'Welcome' Changes? Descriptive and Injunctive Norms in a Wikipedia Sub-Community

JONATHAN T. MORGAN, Wikimedia Foundation, USA
ANNA FILIPPOVA, GitHub Inc., USA

Open online communities rely on social norms for behavior regulation, group cohesion, and sustainability. Research on the role of social norms online has mainly focused on one source of influence at a time, making it difficult to separate different normative influences and understand their interactions. In this study, we use the Focus Theory to examine interactions between several sources of normative influence in a Wikipedia sub-community: local descriptive norms, local injunctive norms, and norms imported from similar sub-communities. We find that exposure to injunctive norms has a stronger effect than descriptive norms, that the likelihood of performing a behavior is higher when both injunctive and descriptive norms are congruent, and that conflicting social norms may negatively impact pro-normative behavior. We contextualize these findings through member interviews, and discuss their implications for both future research on normative influence in online groups and the design of systems that support open collaboration.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**;

Keywords: social norms; open collaboration; online communities; Wikipedia; virtual teams; intra-group processes; socialization

1 INTRODUCTION

Open online collaborations have gained both public and scholarly attention for their ability to channel the effort of thousands of volunteer contributors into the creation of valuable products and services. However, key features that contribute to the success of open collaborations, such as low barriers to entry and porous boundaries, can also make them vulnerable to disruption [20]. High membership churn and low member commitment create challenges for socializing new members and enforcing standards of behavior. Furthermore, open collaborations' dependence on voluntary, and often anonymous, contributions limits their ability to directly sanction rule-breakers. These challenges can lead to conflict, burnout, and quality control issues that threaten the community's survival [10]. Developing and maintaining strong social norms is one strategy that these communities employ to sustain themselves in the face of such challenges [1].

Prior work has recorded various sources of normative influence in online communities—through explicitly stated rules and guidelines [41], implicit inference of the behavior of others [60], participants' prior experiences with other social groups [68], and through corrective mechanisms

Authors' addresses: Jonathan T. Morgan, Wikimedia Foundation, San Francisco, CA, USA, jmorgan@wikimedia.org; Anna Filippova, GitHub Inc. San Francisco, CA, USA, annafil@github.com.

such as moderation [31]. However research thus far has mainly examined one source of normative influence at a time and has not distinguished between different types of social norms. We argue that for better community management and design of open collaborations, it is important to examine distinct sources of normative influence and their interactions. Some sources of normative influence within a community may have a greater impact on behavior than others in particular contexts and/or for particular individuals. Furthermore, norms may conflict with one another, leading to unpredictable—and often, undesirable—outcomes [5]. This normative conflict may in turn lead to greater overall membership turnover and other risks to community success, such as the departure of core members [17, 50].

The present work offers both theoretical and practical contributions to the growing body of literature on normative influence and behavioral regulation in open online communities. We contribute to theory development by drawing on the Focus Theory of Normative Conduct [6] to compare several factors that influence normative behavior within online communities: local *injunctive norms* that set behavioral expectations by prescribing and correcting accepted community practices, local *descriptive norms* that support experiential social learning through observation of other community members's behavior, and norms imported from related communities that share members in common with a target community. We apply the Focus Theory to an open collaboration context, and find evidence that supports and extends previous findings from both offline experimental studies and studies of online communities that examined single sources of normative influence in isolation. Our large-scale empirical examination of the underlying mechanisms of normative behavior in an active Wikipedia sub-community, the Teahouse, also yields practical recommendations for system designs that support community success through greater norm coherence. We supplement our quantitative analysis with interviews with community members to interrogate our analytical assumptions and contextualize our findings.

2   RELATED WORK

2.1   The Focus Theory of Normative Conduct

Social norms are shared beliefs within a group, community, or culture about what conduct is customary or socially appropriate [7]. Strong shared norms help social groups maintain cohesion by articulating shared values and differentiating them from other groups, and by helping establish a distinct group identity that motivates participation [5]. Shared norms also serve important regulatory functions by providing a decentralized mechanism for setting behavioral expectations and sanctioning those who deviate from them, reducing the need for formal, enforceable rules [9, 39].

The Focus Theory of Normative Conduct developed by Robert Cialdini and his collaborators addresses two key ways that norms influence behavior: injunctively, by prescribing behaviors that are approved or disapproved by the group (e.g. what group members *ought* to do); and descriptively, by characterizing the prevailing behaviors within the group (e.g. what group members *actually* do). In a series of experiments [7], Cialdini and colleagues demonstrated they could increase or decrease the likelihood that subjects would litter in public spaces by presenting simple social and visual cues that made particular norms more *salient*—that is, these cues communicated whether littering was common (descriptive norms), and/or whether it was socially acceptable to litter (injunctive norms). These studies demonstrated that while both injunctive and descriptive normative cues influence behavior, injunctive norms may have a stronger effect overall. In absence of specific evidence that littering is disapproved of people generally followed what they saw others doing (a salient descriptive norm) regardless of whether others were littering or not. However, the researchers were able to discourage littering, even in the face of evidence that others commonly littered in the public space, by presenting cues that *focused* participants attention on an anti-littering injunctive

norm—e.g. by showing litter swept into a pile, rather than left scattered about where it apparently fell.

Cialdini and colleagues also found that injunctive norms and descriptive norms were influential in different ways. Descriptive norm compliance tended to be strongly mediated by the prevalence of norm compliance in the immediate environment: increasing the amount of visible litter led to relatively more littering. Injunctive norm exposure, on the other hand, tended to persist across time, even without repeated exposure, and to transcend situational boundaries: people exposed to an anti-littering injunctive norm in one environment littered less even after entering a separate environment.

Cialdini's research indicates that accounting for how people behave in a given social context requires the identification of both injunctive and descriptive sources of normative influence, their relative salience, and their interactions. We believe that this insight is especially important for the design and management of online collaborative communities. The design of technological platforms that power online communities and the backgrounds of the community members themselves may influence behavior in unexpected ways by making different sources of normative influence more or less salient [14]. However, we know relatively little about the interrelationship between descriptive and injunctive norms and their effect on the social dynamics of open online communities.

2.2 Social norms in online collaborative communities

A significant amount of work has focused on understanding the role of normative influence in online collaborative communities. However, most studies do not specify the source of normative influence—i.e. injunctive or descriptive norms—they are examining. We motivate a set of hypotheses and a research question for the current study with a review of this body of research through the lens of the Focus Theory's operationalization of *injunctive* and *descriptive norm exposure* and how such exposure may mediate behavior under conditions of *normative congruence* and *normative conflict*.

2.2.1 Injunctive norm exposure. Studies of online collaborative communities can generally be said to be examining injunctive norms (whether or not the authors themselves frame their work this way) when they address how social expectations are communicated and enacted through mechanisms such as community-created FAQs and policies [3, 35, 46], or through distributed moderation practices such as decentralized gatekeeping [57] and up/down-voting [36]. Taken as a whole, this body of work suggests that increasing the relative visibility of, or exposure to, injunctive norms—whether by explicitly documenting normative expectations in community spaces, or by engaging in public corrective practices to discourage deviants and/or demonstrate acceptable modes of behavior—is a widely used mechanism for encouraging pro-normative behavior.

Cialdini's experiments show that anti-littering injunctive norms can be made salient by priming experimental subjects with explicit anti-littering injunctions (placing a "please do not litter" flyer on the subject's car windshield), and by showing evidence of corrective action (showing litter swept into a pile). Furthermore, making the corrective action more immediate and explicit (showing an experimental confederate picking up a piece of litter) led to an even higher level of compliance, even when the immediate environment was unswept and heavily littered. As noted above, there is relatively little experimental research on the impact of increasing injunctive norm salience in online communities. However, a 2016 study by Matias [41] showed that posting rules on Reddit threads increased normative compliance among new commenters. Therefore, we propose that:

H1: *Exposure to an injunctive norm either through explicitly documented rules (H1a) or through corrective practices (H1b) will be associated with greater likelihood to perform pro-normative behavior, regardless of whether a corresponding descriptive norm is salient.*

2.2.2 Descriptive norm exposure. The design of many online collaboration platforms can also increase the salience of a particular community's descriptive norms. For example, in many text-based discussion forums records of individual actions and social interactions—in the form of recent discussion threads—are prominently displayed, and searchable archives of past discussions are publicly available. Newcomers pay attention to these records of past behavior and draw on them when 'lurking' to assess how well the community aligns with their needs and interests [53] and when learning how to participate productively [47]. Studies can be said to focus on descriptive norms when they focus on how people draw on these observations of behavior to determine how to act in a given situation (social proof) [54], or evaluate themselves against others (social comparison) [4]. Studies that examine the evolution of implicit norm-associated phenomena within communities, such as linguistic conventions [11, 34, 49], can also be said to focus primarily on descriptive norms.

Experimental evidence also suggests that descriptive norms mediate behavior in online communities. Postmes et al. demonstrate that members of online chat groups were able to infer descriptive norms by observing particular linguistic conventions within group conversation, and that group members' language use could be manipulated by varying these conventions [52]. Sukumaran and colleagues demonstrate that by manipulating elements of a discussion interface to draw attention to different descriptive norms (such as the length of others' visible replies), it was possible to encourage more thoughtful interaction [60]. Furthermore, Cialdini's experiments show that the salience of these descriptive norms is mediated by volume of available evidence: increasing the amount of litter in an already-littered environment increases the likelihood that an experimental subject will litter [6]. Therefore, we also propose that:

H2: *Greater salience of descriptive norms in past behavior examples will be associated with a greater likelihood to perform the behavior observed, regardless of prior exposure to an injunctive norm about the behavior.*

2.2.3 Normative congruence. To the best of our knowledge, no previous research has examined the interplay of descriptive and injunctive norms in online communities. However, we can draw on prior work in other domains to understand what may occur when both descriptive and injunctive norms are at work. First, we examine the condition of normative congruence, that is, when both descriptive and injunctive norms are salient and not in conflict [37].

On Reddit forums, for example, community rules (injunctive norms) are often displayed in a sidebar next to a comment board [27] or in a 'sticky' comment at the top of a thread [41], and active discussions (descriptive norms) appear near the top of the page. If the behaviors exhibited in these active discussions reflect the injunctive norms displayed in the community rules, the injunctive and descriptive norms of the forum can be said to be congruent. Consistent with Cialdini's experimental results [6] and other work in offline contexts [55], we expect that multiple, aligned sources of normative influence will reinforce one another—that is, the presence of explicitly-stated injunctive norms will make congruent descriptive norms more salient, and evidence that individual behaviors are aligned with community rules will suggest that those rules more important. Therefore, we propose that:

H3: *Exposure to congruent descriptive and injunctive norms will increase the likelihood to perform behavior significantly more than either type of normative exposure individually.*

2.2.4 Normative conflict. Online communities are particularly vulnerable to scenarios where norms may come into conflict—such as because of a misalignment between injunctive and descriptive norms, or between competing sets of injunctive norms [15]. For example, porous boundaries and high membership turnover may lead to the "eternal September" problem [31]—in which an established community lacks the resources efficiently socialize a sudden influx of newcomers who

are inadvertently violating community norms. In this case, actual behavior (descriptive norms) may begin to deviate from sanctioned behavior (injunctive norms) due to newcomers' lack of knowledge about behavioral expectations [38], resulting in normative conflict.

Normative conflict may also occur within a community when people join from other, similar communities. When someone enters an unfamiliar social situation, the salience of local norms is heavily mediated by their sense of *behavioral ambiguity*—the degree to which they believe they already understand local standards of behavior [37]. Open collaboration platforms such as Wikipedia, Reddit, and GitHub often exhibit a significant degree of membership overlap among their various sub-communities, each of which has distinct, local injunctive norms [29, 65, 67]. New arrivals to these platforms are likely to experience a relatively high level of behavioral ambiguity when they first visit one of these sub-communities. They actively seek out cues in the environment around them for how they should act, which focuses their attention to local norms [63]. However, joiners with prior experience participating in similar communities will likely experience relatively less behavioral ambiguity than true newcomers and therefore may tend to act according to the behavioral repertoires they developed in those communities [37], whether or not their behavior complies with local norms.

These normative conflicts can have negative consequences: Filippova and Cho, in a series of studies of FLOSS communities, showed that some forms of normative conflict had a negative impact on members' identification with the project and their intention to continue participating [16, 17]. However, we know relatively little about the conditions under which different sources of norms— injunctive vs. descriptive, local vs. imported—are most influential in online communities. Because there is insufficient prior work for us to generate a hypothesis about the specific nature of these influences, we ask a more general research question about normative conflict:

RQ1: *In situations of normative conflict, which norms are more likely to influence behavior, and under what conditions?*

2.3 Research site

We test our hypotheses and research question about normative influence in a Wikipedia sub-community called the Teahouse. In the section below, we provide a review of research on the social norms of Wikipedia, and an overview of the relevant sociotechnical characteristics of the Teahouse in order to motivate our choice of research site and our study design.

2.3.1 Social norms and sub-communities on Wikipedia. Wikipedia [1], an online community with tens of thousands of active participants, is governed by a complex and sophisticated set of community-created policies and practices that are developed [18], interpreted [30], and enforced [2] in a highly contextual, contingent, and decentralized fashion.

The complexity of these written and unwritten rules has historically created challenges for new editors. However, Wikipedia as a whole lacks robust institutional socialization mechanisms for onboarding new contributors and teaching them how to participate in a consistent and effective manner[12, 48]. Compounding the socialization issue, the way many experienced Wikipedia editors interact with newcomers can feel hostile, which discourages inexperienced editors who are attempting to learn how to participate productively [8, 13, 23, 25].

The Wikipedia Teahouse was created in 2012 in order to address these issues by providing a community-managed space where new editors could receive answers to common questions that were encouraging, detailed, and easy to understand [43]. Wikipedia contains a rich ecology of sub-communities, including content review boards [62], topic or task-based working groups

---

[1]All references to Wikipedia refer to the English language edition.

Fig. 1. A question from a new editor at the Teahouse, with answers that reflect local injunctive norms.

called WikiProjects [45], and questions and answer (Q&A) forums like the Help Desk—a sub-community established in 2004 to answer questions from the general editing population. Like Reddit, StackExchange, and other open collaboration platforms that feature nested, networked sub-communities, Wikipedia sub-communities are self-organized, open to anyone without restriction, and exist in dynamic relationships to one another. Communities grow and shrink as the nature of work on Wikipedia changes [45]. Communities split and specialize [21]. Although Wikipedia sub-communities are not allowed to impose limits on who can participate or develop enforceable rules to govern how they perform their work and interact with one another, projects do develop and document local norms to guide their collaborative work [21]. Many Wikipedia editors inhabit multiple communities simultaneously [64], and these shared members can transfer successful work practices—potentially including behavioral norms—between communities that share similar task focuses [68].

2.3.2   Design decisions and normative exposure. The Teahouse was designed in direct response to the Help Desk and other spaces on Wikipedia where new editors are likely to experience less polite, positive, and personalized opportunities for socialization [42]. An important feature of this design was the ordering of questions on the Q&A board: recent questions were posted at the top of the page, unlike other discussion forums on Wikipedia, in order to make it easier for new participants—question askers and answerers alike—to learn how to participate by viewing examples of recent questions and answers (descriptive norms) [42]. Another key feature of the initial design was the documentation of the "Methods of the Teahouse"[2], a set of five local injunctive norms that were intended to set behavioral expectations for the experienced Wikipedians who answered new editors' questions [42].

The current study focuses on two normative behaviors that set the Teahouse apart from other discussion spaces on Wikipedia: 1) welcoming behavior and 2) policy citation behavior. The Teahouse's injunctive norm around welcoming behavior, as articulated in the Methods, enjoins hosts to *"Welcome everyone with a friendly hello when you answer questions."* The Methods's injunction about proper policy citation behavior instructs hosts to *"Avoid over-linking when responding to questions. Wikipedia policies, procedures, and documentation are overwhelming to many editors. Do your best to explain processes and policies and answer the guest's particular question rather than pointing them to more documentation as a first resort."* This injunction is intended to prompt answerers to make an effort to interpret and explain Wikipedia rules and technology, rather than

---

[2]https://en.wikipedia.org/wiki/Wikipedia:Teahouse/Host_lounge/Expectations
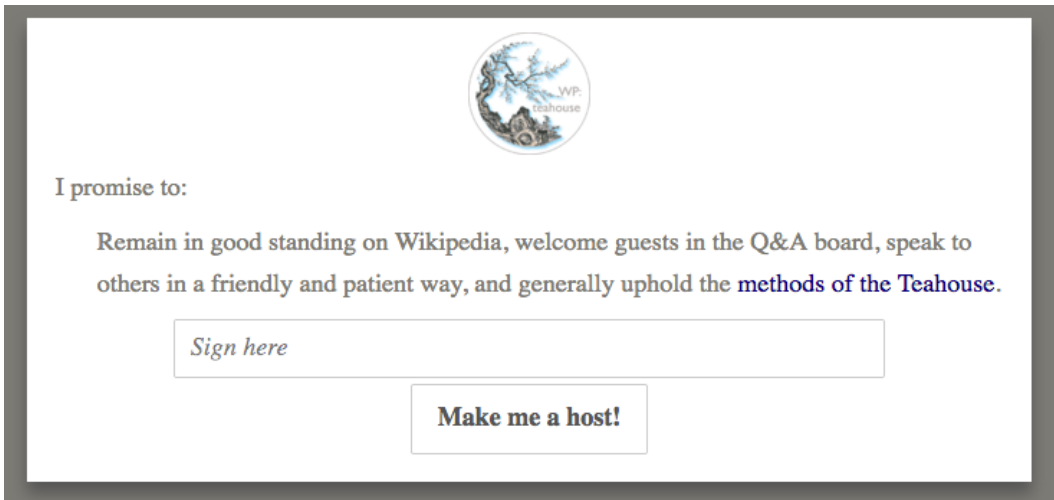
Fig. 2. When creating a host profile, editors are exposed to the Teahouse injunctive norms and asked to follow them.

merely regurgitating them to newcomers or telling them to go "RTFM" [3]. **Figure 1** shows an example of an answer on the Teahouse that reflects both of these injunctive norms.

We selected these injunctive norms for two specific reasons: 1) as we will describe below, proper welcoming behavior is likely to be easier to infer from descriptive norms than proper policy citation behavior, and 2) these examples also afford useful case studies to contrast with conflicting norms on the Help Desk.

Unlike rules in Reddit sub-communities [38], the Teahouse Methods themselves cannot be enforced by banning users or deleting content. Also unlike Subreddit rules [27], the Methods are neither not posted on or linked to from the Q&A board itself. Instead, the Teahouse Methods are only shown to experienced Wikipedians who choose to sign up as project members, or "hosts". Host is a purely informal, elective designation that entails no formal duties and affords no special privileges, and an answerer's host status is not apparent to anyone reading their posts on the Q&A board.

The decision to surface the Methods through the host signup process (**Figure 2**), rather than posting them publicly for all questioners and answerers to see, had the consequence of limiting exposure to the injunctive norms of the Teahouse to a subset of community participants. This may not have been intentional: according to Morgan et al. [43], while the founding members of the Teahouse were all declared hosts, within the first 6 months of the Teahouse's existence many experienced Wikipedia editors began answering questions without signing up as hosts—a pattern which of behavior that is also attested in our dataset.

This design decision created a potential informational asymmetry between host and non-host answerers. While all new answerers have the same opportunity to learn how to answer questions by observing recent behavior on the Q&A board (descriptive norms)—including both behaviors that are congruent with the Methods and behaviors that conflict with the Methods, as well as corrective actions intended to draw attention to norm violations—declared hosts are much more likely to have been exposed to the injunctive norms of the Teahouse directly.

---

[3]"Read the fucking manual," http://www.catb.org/jargon/html/R/RTFM.html

2.3.3   Study design and contribution. We argue that, taken together, the design decisions to limit exposure to the Teahouse Methods to declared hosts and to prominently display plentiful examples of recent answering behavior to all participants, combined with the natural variability in norm compliance expected in any open community with high membership churn and no mechanisms for rule enforcement, presents an opportunity for a natural experiment in the relative influence of descriptive and injunctive norms on behavior that reflects many of the experimental conditions in Cialdini's studies of littering behavior.

The information asymmetry between new non-hosts and hosts allows us to examine individual impact of injunctive norm exposure through documented rules (**H1a**); natural variation in the frequency of recent examples of corrective practices allows us to examine the relative salience of this source of injunctive norms on new hosts and non-hosts (**H1b**); natural variation over time in the frequency of pro-normative behaviors allows us to examine the individual impact of these sources of descriptive norms on behavior of all new answerers under different environmental conditions (**H2**). The combination of information asymmetry and natural variation allow us to examine whether, as Cialdini's findings suggest, Teahouse answerers will be most likely to perform pro-normative behavior when injunctive and descriptive norms are both salient and congruent (**H3**).

The similar focus and shared membership between the Teahouse and the Help Desk also allows us to examine the influence of local and imported norms on behavior. Specifically, we compare normative compliance among Teahouse answerers who also participate in the Help Desk—where answering behavior reflects different descriptive and injunctive norms—with those who only participate in the Teahouse and therefore are unlikely to have been exposed to conflicting norms about how to answer a question (**RQ1**). Among answerers who have participated in both forums, we examine whether compliance is influenced by the order of exposure: Teahouse first, or Help Desk first.

We believe that our study represents the first attempt to apply the Focus Theory to a long-term field study of an active online community and to directly measure the influence of multiple sources of injunctive and descriptive normative influence—individually and jointly—on naturally occurring computer-mediated social behavior. Any attempt to replicate findings from a controlled study under naturalistic setting involves assumptions and variables that cannot be controlled for. We attempt to address some anticipated limitations of our study design by taking a mixed-methods approach: in section **5.1** we interrogate several of our assumptions and contextualize our findings through interviews with seven Teahouse community members. We also highlight several additional limitations of our study (Section **5.2**), which we hope will serve as guideposts for further research—including controlled experiments, surveys, and ethnographic investigations—into the interplay between different sources of normative influence in online communities.

In the Methods section below, we provide a detailed rationale for our study design decisions and describe how we attempt to control for several potential sources of error in sampling, analysis, and the operationalization of key study variables.

3   METHOD

In this section, we first define how we operationalize our hypotheses and research question in the context of two specific injunctive norms of the Teahouse. We then describe our method for collecting and identifying instances of normative behavior within a corpus of Teahouse discussion posts made between 2012 and 2016, and the selection and development of our statistical models.

3.1   Operationalization of norm exposure and normative behavior

Fig. 3. A question from a new editor at the Help Desk, with an answer that reflects local descriptive norms.

3.1.1   Welcoming behavior. We chose to examine welcoming behavior on the Teahouse because it affords us the ability to examine two distinct mechanisms for injunctive norm salience: through exposure to an injunctive norm though documented behavioral guidelines, and through exposure to examples of corrective behavior. We model each of these injunctive mechanisms separately, as *welcome in first reply* and *welcome in subsequent replies* and discuss their operationalization in greater detail below.

Focusing on welcoming allows us to track how well actual behavior (descriptive norms) aligns with social expectations at any given time. Welcoming is a very visible behavior that follows a regular pattern. Even someone who has not been directly exposed to the injunctive norms of the Teahouse can likely infer that welcoming is customary—if not obligatory—by observing the behavior in recent answers. Our focus on welcoming behavior allows us to examine potential impacts of conflict between the descriptive norms of the Teahouse and the Help Desk. Wikipedia's community culture features a high degree of conflict and criticism [8], and the Teahouse was designed to provide a space where new editors were somewhat shielded from these phenomena. Conversations between new editors and answerers have been shown to be longer and more interactive at the Teahouse than at the Help Desk [43], indicating the presence of different descriptive norms around answering. Furthermore, unlike the Teahouse, the Help Desk lacks an explicit injunctive norm about welcoming newcomers. Therefore it is reasonable to assume that welcoming will be relatively less common at the Help Desk. **Figure 3** shows a typical Help Desk interaction between a new editor and an experienced Wikipedian.

3.1.2   Policy citation behavior. We chose to examine policy citation behavior because it affords an instructive contrast to welcoming behavior. Proper policy citation behavior is less easy to discern than welcoming—it is relatively more difficult to infer from examples of past behavior (descriptive norms) that including fewer links to policy pages in a reply is common practice in this sub-community. Therefore, we do not expect to see that recent behavioral examples *alone* will have as strong an effect in predicting answerers' policy citation behavior as in the case of welcoming behavior. This contrast is valuable because it allows us to measure the effect of visibility on the salience of a descriptive norm.

Furthermore, compared with the Teahouse the Help Desk has an active and opposite injunctive norm concerning policy citation. Help Desk answerers are *encouraged*, when in doubt, to provide additional citations to policy and help documentation in their replies because "directing questioners to these pages ensures they get the most accurate instructions" [4]. Thus with policy citations we can potentially observe a different type of normative conflict compared to welcoming: while in the above context of welcoming behavior there was a conflict between an injunctive norm on the Teahouse and a descriptive norm on the Help Desk, in the context of policy citation, there are two conflicting injunctive norms.

---

[4]https://en.wikipedia.org/wiki/Wikipedia:Help_desk/How_to_answer

3.1.3 Operationalization of welcoming and policy citation. To evaluate whether a reply to a question contains a welcome, we perform a pattern match that looks for the phrase "Welcome to the Teahouse" and/or a combination of a greeting "Hi/Hello/Hey/Howdy" and a link back to the answerer's user name, e.g. "[[User name]]" at the beginning of the answer. All comparisons are case-insensitive. This is a binary variable: 1 indicates a matching welcome string, 0 indicates no welcome in the reply. Approximately 52% of first responses to a question contained a welcome in our dataset.

To evaluate over-linking of policy citations, we parse all the relevant hyperlinked citations in the text of each reply (links that point to the WP, Wikipedia, MOS or Help namespaces) and resolve all redirects to generate a canonical list of URLs. If the canonical URL matches a page in the policy, guideline, essay, or help page categories, we count this as a policy citation. Teahouse replies to questions vary in length. We therefore normalize the raw count of policy citations by the number of words in the reply. To obtain an accurate word count, we exclude signatures, date stamps, and any hyperlinks in the body of the reply. An average response to a question in our dataset contains 1 policy citation per 100 words, with a standard deviation (SD) of 4 (rounded to the nearest whole number).

3.1.4 Operationalization of injunctive norm exposure. Because this is a field study, we are interested in operationalizing only conditions in which we are reasonably certain that a prospective answerer was exposed to the injunctive norms of the Teahouse. Thus we operationalize exposure to injunctive norms via two different mechanisms: 1) exposure to explicitly stated injunctive norms, and 2) exposure to recent examples of corrective action that highlight a specific injunctive norm.

The first mechanism is exposure to the Methods of the Teahouse. As described above, Teahouse participants are asked to read and abide by these social expectations when they sign up as a host (**Figure 2**). Although it is possible for a Teahouse answerer to seek out or stumble upon the Methods without first attempting to join, the document is not easily found—it is not prominently linked from the Q&A board itself nor commonly referenced in discussion threads. Thus we operationalize the first mechanism for injunctive norm exposure as the presence of a host profile for a given answerer in the project member list. We created a binary variable with 1 representing users who had a host profile (206 unique users) and 0 for users who did not (6854 unique users) who we will refer to as non-hosts. We excluded from analysis users who attempted to add their profile to the member list but were subsequently reverted [5]. This removed 114 distinct users, with 280 posts between them.

The second mechanism for injunctive norm exposure is the prevalence of recent examples of corrective action. Welcoming a questioner in the very first response sets the tone for the rest of the discussion thread, and signals that before anything else, the asker is in a place where their questions are encouraged. Answerers cannot be blocked or otherwise penalized for neglecting to welcome. However, other answerers may choose to perform corrective actions in order increase the salience of this injunctive norm to the first answerer (and anyone reading the thread): should the first respondent fail to welcome the asker with a friendly greeting, a second respondent may "correct" this behavior by following up with a reply that includes an explicit welcome together with any additional comments on the question at hand.

Thus we operationalized the second, corrective, mechanism as exposure to welcoming replies in subsequent posts in recent threads, if the first response to a thread was not welcoming. Approximately 10% of subsequent responses in our dataset contained a welcome if the first reply was not welcoming. Because pro- or anti-normative policy citation behavior is difficult to discern, we

---

[5]Further inspection suggests that the majority of these rejected hosts were very new editors who had recently been directed to the Teahouse to get help. We chose to exclude these because such users are likely to exhibit different patterns of behavior from both accepted hosts and non-host answerers who never created a profile.

believe that it is unlikely that there will be similar forms of corrective action associated with it. Therefore we do not analyze the relative frequency of policy links in subsequent replies.

3.1.5   Operationalization of descriptive norm exposure. We operationalize descriptive norms as the prevalence of each of the behaviors examined in the study within one archive page of a given answer. Since the Teahouse is a public Q&A board and question threads are listed in reverse-chronological order, new answerers will see recent behavioral examples when they join. When the front page reaches a threshold of content (usually about once per week), older threads are automatically moved to a numbered archive. An average archive page contains 64 posts (SD=15; both questions and replies), and 17 threads (SD=5). Archive links are also prominently displayed in the sidebar of the Q&A board, sorted by recency, allowing any editor to easily browse the most recently archived threads.

Previous research has suggested that potential participants in online communities may lurk for 10 or more days before participating [58]. Therefore, to evaluate the effect of descriptive norms, we examine the archive pages directly preceding an answer and count the average occurrence of examples of the social behavior we are interested in on each archive page. Preliminary analysis has found no effects beyond two pages of archives directly preceeding an answer, thus our models presented in the Findings section only examine the prevalence descriptive norms in the previous archive page and two archive pages back.

3.1.6   Operationalization of normative conflict. To evaluate manifestations of normative conflict, we contrast the behavior of Teahouse users who have no direct experience with the Help Desk at all (*Teahouse-only*) with two different kinds of membership overlap: 1) shared members who first began contributing to the Teahouse and were therefore likely exposed to (injunctive and/or descriptive) Teahouse norms first (*Teahouse-first*), and 2) shared members who were exposed to Help Desk norms first (*Help-Desk-first*), and who are likely to draw on that prior experience when determining how to answer questions at the Teahouse because they experience less behavioral ambiguity. There are 6433 unique Teahouse-only users in our dataset, 230 unique Teahouse-first users, and 397 unique Help-Desk-first users.

3.2   Data collection

Using the Wikipedia API and the WikiChatter Python library (version 0.2.3), we collected and parsed 15435 threads from the Teahouse Q&A archives [6]. Each thread in our dataset consists of a question that was asked between February 27th, 2012 (when the Teahouse was founded) and June 30th, 2016, and which received one or more replies.

We are specifically interested in how answerers respond to new Wikipedia editors, as newcomer support is the focus of the Teahouse. Therefore, we excluded 3954 threads from our sample that were asked by experienced Wikipedia editors, as the patterns of welcoming behavior and policy citation are likely to be different. Following Panciera et al. [51], we consider an editor to be experienced if they had made at least 250 edits at the time they asked their question. It is also possible for Wikipedia editors to change their Wikipedia username, potentially confounding our examination of behavior over time. As we do not have a reliable way of connecting old and new user names by the same author, we excluded 921 posts across 733 questions by authors for whom the date of their first Wikipedia edit (according to the API) was later than the timestamp of their first post in the archive page.

We performed additional data cleaning, removing duplicate posts within question threads that occurred due to an error with the Teahouse archiving bot (77 posts in total across 26 threads); posts

---

[6]https://en.wikipedia.org/wiki/Wikipedia:Teahouse/Questions/Archive_Index

(either questions or replies) where the author did not have a valid user ID we could associate with an account (3488 posts in total across 52 threads); or a first post date available in the user database (57 posts across 38 threads). Our final dataset contained 10206 threads, and 29847 replies by 7012 repliers (206 hosts and 6854 non-hosts).

3.3 Data analysis

As our dataset spans 4 years, we first report on some time-specific trends we observed in the data: an average of 58 (SD=95) new Teahouse answerers per 6 month time period beginning in August 2012. In each time period, there were significantly more new answerers active on the site than answerers who joined in previous time periods: an average 85.6% were new (SD=1.8%). This supports our earlier assumption about significant membership churn in this Wikipedia sub-community.

To address our hypotheses, we ran a series of multiple regressions predicting 1) the likelihood of welcoming in a first reply (using a binary logistic regression), 2) the likelihood of welcoming in a subsequent reply if the first reply was not welcoming (using a binary logistic regression) and 3) the proportion of policy citations to post length (using a negative binomial regression). We employed a binary logistic regression available in the "base" package in R in the first two models because the welcome variable has only two states: 0 or 1 (welcome absent, or welcome present). In the third model, we employed a negative binomial regression (from the MASS package) rather than a traditional linear ordinary least squares regression because the number of policy citations in a post is a count variable that is over-dispersed (in other words, the variance is significantly larger than the mean). We ran a dispersion test from the R package 'AER' [33], and found significant dispersion (Dispersion=10.36; z=2.41; p<0.001). We also compared our negative binomial model to a Poisson model that does not take into account over-dispersion, and found the negative binomial model to have significantly better fit (χ^2=16692;p <0.001).

For all three dependent variables, we used the same independent variables. In the first step, we entered only control variables: age of post (in days since the founding of the Teahouse), total number of posts by the author on the Teahouse (M=8, SD=51), total number of posts by the author on the Help Desk (M=8, SD=184). In the second step, we entered our variables of interest: presence of pro-normative behavior in the two most recent archive pages (exposure to descriptive norms), host status (exposure to injunctive norms), and the first site the user posted to—Teahouse first, Help Desk first, or only posted on the Teahouse (exposure to conflicting norms). In the third step, we entered an interaction term between each of the descriptive and injunctive norms.

To afford direct comparison of effect sizes of different variables, we standardized (M=0, SD=1) all continuous variables in all the models we ran because of the different scales on which our independent variables are measured. Standardization of continuous variables used in interaction calculations also affords interaction effects that are orthogonal to main effects, helping us to interpret main effects in the presence of interaction effects in our models. The output of our models is presented in **Table 1**.

4 FINDINGS

We describe our findings below, and provide a summary of broken down by hypothesis in **Table 2**.

4.1 H1: Effects of injunctive norm exposure

We expected to see that across both welcoming in first reply and policy citations, prior exposure to explicitly stated injunctive norms (through setting up a Host profile) would be more strongly associated with pro-normative behavior, that is greater welcoming and lower citation proportions in an given answer respectively (**H1a**). Our findings support this hypothesis: all other things being equal, Hosts are 2.3 times more likely to welcome a newcomer in a first reply, and cite 14% fewer

	Welcome in 1st reply (Odds Ratio)			Welcome in subs. reply (Odds Ratio)			Policy citation (Coefficients)		
	Step 1	Step 2	Step 3	Step 1	Step 2	Step 3	Step 1	Step 2	Step 3
Age of post	0.83***	0.93**	0.92**	0.92**	1.01	1.01	0.16***	0.12***	0.12***
# of Help Desk posts	0.72***	0.82***	0.82***	0.72***	0.83***	0.83***	0.13***	0.08***	0.08***
# of Teahouse posts	2.01***	2.14***	2.14***	1.54***	1.63***	1.63***	0.10***	0.03	0.03
Host (vs non-Host)		2.26***	2.25***		2.78***	2.79***		-0.14**	-0.14**
Previous archive page		1.13***	1.05***		1.10***	1.11**		0.01	-0.01
2 archive pages back		1.12***	1.12***		1.06*	1.06*		0.03	0.03
Teahouse only (vs Help Desk first)		1.23**	1.24**		1.18 p=0.07	1.18 p=0.07		-0.51***	-0.52***
Teahouse first (vs Help Desk first)		1.37***	1.38***		2.24***	2.23***		-0.22***	-0.23***
Host * Previous archive page			1.20***			0.98			0.05

Note: *** = p < 0.001, ** = p < 0.01, * = p < 0.05

Table 1. Three kinds of pro-normative behavior on the Teahouse predicted by salient injunctive and descriptive Teahouse norms, their interaction, and conflicts with norms of the Help Desk.

policies than non-Hosts. Similarly, Hosts were 2.8 times more likely to welcome in a subsequent reply (correcting a lack of welcoming behavior) than non-Hosts (**H1b**). All the above effects were significant at least at p < 0.01.

4.2 H2: Effects of descriptive norm exposure

We expected to see that the prevailing descriptive norms would be a significant predictor of future normative behavior for all answerers when they are easy to observe (as in welcoming behavior), but not when they are relatively more difficult to infer (as in policy citation behavior).

Our findings fully support H2. Participants were 13% more likely to welcome in first replies for every SD increase in the amount of welcomes contained in the archive page directly before the post, as well as 12% more likely for every SD increase in welcomes two archive pages back. Similarly, regardless of "host" status, participants were 11% more likely to welcome in subsequent replies for each SD increase in the amount of welcomes in subsequent replies in a previous archive page, and 6% more for each SD increase in welcomes in subsequent replies two archives back. All the above effects were significant at least at p < 0.05. By contrast, and in line with our expectations, we do not see a significant effect of either exposure to the previous archive page, or two archive pages back, on the number of policy citations an answer contains.

4.3 H3: Effect of exposure to congruent injunctive and descriptive norms

We expected to see that when descriptive and injunctive norms are both salient and congruent, they will interact such that the likelihood to perform the behavior will be greater when both are present than when only either one is present.

Our results partially support H3. We see that for welcomes in first replies (a condition in which both descriptive and injunctive norms are salient), there is a significant and positive interaction. **Figure 4** demonstrates the interaction in greater detail—hosts who are exposed to pro-welcoming descriptive norms in the previous archive page are significantly more likely to welcome than non-hosts who are exposed to the same amount of welcoming behavior.

In both policy citation and welcome in subsequent reply conditions, no interaction is observed. This is expected in the case of policy citations because descriptive norms are not salient and therefore should not interact with injunctive norms. In the case of welcomes in subsequent replies,
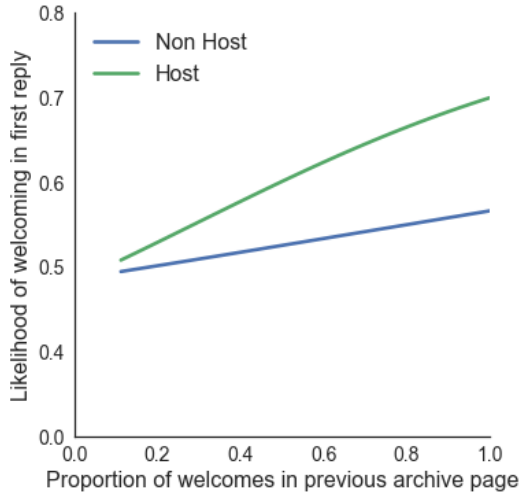
Fig. 4. Interaction effect between injunctive and descriptive norms.

we expected hosts who are exposed to welcomes in subsequent replies (when the first reply lacked a welcome) in previous archive pages to also be more likely to correct the behavior by welcoming in a subsequent reply, compared to non-hosts. We did not see a significant increase in this behavior, possibly because there are relatively fewer welcomes in subsequent replies in our dataset, and therefore the likelihood that examples of this behavior will be present in a previous archive page is fairly low.

We did not observe any interactions between host status and behavior examples two archive pages before the answer in any of the models, thus for the sake of brevity we do not report the details of these results.

4.4 RQ1: Effects of normative conflict

We operationalized normative conflict as membership overlap between the Teahouse and the Help Desk. We looked at whether joining the Teahouse or the Help Desk first would impact editor behavior on the Teahouse. We observe a number of interesting effects that are in line with Cialdini's findings and the tenets of the Focus Theory.

We found that there is a greater likelihood of welcoming in a first reply among answers who participated only in the Teahouse and answers who participated in the Teahouse first, compared to answers who participated in the Help Desk first. Interestingly, we found that answers who began on the Teahouse but have also contributed to the Help Desk have greater odds of welcoming a newcomer (37% more likely) than those who only contribute to the Teahouse (23% more likely), relative to Teahouse editors who began on the Help Desk. Both effects are moderately strong and significant at p < 0.001.

We observe a similar pattern for welcomes in subsequent replies. Again, there is a greater likelihood of welcoming in a subsequent reply for Teahouse-first and Teahouse-only answers. Also similarly, answers who joined the Teahouse first but also contributed to the Help Desk have greater odds of welcoming a newcomer (2.24 times, or 124% more likely) than those who only contribute to the Teahouse (18% more likely). While the latter is a fairly weak effect and has a

marginally significant p value of 0.07, the former is a strong predictor (the second most strong in the model behind injunctive norms) and significant at p < 0.001.

Cialdini's Focus Theory may offer an explanation for these findings. In both cases, there is a conflicting descriptive norm in the Help Desk of less friendly and welcoming behavior (see **Figure 3**). In contrast, the Teahouse has an active injunctive norm that encourages welcoming newcomers, and descriptive norms that generally reinforce this injunctive norm. It is possible that Teahouse answerers who also contribute to the Help Desk become more sensitized to the need to be welcoming by witnessing how new editors are treated on the Help Desk—thus leading to a greater likelihood to welcome at the first reply and in subsequent replies on the Teahouse. In other words, because they are exposed to a conflicting descriptive norm on the Help Desk, they are more likely to enforce the pro-welcoming norm in the Teahouse community. Our follow up interviews (Section **5.1** below) provide some evidence to support this interpretation.

Second, we look at the number of policy citations in replies. In this case, there are conflicting injunctive norms between the Teahouse and Help Desk. The Teahouse has a strong anti-linking injunctive norm, asking editors to offer accessible and personalized explanations specific to the question, while the the Help Desk has a pro-linking injunctive norm. Consistent with this, we see that both Teahouse only, and Teahouse first participants are significantly less likely to cite policies than Help Desk first participants. However, unlike the previous two behavior examples we observed, in this injunctive<>injunctive conflict we see a reversed pattern: Teahouse only and Teahouse first editors both cite relatively fewer policies than Help Desk first editors, but Teahouse only editors cite the fewest (51% and 22% fewer than Help Desk first, respectively). The former is a moderately strong effect, and the later a moderate effect, both significant at p < 0.001.

One possible explanation for this finding may be that when two injunctive norms are in conflict, such as when a Teahouse first editor begins contributing to the Help Desk, the original norm they are exposed to gets diluted. Teahouse editors who later join the Help Desk from the Teahouse may notice that citations are actively encouraged there, reducing the impact of the original injunctive norm of "avoiding over-linking"—if others are doing it this way, perhaps it is less of a big deal.

5 DISCUSSION

Our study is predicated on the assumption that answerers at the Teahouse perform normative behavior because they identify the Teahouse as a social context with distinct norms, and that the primary mechanisms by which they are made aware of these norms are through observation of recent behavior, exposure to the Methods document, or exposure to examples of corrective action. Above, we present evidence that answerer behavior indeed is mediated by such an understanding, and that the mechanisms by which people are exposed to local norms both individually and jointly mediate likelihood of performing pro-normative behavior. Below, we describe a follow-up interview study we undertook to triangulate the results of our quantitative analysis, identify limitations and opportunities for future work, and propose a set of theoretical and practical implications of the current work.

5.1 Follow-up interviews

In order to contextualize our main findings and interrogate the assumptions of normative salience and exposure mechanisms implicit in our study design, we performed interviews with Teahouse answerers. Interview candidates were selected by performing a query of the most recently-active Teahouse answerers using Quarry, a public SQL interface maintained by the Wikimedia Foundation for accessing publicly-available Wikipedia edit metadata.The resulting list was then filtered using other public edit history metadata to contain only editors who had participated in the Teahouse for at least 1 year, and had posted at least 10 answers within the previous month. The first author—who

H1	Injunctive norm exposure	Support	Hosts welcome more in first replies (*H1a*), cite fewer policies (*H1a*), and welcome more in subsequent replies (*H1b*) than non-hosts.
H2	Descriptive norm exposure	Support	When pro-normative behavior is more prevalent, all answerers welcome more in first and subsequent replies (*high-salience descriptive norm*) and do not cite fewer policies (*low-salience descriptive norm*).
H3	Normative congruence	Partial support	When descriptive norms reflect injunctive norms, hosts welcome more in first replies and do not cite fewer policies (*H3 supported*); Hosts do not welcome more in subsequent replies (*H3 not supported*).
RQ1	Normative conflict	n/a	Teahouse-only and Teahouse-first answerers welcome more in first replies and subsequent replies (*injunctive<>descriptive conflict*), and cite relatively fewer policies (*injunctive<>injunctive conflict*).

Table 2. Summary of findings

	Is host	*Teahouse join date*	*Total Teahouse edits*
P1		07-2015	2955
P2	x	06-2012	2774
P3	x	03-2012	5400
P4		06-2016	1132
P5		07-2013	1115
P6	x	01-2014	1813
P7	x	03-2017	628

Table 3. Interviewee host status and contribution history

participates in the Teahouse in a 'backstage' role, as a volunteer maintainer of the Teahouse's technical infrastructure under their real name—reviewed the candidate list and reached out to 12 answerers (6 hosts and 6 non-hosts) over email or user talkpage message to request an interview, disclosing themselves as a researcher and stating their affiliations.

Invited candidates were informed that the interviews would focus on the way that they approached their work at the Teahouse, their history of participation, and their understanding of what kinds of behavior differentiated a 'good' answer from a 'bad' one. The interviews were semi-structured in nature in order to allow the interviewer to follow up on emerging lines of inquiry. The focus of the interview was broad and explicitly non-directive in order to elicit rich descriptions of norm-related attitudes and behaviors without biasing participants towards the epistemological stance of the researchers, i.e. without directly asking participants to describe their behavior in etic terms ('injunctive', 'descriptive', 'norms', 'salient').

Six interviews (**Table 3**) were conducted over voice and/or video chat services, and one interview was conducted over email. Interviewees were generally very experienced answerers, a mix of declared hosts and non-hosts, with very different histories of participation on the Teahouse and (in some cases) the Help Desk as well. Interview participants were not compensated. Prior to each interview, all participants signed consent forms indicating whether or not they were willing to publicly share direct quotes from the interview session. Spoken interviews were audio recorded and transcribed, and all transcripts were subsequently analyzed for recurring patterns and themes [59] related to the stated interview focus, with special attention paid to utterance that provided insight into personal motivations, initial socialization experiences, and attitudes towards other community members.

Below, we describe and discuss how interviewees characterized the overall normative character of the Teahouse, whether they experience Teahouse norms as primarily descriptive or injunctive, and how they understand the distinction between hosts and non-hosts as it relates to the performance of pro-normative behavior.

5.1.1   Do answerers perceive the normative character of the Teahouse? Responses from all interviewees demonstrate a shared understanding that the Teahouse possesses distinct local norms, and that these norms were generally followed. When asked about what distinguishes a 'good' answer from a 'bad' one, all interviewees pointed to the importance of welcoming guests or avoiding over-linking, and often both. These norms distinguish the Teahouse from other parts of Wikipedia. According to P3:

"Whenever a new question comes up and I'm the first person to respond I always say 'welcome to the Teahouse'... I just think that that's the thing that sets the Teahouse off from other more confrontational spaces on Wikipedia."

P6 and P7 further identified these specific answering norms as differentiators between the Teahouse and the Help Desk. P2 also observed that at the Teahouse *"hosts follow the policy of trying to answer in words rather than give links, [and] to some degrees that leaks over to the Help Desk as well."* This observation suggests an intriguing direction for future research. Although investigating how Teahouse participation might influence the descriptive and injunctive norms of the Help Desk is outside the scope of our current study, this observation suggests that the normative influence between the Help Desk and Teahouse may in fact be both reciprocal and dynamic.

5.1.2   How do answerers become aware of the Teahouse norms? Because our study involves analysis of historical data from an existing online community, we are unable to control for all variables that may bear on our findings, such as the influence of alternate mechanisms of normative exposure. Although we suspect these mechanisms exist, our data do not attest to their prevalence. In order to probe whether our operationalizations of descriptive and injunctive norm exposure reflect common lived experience, interviewees were asked a series of questions aimed at eliciting, in a non-directive way, how they first became aware of the Teahouse norms.

Overall, our interviewees confirmed that the mechanisms by which they became aware of Teahouse norms were through either observation of past behavior or direct exposure to the "Methods". P4 noted specifically that the public, shared Q&A board provides a *"mentoring opportunity for hosts, as well as newcomers"* because it allows them to learn how to answer questions by reading previous responses. P6 and P7 (both hosts) also mentioned that they saw the "Methods" very early in their tenure.

However, P2 related they were contacted on their personal userpage after leaving their first (terse, link-filled) answer by a Teahouse host who pointed them to the Methods. This story indicates that some degree of direct interaction does exist between answerers, and that these personal interactions sometimes serve as an alternate mechanism for injunctive norm exposure that is not accounted for

in our study. However, additional observations from this interviewee and several others suggest that such direct discussions of injunctive norms among answerers, or indeed direct interaction in general, is a relatively rare occurrence. P7 characterized the relationship among the answerers as an *"invisible camaraderie"*, noting that while answerers are almost unfailingly collegial in their interactions on the Q&A board itself, he did not feel *"a sense of community among the hosts. We sit in chairs side by side talking to an audience. We should talk more amongst ourselves about how best to do something."*

None of the interviewees mentioned that they had been made aware of the welcoming norm by being soft-sanctioned in a subsequent reply *(H1b)*. However P1, P6, and P7 confirmed that they themselves had included welcomes in replies when the first reply lacked one. As P1 explained:

"Sometimes you get slightly bite-y responses from some of the hosts... [other guests] see that... and I wouldn't want them to be put off of answering questions, probably why I go out of my way to reach out and welcome people."

While this remark does not establish whether welcoming in subsequent replies is intended (by the performer) or perceived (by the target) as an explicit injunctive cue, it does demonstrate that the performance of that action can be driven by normative concerns.

5.1.3   How do answerers experience the host/non-host distinction? By design, it is not necessary to create a host profile in order to answer questions at the Teahouse. Only about 10% of answerers in our dataset chose to do so, and (as **Table 3** shows) both hosts and non-hosts may become highly active and committed answerers. This raises an important question about whether the host/non-host distinction reflects meaningful, persistent differences in answerers' orientation towards the Teahouse norms beyond the initial exposure/socialization period. Are those who sign up as hosts more likely to continue to perceive answering norms as injunctive (and therefore in some sense obligatory) even years later? Are veteran non-hosts still more likely to consider answering norms to be non-binding reflections of common practice rather than social injunctions—even if they are subsequently made aware of the Methods?

We asked a series of open-ended questions aimed at learning more about what this distinction means in an emic sense. We found that most of our interviewees used the term 'host' to refer to any consistently-active answerer, whether or not that individual had created a host profile (a tendency reflected in the quotes presented above).

However, P6 (a host according to our operationalization of the term) was one of several interviewees who explicitly described *"a distinction between someone who answers and someone who is quote formally a host... for a lot of people, signing up means a certain positive attitude to the job, and stronger commitment."*

This suggests that, at least for this participant, the host/non-host distinction is meaningful in ways that are relevant to the performance of normative behavior—i.e. that creating a host profile makes injunctive norms salient through a combination of direct exposure to the Methods and a public commitment to abide by them. We interpret this statement as evidence that the norms of the Teahouse are construed as mutually-understood social expectations by *de jure* hosts which are likely to influence their actions in different ways than *de facto* hosts of similar tenure, consistent with our findings from *H1* and *H3*.

We further speculate that the practice of referring to all active answerers as 'hosts' (which this participant did as well, elsewhere in the conversation) may function as a convenient shorthand, or as a rhetorical device for communicating that the Teahouse's welcoming and inclusive ethos extends to interactions with other experienced editors as well as newcomers who ask questions.

We also observed an additional intriguing, although ultimately not conclusive, indication that the host/non-host distinction is relevant to whether Teahouse norms are experienced as social

obligations rather than mere personal preferences in the way interviewees spoke about pro-normative behaviors they observed or performed. All four hosts interviewed explicitly framed pro-normative answering behavior in injunctive terms at at least one point during the conversation.

For example, P3 said: *"the ethos [of the Teahouse] is that even if somebody's a jackass, that you should greet them, be friendly with them, and explain the policies in a fairly friendly and informative/helpful way... I try to lead by example."*

In contrast, none of the non-hosts interviewed characterized pro-normative behavior this way. P5, a non-host who participated at the Help Desk first, acknowledges that he behaves differently at the two spaces but frames that distinction as a personal choice rather than a obligation: *"I participated at the Teahouse as I had learned to do at the Help Desk. In the Teahouse... I'll try harder to be polite and welcoming."*

In fact, the only non-host (P1) who referred to the Methods at all in their interview characterized them in a way that de-emphasizes their obligatory nature: *"I'm vaguely aware that there's some advice somewhere that I haven't read for a long time... but I do generally provide at least 1 link to a relevant guideline or policy [along with] some brief summary in layman's terms and how it relates to the particular question."*

5.2 Limitations

The fact that experienced non-hosts still frame their answering behavior in different ways than equally experienced hosts bolsters our interpretation of our findings from *H1* and *H3*—that hosts abide by the welcoming and policy-linking norms because they experience them as social expectations, rather than because they are more experienced with or invested in the activity of answering new editors' questions. However, we cannot eliminate the possibility that individual factors antecedent to an answerer's participation in the Teahouse may influence the decision to create a host profile in the first place. We acknowledge the possibility that overall greater norm adherence we found among the host cohort may be partially attributable to such factors.

One possible confounding factor is *personal norms* [56]—standards of behavior that are based on an individual's internal values, rather than social expectations. An experiment by Cialdini et al. [6] found that injunctive norm exposure resulted on a strong and consistent reduction in littering irrespective of an individual's personal littering norms. However, the design of the Teahouse, and the design of our study, make it difficult to measure the distinct influence of personal norms on behavior.

Injunctive normative compliance is also closely associated with group identification. People who chose to add a profile to the host list may be signaling a stronger sense of identification with the Teahouse community, and may therefore be more likely to enforce the social norms of the group [61]. Member lists are a common feature among Wikipedia sub-communities, and these lists have been used as a convenient sampling strategy for examining how Wikiproject communities mediate article editing activity [32]. However, previous studies of the relationship between declared membership and behaviors related to group identification have found equivocal results. Analysis across dozens of WikiProject performed by Gilbert [24] and Morgan [44] showed few significant differences between the participation patterns of members and non-members in the project discussion forums, and an interview-based study by Forte [19] found evidence that, at least in some cases, active and invested sub-community participants do not feel the need to declare membership. Nevertheless, we cannot exclude the possibility that act of joining the Teahouse as a host provides additional incentives for performing normative behaviors beyond the impact of exposure to the Teahouse Methods. Future work could attempt to tease these sources of normative influence apart by surveying new community participants upon arrival, and measure the influence of these factors on behavior under different injunctive and descriptive norm exposure conditions.

Our work, like Cialdini's, explores pro-social behavior—that is, compliance with behavior that is socially desirable. Previous research on factors that influence anti-social behavior in both online and offline contexts (such as college drinking or digital piracy) shows the in some cases descriptive norms are more influential than injunctive norms [28, 55]. One explanation for this difference with our findings may be that when behavior is *not* considered acceptable in society at large, but very common in a local social context (as in college drinking), descriptive norms can indirectly indicate some level of social acceptance from an individual's more immediate social circle—if everyone I know is doing it, it must be OK. The phenomenon may also translate to behavior like trolling that is generally considered anti-social in most online communities—despite the presence of injunctive norms that cast this behavior is anti-normative, if the behavior is very common on the platform descriptive norm may be a more powerful social influencing mechanism. Contrasting the relative power of injunctive versus descriptive norms in influencing pro- and anti-social behavior is a fruitful area for future experimental work on normative influence in online communities.

5.2.1 Analytical assumptions. Our analysis assumes that any time effects observed over 4 years of Teahouse existence are linear (we model these effects with post age). This is a reasonable assumption because initial analysis performed at 6 month intervals showed relatively consistent patterns of membership, post volume, and turnover. The host to non-host ratio varies through time, but the amount of overall welcoming does not, thus we do not believe this to significantly impact our findings.

We also assume that the time of an editor's first post on the Teahouse is within a few weeks (at most) of their first exposure to the descriptive norms of the sub-community, allowing us to examine their behavior in the context of the prevailing descriptive norms at the time of their first (and subsequent) posts. Given the size of our sample and the duration of the study period, it is likely that some answerers in our sample lurked for longer periods before posting their first answer. Unfortunately, we lack access to historical data on Wikipedia user browsing behavior. However, we believe modeling active participation within a consistent exposure window provides a reasonable approximation, especially over a large sample of answerers.

5.3 Implications

The present work examined several types of normative influences on behavior in a Wikipedia sub-community: explicitly stated behavioral expectations, visible behavior examples from other community members, and conflicting examples of expectations and behavior from a similar community with significant membership overlap. In doing so, we leveraged the Focus Theory of Normative Influence in an open collaboration context to empirically examine the circumstances in which multiple sources of normative influence become salient, and mechanisms by which they may work in both congruent and conflicting ways. Our work has both theoretical implications and practical recommendations for system design.

5.3.1 Theoretical implications. The present work finds support in this online setting for many predictions of the Focus Theory. Taken together, our findings support theory-based expectations that normative compliance is the product of multiple interrelated sources of influence, and that there are important differences in the mechanisms by which various social norms influence behavior.

We found evidence that the descriptive and injunctive norms of the Teahouse represent distinct sources of influence, and that when both sets of norms are congruent and salient, they produce a greater likelihood to perform behavior compared to each set of norms individually. We also found evidence that exposure to injunctive norms trumps descriptive norms even when the injunctive norms are not prominently displayed on the site at all times. These findings are consistent with research that has applied the Focus Theory of Normative Conduct to offline settings. For example,

Cialdini finds that once someone has been made aware that an injunctive norm exists in a particular setting, that "activated" norm can remain salient to them even without repeated exposure. Our findings suggest that this can be true for online settings as well, reflecting Harrison and Dourish's assertion that online environments can have a sense of 'place' similar to physical locations [26], and that this sense of place is heavily mediated by the salience of normative cues—especially injunctive ones. While participants also learn from descriptive norms in online groups, our findings provide evidence for the contingency and transience of descriptive norms. On their own, descriptive norms can be difficult to recognize as such (e.g. determining the 'normal' number of policy citations to include in your answer), and their influence does not persist as long—in this study, we do not observe any effects of descriptive norm exposure beyond cues present two archive pages back.

Our work also examines conditions of normative influence, group membership overlap, that have not yet been addressed within the context of the Focus Theory. Existing work on transfer of practices in online communities (e.g. [65, 67, 68]) has largely focused on the positive implications of membership-overlap, such as more efficient onboarding. The current study provides evidence that membership overlap between similar communities can also cause normative conflict, and that in certain situations this conflict can result in lower overall normative compliance. In particular, we show that the type of norms that come into conflict (injunctive<>injunctive versus injunctive<>descriptive) matter, as do which norms a user was exposed to first. Our findings suggest more work needs to be done to understand the potential negative effects of membership overlap in online volunteer communities, particularly from a community health point of view.

Taken together, our findings suggest that studies of social norms in online communities should attempt to account for multiple sources of normative influence within a community, should differentiate between descriptive and injunctive norms, and should attend to the potential interactions between local and imported norms on the behavior of individuals and on the evolution of a community's normative character over time.

5.3.2 Practical implications. We believe that our work also has practical implications for online community design. In particular, our findings show that setting concrete behavioral expectations can be valuable even if those expectations are not enforceable. Open collaborations often lack the means to sanction bad behavior, or are reticent to use them because they rely on voluntary contribution and don't want to drive people away through overzealous moderation. Researchers have argued that posting rules or mission statements may be unnecessary for community self-regulation as long as platforms are designed to showcase examples of recent behavior [40]. While this may be the case with stable, close-knit groups with well established norms, communities with a high member turnover are likely to benefit from publicly posted rules and other explicit cues that make injunctive norms salient to newcomers.

Posting rules may also increase newcomer participation. Wise et al. [66] showed that non-participants expressed a greater intent to participate in communities that displayed strong textual and interface cues to the nature of community expectations around appropriate posting behavior, as well as cues that indicated the consequences for non-compliance (in this case, the presence of a 'report this comment' button). On the other hand, posting rules (or other injunctive normative cues) may not be enough. Findings from a study by Leavitt & Clark [38] of a disaster-focused subReddit suggest that posting insufficiently detailed and prescriptive rules about what kind of content to post in a context with high membership turnover and conflicting descriptive norms may increase moderator workload without preventing disruption, even when community members attempt to intentionally manipulate descriptive norms by up-voting posts that conform with posted rules.

Newcomers, who often lurk for a period of time before participating actively [53], may be less likely to comply with a community's injunctive norms if they cannot readily infer them from

observing the behavior of others—i.e. if most people are not following the norm, or it is difficult to identify 'normative' behavior from reading what other people write. We believe that this challenge may explain non-host answerers' lack of compliance with the Teahouse's "avoid over-linking": since it is not easy to tell what is the proper number of links to include in a response just by eyeballing it, people who had not been explicitly enjoined to reign in their policy-linking behavior tended to include more links than they should. In cases like this, more heavy-handed UI interventions—such as priming commenters with CAPTCHAs that contain pro-social key words [60] or using leaderboards that rank participants according to desired behaviors [54]—may be necessary.

Overall, our study further underscores the importance of successful norm activation and suggests a multi-pronged strategy for effective newcomer socialization: post clear and detailed behavioral expectations; increase the visibility of behaviors that align with those rules; and highlight the consequences of non-compliance in a way that provides injunctive cues to observers, even if it is not possible to sanction the rule-breakers themselves.

People who migrate from a related community with different norms, on the other hand, may need a different kind of socialization. In this case, it is important to focus attention on the injunctive norms that set this particular community apart from similar communities in order to mitigate disruption. For example, when managers of the social Q&A site Fluther noticed that they were receiving a large number of new contributors from a similar site, AnswerBag, and that some of these new contributors were attempting to re-create "AnswerBag traditions and interaction patterns" that conflicted with Fluther's own social norms, they set up a dedicated FAQ to help AnswerBaggers "become acclimated to the site" [22].

Intriguingly, we found that Teahouse participants who were subsequently exposed to the normative regimen of the Help Desk tended to behave in a more pro-normative fashion on the Teahouse. Although our study design does not allow us to assert a causal relationship between Help Desk exposure and increased normative compliance on the Teahouse, the possibility of exposing community participants to the negative consequences of alternative normative regimens (which may activate "personal norms", discussed above) is an intriguing area for further research and design.

6 CONCLUSION

In this paper, we drew on the Focus Theory to examine interactions between different sources of normative influence on answering behavior of 7060 contributors to a Wikipedia Q&A forum for new editors, the Teahouse. Specifically, we examined local descriptive and injunctive norms, their congruent effects, and conflict between different sources of normative influence due to membership overlap in a similar Q&A community, the Help Desk. Our findings across 4 years of data suggest that exposure to both explicitly defined injunctive norms and indirect exposure to injunctive norms through corrective practices have more persistent association with pro-normative behavior than exposure to behavioral examples (descriptive norms) alone; that exposure to descriptive norms is only associated with pro-normative behavior when norm-associated phenomena are easily identifiable as such; and that the highest level of normative compliance is associated with exposure to both explicitly defined injunctive norms and plentiful recent examples of pro-normative behavior.

Additionally, we find that experience in a related sub-community with conflicting norms can be associated with either less or more pro-normative behavior, depending on the type of norm conflict experienced. We contextualized our findings with interviews with Teahouse participants, and presented implications both for theory building on normative influence in online collaborative communities, as well as practical implications for the design and management of these communities.

ACKNOWLEDGMENTS

REFERENCES

[1] Nancy K Baym. 2000. *Tune in, log on: Soaps, fandom, and online community*. Vol. 3. Sage.

[2] Ivan Beschastnikh, Travis Kriplean, and David W McDonald. 2008. Wikipedian Self-Governance in Action: Motivating the Policy Lens.. In *ICWSM*.

[3] Gary Burnett and Laurie Bonnici. 2003. Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. *Library & Information Science Research* 25, 3 (2003), 333–351.

[4] Yan Chen, F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li. 2010. Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens. *American Economic Review* 100, 4 (September 2010), 1358–98. https://doi.org/10.1257/aer.100.4.1358

[5] Robert B Cialdini and Noah J Goldstein. 2004. Social influence: Compliance and conformity. *Annu. Rev. Psychol.* 55 (2004), 591–621.

[6] Robert B. Cialdini, Carl A. Kallgren, and Raymond R. Reno. 1991. A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. Advances in Experimental Social Psychology, Vol. 24. Academic Press, 201 – 234. https://doi.org/10.1016/S0065-2601(08)60330-5

[7] Robert B Cialdini, Raymond R Reno, and Carl A Kallgren. 1990. A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology* 58, 6 (1990), 1015.

[8] Benjamin Collier and Julia Bear. 2012. Conflict, Criticism, or Confidence: An Empirical Examination of the Gender Gap in Wikipedia Contributions. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 383–392. https://doi.org/10.1145/2145204.2145265

[9] Sue ES Crawford and Elinor Ostrom. 1995. A grammar of institutions. *American Political Science Review* 89, 3 (1995), 582–600.

[10] Kevin Crowston, Kangning Wei, James Howison, and Andrea Wiggins. 2008. Free/Libre Open-source Software Development: What We Know and What We Do Not Know. *ACM Comput. Surv.* 44, 2, Article 7 (March 2008), 35 pages. https://doi.org/10.1145/2089125.2089127

[11] Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities. In *Proceedings of the 22Nd International Conference on World Wide Web (WWW '13)*. ACM, New York, NY, USA, 307–318. https://doi.org/10.1145/2488388.2488416

[12] Rosta Farzan, Robert Kraut, Aditya Pal, and Joseph Konstan. 2012. Socializing Volunteers in an Online Community: A Field Experiment. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 325–334. https://doi.org/10.1145/2145204.2145256

[13] Ryan Faulkner, Steven Walling, and Maryana Pinchuk. 2012. Etiquette in Wikipedia: Weening New Editors into Productive Ones. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration (WikiSym '12)*. ACM, New York, NY, USA, Article 5, 4 pages. https://doi.org/10.1145/2462932.2462939

[14] Casey Fiesler, Shannon Morrison, and Amy S. Bruckman. 2016. An Archive of Their Own: A Case Study of Feminist HCI and Values in Design. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2574–2585. https://doi.org/10.1145/2858036.2858409

[15] Anna Filippova. 2016. *Mudslinging and Manners*. Ph.D. Dissertation. National University of Singapore. https://scholarbank.nus.edu.sg/bitstream/10635/124195/1/AFilippova-thesis.pdf.

[16] Anna Filippova and Hichang Cho. 2015. Mudslinging and Manners: Unpacking Conflict in Free and Open Source Software. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 1393–1403. https://doi.org/10.1145/2675133.2675254

[17] Anna Filippova and Hichang Cho. 2016. The Effects and Antecedents of Conflict in Free and Open Source Software Development. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 705–716. https://doi.org/10.1145/2818048.2820018

[18] Andrea Forte and Amy Bruckman. 2008. Scaling Consensus: Increasing Decentralization in Wikipedia Governance. In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*. 157–157. https://doi.org/10.1109/HICSS.2008.383

[19] Andrea Forte, Niki Kittur, Vanessa Larco, Haiyi Zhu, Amy Bruckman, and Robert E. Kraut. 2012. Coordination and Beyond: Social Functions of Groups in Open Content Production. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 417–426. https://doi.org/10.1145/2145204.2145270

[20] Andrea Forte and Cliff Lampe. 2013. Defining, Understanding, and Supporting Open Collaboration: Lessons From the Literature. *American Behavioral Scientist* 57, 5 (2013), 535–547. https://doi.org/10.1177/0002764212469362 arXiv:https://doi.org/10.1177/0002764212469362

[21] Andrea Forte, Vanesa Larco, and Amy Bruckman. 2009. Decentralization in Wikipedia Governance. *Journal of Management Information Systems* 26, 1 (2009), 49–72. https://doi.org/10.2753/MIS0742-1222260103 arXiv:https://www.tandfonline.com/doi/pdf/10.2753/MIS0742-1222260103

[22] Rich Gazan. 2011. Redesign As an Act of Violence: Disrupted Interaction Patterns and the Fragmenting of a Social Q&A Community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2847–2856. https://doi.org/10.1145/1978942.1979365

[23] R Stuart Geiger, Aaron Halfaker, Maryana Pinchuk, and Steven Walling. 2012. Defense Mechanism or Socialization Tactic? Improving Wikipedia's Notifications to Rejected Contributors. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*.

[24] Michael Gilbert, Jonathan T. Morgan, David W. McDonald, and Mark Zachry. 2013. Managing Complexity: Strategies for Group Awareness and Coordinated Action in Wikipedia. In *Proceedings of the 9th International Symposium on Open Collaboration (WikiSym '13)*. ACM, New York, NY, USA, Article 5, 10 pages. https://doi.org/10.1145/2491055.2491060

[25] Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist* 57, 5 (2013), 664–688. https://doi.org/10.1177/0002764212469365 arXiv:https://doi.org/10.1177/0002764212469365

[26] Steve Harrison and Paul Dourish. 1996. Re-place-ing Space: The Roles of Place and Space in Collaborative Systems. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work (CSCW '96)*. ACM, New York, NY, USA, 67–76. https://doi.org/10.1145/240080.240193

[27] B. D. Horne, S. Adali, and S. Sikdar. 2017. Identifying the Social Signals That Drive Online Discussions: A Case Study of Reddit Communities. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)*. 1–9. https://doi.org/10.1109/ICCCN.2017.8038388

[28] Ruud S. Jacobs, Ard Heuvelman, Maurice Tan, and Oscar Peters. 2012. Digital movie piracy: A perspective on downloading behavior through social cognitive theory. *Computers in Human Behavior* 28, 3 (2012), 958 – 967. https://doi.org/10.1016/j.chb.2011.12.017

[29] Corey Jergensen, Anita Sarma, and Patrick Wagstrom. 2011. The Onion Patch: Migration in Open Source Ecosystems. In *Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering (ESEC/FSE '11)*. ACM, New York, NY, USA, 70–80. https://doi.org/10.1145/2025113.2025127

[30] Elisabeth Joyce, Jacqueline C. Pike, and Brian S. Butler. 2013. Rules and Roles vs. Consensus. *American Behavioral Scientist* 57, 5 (2013), 576–594. https://doi.org/10.1177/0002764212469366 arXiv:http://dx.doi.org/10.1177/0002764212469366

[31] Charles Kiene, Andrés Monroy-Hernández, and Benjamin Mako Hill. 2016. Surviving an "Eternal September": How an Online Community Managed a Surge of Newcomers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1152–1156. https://doi.org/10.1145/2858036.2858356

[32] Aniket Kittur, Bryan Pendleton, and Robert E. Kraut. 2009. Herding the Cats: The Influence of Groups in Coordinating Peer Production. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration (WikiSym '09)*. ACM, New York, NY, USA, Article 7, 9 pages. https://doi.org/10.1145/1641309.1641321

[33] Christian Kleiber, Achim Zeileis, and Maintainer Achim Zeileis. 2017. Package 'AER'. (2017).

[34] Farshad Kooti, Haeryun Yang, Meeyoung Cha, P Krishna Gummadi, and Winter A Mason. 2012. The Emergence of Conventions in Online Social Networks.. In *ICWSM*.

[35] Travis Kriplean, Ivan Beschastnikh, David W. McDonald, and Scott A. Golder. 2007. Community, Consensus, Coercion, Control: Cs*W or How Policy Mediates Mass Participation. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work (GROUP '07)*. ACM, New York, NY, USA, 167–176. https://doi.org/10.1145/1316624.1316648

[36] Cliff Lampe and Paul Resnick. 2004. Slash(Dot) and Burn: Distributed Moderation in a Large Online Conversation Space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 543–550. https://doi.org/10.1145/985692.985761

[37] Maria Knight Lapinski and Rajiv N. Rimal. 2005. An Explication of Social Norms. *Communication Theory* 15, 2 (2005), 127–147. https://doi.org/10.1111/j.1468-2885.2005.tb00329.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2885.2005.tb00329.x

[38] Alex Leavitt and Joshua A. Clark. 2014. Upvoting Hurricane Sandy: Event-based News Production Processes on a Social News Site. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1495–1504. https://doi.org/10.1145/2556288.2557140

[39] Lawrence Lessig. 1999. *Code and Other Laws of Cyberspace*. Basic Books, Inc., New York, NY, USA.

[40] Diane Maloney-Krichmar and Jenny Preece. 2005. A Multilevel Analysis of Sociability, Usability, and Community Dynamics in an Online Health Community. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 201–232. https://doi.org/10.1145/1067860.1067864

[41] J. Nathan Matias. 2016. Posting Rules in Online Discussions Prevents Problems & Increases Participation. https://civilservant.io/moderation_experiment_r_science_rule_posting.html

[42] Jonathan T. Morgan. 2013. *Coordinating the Commons: Diversity & Dynamics in Open Collaborations*. Ph.D. Dissertation. Seattle, WA, USA. Advisor(s) Zachry, Mark R. AAI3608999.

[43] Jonathan T. Morgan, Siko Bouterse, Heather Walls, and Sarah Stierch. 2013. Tea and Sympathy: Crafting Positive New User Experiences on Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 839–848. https://doi.org/10.1145/2441776.2441871

[44] Jonathan T. Morgan, Michael Gilbert, David W. McDonald, and Mark Zachry. 2013. Project Talk: Coordination Work and Group Membership in WikiProjects. In *Proceedings of the 9th International Symposium on Open Collaboration (WikiSym '13)*. ACM, New York, NY, USA, Article 3, 10 pages. https://doi.org/10.1145/2491055.2491058

[45] Jonathan T. Morgan, Michael Gilbert, David W. McDonald, and Mark Zachry. 2014. Editing Beyond Articles: Diversity & Dynamics of Teamwork in Open Collaborations. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 550–563. https://doi.org/10.1145/2531602.2531654

[46] Jonathan T. Morgan and Mark Zachry. 2010. Negotiating with Angry Mastodons: The Wikipedia Policy Environment As Genre Ecology. In *Proceedings of the 16th ACM International Conference on Supporting Group Work (GROUP '10)*. ACM, New York, NY, USA, 165–168. https://doi.org/10.1145/1880071.1880098

[47] Gabriel Mugar, Carsten Osterlund, Katie DeVries Hassman, Kevin Crowston, and Corey Brian Jackson. 2014. Planet Hunters and Seafloor Explorers: Legitimate Peripheral Participation Through Practice Proxies in Online Citizen Science. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 109–119. https://doi.org/10.1145/2531602.2531721

[48] Sneha Narayan, Jake Orlowitz, Jonathan Morgan, Benjamin Mako Hill, and Aaron Shaw. 2017. The Wikipedia Adventure: Field Evaluation of an Interactive Tutorial for New Users. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1785–1799. https://doi.org/10.1145/2998181.2998307

[49] Dong Nguyen and Carolyn P. Rosé. 2011. Language Use As a Reflection of Socialization in Online Communities. In *Proceedings of the Workshop on Languages in Social Media (LSM '11)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 76–85. http://dl.acm.org/citation.cfm?id=2021109.2021119

[50] Dominic J. Packer. 2008. On Being Both With Us and Against Us: A Normative Conflict Model of Dissent in Social Groups. *Personality and Social Psychology Review* 12, 1 (2008), 50–72. https://doi.org/10.1177/1088868307309606 arXiv:https://doi.org/10.1177/1088868307309606 PMID: 18453472.

[51] Katherine Panciera, Aaron Halfaker, and Loren Terveen. 2009. Wikipedians Are Born, Not Made: A Study of Power Editors on Wikipedia. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work (GROUP '09)*. ACM, New York, NY, USA, 51–60. https://doi.org/10.1145/1531674.1531682

[52] T Postmes, R Spears, and M Lea. 2000. The formation of group norms in computer-mediated communication. *Human Communication Research* 26, 3 (2000), 341–371. https://doi.org/10.1111/j.1468-2958.2000.tb00761.x

[53] Jenny Preece, Blair Nonnecke, and Dorine Andrews. 2004. The top five reasons for lurking: improving community experiences for everyone. *Computers in human behavior* 20, 2 (2004), 201–223.

[54] Chris Preist, Elaine Massung, and David Coyle. 2014. Competing or Aiming to Be Average?: Normification As a Means of Engaging Digital Volunteers. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1222–1233. https://doi.org/10.1145/2531602.2531615

[55] Rajiv N. Rimal and Kevin Real. 2003. Understanding the Influence of Perceived Norms on Behaviors. *Communication Theory* 13, 2 (2003), 184–203. https://doi.org/10.1111/j.1468-2885.2003.tb00288.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2885.2003.tb00288.x

[56] Shalom H Schwartz. 1973. Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental Social Psychology* 9, 4 (1973), 349–364.

[57] Aaron Shaw. 2012. Centralized and Decentralized Gatekeeping in an Open Online Collective. *Politics & Society* 40, 3 (2012), 349–388. https://doi.org/10.1177/0032329212449009 arXiv:https://doi.org/10.1177/0032329212449009

[58] Vladimir Soroka, Michal Jacovi, and Sigalit Ur. 2003. Communities and Technologies. Kluwer, B.V., Deventer, The Netherlands, The Netherlands, Chapter We Can See You: A Study of Communities' Invisible People Through Reachout, 65–79. http://dl.acm.org/citation.cfm?id=966263.966267

[59] James P Spradley. 2016. *The ethnographic interview*. Waveland Press.

[60] Abhay Sukumaran, Stephanie Vezich, Melanie McHugh, and Clifford Nass. 2011. Normative Influences on Thoughtful Online Participation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 3401–3410. https://doi.org/10.1145/1978942.1979450

[61] Henri Tajfel and John C Turner. 1979. An integrative theory of intergroup conflict. *The social psychology of intergroup relations* 33, 47 (1979), 74.

[62] Fernanda B. Viégas, Martin Wattenberg, and Matthew M. McKeon. 2007. The Hidden Order of Wikipedia. In *Online Communities and Social Computing*, Douglas Schuler (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 445–454.

[63] Joseph B Walther and Shawn Boyd. 2002. Attraction to computer-mediated social support. In *Communication technology and society: Audience adoption and uses*. Hamptom Press, Cresskill, NJ, USA, 153—188.

[64] Loxley Sijia Wang, Jilin Chen, Yuqing Ren, and John Riedl. 2012. Searching for the Goldilocks Zone: Trade-offs in Managing Online Volunteer Groups. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 989–998. https://doi.org/10.1145/2145204.2145351

[65] Xiaoqing Wang, Brian S Butler, and Yuqing Ren. 2013. The impact of membership overlap on growth: An ecological competition view of online groups. *Organization Science* 24, 2 (2013), 414–431.

[66] Kevin Wise, Brian Hamman, and Kjerstin Thorson. 2006. Moderation, Response Rate, and Message Interactivity: Features of Online Communities and Their Effects on Intent to Participate. *Journal of Computer-Mediated Communication* 12, 1 (2006), 24–41. https://doi.org/10.1111/j.1083-6101.2006.00313.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1083-6101.2006.00313.x

[67] 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)*. ACM, New York, NY, USA, 281–290. https://doi.org/10.1145/2556288.2557213

[68] Haiyi Zhu, Robert E. Kraut, and Aniket Kittur. 2016. A Contingency View of Transferring and Adapting Best Practices Within Online Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 729–743. https://doi.org/10.1145/2818048.2819976