, refer to subs as "communities" (*e.g.*, [16, 99]), whereas others use terms like "content feeds" [39] or "discussion fora" [57]. Clearly, there exist important differences in the language people use to describe subs, whereas high variance in sub features and types suggests that they may not all function equivalently as "community" spaces. Moreover, definitions of community have been debated for decades. Some definitions focus on *structural* aspects of what constitutes a community, whereas others focus on people's psychological *sense* of community (SOC)—*i.e.,* the perception of a particular geographical or relational group existing as a community [17, 21, 35, 85]. Structurally, most HCI research posits individual subs as community units—a debatable but reasonable assumption that our work also begins from. However, in order to highlight how the intrinsic properties and interactions of governance bots can impact users' experiences within specific subs, we focus on users' experienced *sense of virtual community* (SOVC) [8–10, 51] in this paper.

Adopting a human-centered design perspective, it is important to evaluate how governance bots are impacting users' SOVC. A growing body of work seeks to *quantitatively* detect bots and their behaviors [31, 39, 63] or to *qualitatively* examine attitudes towards bots in online communities [19, 42, 57]. These studies provide excellent insights on externally observable impacts of bots on human behavior, and on attitudes of small user groups towards them. However, no instrumentation currently exists that allows for precise measurement of how online community users internally assess the impacts of these bots on their communities. **Without such a tool, it is not possible to measure and compare how users perceive beneficial or harmful impacts of governance bots within their communities.** Thus our work has the dual goals of: (1) understanding how users experience SOVC on Reddit, and how bots impact that, and (2) increasing the capability of users, moderators, and researchers to assess user perspectives on bots in online communities. A scientifically validated tool can help researchers to expand the types of questions that can be answered through survey-based study designs, and bot designers to build and evaluate bots that better serve community needs. We use the following research questions to serve these goals:

RQ1: How can we measure users' cognitive assessments of bots in subreddits?

We address RQ1 by developing a new psychometric scale called GOVernance Bots in Online communiTies (GOV-BOTs) that measures two constructs: (1) *bot governance* and (2) *bot tensions.* Bot governance refers to users' perceptions of the beneficial impacts of governance bots within the subreddit, for example, by providing resources, generating positive interactions, or removing violating content. Bot tensions refers to users' perceptions of how bots can cause conflicts or problems in the subreddit, such as manipulating, upsetting, or startling users.

Psychology researchers argue that SOVC is an independent construct that can be influenced by a variety of experiential *antecedents*, and that certain measurable *outcomes* should be observed if SOVC does exist [9]. Therefore, along with our novel GOV-BOTs scale, we also selected previously validated instruments to measure SOVC, two well-established antecedents, and one known outcome of SOVC. Using these scales, we deployed two phases of surveys on Reddit. Following data collection, we used the sibling methods of confirmatory factor analysis (CFA) and structural equation modeling (SEM) to demonstrate that: (1) both new and pre-existing scale items load appropriately onto the correct factors, and (2) the measured constructs relate to each other in theoretically expected ways. CFA and SEM confirmed that GOV-BOTs meets adequate psychometric and theoretical criteria, so we conclude that it is a valid instrument.

RQ2: How do bots impact users' sense of virtual community in subreddits?

The experience of SOC/SOVC is a *cognitive state* that leads to positive outcomes, both for individual community members, as well as the community overall. For instance, SOC leads to better conflict resolution, continued participation in, and higher commitment to the community [10, 11, 61, 113]. Because of primal human needs for connection and community, it is vitally important to understand how we can sustain SOC/SOVC, both online and offline. Yet unlike offline communities, virtual environments have created a new mix of possible community "members." Whereas communities have historically been composed entirely of humans (plus, some would argue, their animal companions), these *non*-human bot technologies are not only present, but are also participating socially [92]. *How does this inclusion of novel social technology evolve our perceptions and experiences of community? And what are the philosophical, psychological, and ethical implications and repercussions?* These are crucial questions without simple answers, yet as online communities continue to grow and permeate society, it becomes more and more important to consider them conscientiously.

No prior research has directly investigated the relationship between bots and SOVC, yet bots can have complex and mixed impacts on communities—sometimes for the worse, and sometimes for the better. Consequently, addressing the *psychological* implications of bots in online communities, we predict that users' assessments of bot governance and bot tensions should function as antecedents to SOVC. For instance, bots can impact information spread [94] and manipulate online discourse [24], with case studies demonstrating that bots can cause psychological, legal, economic, social, or democratic harms [19]. These concerns led to our first hypothesis. ***H1: Higher ratings of bot tensions will negatively contribute to SOVC.*** A growing body of research also suggests that when bots are carefully built, designed, and evaluated with community values in mind, they can have beneficial community impacts [97, 119]. Consequently, attitudes are shifting toward a view of bots as contributors of useful, necessary, or even playful and creative functions [12, 64], leading to our second hypothesis. ***H2: Higher ratings of bot governance will positively contribute to SOVC.*** While these hypotheses only scrape the surface of the major questions posed earlier, they do provide preliminary knowledge to support future inquiry, as we discuss at the end of the paper.

SEM addresses RQ2 by confirming that bot governance and bot tensions do function as antecedents to SOVC. We find evidence in support of H2—positive assessments of bot governance lead to increased SOVC. H1 is more complex; bot tensions were not observed to significantly impact SOVC in our full sample, however our analysis also revealed that users' awareness of bots, and their self-reported membership in the sub, may act as moderating variables.

RQ3: How do users' experiences of SOVC vary across Reddit and other platforms?

Whereas much research (including this paper) begins from an assumption that individual subreddits on Reddit comprise the structural units of community, we wanted to understand to what extent this aligns with users' experienced SOVC. We find that most Reddit users experience a greater SOVC across certain *groups* of subs, and that platforms are perceived as community spaces to differing degrees, suggesting future methodological directions. Overall, this paper contributes GOV-BOTs as a new research tool (RQ1), and empirical results that advance our understanding of how bots impact online communities (RQ2) and how users experience communities differently across Reddit (RQ3). We provide describe prior work that informs this research approach.

RELATED LITERATURE

We begin by briefly summarizing decades of theory from organizational and community psychology on psychometric instrumentation, *sense of community* (SOC), and *sense of virtual community* (SOVC). Next, we describe governance bots, and how they can impact users' community experiences. Finally, we provide more detail on Reddit bots.

Psychometrics

Psychometric scales are survey-based instruments that provide a gold standard for measuring people's internal psychological states or experiences. For each scale, researchers define specific psychological "constructs" or "factors" that relate to discrete cognitive concepts. Each construct is comprised of a list of "items" – short, colloquial phrases about attitudes or experiences that can be rated on a Likert scale.

Developing and validating psychometric scales. Psychometricians use *factor analyses* to validate that a scale accurately measures its intended target [38]. When a list of items exists, but it is unknown which items relate to which factors, *exploratory factor analysis* can be used to suggest item-factor pairings. However, our work uses previously validated scales in which all item-factor relationships are known, and we create new constructs with items derived from prior work. Thus, we instead use *confirmatory factor analysis* (CFA) to demonstrate that items load correctly onto pre-specified factors. After establishing appropriate item-factor loading, *structural equation modeling* (SEM) can be used to show how constructs have pathways of influence on each other.

Psychometric scale development in HCI. Along with prior scales relevant to CSCW (*e.g.,* self-efficacy in Human-Robot- Interaction [77] or secure software [105], user reflection while using technology [6], digital literacy [83], or gamers' motivations [114] and experiences [103]), this scale development work falls within the purview of human-computer interaction because it advances knowledge and methods for studying bots and online communities. While the goals of this paper relate to SOVC, our work is also motivated by a need to help address other open questions from a *user-centered* perspective. Combining measures of internal assessments of bots with externally observable behaviors will improve future researchers' ability to answer questions like: How do users' perceptions of governance bots impact their survival time [116], the prevalence of negative (*e.g.,* trolling/abuse/harassment) or positive behaviors (*e.g.,* social support, mentoring) [4], or their attitudes toward moderation or platforms [41]? How well do users feel that bots are serving their intended functions? How do qualitative insights about bots derived from small populations scale to larger populations? For such questions, GOV-BOTs can provide pragmatic insights.

Selection of psychometric constructs for GOV-BOTs. Bots may potentially function as either stressors or boons to online communities. Therefore, we modeled our selection of GOV-BOTs constructs after classical psychology studies on stress [28, 53]. In the stress process, a person must first "notice" a stressor and next "appraise" it as such; if someone neither sees the stressor, nor thinks it's a problem, then they do not become stressed. Researchers can measure the degree to which people have first noticed, and then appraised, stressors through psychometric instrumentation that uses carefully validated item text to *separately* assess both constructs. Analogously, although bots may be present in any subreddit, users may or may not know that they are there. If users do *not* notice bots, they might not appraise them any particular way. If they *do* notice bots, then we can separately measure the degrees to which they appraise bots as either beneficial or detrimental to the community. Thus, for the development of our new scale, we defined the following three constructs of interest:

- **Bot Awareness (BA):** The degree to which users are aware of bot activities in the subreddit.
- **Bot Governance (BG):** How users evaluate beneficial impacts of governance bots within the subreddit.
- **Bot Tensions (BT):** The degree to which governance bots introduce conflicts or problems in users' experience in the subreddit.

In Methods, we provide more detailed information about how we derived items for these constructs based on related literature that will be presented later in this section.

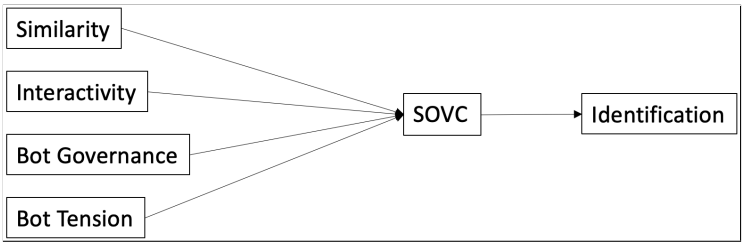


Fig. 2. Theoretical nomological network with antecedents (Similarity, Interactivity, Bot Governance, Bot Tensions), Sense of Virtual Community (SOVC), and one outcome of SOVC (Identification).

Theories and Measures of Sense of Community

Historical perspective on sense of community. Theories and measures of sense of community (SOC) have been debated for decades. Communities can be geographical (neighborhoods, cities, etc.) and/or relational (pertaining to characteristics of human relationships, unrelated to location) [35]. Early SOC scales focused on geographical communities [21, 33]; later scales measured SOC more generally. Most notably, McMillan and Chavis' SOC theory [67] suggests four factors of importance (membership, influence, fulfillment of needs, and shared emotional connection). Despite its popularity, this theory (and its original scale [17]) has not been empirically validated. Shifting to the nuances of online communities, researchers in the 2000s developed theory [10] and scales [8, 51] on sense of *virtual* community (SOVC). However, these scales tend to conflate psychological *antecedents* or *outcomes* with SOVC itself, making them less interpretable.

A modern cognitive and category-based approach to sense of virtual community. As Seymour Sarason ("father of community psychology") wrote of SOC in 1974, *"you know when you have it and when you don't."* [85, p. 157] In other words, people make intuitive cognitive assessments about their community experiences, or lack thereof. In 2006, Bruckman argued that HCI scholars should adopt a *category theory* perspective from cognitive science when thinking about how people experience online communities [13]. Modern approaches in community and organizational psychology also adopt a more cognitive and category-based perspective. Notably, Blanchard, *et al.* (2020) define entitativity as *"an individual's cognitive assessment of a social unit as a group"* and introduce a short 3-item scale to measure it [9]. In this paper, we adapted the entitativity measure by substituting the word "community" for "group" to capture users' cognitive assessment of "sense of virtual community" (SOVC) (see Table 1).

Building on extensive prior theory, Blanchard, *et al.* also distinguish antecedents (similarity, common goals, interactivity, boundaries, shared history of interactions) and outcomes (identification with the group, group cohesion) from entitativity itself, and confirm theoretical relationships between these constructs [9]. Rather than earlier scales which *conflated* antecedents and outcomes with SOVC (*e.g.,* [8, 51]), this modeling approach enables us to test and confirm well-established community theories and concepts relative to our new measures of governance bots.

In particular, the goal of our present study is *not* to comprehensively assess all possible antecedents or outcomes of SOVC, but rather to develop GOV-BOTs and ensure that the constructs it measures perform in a theoretically appropriate manner. To do this, we require deployment of GOV-BOTs along with a minimal set of previously validated instruments. That is, if we can replicate prior well-established theoretical relationships using prior instruments, *and* constructs in GOV-BOTs also perform in a theoretically consistent way, then we can more confidently make valid claims about the constructs in GOV-BOTs [14]. We selected *similarity* and *interactivity* [9] as

Existing Scales for Community-Related Constructs		
Sense of Virtual Community	SOVC1	We are a unit.
	SOVC2	We are a community.
	SOVC3	We feel like a community to me.
Similarity (Antecedent of SOVC)	SIM1	We are alike.
	SIM2	We have similar attitudes.
	SIM3	We have similar values.
	SIM4	We see things much in the same way.
Interactivity (Antecedent of SOVC)	INT1	We respond to each other's messages.
	INT2†	We interact with each other in these messages.
	INT3	We communicate with each other.
	INT4	We spend time interacting.
	INT5‡	We build on each other's thoughts and ideas.
Identification (Outcome of SOVC)	IDPG1†	When someone criticizes this sub, it feels like a personal insult.
	IDPG2	This sub's successes are my successes.
	IDPG3	When someone praises this sub, it feels like a personal compliment.
	IDPG4	I am very interested in what others think about this sub.
	IDPG5	When I talk about this sub, I usually say "we" rather than "they".

Table 1. Existing Psychometric Instrumentation. † indicates items removed through first round of confirmatory factor analysis. ‡ indicates items removed through second round of confirmatory factor analysis.

antecedents, and *identification with the community* [60] as an outcome, because these are the most important (or at least most studied) constructs that also have repeatedly validated psychometric instruments available [15, 55]. We define these terms as follows:

> **Similarity (SIM):** The degree to which community members view one other as similar (attitudes, values, views).
> **Interactivity (INT):** The degree to which community members interact online (responding to messages, building on ideas, interacting).
> **Identification with a Psychological Group (IDPG):** "Definition of self in terms of one's group membership" [2, 60]. In all IDPG items, we substituted the words "this sub" for "group." When people begin to feel that a group or community exists, this then leads to a sense of identity tied to it.

Table 1 contains all verbatim items as adapted from existing scales, and Figure 2 shows our predicted theoretical model of relationships between measured constructs: we expect that higher similarity and interactivity should lead to higher SOVC, and that higher SOVC should contribute to greater identification with the sub. Furthermore, we predict that users' assessments of bot governance and tensions will function as antecedents to SOVC. In the next section, we now provide more detail about governance bots, and the ways in which they can impact online communities.

Governance Bots

Governance in online communities relies upon complex interactions between users, moderators, bots, technical affordances, and platform policies. Schneider argues that early technical conditions

in online communities led to a governance style of *implicit feudalism* that grants user-administrators (*i.e.* moderators) a high degree of power and control [87]. Although moderators can abandon their role at any time (and sometimes do so due to burnout, not having the time anymore, or other factors), it is difficult for the users of their subreddits to remove or contest them. Furthermore, as online groups grow in size, moderators rely more upon automated tools due to the impossibility of managing an increasing volume of content through purely human moderation [7, 32, 46, 82, 97]. For example, bots on Reddit [16, 42] and Wikipedia [30, 36, 97] are often used in a moderation capacity to evaluate and remove content or users that violate community norms, rules, or policies. `u/Automoderator` is an especially pervasive bot available to all Reddit moderators that provides the ability to enforce customizable, syntactic rules in the subreddit [42] and which can reduce abusive language use [115]. However, bots perform many more actions of social and community relevance beyond moderation and rule enforcement. For example, in addition to moderation, bots on Twitch contribute content related to user engagement, information sharing, mini-games, or promotion/advertising of specific users [91]. In one charming example, Seering et al. deployed `BabyBot` in a gaming community on Twitch, so that the community could raise it and teach it effective discourse. The process of caring for `BabyBot`—eventually renamed `PeteBot` when it "grew up"—strengthened community engagement and identity in the process [92]. On Wikipedia, bots are also important, completing prolific editing activities across the platform in roles such as fixers, generators, taggers, and more [116].

Because of the diversity of bot activities present across numerous platforms, we refer to **governance bots** broadly as bots that can assist moderators and users through activities such as posting, commenting, sending direct messages, removing content, or banning users. The question copy in our GOV-BOTs scale instructs participants to consider these types of governance bots rather than, for example, fake user accounts masquerading as real people, or conversational one-on-one chatbots, such as those that are common on commercial websites for assistance with product queries, or `ELIZA`, a famous chatbot from 1966 that mimics a Rogerian psychotherapist [110]. Unlike malicious bots impersonating real humans, our work assumes that governance bots are generally designed with pro-community intent. For instance, it is certainly the case that some online communities are toxic and hateful; people could thus design their governance bots in accordance with toxic and hateful norms, yet those bots would still be intended to support the community. Or, a bot may be designed with pro-community intent in pro-social communities, but if the bot is badly designed, it can still nonetheless be damaging. One example is `Patrobot`, a bot deployed on Spanish Wikipedia for the pro-community purpose of automatically reverting bad edits to the encyclopedia [97]. Unfortunately, `Patrobot`'s default prediction model was configured with low precision (25%), and was therefore wrong 75% of the time. For weeks, no one could turn off the bot, so it wreaked havoc on the community's ability to grow their content. Regardless of whether bots are, themselves, "good" or "bad", GOV-BOTs provides a way to assess users' psychological assessments of their beneficial or detrimental impacts.

Researchers have suggested guidelines for designing bots in communities, although the degree to which bot builders follow them is unknown. For example, Marechal suggests that normative frameworks for using bots across a variety of platforms should include: obvious labeling of bots, contacting users only with consent, and using information according to disclosed purposes [62]. Smith *et al.* present five community-derived values for Wikipedia AIs and bots: reducing community maintenance efforts, maintaining human judgement as the final authority, supporting differing workflows, encouraging positive interaction, and establishing trustworthiness of people and algorithms [97]. On Reddit, bot designers can choose to follow crowdsourced "Botiquette" such as including "bot" in the account username, limiting number of replies per post or usage of a given phrase, or listing a contact account for the bot's creator [1].

The Diversity of Governance Bots on Reddit. Users request bots on Reddit for goals related to administration, community maintenance, improving content quality, archiving content [57], play and humor [64], or political discourse [27, 39, 115]. In order to achieve these goals, bots use technical features such as posting, commenting, private messaging, being "called" by users, or creating and responding to queries [56, 57]; see Figure 1 for example screenshots of bot activities in public posts and private direct messages. In addition to the famous example of `u/Automoderator`, examples of specific instances include bots that can detect and remove other bots (`u/Botwatchman`), modify text (*e.g.*, improve grammar, provide summaries (`u/autotldr`), translate between languages or writing styles (`u/tyo-`), link to and modify external sources (`u/JiffyBot`) [56, 57], or play interactive political games (`u/TrumpTrainBot`) [27]. These examples represent only a tiny fraction of the diversity of bots on Reddit.

Community Impacts of Bots on Reddit

Bots contribute to enforcement of norms. Norms vary widely from sub to sub, if they are present at all [26]. Even within a given sub, enforcement of norms can be opaque and inconsistent, especially when moderators operate upon differing interpretations or operationalizations of rules. Many users report frustrations regarding lack of clarity or communication about post removal, unjust moderation, and difficulty complying with rules [41]. However, bots can make norms more visible and enforce them more consistently. For example, when users receive a highly informative explanation of a moderation decision from a bot rather a human, they are more likely to post again successfully in the future [43], and over 50% of the time, users report receiving such explanations from bots [41]. After moderators adopt `u/Automoderator`, new content contributions to their sub increase [37], possibly because a better regulated community environment is safer and more inviting [66].

Bots change the nature of moderation work. At the same time, using `u/Automoderator` changes the nature of moderators' work and relationship with the community. It creates new time-consuming technical tasks and overhead [42], while also causing moderators to engage in fewer casual and more moderation-related interactions with users [37], likely cementing the implicit hierarchical power of moderators [87]. Moderators can even become so reliant on automated tools that they attempt to replicate these tools across platforms. For example, Reddit moderators who use associated Discord servers often use bots to make their Discord operate more similarly to Reddit [47].

Bots impact user perceptions and behaviors. The prevalence of bots in social media has led researchers to develop automated or semi-automated techniques for detecting them [24]. Along the lines of Turing Tests, research has also explored factors that impact users' "perceived humanness" of bots. For example, grammatical mistakes (*e.g.*, presence of typos) cause users to perceive that messages were more likely written by bots than humans [112]. When uncivil comments are moderated by bots rather than humans, participants may perceive news content to be less biased [108]. Beyond moderation-related impacts such as removing abusive or norm-violating content or users [16, 37, 115], bots can also have meaningful impacts on user behavior like language use and activity levels. For example, a case study of `u/bobby-b-bot` (Fig. 1) shows that the sentiment of bots' comments positively impacts subsequent comment sentiment, and that human language then mirrors the bot's (*lexical entrainment*, a phenomenon that also occurs in-person conversation) [58]. Wang *et al.* deployed a bot in an online mental health support community that left comments on posts which did not organically receive comments from human users within 10 minutes; the bot's activities increased levels of human participation across the whole community [107]. Given powerful social impacts like these, Seering suggests that *"there remains much work to be done in analysis of bot roles in social spaces*" [91]—a line of work we contribute to in this study.

THREE-PART METHODOLOGICAL APPROACH

In this section, we describe the methods and results of the three parts of our study. First, we designed a survey and deployed it on Reddit in two phases. Second, we conducted confirmatory factor analysis (CFA) in order to ensure appropriate item-factor loadings in both phases. Third, we used structural equation modeling (SEM) to understand relationships between constructs. This study was reviewed and deemed exempt by our university's IRB.

Part I: Survey Design

In related literature (see section titled *"Theories and Measures of Sense of Community"*), we described our theory-informed selection of validated scales from [9]. We reiterate that the scale items displayed in Table 1 are not novel, but have been all been directly adapted from well-validated prior instruments. On the other hand, the GOV-BOTs scale (see Table 2), including its two subscale constructs of *bot governance* and *bot tensions*, is a novel contribution of this work. Here, we now describe our item generation for these two new constructs, as well as the design and deployment of our full survey in two phases.

Development of GOV-BOTs Scale. In order to measure the three prospective constructs of *Bot Awareness*, *Bot Governance*, and *Bot Tensions*, we began by exploring the literature on users' perceptions of algorithms [20, 23, 34, 97] and bots [19, 42, 57]. Psychometric instruments require highly specific constructs, therefore we decided to limit the language of our instrument to bots because bots are often visible to users, whereas algorithms are usually invisible. Literature on perceptions of algorithms nonetheless remains relevant, since bots can rely on user-specified algorithms. Furthermore, unlike platform-owned algorithms, bots are typically configurable by community members. Thus, providing a measurement tool to assess bots (rather than algorithms) provides more agency for users who stand to benefit from its use. For example, consider the *bot governance* construct. If this construct receives high scores, mods or developers can empirically demonstrate beneficial community impacts of bots, whereas low scores can indicate what might be preventing those benefits from occurring–*e.g.,* item BG5, "bots help this sub to be a safer place." If bots make users feel unsafe, bot developers or researchers then know to investigate how or why that might be the case. Similarly, high or low scores of *bot tensions* highlight negative or unwanted impacts, and provide actionable pointers toward why.

To generate item text, we began by extracting a list of over 200 seed words and phrases directly from prior literature, and organized them according to our three constructs. For example, the seed phrases "help users understand the social norms of the community" and "enforce compliance with existing guidelines" from [42] relate to *bot governance*. Based on these seeds, we then brainstormed and discussed several dozen possible item wordings per construct before narrowing down to the 13-item initial scale in Table 2. Drawing upon substantial prior expertise in psychometric instrument development, we chose phrases that aligned with the literature, adapted these phrases to the lay vernacular of Reddit, and ensured that all major concepts from our seed list were reflected in at least one or two items. Moreover, we did not use any negations in our item text (e.g., "bots **do not** help this sub to be a safer place") because psychometricians have discovered that negatively worded items tend to load together simply because they are negative, rather than because they refer to the same underlying construct [89].

In our pilot, we used a conditional forked structure for GOV-BOTs. We first presented items for *bot awareness* (BA1-3). If a user's total score for BA indicated awareness of bots in the sub, *bot governance* (BG1-7) and *bot tensions* (BT1-3) items were delivered verbatim as in Table 2. If a user's score indicated that they were *not* aware, BG and BT item text was modified to describe

GOVernance Bots in Online communiTies (GOV-BOTs)		
Verbatim Question Text: On some subreddits, moderators use bots to help remind users of rules, resources, or weekly discussions by making posts, comments,or DMs, or to remove disallowed content or malicious users. The following statements are only referring to these kinds of bots that are intended to help keep subreddits running smoothly. (These questions do not refer to "fake" automated accounts or chatbots, for example.) In r/[subreddit name], I believe that...		
Bot Awareness†† (Antecedent of SOVC)	BA1††	I am aware of bots in this sub.
	BA2††	I have noticed that bots are active in this sub.
	BA3††	I understand what bots do in this sub.
Bot Governance (Antecedent of SOVC)	BG1†	bots help users to follow the rules in this sub.
	BG2†	bots help moderators manage the large amount of content posted in this sub.
	BG3	bots provide resources that are relevant to this sub.
	BG4	bots help to generate positive interactions between users.
	BG5	bots help this sub to be a safer place.
	BG6‡	bots help to keep trolls, spammers, and a**holes off this sub.
	BG7	bots help to remove content that isn't appropriate for this sub.
Bot Tensions (Antecedent of SOVC)	BT1	bots manipulate this sub for the worse.
	BT2	bots frequently upset users in this sub.
	BT3	it is startling when bots respond to issues rather than humans.

Table 2. GOV-BOTs Scale. † indicates items removed through first round of confirmatory factor analysis (CFA). ‡ indicates items removed through second round of CFA. †† indicates preliminary construct and items eliminated due to Kaiser-Meyer-Olkin's (KMO) test indicating that items were not factorable.

hypothetical impacts. For example, BT1 is "bots manipulate this sub for the worse." For users who were unaware of bots, BT1 became "bots *would* manipulate this sub for the worse."

Complete Survey Design. We built our survey using Qualtrics software [78]; the full survey is available as a Qualtrics export at https://bit.ly/gov-bots_qsf, or as a pdf at https://bit.ly/gov-bots_pdf. Participants answered four sets of questions about *one specific subreddit*: (1) prior use of the sub (length and frequency of use, roles, behaviors, *etc.*); (2) all scales in Tables 1 and 2; (3) demographic information; (4) free-response questions about how they experience SOVC on Reddit. We used a 5-point bipolar Likert scale (1: Strongly Disagree to 5: Strongly Agree) for all scale items.

Pilot Phase. We deployed a pilot survey in June 2021 on `r/samplesize`, a subreddit devoted exclusively to surveys. In order to ground responses within a particular subreddit, participants began by selecting from a list of large, popular subs[1] (or `r/samplesize`, if they had never visited other options), which allowed sufficient options without providing so many that responses would be spread too thin across a huge selection of subs. With an incentive to win one of five $50 gift cards, we collected approximately 1.2K responses. Inspection of the data revealed that a large majority of responses were fraudulent or (ironically) completed by bots. This may have resulted from two issues: (1) Because `r/samplesize` offers a convenient location for incentivized surveys, it may be a target

[1] `r/science`, `r/askscience`, `r/pics`, `r/funny`, `r/aww`, `r/explainlikeimfive`, `r/changemyview`, `r/wallstreetbets`, `r/atheism`, `r/politics`, `r/worldnews`, `r/offmychest`, `r/ADHD`, `r/depression`, `r/samplesize`

for users looking to game the system; (2) We used a high incentive ($50 gift card drawing) to attract participation, which may have attracted dishonest users. To ensure that we included only high quality data, we eliminated low quality responses based on: Qualtrics' built-in recaptcha scores and fraud flags; overly rapid completion of key survey questions[2]; and close reading of all free-response questions by two members of the research team to check for nonsensical or copy/pasted text.

Following elimination, 250 valid responses remained. Of these, 212 (84.8%) participants indicated awareness of bots in their selected sub. We excluded the 38 "unaware" responses in our first CFA model due to the wording variation in the BG and BT items, and we did not create a second CFA model for the "unaware" due to insufficient sample size. Demographic composition of the final 212 responses is consistent with user demographics on Reddit, which trends younger, white, and male [80, 86]. See Appendix Section A for complete pilot demographics.

Main Survey Deployments. Building on insights from our pilot, we refined and deployed our main survey. First, we used CFA on our pilot data to eliminate some items (see next section for complete CFA details). Bot Awareness items were not factorable, but most respondents (84.8%) indicated that they *were* aware of bots, thus we decided that a forked survey structure may be unnecessarily complex. Consequently, all GOV-BOTs items were delivered verbatim as in Table 2, except those that were eliminated (BA1-3 and BG1-2). We replaced the 3-item Bot Awareness construct with a single multiple choice question:

- ***Simplified* Bot Awareness:** *"In r/[subreddit name], have you ever noticed bot activities such as making posts or comments, or removing users or posts?"* (Yes | No | I'm not sure)

We inserted attention checks[3] in the psychometric instruments to provide an additional mechanism for checking data quality. After reading the pilot free responses, we also added several new questions to *quantify* some of the trends discussed by respondents that could demonstrate how community experiences vary across Reddit (see Table 5):

- **Perceived membership:** *"Regardless of how you have used* `r/[subreddit name]` *in the past, do you feel that you are a member of this sub?"* (Yes | No | I'm not sure)
- **Connection to offline life:** *"How is* `r/[subreddit name]` *related to your offline life activities?"* (Multiple selection, *e.g., "This sub is related to my offline hobbies or interests"* or *"This sub has no relationship at all to my offline life."*)
- **Wellbeing**: "How does spending some time on `r/[subreddit name]` affect how you feel?" (I usually feel worse afterward | better | about the same | it really depends – sometimes worse, sometimes better)

Finally, we modified our recruitment strategy. We reduced gift card drawing value to $10, and we decided to post customized survey links directly to a variety of subreddits rather than `r/samplesize`. We contacted moderators to request collaboration to run the survey. To protect their privacy, we do not include sub names, however Table 3 contains anonymized sub descriptions, as well as selected demographic information of respondents.[4]

Survey links were either "pinned" or marked "approved by moderators" for a few days during the timeframe of October to December 2021. We collected 1,465 additional responses directly from 13 subs–an approach that improves the ecological validity of our results over our pilot. For one sub (S13), we were informed by moderators that a troll had quickly spammed the survey link to a

[2]We used Gaussian mixture models to identify bimodal distributions in response times, and eliminated responses for which key scale rating questions were completed in less than 10 seconds (approximately one SD above the first means).
[3]For example: "Please select *disagree* to this statement to ensure that you are a human."
[4]See Appendix Section B for complete demographics of main survey. De-identified data (with subreddit names, usernames, emails, and free responses removed) from the main surveys is available at https://bit.ly/GOV-BOTS_data (with data dictionary at https://bit.ly/gov-bots_pdf).

ID	Description	Size	% Female	Age*	% White	Education*	*N*
S1	posts with niche style of absurdist humor	10-50K	6.20%	18-24	75.00%	Post-bach.	16
S2	corrections of poor scientific reporting	10-50K	25.00%	25-34	75.00%	Bachelors	28
S3	discuss fundamental elements of music	100K-1M	18.70%	25-34	81.10%	Bachelors	77
S4	help with professional application materials	100K-1M	43.80%	25-34	57.40%	Bachelors	48
S5	support re: issues with pregnancy	<10k	46.20%	25-34	76.90%	Bachelors	39
S6	POC discuss race & intersectionality	10-50K	36.60%	25-34	45.60%	Bachelors	41
S7	discuss ironic or culturally appropriative matters re: race	<10k	57.10%	25-34	47.60%	Bachelors	21
S8	support for a mental health issue	50-100K	69.20%	18-24	92.30%	High school	13
S9	A US university and geographical area	10-50K	37.20%	18-24	Bachelors	46.50%	43
S10	Q&A with experts on an academic topic	1-5+M	28.00%	25-34	92.00%	Post-bach.	25
S11	screenshots of tweets from POC users on humorous or interesting insights re: race	1-5+M	33.10%	25-34	41.70%	Bachelors	250
S12	share thoughts & learnings on indigenous spiritual beliefs	50-100K	28.60%	18-24	71.40%	High school	7
S13	combat hate speech & radicalization	100K-1M	-	-	-	-	-

Table 3. Descriptions of subreddits included in Phase 2 survey deployment. * indicates columns in which *most common* response bucket is presented. See Appendix Section B for complete aggregate demographics. "POC" is an acronym for People of Color.

variety of hate-based subs; because of compromised recruitment and overwhelming toxicity, we eliminated S13 completely (217 responses) from the present analysis. We subjected responses from all other subs to similar quality checks as the pilot. Following elimination, 608 valid responses remained. (See Table 6 in Appendix for additional details of survey cleaning.)

Part II: Confirmatory Factor Analysis Overview

Our goal was to confirm the proposed item-factor relationships and achieve the most parsimonious model. Our strategy was to eliminate items that improperly loaded onto multiple scales, were too highly correlated with other items of their own scale, and to evaluate the scales' distinctiveness from other constructs in our study. Taking advantage of the fact that we knew which items should load

on which scales, we conducted confirmatory factor analysis (CFA). During this step, we examined collinearity and item uniqueness by eliminating items which correlated too highly with items of their own scale, or items that loaded on a different scale than predicted. Throughout this iterative process, theory trumped empirical findings, thus we never considered eliminating core items of each construct (*e.g.*, SOVC2, "We are a community"). This ensured that we valued the meaning of each construct over empirical quirks or fluctuations within the sample. Across both pilot and main survey phases, CFA analyses and checks were done using MPlus [69] and R/Rstudio software [100] and the `lavaan` package [84], with results reported according to recommendations from [50].

Data Distributions, Assumption Checks, and Modeling Details. Item-level univariate/multivariate normality, skewness, kurtosis and descriptive summary statistics were checked using the MVN package [52]. Additionally, univariate (data distribution and frequency histograms) and multivariate (string responses and Mahalanobis distance) outliers were also checked. Multicollinearity was examined using variance inflation factor (VIF) and inter-item correlations using the `car` [29] and `psych` [81] packages respectively. Sampling adequacy was determined using the Kaiser-Meyer-Olkin's (KMO) test from the `psych` package [81] and Inter-item-polychoric correlations using the `corrplot` package [109]. To determine final CFA models, we selected a maximum-likelihood (ML) estimator and used a combination of pattern coefficients, fit indices (chi-square (χ^2), comparative fit index (CFI), root-mean square error of approximation (RMSEA) and standardized root mean squared residual (SRMR), modification indices and inter-item correlation residuals. We imposed a cutoff value of 0.5 when deciding items to keep or eliminate. Internal consistency (*i.e.*, Cronbach's alpha) was checked using the `psych` package for each subscale. We used current best practices for determining appropriate fit of our CFA to the data: RMSEA < 0.06, CFI > 0.95, SRMR < 0.08 [111].

Pilot CFA Results. We first conducted CFA on all valid responses from the pilot ($N = 212$). The KMO test indicated that the 3 items in the Bot Awareness (BOTA) construct were not factorable (<0.6) and had poor reliability ($\alpha = 0.33$), resulting in its elimination prior to factor analysis. All other items had skewness and kurtosis below |2.0|, but Mardia's test showed significant multivariate skewness and kurtosis, indicating non-normality overall. VIF was less than 10 and inter-item correlations ranged from -0.25 to 0.64.

Initial fit statistics for the 27-item, 6 factor model were: $\chi^2 = 640$, $p < 0.05$; RMSEA = 0.067, 90% CI [0.059, 0.075]; CFI = 0.852, SRMR = 0.073. In addition to fit indices, we examined modification indices to diagnose possibilities of model misfit, which suggested removal of three items. Items BG1 ("bots help users to follow the rules in this sub") and BG2 ("bots help moderators manage the large amount of content posted in this sub") correlated with four other constructs in addition to its own, while IDPG1 ("When someone criticizes this sub, it feels like a personal insult") correlated with five other constructs in addition to its own. Correlation residuals indicated that BG1 and BG2, and to a lesser extent IDPG1, had correlation values greater than |0.10| with several non-related items. For these reasons, we eliminated these items and re-ran the CFA. In our second iteration, INT2 ("we interact with each other in these posts and comments") did not fit well and was removed. Model fit statistics for our third and final pilot 23-item model are: $\chi^2 = 347.18$, $p < 0.05$, RMSEA = 0.054, 90% CI (0.043, 0.064), CFI = 0.918 and SRMR = 0.06. All six factors demonstrated good reliability, where Cronbach's alpha ranged from $\alpha = 0.59$ (BT) to $\alpha = 0.83$ (IDPG).

Main CFA Results. We conducted CFA similarly on our new, independent sample ($N = 608$). Our second CFA closely replicated the first round of results, confirming the validity of our selected and new constructs, and our methodological approach overall. Two additional items were eliminated in the second phase: BG6 was eliminated because of too high of an overlap with BG5 and BG7; INT5 was eliminated because it loaded inappropriately onto SOVC. All other items loaded appropriately.

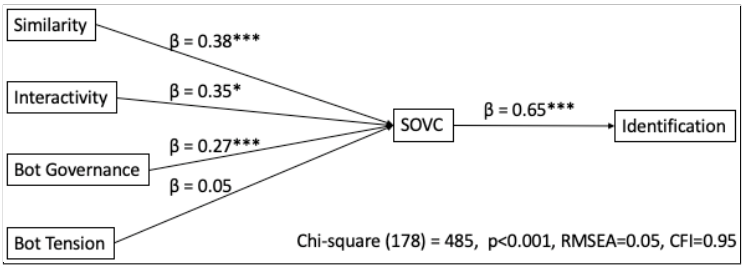


Fig. 3. Main SEM Model: Nomological network with antecedents, SOVC, and outcome of SOVC. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

	Unaware of Bots	**Aware of Bots**	**Non-members**	**Members**
N	341	267	285	323
χ^2	371.86^{***}	350.54^{***}	351.46^{***}	297.26^{***}
RMSEA	0.057	0.06	0.058	0.046
90% CI	(0.048, 0.065)	(0.051, 0.070)	(0.049, 0.067)	(0.036, 0.054)
CFI	0.944	0.936	0.939	0.955
$\beta_{SIM} \rightarrow$ SOVC	0.393^{***}	0.350^{***}	0.418^{***}	0.343^{***}
$\beta_{INT} \rightarrow$ SOVC	0.367^{***}	0.317^{***}	0.154^{**}	0.539^{***}
$\beta_{BG} \rightarrow$ SOVC	0.17^{***}	0.428^{***}	0.344^{***}	0.255^{***}
$\beta_{BT} \rightarrow$ SOVC	0.031	0.14^{*}	0.142^{*}	0.087
$\beta_{SOVC} \rightarrow$ IDPG	0.57^{***}	0.73^{***}	0.482^{***}	0.662^{***}

Table 4. Additional SEM Models demonstrating moderating effects of membership and bot awareness. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Fit statistics of the final 21-item model are as follows: $\chi^2 = 427.31$, $p < 0.05$; RMSEA = 0.049, 90% CI (0.043, 0.055); CFI = 0.959 and SRMR = 0.043. **Therefore, we conclude that GOV-BOTs meets adequate psychometric criteria and can be used to measure users' internal evaluations of *bot governance* (4-item construct, BG3-5 and BG7) and *bot tensions* (3-item construct, BT1-3) on Reddit (RQ1).**

Part III: Structural Equation Modeling

Main SEM Model. Structural Equation Modeling (SEM) is a sibling method to CFA that allows researchers to understand how constructs measured with psychometric instruments can have either direct or indirect pathways of influence on each other; β values show the strength of influence. We conducted SEM path analysis of the expected relationships in our 21-item model using MPlus software [69].[5] Figure 3 depicts our main model, including all users undifferentiated by moderating variables. As anticipated by our theoretical model (Figure 2), all prior constructs behaved as predicted. Two antecedents, *similarity* ($\beta_{SIM} = 0.38$) and *interactivity* ($\beta_{INT} = 0.35$), contribute positively as antecedents to SOVC. That is, when users perceive that others in the sub are similar to them, and are interacting frequently, they may be more likely to experience SOVC.

[5]Note: Demographic control variables were not included in our SEM model (Figure 3) due to current psychometric recommendations to only include control variables theoretically linked to the outcomes of interest [5, 98]. As fundamental human processes, SOVC and IDPG have not been determined to develop differently or to be differentially experienced by diverse populations. To include them likely creates incorrect inferences about the model relationship [98].

Likewise, when users self-report SOVC, they may be more likely to identify with the community ($\beta_{IDPG} = 0.65$). These relationships are well-described by prior literature, thus we will not focus much discussion upon this set of results. However, this high quality replication supports the quality of our final data set and the appropriate use of these data for validating a new instrument.

We find evidence that our new constructs, *bot governance* ($\beta_{BG} = 0.27$) and *bot tensions* ($\beta_{BT} = 0.05$), also function as antecedents to SOVC (RQ2), although pathways associated with bot tensions are not statistically significant in our main model. In support of H2, higher ratings of bot governance contributes to higher SOVC. That is, when users perceive that bots are contributing effectively through activities such as providing relevant resources, generating positive interactions, and removing inappropriate users or content, users are likely to experience a stronger sense of virtual community. On the other hand, we predicted that higher bot tensions (concerns related to bots manipulating, upsetting, or startling users) would potentially damage users' SOVC (H1); this hypothesis was not supported. Rather, our main model shows a small, positive influence on SOVC that is not statistically significant.

Additional SEM Models accounting for moderating variables. We wanted to understand if there were any moderating effects of 2 new variables in phase 2: ***Simplified* Bot Awareness** and **Membership**. In our main survey collection, a smaller overall proportion of users (43.9%) reported awareness of bots in their sub than in the pilot (84.8%). This could be because the sub options in the pilot (footnote 3) had larger subscriber bases (and consequently more noticeable bot activity) than smaller subs included in the second phase, or it may be a relic of changing how we asked and scored the question.[6] Furthermore, about half (53.1%) of phase 2 respondents indicated that they are *members* of the community; the remainder indicated "No" or "I'm not sure." Because people's feelings of membership are of strong and well-studied theoretical importance to their sense of community, it is also possible that membership could impact their relationship with bots in their community.

For both variables, we split the data into two groups, resulting in four models (Table 4). Our main model retains the best overall fit statistics, however each subsequent model also has acceptable fit (albeit with a smaller N). We observe that all pathways retain the same *direction* of influence, however some β and p values shift. Considering prior constructs, *similarity* (β_{SIM}) is relatively stable across all four models. Although *interactivity* (β_{INT}) is similar for those who are aware or unaware, more interactivity is *much* more likely to contribute to SOVC for members than non-members.

Regarding new constructs in GOV-BOTs, the effects of *bot governance* are mediated (unsurprisingly) by awareness. β_{BG} is higher for those who are aware ($\beta_{BG.aware} = 0.428$) than unaware ($\beta_{BG.unaware} = 0.17$), suggesting that the SOVC of users who know about bots is more impacted by them; differences between members ($\beta_{BG.members} = 0.255$) and non-members ($\beta_{BG.nonmembers} = 0.344$) are less stark. Values for *bot tensions* (β_{BT}) vary within a smaller range (0.031 – 0.142). However, unlike the main model, for users who *are* aware of bots, and for non-members β_{BT} values are significant. One possible interpretation is that non-members view bot activities as "governance" that makes the sub appear more community-like to outsiders. However, complex relationships seem to exist between bot tensions and SOVC; we explore this further in the discussion.

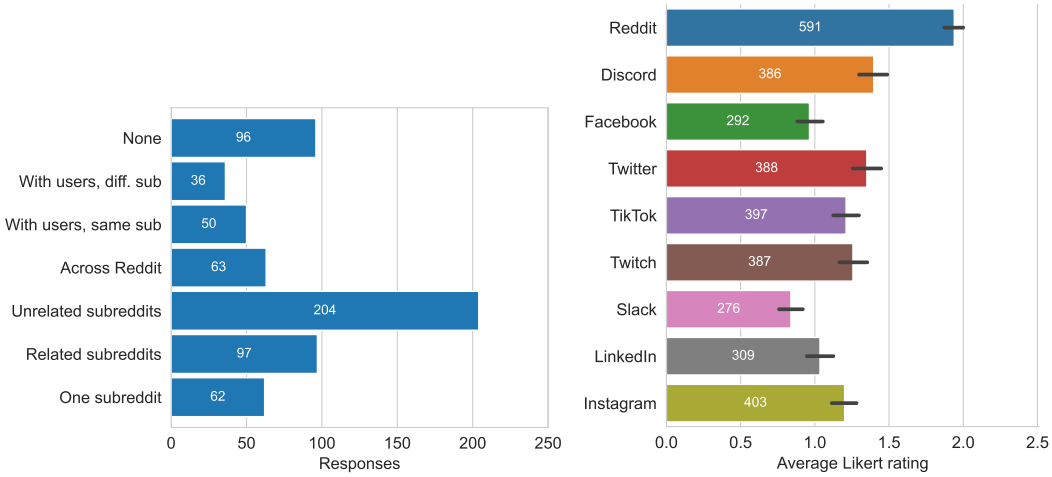
Variations in Community Experiences Across Reddit and Other Platforms

Our results demonstrate that subreddits vary in the degree to which users perceive them as "communities", and that Reddit users view different platforms as spaces where the experience of "community" is more or less important to them (RQ3). For instance, Table 5

[6] Pilot BA construct had 3 items (see Table 2). If users selected "agree" or "strongly agree" to *any* of these, they were marked "aware" of bots, otherwise "unaware." The single *Simplified* Bot Awareness question in phase 2 provides only one opportunity to answer (Yes | No | I'm not sure), providing few opportunities for users to be marked "aware."

	New Questions Added Phase 2				Psychometric Construct Scores					
ID	Aware	Members	Offline	Wellbeing*	SIM	INT	BG	BT	SOVC	IDPG
S1	6.2%	81.2%	50.0%	Better	3.45	3.65	2.72	1.83	3.08	2.81
S2	39.3%	53.6%	82.1%	Same	3.53	3.42	2.97	2.61	2.89	2.54
S3	45.5%	72.7%	96.1%	Better	3.04	3.90	3.22	2.82	3.35	3.16
S4	64.6%	70.8%	89.6%	Better	3.46	3.97	3.5	2.71	3.78	3.56
S5	69.2%	76.9%	81.6%	Better	3.66	3.87	3.55	2.74	3.72	3.51
S6	51.2%	65.9%	78.0%	Depends	3.38	3.54	3.23	2.92	3.33	3.25
S7	74.6%	71.4%	90.5%	Better	3.58	3.59	3.17	2.78	3.22	3.39
S8	38.5%	38.5%	23.1%	Depends	3.44	3.77	3.29	2.38	3.46	2.54
S9	27.9%	60.5%	93.0%	Same	2.57	3.78	2.92	2.43	3.00	2.10
S10	76.0%	68.0%	88.0%	Same	3.22	3.65	3.64	1.87	3.40	2.82
S11	38.0%	32.8%	43.5%	Better	3.27	3.72	3.08	2.80	3.35	2.48
S12	0.0%	42.9%	85.7%	Same	2.61	3.48	3.11	2.52	3.05	2.75

Table 5. Results per sub. * indicates column in which *most common* response bucket is presented. Average Likert ratings (1: Strongly Disagree to 5: Strongly Agree) are reported per psychometric construct.



(a) Perceived SOVC across Reddit.

(b) Perceived SOVC across platforms.

Fig. 4. SOVC across Reddit and other platforms. (a) Question text: *Where do you personally experience the greatest sense of community on Reddit?* Multiple choice, single response. Bar plots indicate total counts of respondents selecting each option. (b) Question text: *To what extent do you feel that a sense of community is important to your experiences on each of these social media?* Responses: (0) Not at all; (1) Somewhat; (2) Extremely important. Bar plots indicate average ratings with standard deviation. Each plot is labeled with N responses; users who selected "I do not use this platform" did not contribute ratings to that platform.

depicts several key results. SOVC scores range from 2.89 to 3.78. Most subs (10/12) had average SOVC>3.0, suggesting that users tend to "agree" that these subs *do* feel like communities–some moreso than others. User awareness of bots ranges from 0–76%, while 32.8–81.2% of respondents consider themselves members. Users of most subs (9/12) rated bot governance positively on average (BG>3.0), while all 12 subs tended to "disagree" that bots cause tensions by startling, manipulating,

or upsetting users (BT<3.0). Consequently, our data suggests that respondents in our sample usually view subs as communities in which governance bots are beneficial, without causing high tensions.

These subs also differ in important experiential ways. For example, in S8 (a sub for a stigmatized mental health issue), only 23.1% of respondents reported that the sub had a connection to offline communities or activities, whereas in S3 (a sub for in-depth discussions of music), that proportion leaps to 96.1%; thus, these communities are experienced as "virtual" to differing degrees. Moreover, spending time in these communities may differentially impact wellbeing. No subs in our sample reported a plurality of respondents who generally felt worse after visiting. Half stated they felt *better* (6/12), whereas the other half indicated that they felt the *same* (4/12), or that it *depends* (2/12).

Are *individual subreddits* the most appropriate *units* of community, however? Figure 4a shows that only 62 respondents (10.2%) experienced the *greatest* SOVC within a single sub–a similar proportion to those who selected "across the Reddit platform" (10.4%). Rather, the plurality of respondents experienced the greatest SOVC across groups of unrelated (33.6%) or related (16%) subs. 15.8% reported no SOVC anywhere on Reddit, while still others reported their greatest SOVC with a specific group of users from the same (8.2%) or different (5.9%) subs. Situating our results within the broader set of possible community spaces online, Figure 4b shows that Reddit users view Reddit as a platform where experiencing SOVC is most important to them; the importance of experiencing SOVC wanes, and also varies substantially, across platforms. In our discussion, we return to these results to share implications for future research.

DISCUSSION

In this paper, we have described the development and validation of the GOV-BOTs psychometric scale (Table 2). We now conclude by contextualizing our findings on how governance bots impact online communities, research implications, pragmatic applications of GOV-BOTs, and limitations and future work.

The impact of governance bots on sense of virtual community

A *sense of community* (SOC) is fundamentally important to people's life satisfaction [74], social, physical, and mental wellbeing [18], perceived safety [72], ability to problem-solve and cope [3], and social/political participation [71]. However, researchers have documented the gradual erosion of geography-based communities due to the impacts of industrialization and globalization, voicing concerns for how a loss of SOC may damage society. The emergence of *relational* communities online began with early discussion fora such as USEnet in the 80s or America Online (AOL) chatgroups in the 90s, and now continues more rapidly and ubiquitously as social media platforms like Reddit, Facebook, Twitter, etc. continue to proliferate new online spaces where users can gather and interact. These online spaces can provide users with a *sense of virtual community* (SOVC) that may help to meet some of people's basic needs for community, especially in cases when suitable offline communities are unavailable (*e.g.*, communities for rare, stigmatized, or life-threatening diseases [59, 95, 96] or marginalized identities [22, 49]). While some aspects of SOVC are similar to offline SOC, prior work also shows that the unique qualities of online interaction cause SOVC to form differently than SOC [8–10, 51]. In particular, online affordances can allow alternatives to in-person interaction that may impact SOVC; our work focuses on governance bots as one such mechanism. Given the rising prevalence of bots across online spaces, it is important to pay heed to how bots can impact SOVC positively or negatively, so that we can continue to shape their design in ways that benefit rather than harm communities and their users.

Bot governance (BG). Our final GOV-BOTs scale includes one 4-item factor that measures users assessments of beneficial BG. One main finding is that user perceptions of effective BG contributes

positively to SOVC. Geiger argues that bots function as "bespoke code" that runs apart from but upon platforms, and which modifies and extends the functionality of platforms [30]. As "extensions" of the Reddit platform, people may now recognize governance bots as inextricable actors in their subs that play a fundamental role in helping to regulate and norm the community [37, 43] (an essential function for community health and safety [48, 66]).

In most of the subs we surveyed (9/12), users rated bot governance positively on average (BG>3.0). However, this trend may or may not generalize across a broader sample of subs (see Limitations). The heterogeneity of bots in different communities suggests that there could be individual bots that are perceived as beneficial or harmful existing side-by-side in the same spaces, with differential effects on the community. On the one hand, GOV-BOTs allows for holistic measurement of people's perceptions of how well the unique collection of bots in a specific sub are serving a community. On the other, GOV-BOTs item text can also be adapted to evaluate individual bots. For example, BG3 is "bots provide resources that are relevant to this sub." To understand user evaluations of a specific bot under development or scrutiny, BG3 can be adapted to "`u/[bot-name]` provides resources that are relevant to this sub." Overall, measures of users' internal perceptions of BG for groups of bots, or individual bots, can inform design processes and help ensure that bots are serving the community, as we will discuss further in the *pragmatic applications of GOV-BOTs* section below.

Bot tensions (BT). In all 12 of the subs we surveyed, users tended to "disagree" that bots cause tensions by startling, manipulating, or upsetting users (BT<3.0). However, BT could also vary significantly across Reddit from our sample. We hypothesized that higher BT would contribute negatively to SOVC. However, we did not find a reliable relationship between BT and SOVC. In our main model of all users, BT did not have a significant effect on users' SOVC. Considering only those who are aware of bots, however, BT contributes positively to SOVC. One interpretation is that if people are aware of bots that threaten online communities they care about, they may have a defensive or protective reflex that *reinforces* SOVC. Another possibility is that when tensions arise between bots and communities, these tensions surface norms that were previously less explicit; by reinforcing norms, SOVC is increased. This interpretation offers a similar mechanism as described above for bot governance–however, only users who are aware of bots, and the tensions they cause, are influenced. Perceived membership in the sub also mediates the relationship between BT and SOVC. For users who consider themselves members, pathways from BT to SOVC are not significant; bot tensions may be less salient to their perceptions, given that they have already affiliated themselves. However, when *non*-members perceive more bot tensions, this strengthens their perception of the sub as a community–possibly by exposing norms that they had not yet been exposed to, or by creating the perception of the sub as an overly-governed community. Our methods in this paper do not provide casual explanations, however future work can interrogate these possibilities through other methods.

Research implications

Aligning use of the term "community" with user experiences. In physical communities, academic constructions of SOC do not always align with the lived experiences of community members [73, 76, 79]. Thus, Mannarini and Fedi write that *"the problems raised by the use of the notion of community are increased if the gap between academic and lay meaning is not taken into account."* [61] Similarly, our work suggests that researchers may improve and refine methods for studying *online* communities, if we can better align users' experiences of "community" with researchers' uses and operationalizations of the term.

Our results support an interpretation of Reddit as an ecological system of *actual* communities nested within broader spaces that may *not* feel like communities to people. For example, we find

that users perceive individual subs as communities to varying degrees, and that a plurality of users may experience the greatest sense of community across *groups of unrelated or related subs* rather than individual subreddits–*i.e.*, the SOVC construct itself applies *across* subs. Teblunthuis *et al.* interviewed Reddit users to understand why people participate across groups of subreddits [102]. They showed that people create *"handpicked portfolios of communities"* that are better able to meet the following three needs than any single sub could: (1) specific topical discussions; (2) socializing with like-minded others; and (3) getting attention from from the largest possible audience. Hwang and Foote also interviewed users of persistently small subreddits, finding that participants' use of these hyperspecific spaces relates to *broader* strategies for curating their online experiences [40]. They argue that *"online communities can be seen as nested niches: parts of an embedded, complex, symbiotic socio-informational ecosystem."* [40] These qualitative insights align with quantitative trends describing overlapping membership and topics *between* communities. For example, Zhu *et al.* have demonstrated that higher levels of *membership overlap* between communities can improve survival rates [118], and that a moderate degree of *topical overlap* also increases community activity levels, whereas too much or too little overlap can harm it [117]. Building on these studies, recent work by TeBlunthuis and Hill [101] also examines whether relationships between subreddits with overlapping members are mutualistic (*i.e.*, driving mutual growth) or competitive (*i.e.*, growth in one group reduces growth in the other). They show that ecologically-related subreddits are more often mutualistic, providing additional evidence that overlapping subreddits often have complementary roles in their ecosystems.

Our results complement and reinforce these prior works with a perspective derived from rigorous psychometric evaluation of users' community-related experiences. **The combined evidence suggests the research implication that if we aim to model users' lived experiences of community, then individual subreddits are likely *not* the best structural units of analysis.** That is to say, the words "subreddit" and "community" are *not* interchangeable: operationalizations of the term "community" should draw from approaches that model relationships between groups of subs (*e.g.*, [99, 101]). Researchers can improve the precision of their reporting by referring to subreddits simply as subreddits, and using the term "community" only when it truly applies. In our future work, we aim to triangulate constructs measured by surveys with analyses of log data and content available via API, in order to provide specific methodological guidance for defining community structures that better match users' community experiences. Doing so will account for both internal user perspectives and externally observable behavior, resulting in a more human-centered analysis. For example, recent work completed on Twitch has adopted this type of approach [45], combining survey and behavioral trace data to show that Twitch users experience community cohesion and sense of belonging differentially across channels on the platform; both cohesion and belonging impact long-term retention of individual viewers.

Moreover, the same communities of humans often span multiple platforms, such as Reddit and Discord [47], or across different fandom platforms [25]. Yet respondents in our survey viewed different platforms as spaces where experiencing SOVC is more or less central to their experiences. Future research can explore how the different technical affordances and designs of these platforms (including but not limited to bots) impact these perceptions.

Investigating psychological, philosophical, and ethical implications. Our evidence that bots *do* influence the formation of SOVC pinpoints some psychological impacts of bots on users' community perceptions, while leaving many other questions unanswered. We suggest the research implication that this domain must be further investigated: What *other* psychological impacts do bots have on individuals and communities? And what types of philosophical and ethical issues might they raise?

For instance, in one study, Morris *et al.* collected supportive comments from humans, and then re-used them on similar support-seeking posts–presenting half the comments as though from human users, and the other half, from bots [68]. Even though "bot" comments were positively rated overall, they were systematically rated lower than "human" comments, suggesting that people's beliefs that support is coming from a human, rather than a bot, improves the quality of support. Another study by Wang *et al.* used a bot that was programmed to leave an initial comment on a support-seeking post, *only* if that post had not naturally received a comment from any other user within 10 minutes [107]. However, the bot did not disclose that it was a bot. Subtracting the bot's own activity, the authors demonstrated that the bot drove up participation from other humans across the community. These studies suggest the fascinating problem that bots can do beneficial things–even when we fail to recognize them as non-human. Yet they can possibly produce even *better* outcomes for individuals or communities, if we simply let users believe that they are interacting with humans. At the same time, intentionally deploying an *identified* bot, and allowing the community to care for it, can also drive up community engagement and create numerous positive and pro-social outcomes for users [91]. **As system designers, we must ask ourselves, where is the line? And how can we best guide these sorts of design tradeoffs for the true improvement of our communities—while also ensuring that users are aware of the many ways in which bots, and the artificially intelligent systems that drive them, are engaging with them socially?** As non-human entities pervade our communities, it is vital to address these types of questions, both empirically and philosophically.

Pragmatic applications of GOV-BOTs

Using GOV-BOTs to evaluate bots in Reddit communities. Given the implicit feudalism of online communities, wherein moderators can accrue relatively incontestable power, Schneider calls for more modular forms of online governance that *"enable platform operators and their users to build bottom-up governance processes from computational components that are modular and composable, highly versatile in their expressiveness, portable from one context to another, and interoperable across platforms"* [88]. Bots offer governance mechanisms that can meet all of these criteria, however the ways in which they are built and deployed could reinforce existing power structures, and/or help users to shape the community through their own technical agency. For instance, bots like `u/Automoderator` can only be configured by moderators, whereas many other types of bots can be configured directly by users.

GOV-BOTs can become part of a toolkit for moderators and bot designers to better monitor and calibrate bot configurations for community needs, and/or to design new kinds of governance bots. For example, mods could **periodically** post announcements about a GOV-BOTs-based survey (open to responses for a specific, limited timeframe) as a pulse-check to evaluate how bots are performing in the community at any given moment. This technique is especially useful if mods have recently added or modified bots in their sub, and want to understand how this impacted the community in a conveniently quantifiable, interpretable, and replicable manner. Furthermore, a straightforward way to **continually** enable and elicit community feedback on bot performance could be to include a live survey link (with an unconstrained timeframe) containing GOV-BOTs items (ideally, along with a broad free response question such as "What do you think of the bots here in `r/[subreddit name]`?") in the subreddit's sidebar, rules, description, or wiki. If the survey is always accessible in a prominent location, responses can trickle in and accumulate over time for occasional review by mods or bot designers, providing a more longitudinal view of bot performance.

Since governance bots *do* contribute to SOVC, it is also worthwhile for researchers or other external parties to invest in evaluations of their design and community impacts, and GOV-BOTs provides a way to empirically address many types of research questions (see the Related Literature

section titled *"Psychometric scale development in HCI"* for illustrative RQs). For one example, to explore the concept of implicit feudalism, researchers could systematically apply the GOV-BOTs scale across subs that may be experienced as more democratic or more feudal in order to understand how bots can variably cement existing power structures, or enable new forms of bottom-up governance. Another promising research opportunity would be to create a separate website, where users can come from specified social media platforms to provide ad hoc ratings of any bot they have encountered. A workflow that allows users to link to a bot's profile (or select from bots that have already been added by others) and submit a GOV-BOTs evaluation with optional free response could enable organic collection of an informative and growing dataset on bots.

Adapting GOV-BOTs across platforms. As mentioned above, GOV-BOTs provides the capability to assess users' perceptions of either groups of bots, or individual bots, in a specific community setting. For the purposes of this study, we wanted to understand the relationship between bot governance and SOVC on Reddit. In order to understand how perceptions of bot governance differs across platforms, or to study bot governance independently in other contexts, GOVT-BOTs can also be adapted to other platforms like Wikipedia, Discord, Twitter, Facebook, or Twitch. Psychometric scales are considered valid "out-of-the-box" only in the contexts where they were designed. Future research will benefit by beginning with the items presented here and adapting language choices to the context. For example, on Wikipedia, BG3 could readily be adapted to "bots provide resources that are relevant to this [article / Wikiproject / discussion page]." Such language adaptations alone yield a better scale for Wikipedia, and allow for initial cross-platform comparisons. To claim full psychometric validity however, adapted versions of the scale should be subjected to confirmatory factor analysis to ensure item-factor fit.

Recommendations for survey research on Reddit. In general, HCI and CSCW need to do a better job adopting standards for survey design and evaluation to ensure *construct validity*; our work here models effective ways to do so by drawing from psychology methods. Furthermore, we were alarmed by the excessive degree of fraudulent survey responses we received. Even in phase two, Appendix Table 6 shows that we eliminated over half of the responses in 5 out of 12 subs retained in the study–and in the case of s11, the entire sub was eliminated. To assure high quality data collection, we encourage future researchers to run surveys *without* incentives on Reddit, to collaborate with moderators, and to subject all survey data to rigorous scrutiny and quality checks to ensure that all data are valid before being included in analysis. We also encourage reviewers to require evidence that quality checks were performed on Reddit survey data, else including all responses can easily lead to shaky or inaccurate empirical claims.

Limitations

Limited specificity to sophisticated community phenomena: One benefit of GOV-BOTs lies in its breadth of applicability across many different types of subreddits and bots. However, one important limitation is that online community governance involves numerous, ever-evolving mechanisms–combining people, platform affordances, algorithms, bots, norms, rules, laws, and broader world events. SOVC offers just one glimpse into these communities. The constructs of *bot tensions* and *bot governance* are a useful starting point, yet more fine-grained measures will be required in order to capture the apparent complexity and diversity of community phenomena across Reddit–in ways that our data suggest are likely to span eco-systems of subs. It will be valuable for future researchers to continually consider what new measurement tools are required for the spaces they are studying, how to design them with community health in mind, and how to ensure that these tools empower the stakeholders that they impact.

The need for qualitative work: We targeted our statistical methods to developing the GOV-BOTs scale. Future work should also *qualitatively* explore how SOVC forms across groups of subreddits, and the ways in which governance bots impact this. Our work provides strong research implications, while future qualitative work can derive specific *design* implications. Furthermore, we selected a *minimal* subset of constructs relevant to SOVC, but we acknowledge that others are known to exist (see Related Literature). Moreover, as the sociotechnical mechanisms within online communities continue to evolve, *novel* constructs will likely emerge. Future work can certainly evaluate other antecedents and outcomes of SOVC.

Generalizability: Our sample size ($N = 820$) is well-suited for validation of a psychometric instrument, and respondent demographics align with overall Reddit demographics. However, this sample is not necessarily representative of all Reddit or all social media users. We recruited public, unobtrusive, and pro-social subs, rather than private, NSFW[7] or anti-social (*e.g.,* hate- or toxicity-based). While we expect that the theoretical relationships between antecedents and outcomes described in this paper would still hold, future research can explore whether trends in construct scores (Tab. 5) generalize across Reddit and other platforms.

CONCLUSION

GOVernance Bots in Online communiTies (GOV-BOTs) is a new psychometric scale that measures two constructs: (1) We provide a 4-item measure for users' assessments of **bot governance**; (2) We provide a 3-item measure for users' assessments of **bot tensions**. Drawing on extensive theories and methods from community psychology, we validated our instrument and confirmed its appropriateness for use in Reddit communities. We show that bot governance contributes positively to users' sense of virtual community, and suggest that future research in human-bot interaction—and moreover *community*-bot interaction—will benefit by using GOV-BOTs to measure users' attitudes towards governance bots. Stakeholders such as moderators and bot designers can also use GOV-BOTs to empirically benchmark users' evaluations of bots as part of ongoing community-centered design processes; scores on GOV-BOTs constructs or items offer a scientifically validated mechanism for holding designers accountable to community members. Finally, most users experience the greatest SOVC across groups of subreddits rather than individual subs, which are perceived as communities to varying degrees: future work should re-assess operationalizations of the term community to better account for users' lived experiences.

ACKNOWLEDGMENTS

This project would not be possible without the collaboration of volunteer Reddit moderators, and the participation of hundreds of users who took our surveys—we appreciate their time and contributions, and we hope this work can provide useful insights for these communities. Many thanks to Dr. Kristen Guth and the research team at Reddit, who provided feedback on our selection of methods and items, and to our undergraduate research assistants, Tamer Shahwan, Sahand Setareh, Vikki Wong, and Gale Prinster, who directly contributed to recruitment and data collection efforts on this project. In the early stages of framing this work, we benefited from conversations with Jessica Feuston, Joseph Seering, Amy Bruckman, Stevie Chancellor, Loren Terveen, Casey Fiesler, and Eshwar Chandrasekharan, and we thank them for sharing their insights and expertise. Finally, one last shout-out to our colleagues in the #Peer weekly meeting hosted by GroupLens Research at the University of Minnesota who have provided valuable feedback and input across the progression of this work; thanks for the support! This work was funded by NSF Award #1910225.

[7] A common online colloquialism indicating "Not Safe For Work" due to the presence of pornographic or other content that could be deemed socially or professionally inappropriate.

REFERENCES

[1] [n.d.]. Bottiquette. https://www.reddit.com/r/Bottiquette/wiki/bottiquette

[2] Blake E. Ashforth and Fred Mael. 1989. Social Identity Theory and the Organization. *Academy of Management Review* 14, 1 (Jan. 1989), 20–39. https://doi.org/10.5465/amr.1989.4278999 Publisher: Academy of Management.

[3] Kenneth M. Bachrach and Alex J. Zautra. 1985. Coping with a Community Stressor: The Threat of a Hazardous Waste Facility. *Journal of Health and Social Behavior* 26, 2 (1985), 127–141. https://doi.org/10.2307/2136602 Publisher: [American Sociological Association, Sage Publications, Inc.].

[4] Jiajun Bao, Junjie Wu, Yiming Zhang, Eshwar Chandrasekharan, and David Jurgens. 2021. Conversations Gone Alright: Quantifying and Predicting Prosocial Outcomes in Online Conversations. In *Proceedings of the Web Conference 2021*. Association for Computing Machinery, New York, NY, USA, 1134–1145. https://doi.org/10.1145/3442381.3450122

[5] Thomas E. Becker. 2005. Potential Problems in the Statistical Control of Variables in Organizational Research: A Qualitative Analysis With Recommendations. *Organizational Research Methods* 8, 3 (July 2005), 274–289. https://doi.org/10.1177/1094428105278021 Publisher: SAGE Publications Inc.

[6] Marit Bentvelzen, Jasmin Niess, Mikołaj P. Woźniak, and Paweł W. Woźniak. 2021. The Development and Validation of the Technology-Supported Reflection Inventory. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Number 366. Association for Computing Machinery, New York, NY, USA, 1–8. http://doi.org/10.1145/3411764.3445673

[7] Reuben Binns, Michael Veale, Max Van Kleek, and Nigel Shadbolt. 2017. Like Trainer, Like Bot? Inheritance of Bias in Algorithmic Content Moderation. In *Social Informatics (Lecture Notes in Computer Science)*, Giovanni Luca Ciampaglia, Afra Mashhadi, and Taha Yasseri (Eds.). Springer International Publishing, Cham, 405–415. https://doi.org/10.1007/978-3-319-67256-4_32

[8] Anita L. Blanchard. 2007. Developing a Sense of Virtual Community Measure. *CyberPsychology & Behavior* 10, 6 (Dec. 2007), 827–830. https://doi.org/10.1089/cpb.2007.9946

[9] Anita L. Blanchard, Leann E. Caudill, and Lisa Slattery Walker. 2020. Developing an entitativity measure and distinguishing it from antecedents and outcomes within online and face-to-face groups. *Group Processes & Intergroup Relations* 23, 1 (Jan. 2020), 91–108. https://doi.org/10.1177/1368430217743577

[10] Anita L. Blanchard and M. Lynne Markus. 2004. The experienced "sense" of a virtual community: characteristics and processes. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 35, 1 (Feb. 2004), 64–79. https://doi.org/10.1145/968464.968470

[11] Anne E Brodsky and Christine M Marx. 2001. Layers of identity: Multiple psychological senses of community within a community setting. *Journal of community psychology* 29, 2 (2001), 161–178.

[12] Phillip Brooker. 2019. My unexpectedly militant bots: A case for Programming-as-Social-Science. *The Sociological Review* 67, 6 (Nov. 2019), 1228–1248. https://doi.org/10.1177/0038026119840988 Publisher: SAGE Publications Ltd.

[13] Amy Bruckman. 2006. A new perspective on "community" and its implications for computer-mediated communication systems. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06)*. Association for Computing Machinery, New York, NY, USA, 616–621. https://doi.org/10.1145/1125451.1125579

[14] Fred B Bryant. 2000. Assessing the validity of measurement. *Reading and understanding MORE multivariate statistics* (2000), 99–146.

[15] Donald T. Campbell. 1958. Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science* 3, 1 (1958), 14–25. https://doi.org/10.1002/bs.3830030103

[16] Eshwar Chandrasekharan, Chaitrali Gandhi, Matthew Wortley Mustelier, and Eric Gilbert. 2019. Crossmod: A Cross-Community Learning-based System to Assist Reddit Moderators. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 174:1–174:30. https://doi.org/10.1145/3359276

[17] David M. Chavis, James H. Hogge, David W. McMillan, and Abraham Wandersman. 1986. Sense of community through Brunswik's lens: A first look. *Journal of Community Psychology* 14, 1 (1986), 24–40. https://doi.org/10.1002/1520-6629(198601)14:1<24::AID-JCOP2290140104>3.0.CO;2-P _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/1520-6629%28198601%2914%3A1%3C24%3A%3AAID-JCOP2290140104%3E3.0.CO%3B2-P.

[18] David M. Chavis and J. R. Newbrough. 1986. The meaning of "community" in community psychology. *Journal of Community Psychology* 14, 4 (1986), 335–340. https://doi.org/10.1002/1520-6629(198610)14:4<335::AID-JCOP2290140402>3.0.CO;2-T Place: US Publisher: John Wiley & Sons.

[19] Florian Daniel, Cinzia Cappiello, and Boualem Benatallah. 2019. Bots Acting Like Humans: Understanding and Preventing Harm. *IEEE Internet Computing* 23, 2 (March 2019), 40–49. https://doi.org/10.1109/MIC.2019.2893137 Conference Name: IEEE Internet Computing.

[20] Michael A. DeVito, Jeremy Birnholtz, Jeffery T. Hancock, Megan French, and Sunny Liu. 2018. How People Form Folk Theories of Social Media Feeds and What it Means for How We Study Self-Presentation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/

3173574.3173694

[21] Robert J. Doolittle and Donald Macdonald. 1978. Communication and a sense of community in a metropolitan neighborhood: A factor analytic examination. *Communication Quarterly* 26, 3 (June 1978), 2–7. https://doi.org/10.1080/01463377809369297 Publisher: Routledge _eprint: https://doi.org/10.1080/01463377809369297.

[22] Brianna Dym and Casey Fiesler. 2018. Vulnerable and online: Fandom's case for stronger privacy norms and tools. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 329–332.

[23] Motahhare Eslami, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert, and Karrie Karahalios. 2019. User Attitudes towards Algorithmic Opacity and Transparency in Online Reviewing Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300724

[24] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The rise of social bots. *Commun. ACM* 59, 7 (June 2016), 96–104. https://doi.org/10.1145/2818717

[25] Casey Fiesler and Brianna Dym. 2020. Moving Across Lands: Online Platform Migration in Fandom Communities. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 042:1–042:25. https://doi.org/10.1145/3392847

[26] Casey Fiesler, Jialun Jiang, Joshua McCann, Kyle Frye, and Jed Brubaker. 2018. Reddit Rules! Characterizing an Ecosystem of Governance. *Proceedings of the International AAAI Conference on Web and Social Media* 12, 1 (June 2018). https://ojs.aaai.org/index.php/ICWSM/article/view/15033 Number: 1.

[27] Claudia Flores-Saviaga, Brian Keegan, and Saiph Savage. 2018. Mobilizing the Trump Train: Understanding Collective Action in a Political Trolling Community. *Proceedings of the International AAAI Conference on Web and Social Media* 12, 1 (June 2018). https://ojs.aaai.org/index.php/ICWSM/article/view/15024 Number: 1.

[28] Susan Folkman, Richard S Lazarus, Christine Dunkel-Schetter, Anita DeLongis, and Rand J Gruen. 1986. Dynamics of a Stressful Encounter: Cognitive Appraisal, Coping, and Encounter Outcomes. *Personality and Social Psychology* 50, 5 (1986), 12.

[29] John Fox and Sanford Weisberg. 2019. *An R Companion to Applied Regression* (3 ed.). Sage Publications. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

[30] R. Stuart Geiger. 2014. Bots, bespoke, code and the materiality of software platforms. *Information, Communication & Society* 17, 3 (March 2014), 342–356. https://doi.org/10.1080/1369118X.2013.873069 Publisher: Routledge _eprint: https://doi.org/10.1080/1369118X.2013.873069.

[31] R. Stuart Geiger and Aaron Halfaker. 2017. Operationalizing Conflict and Cooperation between Automated Software Agents in Wikipedia: A Replication and Expansion of 'Even Good Bots Fight'. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 49:1–49:33. https://doi.org/10.1145/3134684

[32] Tarleton Gillespie. 2018. Regulation of and by Platforms. In *The SAGE Handbook of Social Media*. SAGE Publications Ltd, 55 City Road, 254–278. https://doi.org/10.4135/9781473984066

[33] Thomas J. Glynn. 1981. Psychological Sense of Community: Measurement and Application. *Human Relations* 34, 9 (Sept. 1981), 789–818. https://doi.org/10.1177/001872678103400904 Publisher: SAGE Publications Ltd.

[34] Jonathan Gruber, Eszter Hargittai, Gökçe Karaoglu, and Lisa Brombach. 2021. Algorithm Awareness as an Important Internet Skill: The Case of Voice Assistants. *International Journal of Communication* 15, 0 (March 2021), 19. https://ijoc.org/index.php/ijoc/article/view/15941 Number: 0.

[35] Joseph R. Gusfield. 1975. *Community : a critical response*. Oxford : Blackwell. https://researchportal.scu.edu.au/discovery/fulldisplay/alma990001992490402368/61SCU_INST:ResearchRepository

[36] Aaron Halfaker and R. Stuart Geiger. 2020. ORES: Lowering Barriers with Participatory Machine Learning in Wikipedia. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020), 148:1–148:37. https://doi.org/10.1145/3415219

[37] Qinglai He, Yili Hong, and Raghu Santanam. 2020. Machine-assisted regulation, online participation and human moderation. In *International Conference on Information Systems, ICIS 2020 - Making Digital Inclusive: Blending the Local and the Global*. Association for Information Systems. https://asu.pure.elsevier.com/en/publications/machine-assisted-regulation-online-participation-and-human-modera

[38] Rick H Hoyle. 2012. *Handbook of structural equation modeling*. Guilford press.

[39] Sofia Hurtado, Poushali Ray, and Radu Marculescu. 2019. Bot Detection in Reddit Political Discussion. In *Proceedings of the Fourth International Workshop on Social Sensing (SocialSense'19)*. Association for Computing Machinery, New York, NY, USA, 30–35. https://doi.org/10.1145/3313294.3313386

[40] Sohyeon Hwang and Jeremy D. Foote. 2021. Why do People Participate in Small Online Communities? *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–25. https://doi.org/10.1145/3479606

[41] Shagun Jhaver, Darren Scott Appling, Eric Gilbert, and Amy Bruckman. 2019. "Did You Suspect the Post Would be Removed?": Understanding User Reactions to Content Removals on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 192:1–192:33. https://doi.org/10.1145/3359294

[42] Shagun Jhaver, Iris Birman, Eric Gilbert, and Amy Bruckman. 2019. Human-Machine Collaboration for Content Regulation: The Case of Reddit Automoderator. *ACM Transactions on Computer-Human Interaction* 26, 5 (July 2019), 31:1–31:35. https://doi.org/10.1145/3338243

[43] Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 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 (Nov. 2019), 150:1–150:27. https://doi.org/10.1145/3359252

[44] Steve Jones. 2015. How I Learned to Stop Worrying and Love the Bots. *Social Media + Society* 1, 1 (April 2015), 2056305115580344. https://doi.org/10.1177/2056305115580344 Publisher: SAGE Publications Ltd.

[45] Sanjay R. Kairam, Melissa C. Mercado-Crespo, and Steven Sumner. 2022. A Social-Ecological Approach to Modeling Sense of Virtual Community (SOVC) in Livestreaming Communities. *CSCW '22: ACM Conference On Computer-Supported Cooperative Work And Social Computing* (2022).

[46] Charles Kiene and Benjamin Mako Hill. 2020. Who Uses Bots? A Statistical Analysis of Bot Usage in Moderation Teams. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382960

[47] Charles Kiene, Jialun Aaron Jiang, and Benjamin Mako Hill. 2019. Technological Frames and User Innovation: Exploring Technological Change in Community Moderation Teams. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 44:1–44:23. https://doi.org/10.1145/3359146

[48] Sara Kiesler, Robert E. Kraut, and Paul Resnick. 2012. Regulating Behavior in Online Communities. In *Building Successful Online Communities: Evidence-Based Social Design*. MIT Press.

[49] Shamika Klassen, Sara Kingsley, Kalyn McCall, Joy Weinberg, and Casey Fiesler. 2021. More than a Modern Day Green Book: Exploring the Online Community of Black Twitter. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.

[50] Eva Knekta, Christopher Runyon, and Sarah Eddy. 2019. One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. *CBE—Life Sciences Education* 18, 1 (March 2019), rm1. https://doi.org/10.1187/cbe.18-04-0064 Publisher: American Society for Cell Biology (lse).

[51] Joon Koh, Young-Gul Kim, and Young-Gul Kim. 2003. Sense of Virtual Community: A Conceptual Framework and Empirical Validation. *International Journal of Electronic Commerce* 8, 2 (Dec. 2003), 75–94. https://doi.org/10.1080/10864415.2003.11044295 Publisher: Routledge _eprint: https://doi.org/10.1080/10864415.2003.11044295.

[52] Selcuk Korkmaz, Dincer Goksuluk, and Gokmen Zararsiz. 2014. MVN: An R Package for Assessing Multivariate Normality. *The R Journal* 6, 2 (2014), 151–162. https://journal.r-project.org/archive/2014/RJ-2014-031/index.html

[53] Richard S. Lazarus and Elizabeth Alfert. 1964. Short-circuiting of threat by experimentally altering cognitive appraisal. *The Journal of Abnormal and Social Psychology* 69, 2 (1964), 195–205. https://doi.org/10.1037/h0044635

[54] Carlene R. Lebeuf. 2018. *A taxonomy of software bots: towards a deeper understanding of software bot characteristics*. Thesis. https://dspace.library.uvic.ca/handle/1828/10004 Accepted: 2018-08-31T18:11:23Z.

[55] Brian Lickel, David L. Hamilton, and Steven J. Sherman. 2001. Elements of a Lay Theory of Groups: Types of Groups, Relational Styles, and the Perception of Group Entitativity. *Personality and Social Psychology Review* 5, 2 (May 2001), 129–140. https://doi.org/10.1207/S15327957PSPR0502_4

[56] Duncan Lock. 2013. A Marvellous & Incomplete Compendium of reddit Automatons & Bots. https://duncanlock.net/blog/2013/06/19/a-marvellous-incomplete-compendium-of-reddit-automatons-bots/

[57] Kiel Long, John Vines, Selina Sutton, Phillip Brooker, Tom Feltwell, Ben Kirman, Julie Barnett, and Shaun Lawson. 2017. "Could You Define That in Bot Terms"? Requesting, Creating and Using Bots on Reddit. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3488–3500. https://doi.org/10.1145/3025453.3025830

[58] Ming-Cheng Ma and John P. Lalor. 2020. An Empirical Analysis of Human-Bot Interaction on Reddit. In *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*. Association for Computational Linguistics, Online, 101–106. https://doi.org/10.18653/v1/2020.wnut-1.14

[59] Haley MacLeod, Grace Bastin, Leslie S. Liu, Katie Siek, and Kay Connelly. 2017. "Be Grateful You Don'T Have a Real Disease": Understanding Rare Disease Relationships. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1660–1673. https://doi.org/10.1145/3025453.3025796 event-place: Denver, Colorado, USA.

[60] Fred A. Mael and Lois E. Tetrick. 1992. Identifying Organizational Identification. *Educational and Psychological Measurement* 52, 4 (Dec. 1992), 813–824. https://doi.org/10.1177/0013164492052004002

[61] Terri Mannarini and Angela Fedi. 2009. Multiple senses of community: the experience and meaning of community. *Journal of Community Psychology* 37, 2 (March 2009), 211–227. https://doi.org/10.1002/jcop.20289

[62] Nathalie Marechal. 2016. When Bots Tweet: Toward a Normative Framework for Bots on Social Networking Sites. (2016), 10.

[63] Franziska Martini, Paul Samula, Tobias R Keller, and Ulrike Klinger. 2021. Bot, or not? Comparing three methods for detecting social bots in five political discourses. *Big Data & Society* 8, 2 (July 2021), 20539517211033566. https://doi.org/10.1177/20539517211033566 Publisher: SAGE Publications Ltd.

[64] Adrienne L. Massanari. 2016. Contested Play: The Culture and Politics of Reddit Bots. In *Socialbots and Their Friends*. Routledge. Num Pages: 18.

[65] J. Nathan Matias. 2019. The Civic Labor of Volunteer Moderators Online. *Social Media + Society* 5, 2 (April 2019), 2056305119836778. https://doi.org/10.1177/2056305119836778 Publisher: SAGE Publications Ltd.

[66] J. Nathan Matias. 2019. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences* 116, 20 (May 2019), 9785–9789. https://doi.org/10.1073/pnas.1813486116 Publisher: National Academy of Sciences Section: Social Sciences.

[67] David W. McMillan and David M. Chavis. 1986. Sense of community: A definition and theory. *Journal of Community Psychology* 14, 1 (Jan. 1986), 6–23. https://doi.org/10.1002/1520-6629(198601)14:1<6::AID-JCOP2290140103>3.0.CO;2-I

[68] Robert R. Morris, Kareem Kouddous, Rohan Kshirsagar, and Stephen M. Schueller. 2018. Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions. *Journal of Medical Internet Research* 20, 6 (June 2018), e10148. https://doi.org/10.2196/10148 Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.

[69] Linda K Muthén and Bengt Muthén. 2017. *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.

[70] Adewale Obadimu, Esther Mead, Samer Al-khateeb, and Nitin Agarwal. 2019. *A COMPARATIVE ANALYSIS OF FACEBOOK AND TWITTER BOTS*.

[71] Patricia Obst, Sandy G. Smith, and Lucy Zinkiewicz. 2002. An exploration of sense of community, Part 3: Dimensions and predictors of psychological sense of community in geographical communities. *Journal of Community Psychology* 30, 1 (2002), 119–133. https://doi.org/10.1002/jcop.1054 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcop.1054.

[72] D. D. Perkins and R. B. Taylor. 1996. Ecological assessments of community disorder: their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology* 24, 1 (Feb. 1996), 63–107. https://doi.org/10.1007/BF02511883

[73] Grace M. H. Pretty, Colleen Conroy, Jason Dugay, Karen Fowler, and Diane Williams. 1996. Sense of community and its relevance to adolescents of all ages. *Journal of Community Psychology* 24, 4 (1996), 365–379. https://doi.org/10.1002/(SICI)1520-6629(199610)24:4<365::AID-JCOP6>3.0.CO;2-T _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291520-6629%28199610%2924%3A4%3C365%3A%3AAID-JCOP6%3E3.0.CO%3B2-T.

[74] Miretta Prezza and Stefano Costantini. 1998. Sense of community and life satisfaction: investigation in three different territorial contexts. *Journal of Community & Applied Social Psychology* 8, 3 (1998), 181–194. https://doi.org/10.1002/(SICI)1099-1298(199805/06)8:3<181::AID-CASP436>3.0.CO;2-4 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291099-1298%28199805/06%298%3A3%3C181%3A%3AAID-CASP436%3E3.0.CO%3B2-4.

[75] Nicholas Proferes, Naiyan Jones, Sarah Gilbert, Casey Fiesler, and Michael Zimmer. 2021. Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social Media + Society* 7, 2 (April 2021), 20563051211019004. https://doi.org/10.1177/20563051211019004 Publisher: SAGE Publications Ltd.

[76] John E. Puddifoot. 1995. Dimensions of community identity. *Journal of Community & Applied Social Psychology* 5, 5 (Dec. 1995), 357–370. https://doi.org/10.1002/casp.2450050507

[77] Astrid Rosenthal-Von Der Pütten and Nikolai Bock. 2018. Development and Validation of the Self-Efficacy in Human-Robot-Interaction Scale (SE-HRI). *ACM Transactions on Human-Robot Interaction* 7, 3 (Dec. 2018), 21:1–21:30. https://doi.org/10.1145/3139352

[78] Qualtrics. 2021. Qualtrics XM. https://www.qualtrics.com/

[79] Mark Rapley and Grace M. H. Pretty. 1999. Playing procrustes: The interactional production of a "psychological sense of community". *Journal of Community Psychology* 27, 6 (1999), 695–713. https://doi.org/10.1002/(SICI)1520-6629(199911)27:6<695::AID-JCOP5>3.0.CO;2-M _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291520-6629%28199911%2927%3A6%3C695%3A%3AAID-JCOP5%3E3.0.CO%3B2-M.

[80] Pew Research. 2016. Reddit news users more likely to be male, young and digital in their news preferences. https://www.pewresearch.org/journalism/2016/02/25/reddit-news-users-more-likely-to-be-male-young-and-digital-in-their-news-preferences/

[81] William Revelle. 2021. psych: Procedures for Psychological, Psychometric, and Personality Research. https://CRAN.R-project.org/package=psych

[82] Sarah T. Roberts. 2014. *Behind the screen: The hidden digital labor of commercial content moderation.* Dissertation. University of Illinois at Urbana-Champaign. https://www.ideals.illinois.edu/handle/2142/50401

[83] Isabel Rodríguez-de Dios, Juan-José Igartua, and Alejandro González-Vázquez. 2016. Development and validation of a digital literacy scale for teenagers. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM '16)*. Association for Computing Machinery, New York, NY, USA, 1067–1072. https://doi.org/10.1145/3012430.3012648

[84] Yves Rosseel. 2012. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* 48, 1 (May 2012), 1–36. https://doi.org/10.18637/jss.v048.i02 Number: 1.

[85] Seymour B. Sarason. 1974. *The psychological sense of community: Prospects for a community psychology.* Jossey-Bass, Oxford, England. Pages: xii, 290.

[86] William Sattelberg. 2021. The Demographics of Reddit: Who Uses the Site? https://www.alphr.com/demographics-reddit/

[87] Nathan Schneider. 2021. Admins, mods, and benevolent dictators for life: The implicit feudalism of online communities. *New Media & Society* (Jan. 2021), 1461444820986553. https://doi.org/10.1177/1461444820986553 Publisher: SAGE Publications.

[88] Nathan Schneider, Primavera De Filippi, Seth Frey, Joshua Z. Tan, and Amy X. Zhang. 2021. Modular Politics: Toward a Governance Layer for Online Communities. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 16:1–16:26. https://doi.org/10.1145/3449090

[89] Chester A Schriesheim and Regina J Eisenbach. 1995. An exploratory and confirmatory factor-analytic investigation of item wording effects on the obtained factor structures of survey questionnaire measures. *Journal of Management* 21, 6 (1995), 1177–1193.

[90] Joseph Seering. 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 (Oct. 2020), 1–28. https://doi.org/10.1145/3415178

[91] Joseph Seering, Juan Pablo Flores, Saiph Savage, and Jessica Hammer. 2018. The Social Roles of Bots: Evaluating Impact of Bots on Discussions in Online Communities. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 157:1–157:29. https://doi.org/10.1145/3274426

[92] Joseph Seering, Michal Luria, Connie Ye, Geoff Kaufman, and Jessica Hammer. 2020. It takes a village: integrating an adaptive chatbot into an online gaming community. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.

[93] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. The spread of low-credibility content by social bots. *Nature Communications* 9, 1 (Nov. 2018), 4787. https://doi.org/10.1038/s41467-018-06930-7 Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Complex networks;Politics;Society;Technology Subject_term_id: complex-networks;politics;society;technology.

[94] Prasha Shrestha, Arun Sathanur, Suraj Maharjan, Emily Saldanha, Dustin Arendt, and Svitlana Volkova. 2020. Multiple social platforms reveal actionable signals for software vulnerability awareness: A study of GitHub, Twitter and Reddit. *PLOS ONE* 15, 3 (March 2020), e0230250. https://doi.org/10.1371/journal.pone.0230250 Publisher: Public Library of Science.

[95] C Estelle Smith, Avleen Kaur, Katie Z Gach, Loren Terveen, Mary Jo Kreitzer, and Susan O'Conner-Von. 2021. What is Spiritual Support and How Might It Impact the Design of Online Communities? *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–42.

[96] C Estelle Smith, Zachary Levonian, Haiwei Ma, Robert Giaquinto, Gemma Lein-Mcdonough, Zixuan Li, Susan O'Conner-Von, and Svetlana Yarosh. 2020. " I Cannot Do All of This Alone" Exploring Instrumental and Prayer Support in Online Health Communities. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 5 (2020), 1–41.

[97] C. Estelle Smith, Bowen Yu, Anjali Srivastava, Aaron Halfaker, Loren Terveen, and Haiyi Zhu. 2020. Keeping Community in the Loop: Understanding Wikipedia Stakeholder Values for Machine Learning-Based Systems. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376783

[98] Paul E. Spector. 2021. Mastering the Use of Control Variables: the Hierarchical Iterative Control (HIC) Approach. *Journal of Business and Psychology* 36, 5 (Oct. 2021), 737–750. https://doi.org/10.1007/s10869-020-09709-0

[99] Chenhao Tan. 2018. Tracing Community Genealogy: How New Communities Emerge from the Old. *Proceedings of the International AAAI Conference on Web and Social Media* 12, 1 (June 2018). https://ojs.aaai.org/index.php/ICWSM/article/view/15003 Number: 1.

[100] R Core Team. 2020. R: The R Project for Statistical Computing. https://www.r-project.org/

[101] Nathan TeBlunthuis and Benjamin Mako Hill. 2022. Identifying Competition and Mutualism between Online Groups. *Proceedings of the International AAAI Conference on Web and Social Media* 16 (May 2022), 993–1004. https://ojs.aaai.org/index.php/ICWSM/article/view/19352

[102] Nathan TeBlunthuis, Charles Kiene, Isabella Brown, Laura Alia Levi, Nicole McGinnis, and Benjamin Mako Hill. 2022. No Community Can Do Everything: Why People Participate in Similar Online Communities. arXiv:2201.04271 [cs.SI]

[103] Vero Vanden Abeele, Lennart E. Nacke, Elisa D. Mekler, and Daniel Johnson. 2016. Design and Preliminary Validation of The Player Experience Inventory. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY Companion '16)*. Association for Computing Machinery, New York, NY, USA, 335–341. https://doi.org/10.1145/2968120.2987744

[104] Onur Varol, Emilio Ferrara, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online Human-Bot Interactions: Detection, Estimation, and Characterization. *Proceedings of the International AAAI Conference on Web and Social Media* 11, 1 (May 2017), 280–289. https://ojs.aaai.org/index.php/ICWSM/article/view/14871 Number: 1.

[105] Daniel Votipka, Desiree Abrokwa, and Michelle L. Mazurek. 2020. Building and Validating a Scale for Secure Software Development Self-Efficacy. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–20. http://doi.org/10.1145/3313831.3376754

[106] Dijana R. Vukovic and Igor M. Dujlovic. 2016. Facebook messenger bots and their application for business. In *2016 24th Telecommunications Forum (TELFOR)*. 1–4. https://doi.org/10.1109/TELFOR.2016.7818926

[107] Liuping Wang, Dakuo Wang, Feng Tian, Zhenhui Peng, Xiangmin Fan, Zhan Zhang, Shuai Ma, Mo Yu, Xiaojuan Ma, and Hongan Wang. 2021. CASS: Towards Building a Social-Support Chatbot for Online Health Community. In *Proceedings of the CSCW*.

[108] Sai Wang. 2021. Moderating Uncivil User Comments by Humans or Machines? The Effects of Moderation Agent on Perceptions of Bias and Credibility in News Content. *Digital Journalism* 9, 1 (Jan. 2021), 64–83. https://doi.org/10.1080/21670811.2020.1851279 Publisher: Routledge _eprint: https://doi.org/10.1080/21670811.2020.1851279.

[109] T Wei and V Simko. 2021. R package "corrplot": Visualization of a Correlation Matrix. https://github.com/taiyun/corrplot original-date: 2011-12-04T15:15:45Z.

[110] Joseph Weizenbaum. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 1 (Jan. 1966), 36–45. https://doi.org/10.1145/365153.365168

[111] Stephen G. West, Aaron B. Taylor, and Wei Wu. 2012. Model fit and model selection in structural equation modeling. In *Handbook of structural equation modeling*. The Guilford Press, New York, NY, US, 209–231.

[112] David Westerman, Aaron C. Cross, and Peter G. Lindmark. 2019. I Believe in a Thing Called Bot: Perceptions of the Humanness of "Chatbots". *Communication Studies* 70, 3 (May 2019), 295–312. https://doi.org/10.1080/10510974.2018.1557233 Publisher: Routledge _eprint: https://doi.org/10.1080/10510974.2018.1557233.

[113] Qingwen Xu, Douglas D Perkins, and Julian Chun-Chung Chow. 2010. Sense of community, neighboring, and social capital as predictors of local political participation in China. *American journal of community psychology* 45, 3 (2010), 259–271.

[114] Nick Yee, Nicolas Ducheneaut, and Les Nelson. 2012. Online gaming motivations scale: development and validation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2803–2806. http://doi.org/10.1145/2207676.2208681

[115] Li-Yin Young. 2018. The Effect of Moderator Bots on Abusive Language Use. In *Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence (PRAI 2018)*. Association for Computing Machinery, New York, NY, USA, 133–137. https://doi.org/10.1145/3243250.3243257

[116] Lei (Nico) Zheng, Christopher M. Albano, Neev M. Vora, Feng Mai, and Jeffrey V. Nickerson. 2019. The Roles Bots Play in Wikipedia. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 215:1–215:20. https://doi.org/10.1145/3359317

[117] Haiyi Zhu, Jilin Chen, Tara Matthews, Aditya Pal, Hernan Badenes, and Robert E. Kraut. 2014. Selecting an effective niche: an ecological view of the success of online communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 301–310. https://doi.org/10.1145/2556288.2557348

[118] Haiyi Zhu, Robert E. Kraut, and Aniket Kittur. 2014. The impact of membership overlap on the survival of online communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 281–290. https://doi.org/10.1145/2556288.2557213

[119] Haiyi Zhu, Bowen Yu, Aaron Halfaker, and Loren Terveen. 2018. Value-Sensitive Algorithm Design: Method, Case Study, and Lessons. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 194:1–194:23.

APPENDIX

A. Pilot Survey Demographics

In our pilot, 54% of respondents identified as male, 42% female, and other categories like non-binary, prefer not to disclose, and self-describe are less than 2% each. 56% were 25–34 years of age, 19% were 18–24, 17% were 35–44, and 6.7% were 45–54. 72% identified as White/Caucasian, 10% as Hispanic, and less than 6% other. 35% have a bachelor's degree, 31% high school diploma or GED, and 23% trade or associate degree. 74% reported being employed and 10% reported being students. 68% live in North America, 17% in Europe, and 7% in South America. 41% reported using their sub for 1–5 months, 25% 6 months–1 year, and 15% less than 1 month. 36% report visiting their sub "almost daily", 35% about once per week, and 12% multiple times per day.

B. Demographics and Data Cleaning Details from Main Surveys

58.8% of respondents identified as male, 33.8% as female, and other categories like non-binary, prefer not to disclose, and self-describe are 1–3% each. 50.1% were 25–34 years of age, 21.3% were 18–24, 18.8% were 35–44, and 7.5% were 45–54. 61.5% identified as only White/Caucasian, 15.1% as only Black/African American, 4.8% as only Asian or Pacific Islander, 3.6% as only Hispanic, and less than 3% in other categories, including multiple races. 41.7% have a bachelor's degree, 20.8% post-bachelor's degree, 19.6% high school diploma or GED, and 16.3% trade or associate degree. 73.9% reported being employed, 14.5% reported being students, and 6.0% reported being unemployed. 81% of respondents reported living in North America, 7.7% in Europe, 6.6% in North America outside the U.S., and less than 3% elsewhere. 50.8% of respondents reported using their sub for 1–5 years, 20.9% 6 months–1 year, and 12.8% 1–5 months. 41.4% of respondents report visiting their sub "almost daily", 27.3% about once per week, and 20.4% multiple times/day.

Data Cleaning Details.

Sub ID	Total Responses	Automatic Removals	Manual Removals	% Removed	Final N
S1	20	5	0	25.00%	16
S2	45	17	0	37.78%	28
S3	174	63	34	55.75%	77
S4	156	88	20	69.23%	48
S5	179	97	43	78.21%	39
S6	229	33	155	82.10%	41
S7	92	59	12	77.17%	21
S8	14	1	0	7.14%	13
S9	46	3	0	6.52%	43
S10	26	1	0	3.85%	25
S11*	260	10	0	3.85%	250
S12*	7	0	0	0.00%	7
S13*	217	-	-	100%	0

Table 6. Reporting on Removals of Invalid Phase 2 Survey Responses. * indicates a shift in recruitment strategy, in which mentions of an incentive lottery were removed from recruitment copy. Automatic Removals refers to responses that were removed due to failing attention check questions or less than 18yo. Manual Removals were removed due to other issues, such as nonsensical or copy/pasted text or overly speedy responses.

Received January 2022; revised April 2022; accepted May 2022