NORTHWESTERN UNIVERSITY

Three Studies on Technological Support for Participation in Volunteer Communities

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Technology and Social Behavior

By

Sneha Narayan

EVANSTON, ILLINOIS

September 2019

Abstract

Many volunteer communities rely on technological systems to help their members connect, collaborate and learn the norms of how to participate in the organization. This dissertation presents research that examines technological interventions designed to support participation in three different volunteer-run communities, all of which have porous boundaries, and allow volunteers to self-select into the tasks they perform. The first study evaluates the impact that a Wikipedia tutorial system has on the contributions of new volunteers. The second study examines the relationship between interpersonal communication and productivity by estimating the impact that a new system for interpersonal communication has on article production on a population of wikis. The third study is an analysis of how promotion processes in a volunteer-run bicycle co-operative are shaped by its physical and technological infrastructure.

Acknowledgements

The work presented in this dissertation has benefitted extraordinarily from extensive and thoughtful feedback from my advisor Aaron Shaw, and committee members Benjamin Mako Hill, Darren Gergle, and Anne Marie Piper. Their advice on this project and others has helped shape me as a researcher.

Much of the research presented here was conducted alongside some amazing collaborators, including Nate TeBlunthuis, Wm Salt Hale, Jonathan Morgan, Jake Orlowitz, Benjamin Mako Hill, and Aaron Shaw. Working with these smart and passionate people over the years has taught me so much about the practice of collaborative scholarship, and I look forward to continuing many of these collaborations in the future.

I have several academic communities and programs to thank for feedback, guidance and camaraderie along the way. They include the Community Data Science Collective, the BYOR workshop, and Northwestern's Dissertation Boot Camps.

Writing this dissertation required not just academic and intellectual mentorship, but also a significant amount of emotional support and cheerleading. Many wonderful friends (and their pets!) rose to this task over the years. They include Sam Taggart,

Whitney Russell, Jake Levinson, Tony Robinson, Alyssa Coffey, Whitney Bennett, Will Marler, Marion Gutwein, Sam Kling, David Bates-Jefferys, Matthew McMunn, Pete Manz, Loren Phillips, Amelia Plunk, Natalia Smirnov, Ben Remsen, Matt Ford, Beatrice Choi, Casey Caldwell, Jack the dog, and many others. My family (K. G. Yegnanarayanan, Janaki Narayan, Shravan and Siddarth Ravi Narayan) has been a constant source of encouragement from the very beginning, when they recognized my passion for learning and supported my choice to spend most of my adult life pursuing my education on the other side of the planet. My partner, Matt Meshulam, was the ultimate cheerleader, especially during the final stretch.

I became interested in this area of research because of the insipiration I have drawn from the many volunteer-run organizations I have been a part of over the years, including the Oberlin Student Co-operative Association, North American Students of Co-operation, MOSAIC Co-op, The Recyclery Collective, and Northwestern University Graduate Workers. Each of these communities taught me so much about the challenges and potential of volunteer-run organizations, and continue to serve as reminders for why the questions I ask in my work are important ones to answer.

Finally, I extend my deepest gratitude to the members of the volunteer communities I've studied in this dissertation. None of this work would have been possible or relevant without their commitment to organizing together to build something greater than themselves.

Contents

List of Tables

List of Figures

Introduction

The last two decades have seen the rise in prominence of a number of projects that are driven largely by volunteer contributions. Wikipedia is an encyclopedia that is freely available online, and maintained by a large group of people who are not paid for their efforts. Similarly, multiple open-source software projects, such as various distributions of the Linux operating system, or the Apache HTTP server, feature many contributions from developers who volunteer their time. Communities like StackOverflow encourage people to voluntarily provide answers to technical questions. Despite the fact that many of these volunteer efforts are structured very differently from most conventional firms, a number of these projects have proven to be very successful and widely relied upon. Rather than using monetary incentives to motivate contributors, many of these volunteer projects choose instead articulate a shared vision that they build a community around, and allow an ever-changing base of volunteers to determine how much and in what way they choose to participate.

My work in this dissertation focuses on understanding how the choices volunteer-run organizations make in designing and building technical systems can impact participation from their volunteer base. Supporting volunteer participation could range from

helping newcomers learn organizational norms, to facilitating smoother interpersonal communication among community members, to ensuring committed volunteers are mentored and promoted to roles of greater responsibility. To this end, I present three studies on technological systems that are designed to support those who contribute to volunteer-run organizations. The first two studies evaluate how particular design interventions impacted volunteer participation along different dimensions within the online communities they were deployed in. The third study analyzes how a volunteer community navigates member promotion into leadership positions through its technological and physical infrastructure, and identifies process breakdowns and potential for design interventions in this space. In each case, I evaluate how technological systems impact the organizations I study, reflect on how my findings inform our understanding of these communities, and identify implications for design in these contexts.

This area of inquiry is particularly important to the success of volunteer-run projects, whether they be large-scale peer production[11] projects like wikis, or smaller, community-run services like bike co-operatives. Since no single volunteer contributes to a project indefinitely, organizations need to understand how to consistently recruit, train and engage volunteers to become long term committed contributors and community leaders in order to ensure organizational sustainability[51].

Having systems in place for volunteer support is important for contributors as well, particularly newer ones. Many new volunteers experience confusion and anxiety when attempting to join a community. Even if they begin contributing to a project with altruistic intentions, they can become discouraged if they do not see a clear path to engaging with the project, or if they perceive a lack of support from the community[39].

Poor onboarding can lead to a decreased sense of self-efficacy, which makes it harder for volunteers to have a sense of ownership of their work and identify themselves as a member of the community they are attempting to join[10].

Much of the research and theory on socialization and commitment within organizations is based on conclusions from studies of conventional organizations, such as firms or universities[10, 83]. While these theories do shed some light on these phenomena in online communities and volunteer-run organizations as well, there are many issues particular to these communities that pose challenges to successfully designing technological systems to support them. In particular, the studies I present in this dissertation focus primarily on designing for organizations that allow contributors to a) self select into the tasks they perform, b) do not provide monetary incentives, and c) have fuzzy criteria for membership. Each of these characteristics contrast with the nature of many conventional organizations, and significantly impact the volunteer experience.

Self selection of tasks

When a new person is hired to join a firm, she usually has a job description that delineates the tasks that she's required to do; i.e. tasks are assigned to her by the organization. Task assignment plays a large factor in shaping the volunteer experience; as in any organization, having clarity about one's role helps the volunteer seek out relevant information, and allows the organization to provide more targeted guidance and training about this role[10]. In peer production communities and other volunteer organizations, task assignment is typically the purview of the volunteers themselves -

they are rarely (if ever) directed by the organization to work on particular tasks. Their choices in task can sometimes be guided by more experienced volunteers, or by organizational needs, but ultimately, the volunteer has the prerogative to choose whether or not they wish to work on a particular task. This poses a particular challenge to peer production communities - if there are many different tasks that a new volunteer could potentially do, how does the community proactively train their newcomers? Given that the paths of volunteers through these communities can look quite different, designing onboarding programs and scaffolding skill building in a way that is sensitive to the self-selection of tasks is a pressing challenge for these organizations.

Fuzzy notions of membership

Another challenge in designing for peer production communities is that the boundaries of membership are very diffuse. Conventional organizations typically have clear criteria for what constitutes membership. In a conventional firm, it is generally clear whether a person is considered a member or not. In many volunteer-run organizations, however, the idea of membership is significantly less clear, and is often a matter of self-identification. For example, many people participate in editing Wikipedia peripherally, but far fewer people call themselves Wikipedians.

Scholarship in this area has tried to approximate what newcomers and veteran members look like based on their digital trace[68], but these definitions are necessarily arbitrary to some extent. When does one become a Wikipedian? After their tenth edit? After their thousandth? Deciding what criteria constitute a veteran, versus a

one-time contributor, versus a budding but green newcomer, is an active problem for researchers and designers who build systems for these communities.

From the perspective of the volunteer base, fuzzy notions of membership present additional challenges. For newer members, there may never be a point where they are formally recognized as a member of the community; rather, they might gradually begin identify as such after having contributed to the community. Figuring out what their status is within a community can present a source of uncertainty for newcomers. For veteran members, fuzzy boundaries present challenges around trust and quality control. When communities have porous boundaries, it is harder to ensure that new or itinerant volunteers are contributing in good faith, or have the training and context to make good contributions. Bringing in new, committed volunteers, while also maintaining the quality of of the product or service that the community is providing, can be a tricky balance to strike[36].

Lack of monetary incentives

Finally, encouraging ongoing participation is a particular challenge in communities where people have no monetary incentive to participate. Due to the entirely voluntary and non-contractual nature of participation in open content production communities and volunteer-run co-operatives, there are typically no reprisals at all when participants in these spaces decide to leave for any reason. Thus, it is incumbent on communities to simultaneously keep volunteers engaged and ensure that they feel like their contributions are appreciated.

While volunteer-run organizations have existed for centuries, there has been significant interest in studying them since the advent of technologies for collaboration on the web. In the wake of the tremendous success and influence of volunteer-run communities like Wikipedia and Linux, many scholars have sought to understand how these communities self-organize, maintain a volunteer base, and act collectively to provide public goods and services[81, 13]. Particular attention has been paid to the role that technological platforms have played in reducing transaction costs associated with these forms of collaboration, making them more viable[12]. Understanding how the design of technological platforms can mediate the success of these kinds of organizations has been an active area of research[51].

To this end, I present three studies that further inquiry in this research area. The first concerns the design and evaluation of an interactive tutorial that helps newcomers learn how to edit Wikipedia. The second evaluates a design change that was introduced in Wikia's discussion software, and measures the effect that this change has on volunteer participation. The third examines the kinds of challenges experienced by a volunteer-run non-profit in designing effective systems to manage volunteers.

Each of these studies examines how technological design can shape and impact volunteer onboarding and participation. Broadly, the purpose of this dissertation is to empirically evaluate existing theories from organizational studies and technological design as they are applied to decentralized volunteer communities, use these empirical results to further refine theory in this space, and generate principles for designing for these communities.

Overall, I argue that the communities I study experience significant challenges balancing centralized goals and systems with a decentralized and diffuse organizational structure. The tensions between these two organizing tendencies manifest in different ways in the various settings I study, yet they remain a persistent challenge for communities of this form. I elaborate on these themes as well as the specific studies presented in this dissertation in the next section.

Chapter Overview

Chapter 1 - The Wikipedia Adventure: field evaluation of an interactive tutorial for new users

The first study in this dissertation concerns the design and evaluation of a tutorial that teaches newcomers how to edit Wikipedia. This work aims to provide an intervention very early on in a new user's tenure on Wikipedia, and asks the question: Would a structured, engaging tutorial, provided soon after a new user joins Wikipedia, positively affect their subsequent contributions to the project?

To this end, I worked with Wikimedia Foundation staff to evaluate The Wikipedia Adventure, a narrativized, gamified tutorial that introduces newcomers to the basics of editing and communicating with others on Wikipedia. The choice to make such an intervention centered around clarifying a path for new members to become Wikipedians by laying out the norms and policies that the community followed. Since there are no monetary or contractual incentives for learning about editing Wikipedia, the developers

used narration and gamification to make the system as engaging as possible.

After designing a learning intervention that taught new users how to contribute to the project, we surveyed a number of Wikipedia editors on whether they found the system engaging, and whether it improved their confidence in editing Wikipedia. Overall, user feedback indicated that The Wikipedia Adventure succeeded on these fronts. We then tested its efficacy by deploying the system to new users on Wikipedia, and ran an experiment that compared how newcomers who were invited to use the system fared on Wikipedia, in relation to those who were not.

This was also an examination of the extent to which theory that is based on research on conventional organizations (i.e. firms where membership is clear, tasks are assigned, and monetary incentives provide contractual obligations to complete tasks) can apply to peer production communities like Wikipedia. The consensus from this body of work is that structured orientations that provide newcomers with a uniform onboarding experience lead to a greater sense of self-efficacy and better retention of newcomers[10]. However, findings from this experiment contradicted received wisdom from conventional organizations. We found that newcomers on Wikipedia did not, on average, exhibit differences in participation subsequent to being invited to play The Wikipedia Adventure, largely because of how Wikipedia is structured to have new users self-select into anything they choose to do on the project, whether it be editing, communicating with others, or participating in an interactive tutorial. This study revealed that the self-selecting behavior of new volunteers in peer production communities makes it much more challenging to provide onboarding interventions at scale to all new users. In particular, I found that the centralized system for onboarding in this

context lay at odds with the decentralized nature of participation on Wikipedia.

This chapter has been published in the Proceedings of the 2017 ACM Conference on Computer Supported Co-operative Work and Social Computing and has been reproduced within this text in its entirety.

Chapter 2 - More connected but not more productive: analyzing support for interpersonal communication in wikis

This study examines the effect that making interpersonal communication easier in peer production communities has on volunteer participation in these projects. An interactionist understanding of how volunteers become embedded in communities of practice centers on the idea of legitimate peripheral participation[56]. In online communities where volunteers arrive sporadically and choose the tasks they are interested in, interactions between newer volunteers and veteran community members provide an important site where volunteers can learn the norms of the community, and develop a sense of self-identification with the community. This led my co-authors and I to ask: if a community makes it easier for its volunteers to communicate with each other, how would that affect their subsequent participation? Working with data from Wikia, a popular meta-community of publicly editable fan wikis, my co-authors and I investigated the effect of a change in discussion software on newcomer and veteran communication and contributions in a population of 275 wikis.

Wikis hosted by Wikia are built on the MediaWiki framework. Talk pages on MediaWiki wikis (including Wikipedia) are typically designed with exactly the same

affordances as article pages, require an understanding of specific community norms to participate, and differ significantly from most discussion software used elsewhere on the web. In 2011, Wikia rolled out a new feature where user talk pages were replaced with threaded discussion boards called 'message walls'. This provided an opportunity to understand how such a design change could influence communication and contribution activities on wikis, and test theories on the importance of interpersonal communication to volunteer participation.

We found that although the introduction of message walls led to an increase in the number of messages that users sent to each other, it did not increase the number of article contributions that users made on the platform. Indeed, new users experienced a small drop in contributions, suggesting that the introduction of message walls might have led them to substitute their contribution activity with more communication. These results suggest that designing systems that facilitate conversations can have positive effects on the number of messages that participants send to each other, but ultimately do not directly increase participation along other dimensions. Indeed, it is possible that the introduction of the system allowed for more diffuse and decentralized conversations to take plaace, thus drawing attention away from the central project of article production on wikis.

Chapter 3 - Technology adoption and tool use in a volunteer-run non-profit

My first two studies revealed the difficulties inherent in designing interventions that affect volunteer participation in online peer production communities. In both cases, the communities of interest introduced design changes that ultimately did not produce overarching changes in patterns of contribution, contrary to what we would expect from existing theory in this space.

To gain a better understanding of the challenges that exist in designing technological systems in volunteer communities, I chose to study how an embodied volunteer organization, namely, a community-run bicycle co-operative, used information systems and physical artifacts to manage volunteers in the organization.

Drawing upon extensive volunteering, participant observation, analysis of artifacts and interviews with organizational stakeholders, I assess how this organization manages volunteers through a variety of technological systems and uses of physical artifacts. I identify process breakdowns and outstanding needs the community faces in this regard, and elaborate on the challenges associated with designing technology for this space. In particular, I found that the primary tension within this organization's processes is between its centralized system for formalizing leadership roles, and its decentralized, informal, volunteer development process. While volunteers largely interact with a single staffmember when scheduling when they work at the co-operative, they can be taught key skills and granted privileged access by any one of a number of different experienced volunteers. Breakdowns occur when scheduling and promotion need to

	Chapter 1	Chapter 2	Chapter 3
Research setting	English Wikipedia	Wikia wikis	Volunteer-run bike co-op
Methods used	Field experiment	Quantitative comparison based on a regression discontinuity design	Participant observation, analysis of artifacts, interviews
Phenomenon studied	Volunteer training	Volunteer communication	Volunteer scheduling and promotion
Organizational boundaries	Fuzzy	Fuzzy	Fuzzy
Co-located members?	No	No	Yes
Task assignment	Self-selected	Self-selected	Mix of self-selection and assignment
Monetary incentives to contribute	None	Mostly none, with a few paid community managers	Mostly none, with a few paid staff

Table 1: Study comparisons

influence one another, which requires reconciling a structured, centralized process with a more informal, decentralized one. My analysis identifies how the conflict between these independent processes within the co-operative presents a compelling area for the organization to develop new tools or institute new routines to address breakdowns that occur in this space, and I make design recommendations to this end.

Overarching research contributions

Overall, I have examined the effects of technological design on multiple facets of volunteer participation in a variety of volunteer-run organizations over the course of this dissertation (See Table 1). All three studies concern designing for the benefit of volunteers in decentralized communities. I argue overall that the specificities of volunteer communities (in terms of how they allow for self-selection of tasks, have fluid and open boundaries, and depend on non-monetary incentives), should help guide the design of technological interventions in these cases. The design and evaluation of the Wikipedia Adventure in Chapter 1 sheds light on the difficulty of reaching new editors on

Wikipedia to provide them with the information they need in order to contribute to the community, emphasizing the challenge that self-selection presents when designing newcomer interventions for peer production communities. The message wall intervention explored in Chapter 2, shows us that while making interpersonal communication smoother on wikis increased the number of messages sent between users, it did not, by itself, increase article contributions made to wikis. In contrast with the approach in Chapter 1, the design intervention in Chapter 2 was intended to pave the way for more informal and individualized types of volunteer socialization that could affect participation. We found that simply providing the means for this process to occur by revamping communication systems on Wikia were not sufficient to detect any subsequent changes in the nature of contributions to these projects within the study period. Finally, Chapter 3 shows that decentralized organizations like bicycle co-operatives can actually exhibit a mix of organizational processes in managing volunteers, some of which, like scheduling, could be more centralized around a particular staff member, while others, like promotion, are distributed among various community members. My analysis showed how the organization's information systems were set up in response to these different processes, and thus broke down at times when these processes needed to influence one another. In contrast with the first two chapters, which were coarser analyses conducted on a large scale, Chapter 3 delves deep within a single, small, community to gain a more detailed sense of how a volunteer organization faciliates promotion among its volunteer base through interaction with technological systems and the physical environment.

The studies presented in this dissertation draw on a wide range of methods, in-

cluding contextual inquiry, causal inference on experimental and observational data, and quantitative and qualitative evaluations of theory-driven design. Rigorous system evaluation is a particularly important facet of my research. While many technological systems are evaluated primarily for positive user experience, I am interested in understanding whether design interventions create measurable changes in the communities in which they are deployed. Chapters 1 and 2 focus on developing clear understandings of what kinds of effects different sorts of technological solutions have on *aggregate* volunteer behavior in a community. This perspective is important to help researchers understand what kinds of interventions are likely to curtail phenomena like volunteer decline and disengagement from these communities. While field deployment and evaluation drove my choice to conduct quantitative studies in Chapters 1 and 2, I chose to use a qualitative approach in Chapter 3, where I examine a bicycle co-operative's use of various technological systems and artifacts to organize volunteer participation. In this case, methods like participant observation and contextual interviews helped me develop a more nuanced understanding of the circumstances within which these tools are built and used, and how the volunteer experience both shapes and is shaped by them. Though quantitative methods are crucial for evaluating impact, qualitative methods have been useful for teasing out the ways systems and processes affect volunteer development and promotion, and building theories of design that are specific to the challenges faced by decentralized volunteer-run organizations.

The Wikipedia Adventure: Field Evaluation of a Interactive Tutorial for New Users

Co-authored with Jake Orlowitz, Jonathan Morgan, Benjamin Mako Hill, and Aaron Shaw[1]

Introduction

Social computing systems and peer production communities that aggregate voluntary contributions depend critically on recruiting and retaining new users [51]. Since no user will contribute indefinitely, online collaborative projects must successfully mobilize newcomers in order to maintain their community. However, in order to successfully make high-quality contributions and avoid censure, new users must quickly learn community norms. As inexperienced users inevitably violate norms, the impetus to recruit newcomers can be in tension with the desire to maintain quality.

Wikipedia provides a well-known case of this dilemma in social computing. Since

[1]This chapter has been published in the Proceedings of the 2017 ACM Conference on Computer Supported Co-operative Work and Social Computing[64]

Figure 1: Welcome page shown to users who arrive at The Wikipedia Adventure for the first time.

a short contribution history is an excellent predictor of vandalism in Wikipedia [69], established community members often delete or "revert" newcomer contributions. This demoralizing experience drives many good-faith newcomers away [36, 39, 60] and has contributed to an overall decline in Wikipedia's active editors since 2007 [36, 82]. Making matters more difficult for newcomers, Wikipedia's norms, procedures, conventions, and policies have expanded considerably since the inception of the community [17]. While a growing body of design research aims to overcome these challenges [37, 60], existing systems frequently rely on the helpfulness of veteran editors (a limited resource), and significant initiative from newcomers themselves.

We present a novel system called The Wikipedia Adventure (TWA): an interactive

tutorial that provides an introduction to editing Wikipedia. Unlike most prior systems designed to socialize new users on Wikipedia, TWA creates a structured, interactive experience that guides newcomers through critical pieces of Wikipedia knowledge: editing using wikimarkup (the code that editors use to format text), communicating with other editors, and learning basic community policies. It also incorporates elements of gamification in an attempt to increase the motivation, engagement, and enjoyment of newcomers as they learn about the community.

After describing TWA, we evaluate it in two ways. First, we report a survey that assesses how new users perceive the system's design and tone. We then conduct a randomized controlled field experiment in which we invite new Wikipedia editors to use the system and measure its effect on their subsequent contributions using multiple parametric and non-parametric techniques. One of these techniques is two-stage least squares regression which we use to estimate the effect of playing TWA (conditional on having been invited to do so) on the number of future contributions.

We find that survey respondents perceived TWA as engaging and well-designed for newcomers. At the same time, our field experiment shows that deploying TWA does not alter newcomer contribution patterns. These results imply that the design principles of TWA are sound but that the system does not produce the expected impact. We propose multiple potential explanations for this null effect and suggest that they may point to a gap in existing literature on newcomer socialization. In addition to the empirical findings, the study also contributes one of the first randomized controlled studies of the effect of a gamified orientation system as well as the first application of two-stage least squares analysis in social computing research.

Background

Newcomer Socialization in Online Communities

The norms and routines of organizations and communities often appear opaque to newcomers. Failing to communicate the skills that a newcomer needs in order to effectively contribute can lead to frustration and alienation for newcomers, while also breeding distrust of newcomers among established members. For online communities that depend on voluntary contributions from their users, socializing newcomers is among the most crucial tasks [51].

Prior research on newcomer socialization in organizations distinguishes between *individualized* and *institutionalized* socialization tactics [51, 84]. Individualized socialization is informal, akin to on-the-job learning, and is directed by newcomers themselves. Institutionalized socialization is more collective and formal, with the aim of providing a uniform set of experiences. In conventional firms, there is a broad consensus that institutionalized forms of socialization are more effective in retaining new members [10]. Institutionalized socialization techniques facilitate newcomers joining organizations by increasing self-efficacy, providing role clarity, and instilling a sense of social acceptance which, in turn, leads to better performance and higher commitment [10].

While several online communities and social computing systems use institutionalized socialization, large peer production communities, including Wikipedia, have relied almost exclusively on individualized socialization techniques, and typically require users to figure out what they need to know in order to contribute to a project [26, 28, 61].

More general, formal, and institutionalized systems for newcomer socialization can complement these existing approaches.

Gamified Onboarding

In considering strategies for effectively onboarding newcomers in online communities, we draw on recent research showing that gamification can support engagement in interactive systems. Gamification has become an increasingly popular approach within interactive system design for improving engagement in learning activities. In a meta-analysis of gamification studies in computer science, HCI, and eLearning, Hamari et al. [40] identified 10 common motivational affordances of gamified systems: points, leaderboards, achievements/badges, levels, stories/themes, clear goals, feedback, rewards, progress, and challenges. They showed that in a majority of cases they reviewed, these features led to positive learning outcomes and enhanced enjoyment among participants.

However, gamification has limits. Work in psychology [23] and behavioral economics [31], as well as studies of gamified systems [41], have highlighted the potential demotivating effect of gamification. In particular, competition-based incentives which are central to gamification affordances like leaderboards have been shown to undermine participants' motivation.

The effects of gamification in non-gamified contexts over time remain largely unstudied. As a result, the question remains open as to whether a gamified tutorial that effectively introduces a particular task will have any effect on participants' motivation

to perform that task outside of the tutorial. On one hand, a gamified introduction to a task might increase positive affect, confidence, and self-efficacy which might in turn increase subsequent participation. On the other hand, the social psychology literature on crowding out [23], has shown that shifting incentives, particularly the removal of extrinsic incentives, can decrease levels of motivation overall.

The Challenges of Becoming a Wikipedian

A mature community like Wikipedia poses particular challenges for onboarding newcomers. Wikipedia has seen a massive decrease in newcomer retention over the past decade. Between 2006 and 2010, Wikipedia's retention rate of newcomers acting in good faith (i.e. those whose initial edits showed a desire to contribute productively) dropped from 25% to 5% [36]. Research has suggested that this drop in retention is due in part to higher rates of negative socialization experiences like receiving warning messages or having an edit reverted [39], the lack of effective socialization in the presence of increasingly formal policies and rules, and an increase in the use of automated quality control tools to enforce rules and sanction new users [17, 36].

Most new editors begin editing without any structured external guidance and, perhaps as a result, quickly adopt behaviors and roles that tend to persist over their Wikipedia careers. Panciera et al. [68] have shown that initial rates of editing are among the strongest predictors of long-term contribution rates. That said, subsequent work has shown that contributions from extremely active users in social computing systems also change dynamically over time in ways that can be influenced by both

users' experiences and technological interventions [45]. Panciera et al. and others have suggested that the creation of systems to support newcomers immediately after joining Wikipedia is important for building long-term commitment [51, 68].

Historically, Wikipedia has relied on user-initiated, individualized forms of newcomer socialization. New editors on Wikipedia frequently learn what to do by requesting help and feedback [63] and by consulting Wikipedia's help and policy pages. They also learn what not to do through warning and advisory messages they receive from experienced editors when they violate a policy or make a mistake [33].

While new Wikipedia editors select their own tasks, the increasing scope of the community policies, distrust from experienced editors, and requirements of specialized knowledge has made it difficult for them to act independently [15, 36]. In a 2010 survey, over 40% of new editors who decided to leave Wikipedia cited a lack of support or an unpleasant social atmosphere.[2] In particular, women reported that they found that contributing to Wikipedia involved a high level of conflict and that they lacked confidence in their expertise [20].

Why Gamify Becoming a Wikipedian?

A gamified tutorial has the potential to increase newcomer participation on Wikipedia for several reasons. At a minimum, tutorials seem to have helped existing active Wikipedians. In a survey of readers and editors of the French Wikipedia, regular contributors reported using tutorials when they began editing at a greater rate than

[2]Based on the Wikimedia Foundation's survey of former contributors, available at: `https://strategy.wikimedia.org/wiki/Former_Contributors_Survey_Results`

occasional contributors [25]. An interactive tutorial that addresses the most critical technical and organizational topics, made available to a large number of new editors shortly after they join, might help more newcomers overcome these challenges. Taking a structured, step-by-step approach with milestones and incremental feedback can increase their sense of self-efficacy [10].

Research on experienced Wikipedians also shows that many active contributors experience several elements of gamified systems through their participation in the community. For example, WikiProjects often set project level "challenges" that encourage editors to complete a certain number of tasks in a short period of time [90]. Editors who achieve high-visibility goals in the community are acknowledged for their efforts with badge-like social awards which confer external recognition of their achievements [53]. Making such gamified elements more transparent to new editors could increase their enjoyment and lead to increased contributions.

Although prior attempts at socialization within Wikipedia have incorporated institutionalized elements, none have explicitly combined these with gamified design. The Wikimedia Foundation previously deployed a tutorial called GettingStarted which introduced participants to basic editing concepts using an interactive interface without game-like features. It showed no effect on retention and was subsequently deactivated.[3] New editors may also receive a generic welcome message from members of the Wikipedia Welcoming Committee soon after they register.[4] Typical welcome messages explain community values and provide a list of policies and resources. Wikipedia's pol-

[3]See: `https://meta.wikimedia.org/wiki/Research:Onboarding_new_Wikipedians/OB6`

[4]An example of a welcome message can be viewed at: `https://en.wikipedia.org/w/index.php?title=User_talk:Krishna_7murari&oldid=643725558`

icy pages are, in this sense, also an effort to provide institutionalized guidance about norms and routines. However, without a more structured introduction to facilitate learning about these policies, many newcomers will never read them.

Although several existing systems provide explicit and personalized direction to Wikipedia newcomers, most require extensive one-on-one interaction and effort from experienced editors. The Wikipedia Teahouse Q&A forum [60] provides a safe space for newcomers to get personalized help from experienced volunteer "hosts" and fellow newbies in a many-to-many setting. The Adopt-a-User program provides opportunities for newcomers to enroll in extended, one-on-one mentoring relationships with experienced editors [63]. The MoodBar allowed new editors to provide instant post-edit feedback that was piped to a feed monitored by experienced editors who were encouraged to step in and provide rapid assistance [19]. These examples demonstrate tensions between scale, information density, and personalization which many efforts at newcomer socialization must confront. The differences between them also suggest how a gamified tutorial might provide a scalable approach that does not feel as impersonal as a message from a bot account or as overwhelming as navigating a thicket of complicated policy pages.

While personalization is a major advantage of socialization efforts such as The Teahouse and Adopt-a-User, these systems are limited by volunteer time of hosts and mentors and the burden is still on the new user to initiate conversations and ask questions. A scalable, institutionalized effort at orientation initiated by the Wikipedia community has the potential to reach a greater number of new users more quickly and provide a framework for understanding what it means to contribute to Wikipedia.

Motivated by prior research on newcomer socialization in online communities and gamification in interactive environments, we designed a system to bridge gaps in the current newcomer onboarding experience in Wikipedia. Our main research questions were:

1. *Would a gamified tutorial produce a positive and engaging experience for new Wikipedians?*
2. *Would playing the tutorial cause newcomers to contribute more?*

System Design: The Wikipedia Adventure

In order to address the challenges of newcomer socialization on Wikipedia, we designed The Wikipedia Adventure, a system that introduces newcomers to the nuts and bolts of editing the encyclopedia. TWA provides a positive and gamified introduction to Wikipedia by taking a new user on a guided journey through the basics of editing, communication, and community norms in order to help them develop the skills to make effective contributions.[5] While playing TWA, the user is asked to perform a series of tasks for which they are provided detailed instructions. The tasks are couched in realistic Wikipedia editing scenarios. The user's input is evaluated and they receive a prompt informing them whether they completed the task correctly. If their response was incorrect, they receive additional instruction and cues.

TWA incorporates institutionalized socialization techniques by providing a stan-

[5] `https://en.wikipedia.org/wiki/Wikipedia:TWA/Story`

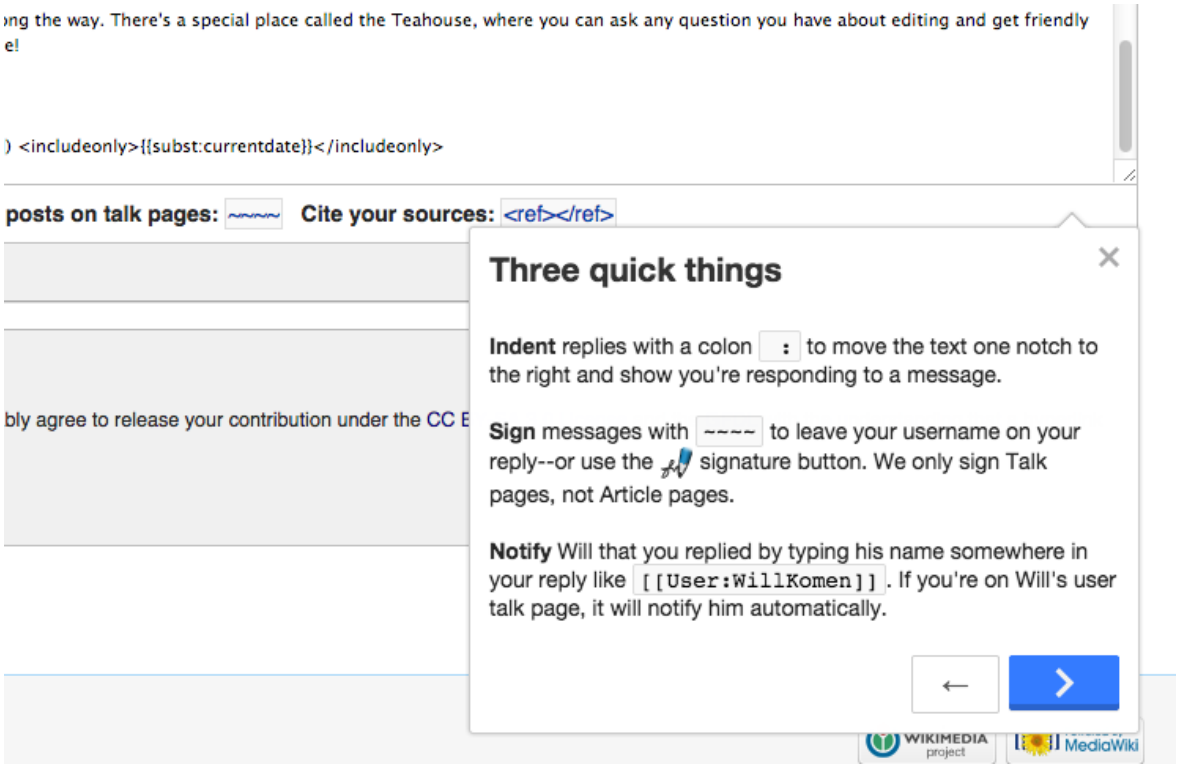


Figure 2: Learning syntax through The Wikipedia Adventure.

dardized, sequential introduction to the norms and policies of Wikipedia. While interactive, it does not depend on the availability, helpfulness, or intervention of existing Wikipedia editors and can therefore scale to support an arbitrary number of newcomers. The Wikipedia Adventure teaches users how to edit using wikimarkup code by using a series of pop-up boxes that point out where to click and what syntax to use in each context. Figure 2 shows how the user is given a lesson on how to edit their talk page through a pop-up box that appears next to their talk page editor.

In order to create a safe space for the user to try new things without fear of scrutiny or reprisal, TWA provides a training experience in a section of Wikipedia that is

separate from existing articles. That said, this design decision also increases the risk that players will consider the edits they make through the Wikipedia Adventure to be inauthentic. We considered an alternative approach in which newcomers edited actual articles on Wikipedia. Support for this approach can be found in the theory of legitimate peripheral participation (LPP) which emphasizes that learners should be able to perceive that their initial contributions, however small, are valuable to their new community of practice [55].

Unfortunately, Wikipedia presents challenges to an LPP-based approach. Research has shown that inexperienced Wikipedia editors who edit "public" articles are likely to make mistakes that elicit powerfully demotivating reactions [39]. As a way of reducing this risk while still creating an effective system for socialization, we drew inspiration from researchers working to reconcile LPP with traditional instructional methods who faced similar challenges.

Drawing on previous work around the concept of authenticity in education [47, 79], Guzdial and Tew [35] argue that learners may derive similar benefits from performing inauthentic tasks as long as they perceive an alignment between the tasks they are assigned and the work of a community of practice they value. Moreover, they suggest that educators can facilitate such alignment through storytelling, by (for example) creating a fictional narrative context in which students can perceive their learning tasks as legitimate peripheral participation within an imagined community of practice. TWA was designed to encourage alignment in both of these ways through a number of game-like elements. While we drew significantly from the gamification literature in making design choices, we eschewed features like leaderboards to avoid potential

demotivating effects associated with competition. We describe three of the gamified elements below and show how they fit with the goals of the project and the context of new users' experience of Wikipedia.

Missions

TWA is split into seven *missions* which accomplish different learning objectives tied to the five pillars of Wikipedia[6] and reflect key community rules and norms [72]. Missions introduce new users to setting up their user page, communicating effectively with other users, making basic edits, maintaining a neutral point of view, evaluating content quality, understanding revisions, and using built-in tools like watchlists and history pages to see how articles can be maintained over time. Additional techniques such as adding sources and formatting sections are also introduced in various missions. Each mission consists of a guided tutorial that explains a policy or editing technique in the context of a specific task and presents a simple challenge that tests the user's understanding of the topic. Although the interface prompts the user to go through the missions in order, they can select missions out of order or exit the tutorial at any point.

TWA introduces the basics of communicating and collaborating with other editors early in the tutorial, thus framing Wikipedia as a community of editors, and not just a repository of articles. Throughout the missions, prompts within the tutorial share key facts about the history and philosophy of Wikipedia as a free, global, open knowledge project. This reinforces the idea that Wikipedia is a collective effort driven by volunteer

[6]See: `https://en.wikipedia.org/wiki/Wikipedia:Five_pillars`

contributions, and attempts to establish a sense that the user is becoming part of an endeavor larger than themselves. In this way, the gamified missions of the tutorial prepare newcomers to identify their contributions in the context of the broader goals of the community.

Badges

The Wikipedia Adventure uses badges to invite newcomers to take on editing Wikipedia not just as a task, but as a part of their identity. Participants who complete all seven missions will earn a total of 15 badges. At the end of each task, the user receives a badge (e.g., Figure 3) and a pop-up message that congratulates them on their accomplishment. The badge titles are designed to reflect new identities the user has taken on (e.g. Copyeditor, Collaborator) as well as competencies around community norms that they have gained (e.g. Verifiability, Neutrality).

These badges are inspired by barnstars, which are awards that Wikipedia editors give each other to acknowledge valued work [53], and userboxes, which are badge-like labels that editors assign to themselves to communicate their achievements, skills, and aspects of their on-wiki identity. Many Wikipedians' user pages contain dozens of badges and barnstars they have accrued during their tenure. When a TWA participant earns a new badge, it is also placed on their user page where it serves a persistent, public reminder that they have been introduced to the values and principles of the community.

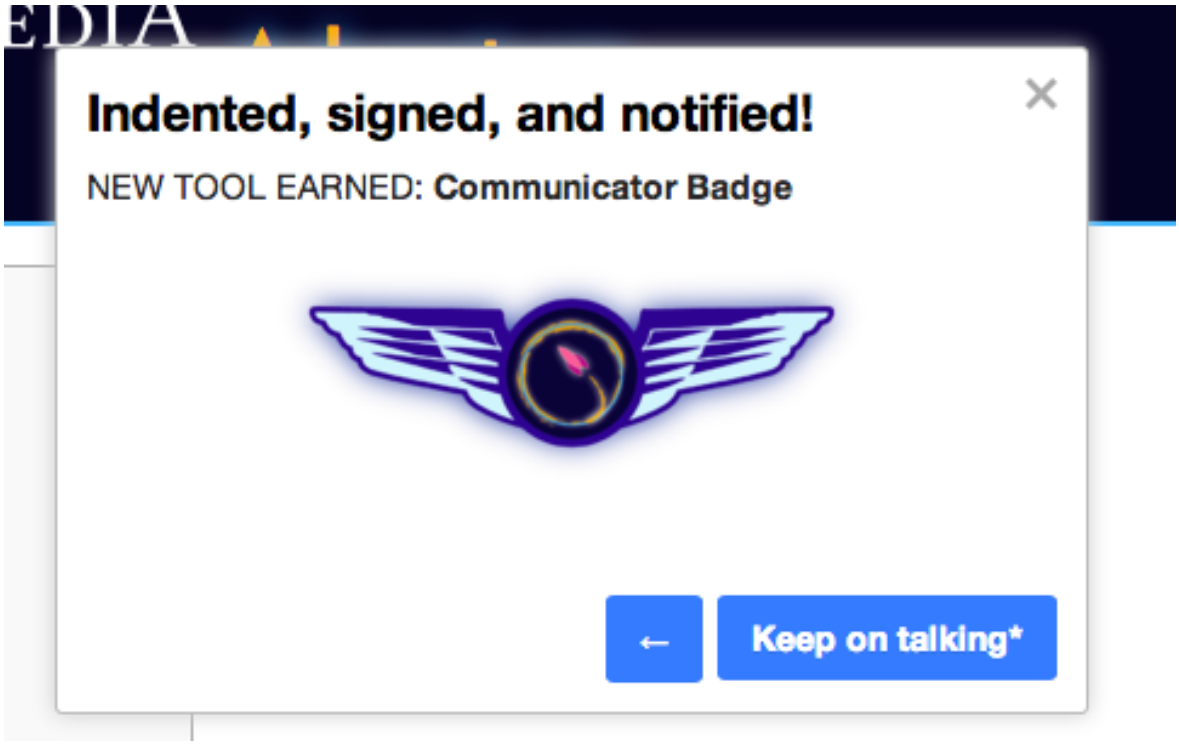


Figure 3: Badge for completing the mission about communicating with editors.

Story and Theme

The tutorial is centered around the scenario of creating the article 'Earth' in collaboration with several fictional editors. Earth was selected as subject for its universal relevance, reflecting Wikipedia's mission as a collaborative project that spans geographic and cultural boundaries. Focusing the missions around a single article allows the tutorial to teach a range of technical skills, and also address community norms such as effective communication and adherence to policy while maintaining continuity and verisimilitude.

The visual theme of TWA is galactic exploration, and the tutorial uses graphical

elements that are lush, colorful, and whimsical. The tone of the guiding prompts is conversational and humorous in order to create a relaxed and friendly atmosphere in which a new user feels welcome and free to make mistakes.

Deployment

The Wikipedia Adventure was written and developed by a team of volunteers and Wikimedia Foundation staff.[7] It is implemented through the Guided Tours extension (a framework for creating interactive tutorials for MediaWiki), and was deployed in an alpha-test in October 2013 to a group of 50 editors.

After an initial round of testing and debugging, TWA was released on English language Wikipedia in beta in November 2013. The decision to release an English version of the system first was prompted by both the language competencies of the development team, as well as the fact that English Wikipedia is the largest language edition and has the most globally diverse contributor base. However, the system itself can be easily adapted to other language editions in the future through translation of the tutorial text.

Study 1: User Survey

After developing The Wikipedia Adventure, we sought feedback and input from new Wikipedia editors who used the tutorial. We conducted a user survey to collect this

[7] See: `https://en.wikipedia.org/wiki/Wikipedia:The_Wikipedia_Adventure`

feedback and evaluate user perceptions of the system design.

Methods

Alongside the initial beta release of TWA, we invited 10,959 editors to use the tutorial via their talk page. We used a large scale deployment so that a diverse group of English Wikipedia users from across the world could participate in the tutorial and provide initial feedback.

We distributed invitations via user talk page messages on a rolling basis to new editors who satisfied the following criteria: they had created their account within the past 24 hours, they had made at least 2 edits, and they had not yet been blocked or received a Level 4 user warning message on their talk page.[8] We distributed the survey invitation to those who used the tutorial after the initial invitation to play the game. The survey invitation consisted of another talk page message with a link to a Qualtrics survey.

Since we sent invitations through talk pages which are visible to the public, it was technically possible for users who were not in our invitation sample to play the game and take the survey. While the likelihood of this is small (new editors see relatively little traffic from other editors to their talk pages), we cannot rule out the possibility that our survey response data may include responses from community members who did not receive the original invitation.

[8]Wikipedia has four levels of warnings given to users suspected of vandalism. Level 4 warnings are the most severe and are reserved for users who commit extreme or frequent vandalism in bad faith.

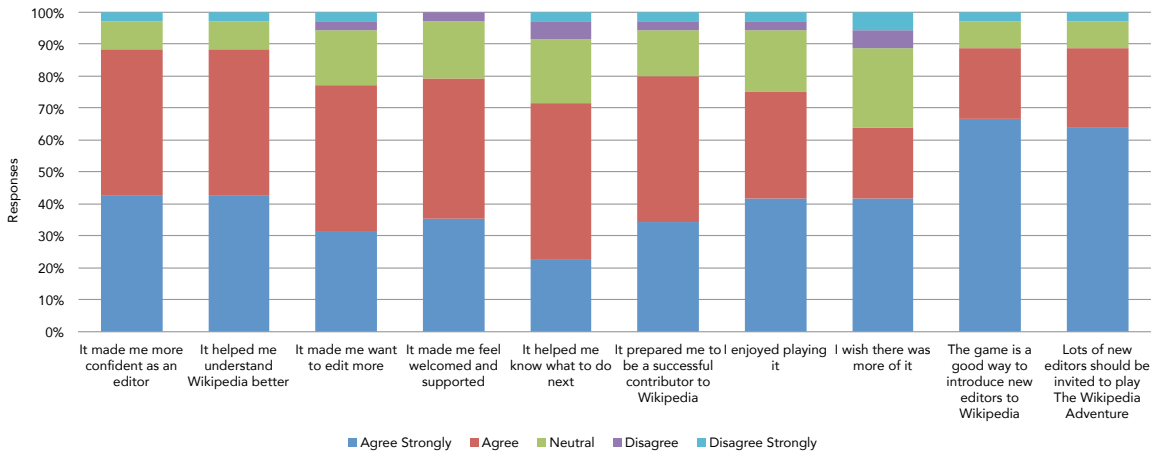


Figure 4: Survey results measuring user confidence, engagement, and satisfaction with The Wikipedia Adventure.

Measures

In keeping with the goals of the system design, the survey examines how survey participants perceived the impact of TWA on their confidence and engagement in editing Wikipedia, whether it communicated effectively, and whether users were satisfied with the design and tone of the tutorial. We also gathered participants' demographic information.

To measure *user confidence*, we asked users to rate on a 5 point scale the extent to which they agreed or disagreed with statements about TWA. Statements included "It made me more confident as an editor," "It helped me understand Wikipedia better," and, "It prepared me to be a successful contributor to Wikipedia." To measure *user engagement*, we asked users to rate on a 5 point scale the extent to which they agreed or disagreed with statements such as "I enjoyed playing it," "It made me want to edit

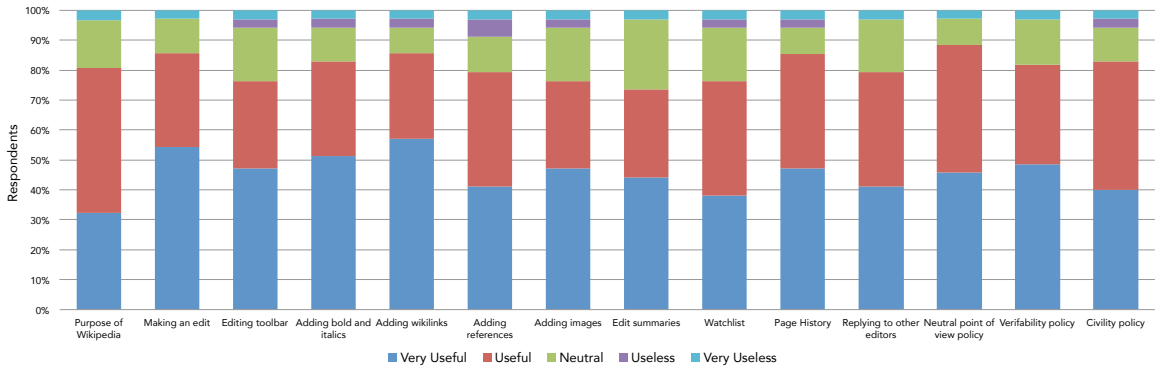


Figure 5: Survey results measuring perceived utility of The Wikipedia Adventure for performing specific editing tasks.

more," and "I wish there was more of it." To evaluate how clearly TWA communicated information related to editing Wikipedia, we asked respondents to rate on a 5 point scale how well the tutorial taught specific skills such as, "making a wikilink," "adding references," and "adding images." To measure *design satisfaction*, we asked users to rate on a 5 point scale the effectiveness of the interactive elements of the tutorial, their degree of satisfaction with its tone and visual design, and their overall satisfaction with the design.

We also asked participants to select all age groups for which they thought the tutorial would be most appropriate and effective from six choices of age groups that ranged between children below the age of 12, and adults over the age of 55. Finally, we also collected a number of open ended responses from the survey participants.

Results

Out of the 10,959 individuals invited to use the tutorial in the initial beta test, 600 (6%) clicked through and completed at least one mission. From the 600 who used the tutorial and then received an invitation to the survey, 42 individuals (7%) responded between December 23rd, 2013 and January 4th, 2014.

Respondents to the survey came from a number of countries: Australia, Bangladesh, Brazil, Canada, Estonia, Hong Kong, India, Ireland, Nigeria, Portugal, Singapore, Sweden, Macedonia, US, and UK. Although globally diverse, the majority of players came from US or the UK. Close to 11% of the survey respondents identified as female. About 94% had made 100 or fewer edits to Wikipedia, suggesting that our sampling successfully targeted newer editors.

The survey responses suggest that participants felt TWA was an effective way to welcome newcomers and teach them about Wikipedia. We found that 90% of respondents agreed or strongly agreed with the statements "It made me more confident as an editor," and "It helped me understand Wikipedia better,", suggesting that TWA could help build confidence among new editors (see Figure 4). We also found that 80% of respondents agreed or strongly agreed with the statements "It made me want to edit more" and "It made me feel welcomed and supported," which suggests that TWA was helping build engagement among newcomers.

In terms of specific Wikipedia-related skills, over 85% of respondents reported that TWA was "useful" or "very useful" in explaining the neutral point of view policy, how to make an edit, how to add wikilinks, and how to view page histories (see Figure 5).

Overall, 91% of respondents found The Wikipedia Adventure "useful" or "very useful" as an introductory tutorial on editing Wikipedia.

Open-ended responses provide additional support for these findings. For example, one respondent noted that "the interactiveness of The Wikipedia Adventure was an easier and better way to learn the basics of Wikipedia versus trying to run around to different pages and just reading about it."

Since The Wikipedia Adventure incorporates a narrative that uses space metaphors and graphics of cartoon aliens, we measured whether these elements appealed to new users. When asked if they would have preferred if TWA had a more 'serious' tone and design, 70% of respondents reported that they liked the tutorial as it was, while 14% wished it were more serious. 76% of participants responded positively to the interactive elements of the tutorial. Respondents suggested that the most appropriate age groups for the tutorial were teens and young adults between the ages of 13 and 22. Overall, 83% of respondents were "satisfied" or "very satisfied" with the design of the tutorial.

One respondent noted:

> It was all beautifully designed. I enjoyed aspects such as the challenges and badges that made it feel more like an educational tool or game rather than a lecture, and [the way that it] recorded your achievement to date.

The positive response to the interactive elements of TWA, seen through both the survey questions as well as the open ended responses, provides validation of our choice to gamify the tutorial. Another respondent said, "I completed the entire game because

it wasn't as dry as other training tools out there," suggesting that features such as the interactivity and narrative structure help maintain interest in learning how to edit Wikipedia.

Overall, the survey participants found TWA an effective and useful tool. In particular, respondents valued the general introduction to Wikipedia provided, and reported that the system improved their understanding of Wikipedia and gave them more confidence to edit.

Study 2: Field Experiment

As we suggested in our discussion of gamification, the nature of the relationship between gamified participation and participants' subsequent engagement in the task they have learned about remains an active area of research in social computing. Even if a gamified tutorial is engaging, enjoyable, and effective at supporting learning, this might not translate into increased participation after the tutorial is over. Will playing a gamified tutorial like TWA lead to more contributions from newcomers on Wikipedia? The potential for increased enjoyment, confidence, and feelings of self-efficacy are reasons to believe it will. To this end, we use a large-scale, invitation-based field experiment on Wikipedia to evaluate the effect of TWA on the contribution activity of newcomers.

While the survey in Study 1 measured the subjective user experience of playing the game, an experiment tests for causal, behavioral effects of the tutorial on users' subsequent contributions outside of the game's learning environment. A *field* experi-

ment that deploys TWA "in the wild" allows us to estimate the effects that such an intervention would have on Wikipedia newcomers at a large scale.

We use an invitation (or encouragement) design in which some users are randomly invited via their talk pages to play TWA. The Wikipedia community has a tradition of allowing low-barrier contributions on any topic without requiring contributors to do so much as register an account. Due to this tradition and the significant length of the tutorial, the plan for deploying TWA required users to opt-in to the system. In other words, self-selection is a part of the system, and the evaluation of the system's impact needs to incorporate this element of the design to account for the fact that many newcomers might choose to never play the game at all. The choice to use an invitation-based experiment design thus supports the analytic clarity and realism of the experimental results.

TWA was designed to reduce skill-related barriers to entry in editing Wikipedia and to provide an institutionalized, gamified, introduction to concepts like making an edit, using wikimarkup, and communicating via talk pages. Following the evidence of prior work discussed above and consistent with the results from Study 1, we hypothesized that newcomers who played TWA (conditional on having received an invitation to do so) would make:

($H1$) *an increased number of contributions overall*,

($H2$) *an increased number of contributions on talk pages*,

($H3$) *contributions of greater average quality.*

Methods

These hypotheses were tested using an experiment that followed the deployment described in Study 1, after the system was no longer in beta. Starting in February 2014 and continuing over a period of three months, we identified accounts on English Wikipedia to be included in our study on a rolling basis using the same criteria used for inclusion in Study 1. Qualifying accounts were identified on a daily basis and randomly sorted into treatment and control groups. For each user in the treatment group, we sent an invitation to play TWA via their talk page within two days of the creation of their account. We designed the invitations to closely resemble the way that such an intervention might be rolled out on Wikipedia at a large scale. The invitation incorporated graphics from TWA which contrast heavily with other text on Wikipedia, and was thus more likely to be noticed. The invitations were sent out via HostBot (which has been used in the past to invite newcomers to the Teahouse [60]) and logged-in users in the treatment group received a notification that they had been sent an invitation.

To ensure that every participant in our sample had a chance to see their invitation, we only included participants who made at least one edit after getting invited in the analysis. To maintain an equivalent sampling procedure in the control group, we kept only those editors who made at least one edit after the time they would have been invited had they been assigned to receive an invitation. We observe no evidence of cross-over between the treatment and control groups—i.e. no users in the control group used TWA.

In total, we identified 1,967 accounts to be included in our study. Of these, 1,751

(89%) were randomly selected to form our treatment group and were invited to play TWA. The other 216 users in our study were placed in a control group and received no invitation. We chose an imbalanced design based on preliminary evidence that the uptake of the invitations to play the game would be low (6% in Study 1). Of the 1,751 users invited, 386 (22%) completed at least some portion of the tutorial. This increase in uptake compared to Study 1 may be due to changes in the invitation text.

We chose to observe the editing behavior of every user over 180 days after their *date of inclusion* in the study. The date of inclusion for a user in the treatment group is either the date that invitation was sent or, for users in our control group, the date the invitation *would have been sent* had the user been in treatment. Although a longer data collection period provides more time to observe systematic variance between the treatment and control groups, it can raise concerns that differences long after the intervention may not be justifiably attributed to the intervention. Our 180 day window was chosen because it is as long as any previous field experiment or system deployment in Wikipedia that we have seen [60, 89].

Measures

Our dependent variables consist of three measures corresponding to each of our hypotheses. To test H1, we measure the overall contributions as the *total number of edits* made by each user in the 180 days after their date of inclusion in the study. This count excludes edits made to the subjects' user pages and user talk pages because TWA automatically generates edits that show up as contributions to these pages. We count

all others edits made to Wikipedia including those that were subsequently reverted or deleted. Our results are not substantively affected by the decision to include reverted or deleted edits.

To test H2, we measure the extent to which each subject interacted with others on Wikipedia as the *total number of edits they made to talk pages* on Wikipedia in the 180 days from the time of their inclusion in the study. This variable reflects the emphasis that TWA places on the community dimension of the system.

To test H3, we measure the *average quality of contributions* for each subject by calculating a measure of content "persistence" for all contributions to article pages using metrics developed in parallel by Adler et al. [1, 2] and Halfaker et al. [38, 71]. We estimate the quality of each edit, e_i, by calculating the number of subsequent edits within a fixed radius of subsequent edits in which each word in e_i persists before it is changed or removed. Our measure is the average persistent word score of the article edits made by the user in the 180 days from their inclusion in the study.

Although other radii have been used in research [27], we adopt a radius of 6 edits because this is what is used by Adler et al. [1, 2] in WikiTrust – the most frequently used and widely validated content persistence implementation. Following WikiTrust, we also collapse sequential edits by the same user. Although this will underestimate the productivity of users who edit very infrequently edited articles, we find no statistically significant difference between the mean size of radiuses for edits made by users in our treatment and control groups.

Our key independent variables are two dichotomous measures that indicate whether

a particular user was *invited* to play TWA, and whether or not they subsequently *played* it. We consider a user to have played the game if they completed any part of the game, regardless of whether they completed one mission or all seven. Our results are not affected by incorporating the number of missions played, and so we report models that use a dichotomous version of this measure.

Finally, we include a categorical measure capturing the date on which the subjects were incorporated into the research sample. Because random assignment took place within these *sample dates*, this works as a blocking variable that controls for unobserved heterogeneity introduced by running the study over several months [34].

Analytic Approach

Our analysis examined two different facets of the intervention: the effect of inviting a user to play TWA on subsequent contributions, and the effect of playing TWA, conditional on having been invited to do so, on subsequent contributions. Because invitation-based designs are uncommon in social computing research, we explain our analytic approach in detail below.

We estimate the effect that an invitation to use TWA has on a user's subsequent contributions by comparing invited users to users who received no invitation. This is known as an *intent-to-treat* (ITT) estimator, because it does not presuppose that all invited users necessarily received the experimental intervention (i.e. the experience of using the system) [34, 62]. Our ITT models provide unbiased estimates of the effect of distributing the invitations. For each dependent variable, the ITT model (M_{ITT})

takes the generic form:

$$(\mathrm{M_{ITT}}) \quad Y \sim invited + sample.date + \varepsilon$$

To test the impact of playing TWA, we estimate the effect of playing the tutorial conditional on being invited to do so.[9] Experiments of this type, also known as "encouragement designs," are analyzed using two-stage least squares regression (2SLS) [5, 24, 34]. In the first stage of 2SLS we estimate the likelihood that an invitation predicts playing TWA. The vector of fitted values from the first stage model then becomes a predictor in a second stage model which estimates the relationship with our outcome variables. Since the invitation was randomly assigned, the fitted values of the first stage model capture the variation in gameplay caused by the treatment. The second stage model produces an unbiased estimate of the causal effect of playing the game conditional on receiving an invitation.[10] In generic form, the first stage ($\mathrm{M1_{2SLS}}$) and second stage ($\mathrm{M2_{2SLS}}$) models we use for 2SLS are:

$$(\mathrm{M1_{2SLS}}) \quad played \sim invited + sample.date + \varepsilon$$

$$(\mathrm{M2_{2SLS}}) \quad Y \sim \widehat{played} + sample.date + \varepsilon$$

[9] Note that subjects who played TWA are a (small) subset of subjects who were invited to do so. As a result, comparing the outcomes for the set of TWA players against the entire control group cannot recover an unbiased estimate of the effect of playing the game. The control group contains subjects who would have played and those who would have not (had they been invited).

[10] We refer interested readers to several key references for formal details and proofs of 2SLS [4, 24, 34, 62]. We are not familiar with prior work in social computing that applies these methods.

Intent-to-treat and two-stage least squares estimators provide unbiased estimates of treatment effects because of the encouragement design of the study. Unlike a more typical lab experiment or A/B test, we observe relatively low uptake of the game by individuals in the treatment group. Although we might observe that users who played TWA contributed more, on average, than users in the control group (an actual relationship in the data), we must account for the fact that the vast majority of invited (treated) users never visited the tutorial. It is the full treatment and control groups that are "equal in expectation" prior to treatment assignment, and it would be misleading to compare the few users in the treatment group who worked through the tutorial to the full control group who, in large part, would never have done so [62, 76].

Participation is highly skewed in Wikipedia (i.e. a tiny percentage of editors make a large proportion of the total edits) and all of our dependent variables are over-dispersed count measures (see Table 2). As a result, we use negative binomial regression models for all of our estimates. This is typical for highly-skewed count variables and has been applied in prior field experiments on Wikipedia [75]. As a part of our ITT estimate of treatment effects, we also conduct a non-parametric Mann-Whitney U test to identify rare effects by estimating whether the dependent variables for the treatment and control groups are drawn from the same distribution [76]. For all regression models, we report heteroskedasticity and cluster-robust standard errors [5].

Measures	Min.	Median	Max.	Std. Dev.
Total edits	0	6	5282	159.97
Talk Page edits	0	0	365	11.84
Avg. edit quality	0	2.06	6	2.34

Table 2: Summary statistics for dependent variables.

Results

Results from the experiment are shown in Tables 2, 3, 4, and 5 as well as Figure 6. In total, 386 (22%) of those invited completed at least some part of the game. Table 2 describes the distributions of our key measures.

Table 3 shows how many users dropped out of the game after each mission. We find that 181 out of 386 (46%) played till the seventh (and final) mission. The highest dropoff, 93 (24%), occurred after the first mission, with smaller numbers dropping out for most subsequent missions. After the first mission many users kept playing all the way to the final mission.

Mission	Topic	Attrition
1	Editing user page	93
2	Using talk pages	40
3	Editing articles	18
4	Neutral point of view	9
5	Verifying sources	11
6	Civil discussion	34
7	Adding sections	181

Table 3: The attrition for every mission is measured as the number of subjects who play some part of the mission but did not go on to play subsequent missions. For example, 93 subjects played some part of Mission 1, but did not proceed to Mission 2. A total of 386 subjects played TWA.

Outcome	Diff. of medians	Test statistic	p-value
Total edits	-1	173492.5	0.05
Talk Page edits	0	190527	0.79
Avg. edit quality	-0.6	181712.5	0.34

Table 4: Results of non-parametric Mann-Whitney U tests for all dependent variables.

Figure 6 plots the distribution of all three outcome measures across treatment and

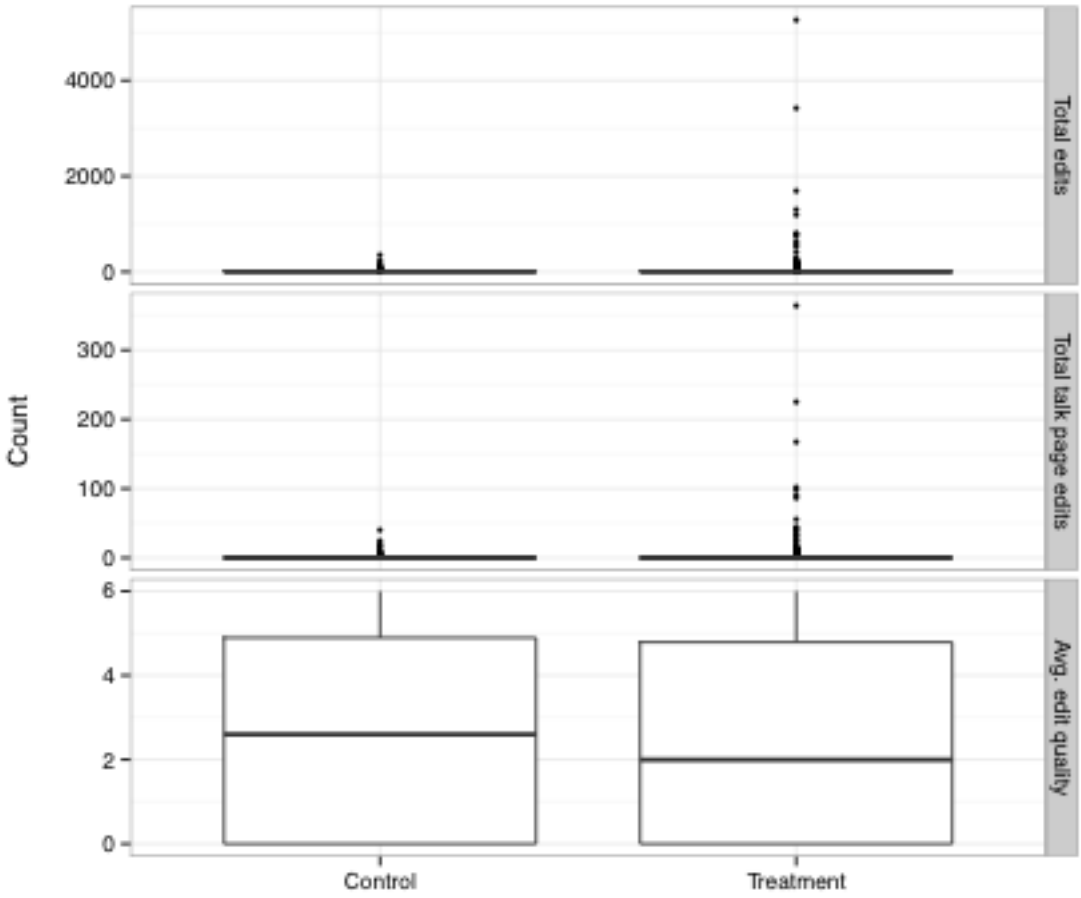


Figure 6: Boxplots showing the distributions of outcomes for total edits (top), talk page edits (middle), and average edit quality (bottom) across subjects assigned to the treatment and control conditions.

control conditions. We note that for the first two outcomes (total edits and total talk page edits) many subjects in both the treatment and control groups registered no further edits of either kind following the intervention. In terms of the distributions, we see that the treatment condition had a longer tail (a handful of extremely high outcomes) along both of these measures. The third boxplot illustrates that outcomes

Estimand	Model	Dependent Variable		
		Total edits	Talk page edits	Avg. quality
Invited to play TWA	Negative binomial	-0.107	-0.146	-0.065
		(0.112)	(0.273)	(0.088)
Playing TWA conditional on invitation	2SLS	-0.545	-0.730	-0.155
		(0.478)	(1.151)	(0.375)

Table 5: Regression results estimating the effects of (1) the invitation to play TWA and (2) playing TWA conditional on having been invited to do so on three measures of newcomer participation. For each dependent variable, we provide coefficients with standard errors in parentheses. The models reported here all include a control (unreported) for the number of days that each participant had edited Wikipedia. The results are substantively unchanged when we drop the control.

for average edit quality are distributed in a nearly identical fashion across the two conditions.

We report the results of non-parametric Mann-Whitney U tests in Table 4. The table includes the difference in medians ($\mu_{Y|treatment} - \mu_{Y|control}$), the value of the test statistic U, and the corresponding p-value for each dependent variable. The results indicate that the distributions of talk page edits and average edit quality are not different. In the case of total edits, the difference of medians is -1 and the p-value reaches conventional levels of significance, suggesting evidence of a negative distribution shift caused by the treatment. We interpret the results of the test as weak evidence of a statistically meaningful variation between the individuals invited to play TWA and those who were not. We suspect that this might be because the time taken by new editors to play TWA cut into the potential time they had to contribute to Wikipedia early on.

In our intent-to-treat analysis using negative binomial regression to estimate the

effect of an invitation to TWA, our models produce small estimates for all our dependent variables that are not statistically distinguishable from zero (see the top part of Table 5).

When we test the effect of playing TWA with 2SLS (see the bottom part of Table 5), we also find null effects with small coefficients for all dependent variables. For all coefficients, the standard errors are relatively large compared with the estimates and none approach conventional levels of statistical significance.

The parameter estimates for our null results represent well-estimated zeroes and suggest that any underlying effect we are unable to estimate with our sample would likely be extremely small. Post-hoc power analysis shows that if a data set of this size displayed even a small effect size (0.2 standard deviations), we would have had a 99% chance of detecting it at the 0.05 significance level. Thus, we can conclude that TWA does not alter the quantity or quality of newcomers' contributions to Wikipedia.

As a robustness check, we also estimated models using measures of our dependent variables computed over both 360 and 60 days following inclusion in the study. We again find null results for all three hypotheses over 360 days as well as for H2 and H3 over 60 days. Echoing our Mann-Whitney U test, we find a small negative relationship in our test for H1 in the 60 day dataset. In these results, we estimate that invitation to the game (ITT) was associated with approximately 1.25 fewer edits and approximately 3 fewer edits in our 2SLS model over the 60 days after inclusion. One potential explanation is that participation in the tutorial may have supplanted other editing among participants in the treatment group but that this effect is "washed out" over time. In

any case, the pattern of results across the three study lengths is not consistent with predictions from previous work that the system would cause new users to contribute more. If anything, there is weak evidence suggesting that TWA might have caused them to make several fewer edits in the period immediately following inclusion in the study.

Discussion

The results from Study 1 show that respondents found TWA to be a useful and satisfying tool for learning how to edit Wikipedia. In particular, study participants valued having a system that provided a general introduction to Wikipedia, and stated that it improved their understanding of the community and gave them more confidence to edit. Study 2, however, shows that despite the perceived effectiveness of the design and the satisfaction of the users, playing TWA did not alter the subsequent behavior of newcomers on Wikipedia.

The survey responses validated the idea of using gamification to introduce an institutionalized form of socialization to Wikipedia. Users found the gamified aspects of the tutorial rewarding and engaging and agreed that a tutorial that provides a broad overview of editing should be shared with new editors on Wikipedia. These findings suggest that we accomplished our system design goals and that the tutorial provided a compelling and enjoyable institutionalized introduction to the skills, norms, and expectations involved in becoming a Wikipedian. We believe these findings validate some of the claims of prior gamification research as well the theoretical justification for pur-

suing institutionalized socialization as a complement to existing onboarding systems in Wikipedia.

Study 1 has several limitations, including the small pool of respondents and the fact that the survey can not capture the reasons why some individuals chose not to play the game. The limited uptake of the game and low response rate of the survey mean that our Study 1 findings might not extend to all the individuals invited to play TWA or even to all users who played it as part the initial deployment. Additionally, the survey does not assess whether respondents who played TWA effectively learned the skills that the tutorial sought to introduce. Finally, the wording of the survey questions could have elicited overly positive responses due to satisficing behavior and social desirability pressures. Nevertheless, the results of Study 1 made us optimistic that some newcomers would elect to play TWA and that doing so could have a positive impact on their subsequent engagement through increased enjoyment in learning how to edit Wikipedia and improved confidence and self-efficacy.

However, contrary to predictions from organizational and social computing theory and design, Study 2 shows that TWA had no measurable impact on newcomer participation. All of our statistical tests for regressions of all three outcome measures fail to reject the null hypothesis of no effects regardless of the model specification or the estimand (ITT or playing TWA conditional on receiving an invitation). Robustness checks conducted on a smaller data collection window suggest that TWA may have even reduced contribution rates in the short term. We conclude that these results demonstrate a null effect.

The limitations of Study 2 include the possibility that the invitation process or deployment of the system might have shaped the outcomes in ways we cannot detect. The visual differences between the outcome distributions in Figure 6 also suggest that we might explore alternative, novel estimation techniques focused on detecting rare, large effects [76]. It is also possible that the treatment impacted some other outcome variable that we do not measure in this study, such as survival rate. Future research might explore these questions.

What factors might explain the null effects in Study 2? The statistical tests we report are appropriate to the design and the study had adequate statistical power to detect treatment effects had they existed. We propose several interpretations, which focus on the system design, the culture of the Wikipedia community, the self-selected and voluntary nature of participation in peer production communities, and the limitations of gamified interactive systems.

System design factors

Shortcomings in the design of TWA offer one possible explanation of the null effects in Study 2. If the tutorial itself was poorly designed, a better implementation may have altered new editor contributions. For example, if TWA users perceived editing within a sandboxed environment instead of a live Wikipedia page as an illegitimate or inauthentic form of participation [55], this might have undermined the system's effects. Although other work has suggested that designs like ours can overcome these effects [35, 47, 79], it remains a possible explanation and a design choice worth revisiting.

The fact that the survey respondents in Study 1 overwhelmingly perceived TWA to be positive and well-designed also suggests that the system design did not have glaring shortcomings. Subsequent to the completion of this study, many more TWA users have also given the tutorial glowing feedback through the comment box on the game's webpage. This implies that limitations of the particular system design do not fully explain the null result.

The temptation to attribute shortcomings visible only after field testing to details of our implementation of TWA points to a larger concern about understanding the impact of any new system. A creative and thoughtful designer can always imagine alternative approaches that *might* transform the effects of an existing system, no matter how carefully planned or executed the existing design may be. Our findings in these two studies indicated that TWA's system design satisfied the criteria and goals of the system's creators, as well as the system's early users. Even so, it did not produce the effects predicted by either theory or preliminary testing. For this reason, we believe limitations in the existing theories as well as the specific conditions of TWA's deployment better explain the observed outcomes.

Cultural factors related to Wikipedia

Prior research points to several ways that Wikipedia's existing culture may have undermined TWA's expected effects. The new editors in our study may have had unpleasant experiences during their initial time on Wikipedia that negated any potential motivational benefits they may have gained from playing TWA. Even for experienced

contributors, the abrasive and hostile tone of interactions among Wikipedia editors deters participation.[11] Many new editors receive multiple warning messages on their talk pages within their first few editing sessions [33]. These warning messages, the majority of which are automated, strongly-worded, and accompanied by a revert, can drive new editors away from the site [36].

Just observing toxic exchanges among other Wikipedians could have convinced the new editors in our study that the Wikipedia community is not a welcoming place. A lighthearted, automated tutorial depicting a collegial collaboration process may not be sufficiently compelling to counteract these negative observations or experiences. We cannot confirm or reject this possibility fully through our empirical analysis because we do not know what perceptions study participants who dropped out of editing may have had.

Limitations of gamification

Most previous studies of gamified systems have focused on subjective measures of engagement with, and enjoyment of, the system. With few exceptions, they have not evaluated the impact of gamified systems on subsequent performance. Our study is one of the first to assess the effect of a gamified learning system on engagement behaviors outside of the system itself. Although some studies have shown that gamification can support learning, meta-analyses have suggested that the few studies that did analyze impact on performance did not reliably show improvements [40]. This may help explain

[11] This is based on the Wikimedia Foundation's survey of former contributors referred to in a previous footnote.

the contrast between our survey results and our field experiment.

Another explanation stems from the shift in incentives between the game and the "real" world of Wikipedia contributing. It is possible that the extrinsic motivation provided by the gamified tutorial was simply not replaced by intrinsic motivations needed to drive subsequent contributions. In a project like Wikipedia that depends heavily on intrinsically motivated members to make contributions, a gamified tutorial may be helpful and fun to use, but ultimately unsuccessful at building long-term commitment and retention. This echoes our earlier point about the possible limitations of sandboxed learning environments. The current study contributes to our understanding of the effectiveness of gamification by presenting both a subjective evaluation of a novel gamified system, as well as a measurement of its subsequent impact in a non-gamified context.

Self-selection and voluntary participation

The voluntary nature of participation and membership in peer production communities like Wikipedia offers another possible explanation. Contributors engaged in peer production self-select into their preferred communities, tasks, and social roles [11, 14, 86]. Institutionalized training programs may be more effective in more formal organizations because newcomers cannot select out of them as easily. TWA provides a general overview of contributing to Wikipedia, but a person who is interested solely in learning how to accomplish one specific task (e.g., fixing a citation) might find the full tutorial burdensome and choose not to play. An institutionalized socialization approach might

also fail because it tries to artificially speed up the process of "becoming Wikipedian", which involves a gradual transformation from participating peripherally to seeing oneself as part of the community [15].

The dynamics of self-selection may best explain the null effect of Study 2. Because TWA depends on users to choose to play the game, those who do so are likely to vary systematically from those who do not. Specifically, the newcomers who received and accepted our invitation may be more motivated, committed, or skilled than those who received the invitation and chose not to play. Playing the game may have given the exceptional newcomers a positive experience without impacting the quantity or quality of their subsequent contributions. The fact that a subset of individuals in both the treatment and control conditions went on to make numerous edits of high quality supports this idea. Our analytic approach with 2SLS supports this inference in that we identify the causal effect of playing TWA conditional on having received the invitation. The null findings in these models indicate that the people who played the game and went on to contribute extensively would have done so anyway. We cannot say more generally whether Wikipedians may be "born," "made," or some combination of the two [45, 68]. We conclude that TWA did not make active editors out of people who would have been inactive in the absence of the game.

This study illustrates the value in evaluating novel systems in "live" field deployments within communities. As discussed above, the findings of Study 1 validate many of the design principles and findings from prior literature on gamification and newcomer socialization. However, Study 2 revealed that these principles and findings cannot explain the empirical impact of the system. Study 2 does not invalidate prior

work on institutionalized socialization or gamification, but it does show that successful newcomer orientation in a volunteer community like Wikipedia remains a compelling design challenge.

While the invitation-based field deployment of TWA yielded no effect on newcomer contributions, it is possible that such a system would work better in contexts where newcomers are required to play it before editing Wikipedia, thus circumventing the self-selection issue. For instance, using the tutorial in a classroom setting where students are required to contribute to Wikipedia (increasingly common through initiatives such as the Wikipedia Education Program) might produce positive results.

Conclusion

We designed and evaluated The Wikipedia Adventure, a gamified, interactive tutorial that extended techniques of institutionalized socialization to newcomers in Wikipedia. The first part of our evaluation, a user survey, validated the principles, theories, and goals of TWA's design. The second part of our evaluation, an invitation-based field experiment, revealed that deploying the system did not alter newcomer contribution patterns over several months. We suggest that the null findings may be due to a combination of factors including the culture of Wikipedia, limitations of gamified systems, and the dynamics of self-selection in voluntary peer production communities. We believe that our second study represents the first invitation-based field experiment and application of two-stage least squares in the social computing literature.

Given the positive results in Study 1 and the strong theoretical support for its design, we were genuinely surprised by the null result in Study 2. The discrepancies between the results in our two studies point to an important secondary contribution of our work. Despite the positive response from users that were surveyed in Study 1, our field experiment in Study 2 demonstrated clearly that subjective perceptions of utility and usability do not necessarily translate into lasting changes in user behavior.

Although null results can be difficult to convincingly establish and interpret, they can play an important role in contributing to our knowledge of social computing theories and systems. TWA's design was informed by previous empirical, theoretical, and systems work, and it performed well according to the types of survey self-report measures used to evaluate the usability of many social computing systems. Post-hoc power analysis suggests that our estimates in Study 2 are well estimated zeros and that our sample size is sufficiently large to detect even small effects. Our work does not provide the final word on institutionalized socialization or gamified tutorials in peer production. That said, we believe it contributes to our understanding of these topics through both what we have been able to show, as well as what we have not.

Finally, we believe that our work shows how any intervention that attempts to assimilate new users into an existing peer production community might be limited when deployed in the wild. Wikipedia has complex norms and rules for participation which are obscure to newcomers. Institutionalized and gamified socialization systems like TWA may inform the design of future orientation systems, but more profound changes to the interface or modes of interaction between editors might also be needed to increase contributions from the targeted groups.

More Connected but not More Productive: Analyzing Support for Interpersonal Communication in Wikis

Co-authored with Nate TeBlunthuis, Wm Salt Hale, Benjamin Mako Hill, and Aaron Shaw

Introduction

Online communities engaged in the co-creation of public information goods rely on extensive collaboration between users. Although Wikipedia is the most well-known example of a community that produces in this way, the model of production that Wikipedia popularized has been emulated in millions of smaller wikis and in numerous other types of online communities.

Studies have found that as these projects mature, members increasingly devote time to coordination through interpersonal communication to better organize their own efforts and to assign tasks [49, 52]. Furthermore, interpersonal communication

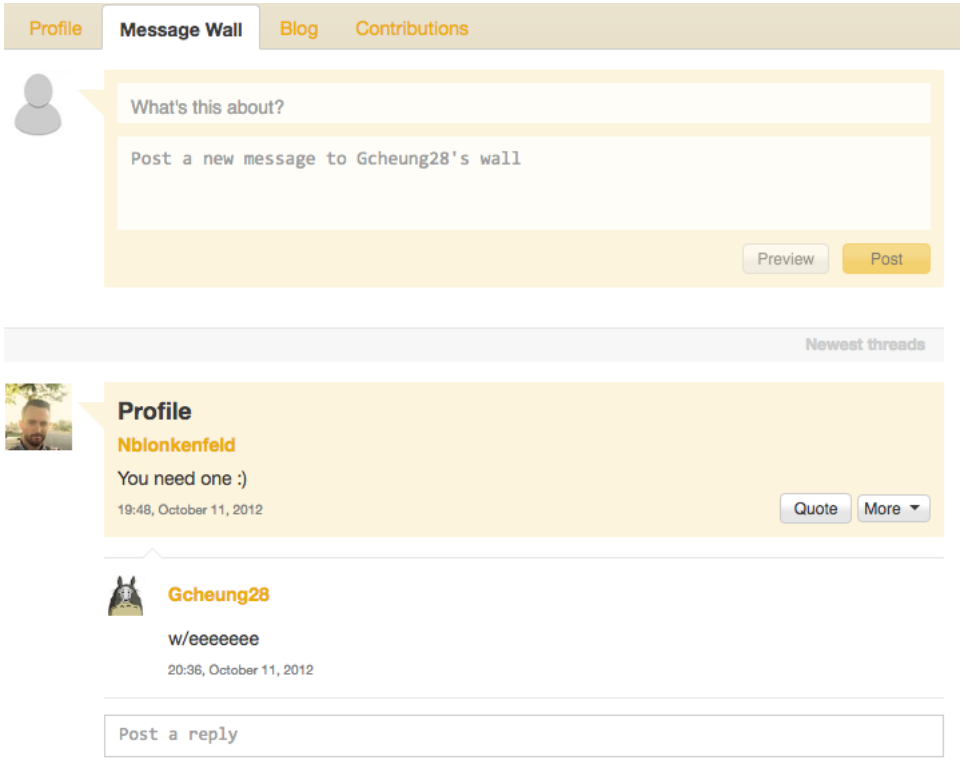


Figure 7: Example of a "message wall" page in the *Recipes* wiki. The top half shows the interface for starting a new discussion thread by entering both subject and body text. The bottom shows an existing thread consisting of a message and a reply.

plays a central role in welcoming and initiating new members into communities by facilitating information seeking and mentorship, enabling newcomers to become better integrated and participate more effectively in the organization [3, 15, 57]. This prior work suggests that better support for interpersonal communication might lead to both increased communication and increased productivity, particularly among newcomers.

Driven by these ideas and earlier findings, community managers and systems designers have invested heavily in supporting interpersonal communication. For example, the Wikimedia Foundation has spent many years building a system called "Flow"

which aimed to improve the ease of interpersonal communication in Wikipedia.[12] Despite such investments, little evidence directly evaluates whether easier interpersonal communication translates into increased productivity in these settings.

In this paper, we provide a large-scale empirical test of whether easier interpersonal communication leads to enhanced productivity and newcomer participation in peer production communities. To do so, we examine a population of 275 wikis hosted by Wikia that made interpersonal communication easier by introducing an interface called "message walls" (shown in Figure 7). We estimate the impact of the new design by analyzing panel data on rates of user participation before and after wikis transitioned to the message wall interface.

We find that the move to message walls is associated with increases in communication both overall and among new contributors. However, we do not find evidence of changes in article editing activity beyond a short-lived initial bump. Contrary to our expectations and prior theory, we find evidence that the transition is actually associated with *lower* rates of contribution to articles from new editors, a group the intervention was designed to impact in precisely the opposite way. Our research contributes to human-computer interaction, social computing, and organizational research by examining how design changes in socio-technical systems impact community-level outcomes and by suggesting that systems facilitating easier communication may not increase productivity.

[12] `https://www.mediawiki.org/wiki/Structured_Discussions`

Background

First introduced by Benkler [11], the term "peer production" refers to the model of organizing the collaborative production of public information goods that has created Wikipedia, GNU/Linux, and a range of other important knowledge resources [14]. Interpersonal communication features prominently in peer production. Haythornthwaite [42] goes so far as to use the presence of interpersonal communication to distinguish between the types of projects that Benkler considers peer production and related efforts like crowdsourcing that, for reasons unrelated to communication, Benkler treats as distinct [13].

Research points to two closely linked reasons that interpersonal communication may play such a central role in successful peer production projects. First, more communication may lead to more efficient work. The lack of top-down decision-making in peer production projects means that volunteers must define collective goals, develop and enforce social norms, and create organizational structures and workflows [14]. Successfully doing so may be facilitated by—and in some cases, may only be possible with—intense and frequent interpersonal communication between participants.

In particular, interpersonal communication frequently underpins effective coordination, which a large body of social computing and organizational research links to greater collaborative productivity [52, 59, 51]. In prior studies of peer production in both Wikipedia and Wikia wikis, higher levels of coordination explain variation in content quality [49, 50]. Among work groups and virtual teams, a similar pattern occurs whereby collaborators with more integrated communication network structures

perform better at complex information seeking and problem solving tasks [9, 22].

Interpersonal communication may also elicit more participation in peer production projects. Increased levels of participation are important because success in peer production is frequently defined—by both researchers and participants—in terms of communities' ability to marshal the volunteer resources needed to sustain production [21, 78, 30]. Prior research suggests several pathways by which interpersonal communication might increase participation. Just the presence of other contributors to a collective endeavor can sustain participation [88] and communicating and interacting with other volunteers motivates many peer production participants [15, 54]. Higher levels of social interaction also support group identification, commitment, and socialization, all of which are associated with increased levels of participation in online communities [51]. Spaces for interaction provide opportunities for feedback, which can elicit additional contributions and increase contribution quality [18, 91].

Interaction also facilitates the formation of collective identities that sustain participants' commitment to communities. For example, sustained engagement with other volunteers leads Wikipedia editors to transition from seeing the project as simply a repository of articles to a collection of people working together to maintain a public good [15]. This process of social integration through communication can build a cohesive "organizational culture" that sustains collaboration [48], and deeper social bonds that encourage commitment to the project [73, 74].

Given this prior work, if a community made it easier for members to communicate with one another, we would expect to see an *increase in the number of messages (H1a)*

that occurs within that community. We would also expect to see an *increase in the number of communicators (H1b)*. Prior work also leads us to expect that increased communication will result in an *increased number of contributions (H2a)* to the project made by a *greater number of contributors (H2b)*.

Interpersonal communication and newcomers

Interpersonal communication supports a second critical process in peer production: the socialization of new contributors to communities through legitimate peripheral participation [51, 57]. Although peer production projects have experimented with formal and structured forms of socialization [64], newcomers to peer production projects are almost exclusively oriented through unstructured individualized processes. In these processes, newcomers typically begin as lurkers [6] and may become occasional and then more active contributors [70, 7]. Interpersonal communication contributes to this model of socialization because newcomers learn norms, routines, and cultures of groups through conversations with more experienced contributors [15, 70].

Discussion can elicit newcomers to pursue sustained participation to peer production projects in several ways. First, newcomers seek information from veteran volunteers with questions about contributing to the project [3]. Asking questions and receiving answers helps newcomers acquire the skills to effectively make contributions [60]. These question-and-answer interactions also demonstrate the existence of distributed mentorship in the community, signaling the community's overall investment in newcomers [58, 91]. Additionally, veteran volunteers often approach newcomers personally

to connect them with tasks that they might like or be well suited for. These overtures help provide role clarity, and encourage newcomers to take on more challenges and responsibility [63].

Interpersonal communication in peer production is typically archived and visible to other participants. As a result, new volunteers read existing discussion threads to learn more about the project and community norms. Past discussions help newcomers gain a sense of what kind of community they are joining and how they may or may not fit into it [87, 6].

Finally, interpersonal communication is the cornerstone for building affective bonds of camaraderie among volunteers. Through conversations (both on and off topic) and discussions of the project's history and vision, volunteers transition from viewing the project as a static information repository to a dynamic collective that is created and maintained by people like them [15].

If a community makes it easier for people to communicate with one another, we would expect *newcomers to increase the number of messages (H3a)* they engage in. We would also expect that easier interpersonal communication would enable more people to interact with one another leading to a *greater number of new communicators (H3b)* in each community. Since communication is an important facet of socialization within a community, we would expect that making it easier for newcomers to communicate with other members would help them navigate the community better, understand what tasks need to be done and how to do them. We anticipate this will lead to an *increase in the number of newcomer contributions (H4a)* as well as an *increase in the number*

of new contributors (H4b).

Empirical Setting

To test the hypotheses derived above, we examine the introduction of a new interpersonal communication interface in a large population of wiki communities engaged in peer production. The wikis are all publicly editable and are hosted by Wikia, a for-profit company that allows anyone to start a wiki on any subject. The wikis themselves span many different topics, though many of the largest relate to fan culture.[13] It is common to see Wikia wikis about gaming, comic book or movie franchises, popular TV shows, and science fiction lore.

Each Wikia wiki is a community of its own, although members may participate in multiple communities. All wikis on Wikia use the MediaWiki software popularized by Wikipedia. As is the case with Wikipedia, content produced on Wikia is distributed freely and released under a Creative Commons Attribution-ShareAlike license.

Intervention: From user talk pages to message walls

In general, wikis built using MediaWiki have discussion pages associated with every page on the site. Discussion pages for articles are called 'talk pages' and those for user profiles are called 'user talk pages.' While talk pages are used for discussion of particular articles, user talk pages are typically used for informal conversations between

[13]On October 4th, 2016, Wikia partially rebranded itself as Fandom, to highlight the prevalence of entertainment and fan communities on its site.

Figure 8: A user talk page on *Recipes* wiki

.

```
== ShortPages ==

Hi Kim, I noticed there seem to be a rather lot of pages with basicly no content.
([[Special:ShortPages|*]]) Would it be worth deleting them all, if so, I can get my bot to do
it, so its hidden on the RC.. --[[User:Lcawte|Lcawte]] 15:38, March 3, 2010 (UTC)
:Hi Lcawte, Can you send me a link and give me an example of these short pages? ----
[[User:Kimberly McCollister|<font color="Blue">Kim</font>]]<staff /> ([[User talk:Kimberly
McCollister|<font color="Blue" size="1">talk</font>]]) 17:56, March 3, 2010 (UTC)
::Click the link in the *... then click a random page on their... --[[User:Lcawte|Lcawte]]
18:05, March 3, 2010 (UTC)
:::Okay, I am going to be building out the ingredients pages so I don't want to get rid of all
of these. Still, some obvious ones like "Daniel" can be deleted.
```

Figure 9: Source code editor for text in Figure 8

.

editors, making introductions, asking questions, seeking information, making requests, and giving feedback. The default MediaWiki interface for all pages (whether articles, policies, or discussion) is exactly the same: pages, no matter their purpose, start blank and are editable through a "Wiki markup" text editor such as the one shown in Figure 9.

The freedom to edit any page in any way is both a hallmark of MediaWiki and a source of confusion. Starting a discussion on a blank page puts the onus on participants to organize their conversation (e.g., label distinct topics, indent or otherwise represent

relationships between comments), sign their messages with their name as they go along, and so on. All of this requires additional time, skill using Wiki markup, and knowledge of community norms and routines. User evaluation studies of the MediaWiki talk page have shown that new editors can take up to four or five minutes to understand and figure out how to edit a talk page [77]—a long period of time in comparison to many similar messaging interfaces. Additionally, the default MediaWiki notification system makes it a challenge to keep track of which conversations have received responses.

In September 2011, Wikia announced plans for a new feature—the message wall—to replace user talk pages.[14] Message walls resembled the discussion threads that are commonly seen on other social media, forums, and blogs (see Figure 7), and was offered to the Wikia user base as a way to make interpersonal conversations among editors easier. All other talk pages, such as those associated with articles and projects, were left unaffected. Editors access message walls through the same (relabeled) tab they used to access user talk pages in the past.

Message walls introduced a number of new features that changed and clarified how users communicate with each other. First, it moved away from the standard markup editor and introduced a text box that a user could directly type into when they landed on a different user's page (see Figure 7). This shortened the time it took to communicate with others, since users no longer had to load a new page or use Wiki markup to write their message. The message walls interface also automatically performs a number of operations that user talk pages previously required editors to do

[14] See: `http://community.wikia.com/wiki/User_blog:Dopp/Communicate_Easily_with_Message_Wall`

by hand. For instance, all messages left on user talk pages needed to be deliberately signed with four tildes ("~~~~") in order to have the user's name show up in the thread. In Figure 8, drawn from a real conversation that occurred on the same *Recipes* wiki shown in Figure 7, editors signed the first three messages in the thread, but the fourth one is left unsigned (a common oversight). Furthermore, the indentations visible in Figure 8 are manually inserted by the editors (represented by colons (":") in the Wiki markup in Figure 9). Message walls automatically associate usernames with posts, provide subject headers, and indent conversations. Since user talk pages require conversations to be threaded manually, the visual organization of user talk pages can vary across communities. For example, some wikis use a convention of displaying more recent comments on top, others do the opposite. Message walls provided a uniform interface that meant that communicating in a new wiki was a much smoother process for any editor.

Message walls also included a revamped notification system. Previously, a user would need to choose to follow an entire page in order to know when their message received a response, meaning that any new post on that page would generate a notification. Instead, the message walls notification system automatically notified users when they received a response on their specific thread.

A beta version of message walls was first released October 5, 2011 on five Wikia wikis. After a period of testing and debugging, it was introduced to five more wikis in November 2011. After adding a few new features (such as the notification system), it was made available to administrators as a wiki-level setting they could change in a dashboard present on all Wikia wikis in January 2012. By 2015, message walls became

the default user discussion option for newly created Wikia wikis, and almost all Wikia wikis were transitioned over to the new system. As a result, the feature was rolled out to wikis at different points in time and at different points in wikis' lifecycles.

Methods, data, and measures

To understand the impact of the introduction of message walls, we built and analyzed an exhaustive longitudinal dataset of interpersonal communication and contribution activity that occurred on all wikis that experienced the intervention. We identified our study population of wikis using records provided by Wikia. The company provided the exact date and time of 6189 events where wikis enabled message walls during the initial period when the feature became available between January 2012 and October 2, 2012. We then used publicly available database "dump" files (posted by an archival group called WikiTeam to the Internet Archive)[15] to generate exhaustive longitudinal trace data for 4380 of these wikis. Data for the remaining 1,809 wikis was not available in the WikiTeam archives.

Since message walls could be turned on and off by administrators, it was not always the case that wikis would transition to message walls permanently. As a result, we consider only the *first* time that each wiki in our sample transitioned to ensure that the transitions in our analysis are comparable. We also limit our analysis to wikis that adopted message walls continuously for at least two months after their first migration.

[15] A list of archives is available here: `https://archive.org/search.php?query=subject%3A%22wikiteam%22`

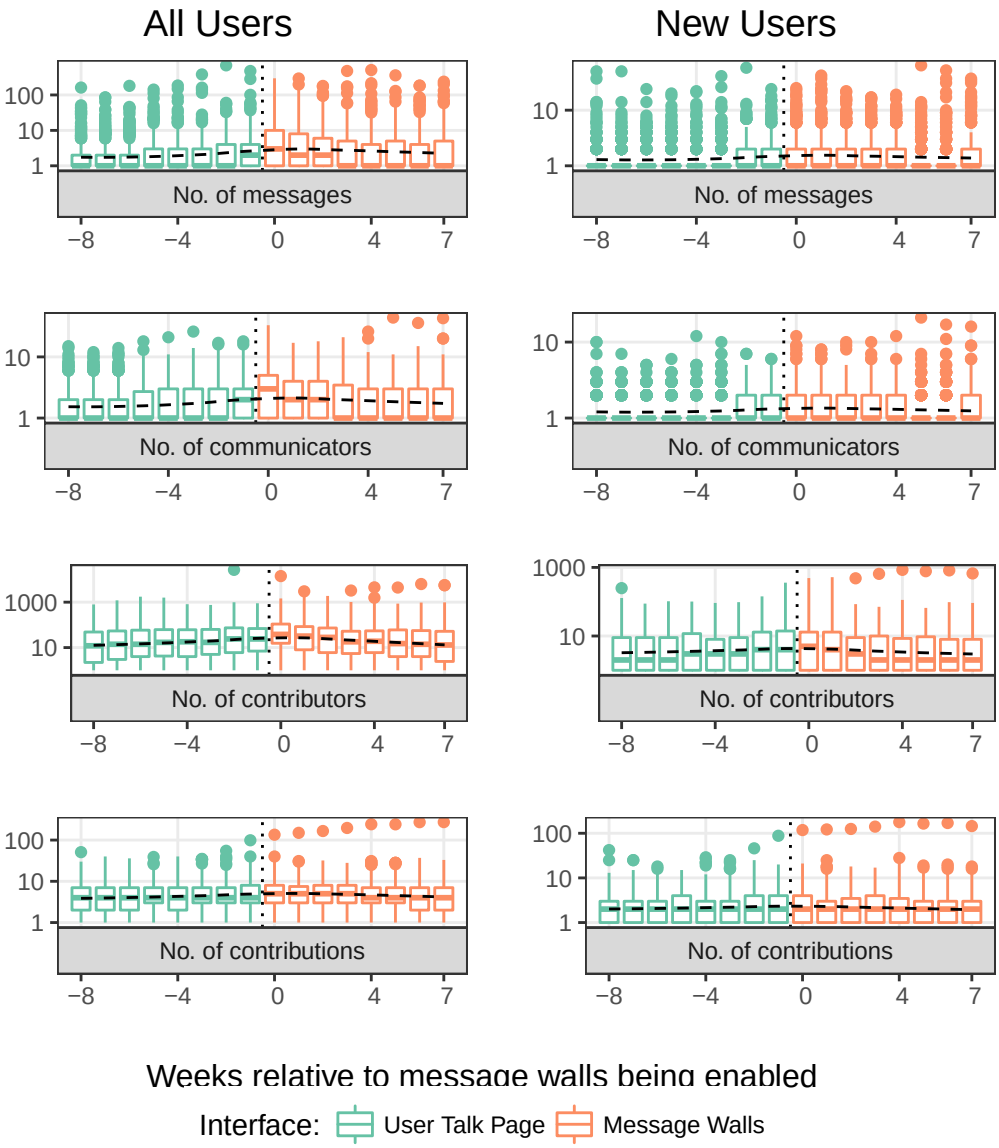


Figure 10: Boxplots showing the distributions of dependent variables for all editors during the analytic window, grouped by week.

We apply several other exclusion criteria. In order to compare only wikis that adopted the new feature for a sustained period, We drop wikis that turned off message

walls for more than 10 minutes during the eight weeks after it was first turned on. We also excluded wikis from our analysis that had no edits or that had not existed for at least four weeks prior to their migration to message walls. This ensured that communities had some period of exposure to user talk pages against which we could draw a comparison. We also exclude inactive wikis by requiring wikis to have at least one edit in at least three quarters of the weeks of the study period in our analysis. Applying all of these exclusion criteria left us with a dataset of 275 wikis.

To estimate the impact of the introduction of message walls, we analyzed activity occurring 8 weeks before and 8 weeks after each wiki moved to the new interface. We believe that this is long enough to allow us to model underlying trends in contribution activity while short enough that factors unrelated to the message walls transition are less likely to impact the results. Within the 16 week analytic window, our unit of analysis is the *wiki week*, meaning that for every measure we aggregate observations from each wiki into week-long bins.

We only include activity from registered editors (user accounts) on Wikia. While unregistered users may edit wikis with attribution made to their IP address, we choose to exclude them in this case. IP addresses are not reliable identifiers, and dropping unregistered users avoids double-counting editors who may both edit from a registered account and edit anonymously. We describe the individual measures in detail below.

Outcome Measures

Number of messages

Our first dependent variable is the aggregate amount of interpersonal communication on the wiki. In the first 8 weeks of our study, users communicated by editing user talk pages, then they abruptly transitioned to message walls, which they used for the second 8 weeks.

Edits to user talk pages and posts to message walls leave different kinds of traces in the data recorded by Wikia. For example, multiple edits may be made to a user talk page to create text that appears as a single comment while a single edit to a message wall is more likely to reflect a single comment. In order to make a single commensurate measure, we collapsed edits to user talk pages into 'edit sessions' [32]. This treats series of consecutive edits by the same user to a given user talk page or message wall as a single communicative act. Our aggregate measure of interpersonal communication is the number of user talk page edit sessions per week in the period before the transition, and the number of message walls edit sessions per week in the period after.

Number of communicators

The count of unique editors who made edits to either a user talk page or a message wall in a given week.

Number of contributions

The total number of edit sessions made on article pages in each week.

Number of contributors

The number of unique editors who made edits to a wiki's article pages in each week.

Impact on newcomers

Since we were interested in the extent to which newcomers in particular were affected by the transition to message walls, we also generated separate newcomer measures for each of the previous outcomes.

We define an editor to be a *newcomer* at time t if they had created their account less than 3 months before t and had made fewer than 20 total edits at t. We define a *newcomer edit* as any edit made by a newcomer. We used this definition to aggregate the total number of newcomer edits made to articles, user talk pages, or message walls during a week to construct newcomer versions of each of the previously introduced dependent variables. For example, in addition to the total number of messages in a particular week, we had a corresponding measure for the *number of messages by newcomers* for that week. In this vein, we also constructed dependent variables for *number of newcomer communicators*, *number of newcomer contributions*, and *number of newcomer contributors*.

Analysis Plan and Models

We constructed longitudinal regression models to compare participation in wikis immediately before and after the introduction of message walls. We adapt the panel regression discontinuity approach introduced by Fe and Hollingsworth [29] and used previously in social computing research by Hill and Shaw [43]. The sudden switch to the message walls feature in software allows us to draw within-wiki comparisons immediately around the intervention and estimate its impact on the outcome variables. In doing so, we draw on the methods and concepts of quasi-experimental techniques of observational causal inference [62].

In quasi-experimental regression discontinuity designs (RDDs), analysts make a strong claim to causal identification by assuming (and providing credible evidence) that an intervention occurred at an "as-if random" point along the distribution of an otherwise smooth and continuous "forcing variable," resulting in a well-defined treatment and control group similar to those created in researcher-designed experiments [62]. We make an analogous, but much weaker claim: the introduction of message walls occurred at a precise moment for the wikis in our sample, allowing us to approximate an RDD by comparing each wiki to itself immediately before and after the intervention. By modeling the relationships between our dependent and independent variables around the intervention, we observe whether and how they changed in the weeks afterwards.

We refrain from making strong causal claims because the way the intervention took place leaves room for confounding factors that may have shaped the outcomes we observe. For every wiki in our study, at least one of the community administrators

requested the intervention and, in some cases, the change was discussed among community members prior to implementation. Wikia had also publicly announced their goals for its impact on communities. This means that even though our estimates precisely capture any shift in each outcome variable that occurred after the migration to message walls, the shifts we observe can not be attributed to the effect of message walls alone.

$$Y = \beta_0 + \beta_1 msgwall + \beta_2 \mathbf{wiki} + \beta_3 \mathbf{wiki} \times week + \epsilon \quad (1)$$

Our dependent variables are over-dispersed counts, so we used negative binomial regression. Our models all take the general form in equation 1. For each model, Y is an outcome measure, $week$ is the wiki week in the study period relative to the intervention, and $msgwall$ is a dichotomous variable indicating whether message walls are present or absent. We also include $\boldsymbol{wiki}$, a vector of wiki-level fixed effects, to fit individual trend lines for each wiki.

When plotting our dependent variables, we observed that the outcomes measuring number of edits to articles and number of article editors exhibited spikes around the week that wikis turned on message walls (see Figure 10). This brief spike led us to adopt two empirical strategies. First, we report the means for all of our outcome variables in the week immediately before the transition and the two weeks immediately after. Second, we decided to drop the week immediately after the transition ($week = 0$ or "week zero") from the datasets that we used to estimate the regression models reported in our main results. We drop week zero because our hypotheses and prior work

Outcome	Week -1	Week 0	Week 1
All Users			
No. of messages	6.39	10.43	8.65
No. of communicators	1.50	2.69	1.93
No. of contributions	73.49	151.05	99.08
No. of contributors	5.62	6.33	5.98
New Users			
No. of messages	0.79	1.31	1.61
No. of communicators	0.45	0.71	0.61
No. of contributions	9.87	10.55	10.76
No. of contributors	2.41	2.59	2.48

Table 6: Mean outcomes across all wikis for the weeks immediately around the transition.

suggested that message walls would lead to substantial and sustained shifts and because doing so results in a more conservative estimate of the change after the intervention.[16]

Preliminary visual exploration of the data also led us to expect that underlying trends in wiki activity might be moderated by the size of the editor community. In response, we estimated models that interacted the *msgwall* indicator with a measure of wiki size, taken as the natural logarithm of the number of editors who had ever edited the wiki at time of the transition. The interaction term did not help explain variation in any of our outcomes so we report models without it.

Results

Overall, we find evidence that the transition to message walls is associated with an increase in communication—especially among newcomers—but very little evidence of sustained change in contribution activity. We observe some evidence of a decline among contributions among newcomers.

Table 6 shows the means of each dependent variable in our sample over a three week period around the intervention. The table captures the short-term impact of the intervention that entails neither parametric assumptions nor adjustments for other potential sources of variation. Although the differences between the weeks varies enormously in magnitude across the outcome variables, the pattern is consistent. In every case, there is an increase from week -1 (before the intervention) to week 0 (immediately following it). In almost every case, the values revert back to lower levels in week 1.

Table 7 reports parameter estimates and standard errors for the *msgwall* term in each of our negative binomial regression models. To assist in interpretation, Figure 11 visualizes the estimates by plotting model predicted values for a prototypical wiki. We generate the trend lines for these plots by fitting a regression model on the predicted values from each of the models reported in Table 7. Our estimate is represented by the magnitude of the discontinuous vertical jump seen at week 0 in each of the plots in Figure 11, with the grey ribbons in the post-intervention period representing the 95% confidence intervals for the change associated with *msgwall*.

[16]We report the results of the same models estimated on the full dataset including week zero in our online supplement. As expected, these alternative specifications produce substantially different estimates of the impact of the intervention. We elaborate on this difference and the implications for our findings in the Discussion section below.

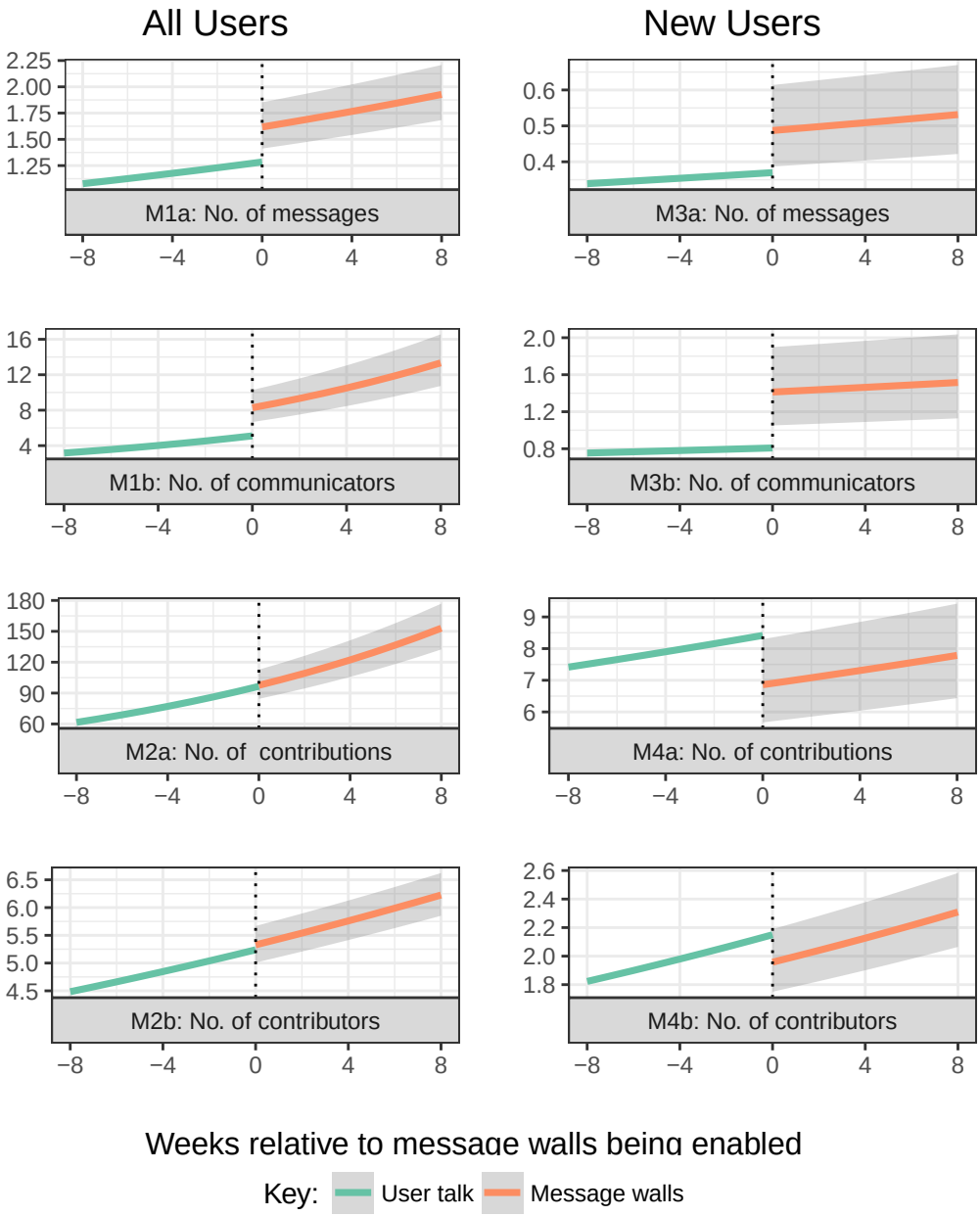


Figure 11: Visualizations of model predicted values for a prototypical wiki. The ribbons in the post-intervention period represent 95% confidence intervals in the estimate of the coefficient for *msgwall*.

Outcome	Estimate	SE
All Users		
M1a: No. of messages	0.48^{***}	(0.11)
M1b: No. of communicators	0.23^{**}	(0.07)
M2a: No. of contributions	0.01	(0.08)
M2b: No. of contributors	0.02	(0.03)
New Users		
M3a: No. of messages	0.56^{***}	(0.13)
M3b: No. of communicators	0.28^{**}	(0.11)
M4a: No. of contributions	-0.20^{*}	(0.10)
M4b: No. of contributors	-0.09	(0.06)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 7: Estimates (with standard errors in parentheses) of the impact of the transition to message walls for all dependent variables.

We find evidence to support H1a and H1b, that making communication easier increases the number of messages between editors in general, as well as the number of people communicating. M1a estimates that the number of messages by all editors rose 62% from 5.1 to 8.3 messages per week in a prototypical wiki at the point they introduced message walls ($\beta = 0.48, \sigma = 0.11$). Similarly the number of editors who communicated in the prototypical wiki increased 26% from 1.3 to 1.6 according to M1b ($\beta = 0.23, \sigma = 0.07$). Aggregating over the eight weeks in the study period following the change, our models predict a total of 162% more messages sent by 126% more communicators.

We do not find evidence to support H2a or H2b, which proposed that the number of contributions, and the number of people contributing would increase with a more intuitive communication interface. Table 7 and Figure 11 show that the estimates for both of these dependent variables are not distinguishable from zero.

For the newcomer related hypotheses, we find evidence to support H3a and H3b. The transition to message walls led to an increase in messages sent by newcomers, as well as the number of newcomers sending messages. M3a estimates that messages sent by newcomers increased by 75% in the prototypical wiki from 0.81 to 1.4 messages per week ($\beta = 0.56, \sigma = 0.15$) as the number of communicating newcomers each week increased by 32% from 0.37 to 0.49 ($\beta = 0.28, \sigma = 0.12$). Over the eight weeks in the study period following the change, our models predict a total of 175% more messages sent by 132% more newcomers. In both cases, the point estimates for the increases in edits by newcomers (and the proportions of edits within our prototypical wikis) are larger than our estimates for all users.

With respect to H4a, we find that the transition to message walls was associated with a decrease in the number of article contributions by newcomers. This outcome contradicted our expectations.M4a estimates that the number of article contributions made by newcomers declined 18% from 8.4 to 6.9 edits per week in the prototypical wiki shown in Figure 11 at the point they introduced message walls ($\beta = -0.2, \sigma = 0.1$). We note some uncertainty around this estimate as the standard error is fairly large relative to the value of β.

We found no evidence to support H4b, which stated that the transition to message walls would lead to an increase the number of new article editors.

Threats to Validity

We evaluated the sensitivity of our results to several potential threats to validity by conducting a number of robustness checks and estimating alternative model specifications. We explain several of these sensitivity analyses below. More detailed descriptions and full results can be found in Appendix 1.

Several aspects of the analysis and variable construction reflect decisions we made in the design of the study. The main results report findings from an analysis of 8 weeks of activity before and after the moves to message walls, we also ran our models on versions of the dataset with 6 and 12 weeks of data around on either side of the transition. These alternate specifications did not substantively alter our results. Furthermore, we tested the sensitivity of our results to the use of 7 day time intervals. Our results were not substantively different when we aggregated our dependent variables in 4-day or 10-day 'weeks'. We also varied the parameters that defined newcomers in this study (i.e. time and number of edits made since account creation). We shifted these thresholds to 2 and 4 months since account creation and 10 and 30 edits made since account creation respectively. Changing the measure in this way did not substantively alter our findings.

An additional concern common to RDDs is that results might be driven by underlying stochastic variation in the dependent variables. This is typically tested by conducting a number of tests on datasets with an incorrect or "dummy" transition date. To do so, we selected 14 different dummy dates either far before or after the actual transition. We ran our eight models on 14 different dummy datasets and found that $\tilde{9}0\%$ of the models produced no significant effects at the random cutoffs.

The inconsistency of some of our measures across the transition to message walls also poses an additional threat to the validity of our results. We have chosen to compare the number of edit sessions on user talk pages made before the transition directly with edit sessions on message walls made after. While we argued earlier that these two measures of communication are comparable, we also acknowledge that these measures are constructed from trace data produced by users interacting with different interfaces. For example, editors frequently organize and edit conversations on talk pages after discussions to clean up threads that are poorly formatted. However, since we hypothesize that message walls would elicit more interpersonal communication in general, any resulting inflation of our measure of communication on user talk pages means that our estimates would be conservative with respect to our hypotheses. We find evidence of an increase despite this potential inflation in the measure of communication on user talk pages.

Discussion

Our results provide evidence that the message walls intervention was associated with more communication among all editors and among newcomers in particular. This suggests that the incumbent user talk page system limited communication on Wikia wikis, as hinted at by previous research on Wikipedia [77], and that message walls succeeded in making communication easier. However, the intervention did not change rates of article editing overall and may have even reduced article editing among newcomers. We believe that the most likely explanation is a substitution effect where the increased

ease of communication led newcomers to converse more and edit articles less. It may also be the case that message walls empowered former lurkers with no intention of editing articles to begin chatting with other editors.

These results underscore that communication and collaborative production are distinct activities. While prior research, theory, and intuition led us to expect that increased communication would bring about corresponding increases in productivity, our evidence contradicts these expectations. The absence of such evidence may reflect aspects of our research design. For example, it is possible that changes unfolded over a longer time horizon than we included in our analytic window or in a manner not captured by our estimation techniques. It is also possible that our measures do not capture other changes brought about by the migration to message walls, such as shifts in the tone of interpersonal communication, in the perceptions of participants, or in their organizational culture. In both respects, alternative measurement and modeling strategies might yield results more consistent with the expectation that more communication leads to increased productivity. Our results may also reflect elements of the message walls transition itself. As we suggested, some administrators and editors of the wikis that adopted message walls were likely aware that the change was coming and may have altered their behavior pre-emptively. It is also possible that the outcomes we observe are not due to the effect of message walls alone.

We find the short-lived spikes we observe in some of the dependent variables immediately after the transition to message walls—visible in Table 6 and Figure 10)—difficult to explain conclusively. The spikes could be due to the novelty of message walls, an increase in administrator activity, or some other, unobserved common attribute that

led these wikis to adopt message walls in the first place. Alternately, they might also be due to the bursty nature of activity on many wikis. As Table 6 suggests, models run on the dataset without dropping the week immediately after the transition provide notably different estimates. In particular, the week 0 bump leads to significant positive estimates for number of article contributions and number of article editors (H2a and H2b), and positive (but insignificant) estimates for number of newcomer article contributions and number of newcomer article editors (H4a and H4b).[17] These estimates are dependent on the burst of activity in the first week when message walls were in use. It is possible that this burst was, in fact, the immediate effect of message walls. Nevertheless, a short-lived effect is inconsistent with both the theory that motivated our study and the goals articulated by Wikia in implementing the new feature. Despite some evidence that the design change may have produced a short-lived flurry of activity, we conclude that the easier interpersonal communication made possible by message walls and the more frequent communication associated with its deployment did not result in corresponding increases in productivity.

Our findings suggest several possible avenues for future research. While our results report the overall impact of message walls on our population of wikis, we encountered enormous variation in our dataset. Conducting analyses of subgroups of wikis (based on factors like age or topic of the wiki) could shed light on whether and how various kinds of communities respond to such feature changes differently. Additionally, further study on the message walls deployment could identify more granular effects of the intervention by analyzing whether certain kinds of activity are more likely before or

[17]Full results of these models are included separately in Appendix 1.

after the transition (i.e., by asking if users ask more questions, or if they are likely to reach out to new editors, and so on).

Conclusion

This study contributes a large-scale observational test of prior work that connected ease of interpersonal communication with increased productivity and newcomer participation. The results contradict previous findings and challenge an important set of long-standing assumptions in the design of social computing systems. Message walls seem to have increased interpersonal communication with no apparent impact on production overall. Among newcomers, the adoption of message walls may have even decreased production. Our results suggest the need for more fine-grained and empirically-tested theories of any link between interpersonal communication and collaborative production.

Both Wikia and Wikipedia have now invested resources in revamping interfaces for communication on their platforms. This study offers an evaluation of one such system, and sheds light on the impact of such a change on communication and productivity. While many projects might consider implementing systems like message walls to facilitate easier communication on their platforms, they might not be sufficient by themselves to encourage productivity among the user base or to increase participation by newcomers.

Technology adoption and tool use in a volunteer-run non-profit

Introduction

Managing a committed volunteer base is a central focus of many non-profit organizations, co-operatives, DIY (Do-It-Yourself) spaces and activist campaigns where very few (if any) individuals receive monetary remuneration for the work they contribute. Volunteers need to be matched with available tasks that suit their expertise, taught new skills so they can make effective contributions, mentored into positions of increasing responsibility, and developed into leaders in their own right.

Each of these processes are associated with complex information needs. Organizations that rely on volunteers frequently need to effectively manage information on who their volunteers are, what they know, what their responsibilities are, and whether the organization's needs are being met through their contributions.

However, many volunteer-run organizations face significant challenges in this regard. Although larger non-profits and political campaigns can procure or develop Client-Relationship Management (CRM) software that meets these needs, many smaller orga-

nizations do not have either the financial or human resources to institute and maintain such systems. Furthermore, these organizations experience high turnover of membership and leadership, which can affect both how effectively a new technological system is adopted, as well as the purpose it is adopted for. These changes in leadership and process might emerge informally, and lead to creative adaptions of existing technologies or partial implementations of new technologies to solve immediate process problems. Creating technological systems that support volunteer management in these contexts thus presents a compelling design challenge.

Understanding the implementation of technological systems and patterns of their use can also help reveal implicit structures and leaders within volunteer-run organizations. Whereas in conventional firms, new technologies are introduced from the top down by the leadership, in organizations where leadership structures are less formal or more diffuse, understanding who is able to introduce new systems, or grant more privileged access, reveals implicit hierarchies within collaborative communities. Additionally, the introduction and use of new technologies to perform this type of work may actively shape and reify organizational structures in these environments.

In this paper, I report on a contextual inquiry to understand how a particular volunteer-run non-profit, namely, a community bicycle co-operative, uses both information systems and artifacts from the physical environment to develop and promote volunteers. Based on extensive volunteering, participant observation, analysis of artifacts, and interviews of volunteers and staff, I find that a central tension in this organization exists between its centralized technological system for formally assigning volunteer roles, and decentralized manner of training volunteers and granting them

increased access to the organization. I describe how different stakeholders in this community developed independent processes for volunteer development and management that made sense in the context that they were being deployed, but were likely to breakdown when one of these processes had to influence the other.

This study was undertaken in order to assess the organization's needs with respect to volunteer management, and understand how the systems and artifacts that it currently uses contribute to its goals. I argue that identifying ways to bridge these separate yet highly related processes in this organization will allow for fewer breakdowns and help it more effectively promote and retain committed volunteers.

Background

My work in this paper bridges prior research on technological adoption in organizations and promotion processes in communities of practice. I argue broadly that the technological adoption in communites of practice has the potential to structure organizational practice, and make these intermediate functional roles that structure promotion visible to participants in the community. I expand on these related literatures in the sections below.

Technology adoption in organizations

Volunteer-run organizations that exhibit open membership structures and distributed leadership are often associated with decentralized organizational processes. Previous

work in this area has discussed the decentralized nature of contributions[13] and volunteer recognition[53] in these communities. However, even within these distributed organizations, particular processes can become centralized. As open and distributed organizations age, and develop more complex needs, they can start formalizing some processes that lead to more centralization[80]. Some processes, such as the socialization of new members, can become centralized, since factors such as uniformity of the experience, or having a single point of contact, can influence its efficacy[84, 10, 64]. Organizations that demonstrate these different organizing practices need information systems that are responsive to those needs.

Prior work in information management systems in non-profit organizations identifies multiple challenges that such entities face in developing technical infrastructure to manage their operations. Like any other organization, volunteer-run non-profits have complex information needs, and are charged with maintaining large databases of volunteers and clients [85]. However, unlike larger, for-profit firms, these organizations are rarely able to invest in enterprise applications designed for such purposes due to financial constraints. Furthermore, staff that are tasked with managing organizational data rarely have formal training or experience with information systems, and can be discouraged by the poor usability of many database applications [46]. Thus, many volunteer-run organizations create 'homebrew databases', cobbled together from a combination of personal office applications, paper records, and email clients. These homebrew databases often are adapated and re-adapted in different ways to suit the organization's present needs, and can produce inefficiencies in workflow, due to the need to migrate and access information across different platforms [85]. Creating effective

information systems under these constraints thus poses a compelling design challenge.

The patchwork nature of tool use and information architecture in these contexts also creates a rich opportunity to delve into the relationship between the social and technical nature of volunteer-run organizations. Creating homebrew databases reflects the 'improvisational model' of managing technological change in organizations[66], where stakeholders respond to and 'satisfice' emerging organizational needs through adaptations of their technology use. Structurational perspectives on technology in organizations emphasize the co-creative relationship between organizational structure and technology adoption. This framework emphasizes the ways that human agency, technological affordances and institutional properties mutually affect one another [65].

While much of the current literature on technology adoption in organizations focuses on the decisions and practices of large, hierarchical firms, this perspective also lends itself well to smaller, more decentralized organizations. In organizations with relatively 'flat' hierarchies that emphasize co-operation (often a feature of volunteer-run non-profits), organizational processes can be fuzzy, informal, and hard to identify. Through the structurational perspective, understanding technology management in volunteer-run non-profits affords us the opportunity to identify how various stakeholders use their 'technological frames'[67] to shape organizational practice in different ways, both implicitly and explicitly.

Promotion in communities of practice

In order to perpetuate itself and ensure its long term existence, every volunteer-run organization must continually mobilize, retain and promote members into leadership roles. Prior work on promotion in communities of practice has utilized the legitimate peripheral participation (LPP) framework to describe the way that newcomers transition from peripheral to core participation [55, 15]. The LPP framework suggests that people's identification as a member of a community is mediated by the kinds of participation they have access to within the community. In the beginning, newcomers might have access to peripheral tasks in the community of practice - Lave and Wenger cite the practice of apprentice tailors, who are initially encouraged to work on simpler, lower stakes tasks like putting 'finishing touches' on garments, and eventually move on to more core, skilled practices such as designing and sewing trousers [55].

Similarly, studies of volunteer communities have indicated that the move from peripheral to core participation is also mediated by intermediate functional roles that volunteers can take on along the way, typically indicated by the level of access or forms of participation that members take on[86, 8]. Making their history of participation along these dimensions visible to the community at large is part of how volunteers have managed to showcase their commitment and engagement with the community, and gained formal access to leadership roles[16].

Setting

The setting for this study is a non-profit educational bicycle co-operative (henceforth referred to as 'the Bike Co-op') located in a large Midwestern city. The organization provides a number of services related to biking and community development, almost all of which are entirely volunteer run. It has 'open shop' hours where bicycle owners can bring in their bikes and learn how to fix and maintain them, and a youth program that teaches teenagers bike repair skills. Volunteers help repair old bicycles that are donated to the co-op, so that they can be donated again (through partner organizations) to low income individuals, refugees, and people experiencing homelessness. Some of these donated bikes are also repaired and sold in order to fund these services. Finally, the Bike Co-op also runs a number of community events related to biking, sustainability and social justice.

Though the Bike Co-op employs five staff members, the services that the organization provides are almost entirely reliant on a team of volunteers. Every month, the Bike Co-op sees around 50 to 60 volunteers contribute hours to running programs, with each volunteer contributing anywhere from 2 to 25 hours a month. The Bike Co-op is located in a racially and socioeconomically diverse neighborhood, and this diversity is reflected in both the volunteer base as well as the patrons who come to the shop. The volunteers also span a wide age-range; while most are in their 20s or 30s, a significant number are adolescents, or older adults.

Volunteers work on a number of different kinds of projects - fixing up bikes that are marked for donation, helping patrons fix their own bikes, selling bikes, sorting

parts, and representing the organization at external events. They start with a range of skills with regards to bike repair - some come in with a lot of prior knowledge of bike mechanics, some have none.

The central mission and most important concern of the Bike Co-op is to regularly provide bike-related services in its community. All public facing services are almost entirely volunteer-run, and thus, the most persistent, ongoing challenge this organization faces is ensuring that it finds volunteers to staff shifts on a regular basis. While there are other activities that the Bike Co-op conducts more sporadically outside of its weekly shifts (such as pop-up sales, or organized group bike rides), my analysis focuses mostly on the core programs that the organization offers every week.

Formal roles at the Bike Co-op

Broadly, the Bike Co-op consists of many volunteers, and a few staff members. One of the staff members employed by the Bike Co-op is a *volunteer co-ordinator*. The volunteer co-ordinator keeps track of which volunteers have committed to work which shifts, and strives to ensure that each shift has enough volunteers. Newer volunteers work with the co-ordinator to identify which shifts need support in any given week, and begin volunteering at their shift alongside more experienced hosts. Experienced volunteers can eventually become hosts in their own right. Other staff roles include the *accountant*, who keeps track of the organization's finances, a *youth co-ordinator* that manages programming for children and youth, and two or more paid *mechanics* who fix up the stock of used bikes that the organization sells. These staff members meet

regularly to discuss and manage day to day operations at the Bike Co-op.

The typical way that volunteers contribute to the Bike Co-op is by working a *shift*. Each week, there are up to eight shifts that the Bike Co-op commits to staffing [18], each of which lasts between two and a half to four and a half hours. Each shift is led by at least one or two *hosts*, who are experienced volunteers who understand what tasks need to be done during the shift, have the necessary levels of access to perform those tasks, and can mentor newer volunteers if necessary. Shift hosts are identified by name in the weekly program schedule, and have typically committed to being present at their shift every week for a three month term.

A role that straddles the line between volunteer and staff is *class instructor*. The Bike Co-op occasionally offers classes on bike repair that last a few sessions, and the class instructor is paid a small honorarium for their contribution. Although the instructor is paid, they are not formally part of the staff.

The Bike Co-op is overseen by a board of directors who operate on a volunteer basis, commonly known as *the collective*. The Bike Co-op's collective members are long-time volunteers or staff who decided to extend their commitment to the organization by taking on leadership roles[19]. Collective members attend meetings every other week to address overarching issues faced by the organization (these meetings are occasionally attended by interested volunteers or shift hosts as well). An experienced volunteer may become a collective member by getting 'voted in' by existing collective members during

[18]This number varies depending on the season or time of the month. For instance, there is an extra sale shift during warmer months to meet the greater demand for bicycles, and once a month there is a dedicated open shop for women and transgender patrons.

[19]Staff who are also collective members are paid for their duties as staff, but not for their participation in the collective.

one of these meetings. This work includes making staff hiring decisions, determining organizational policy, and planning for the future. In addition, most collective members volunteer at least three hours a week at the Bike Co-op, usually as shift hosts.

Methods

In order to understand how the Bike Co-op was currently managing volunteers and identify potential breakdowns in process, I adapted a contextual inquiry process as outlined by Holzblatt et al[44]. My goal was to understand what the process of volunteer management and development looked like for a variety of different stakeholders in the organization (including the volunteer co-ordinator, collective members, and new volunteers) and how these stakeholders used technology and artifacts to meet their goals.

I built up context and connections for this project through multiple years of volunteering for this organization, starting in 2015. I began formally collecting data for this project in late July 2017, which entailed participant observation of shifts, interviews with several organizational stakeholders, and collection and analysis of electronic artifacts. Using all these data sources in combination, I was able to create sequence models that illustrated different stakeholders' workflows, which informed my findings and design recommendations. I elaborate on these data sources and the nature of my analysis below.

Volunteering and Participant Observation

My findings draw upon context I have built up over a period of three years of volunteering at the Bike Co-op. I began regularly volunteering at the Bike Co-op in June 2015, contributing around 12 hours a month. In May 2016, I joined the collective, which led me to increase my participation to approximately 18 hours a month. I estimate that I have spent over 550 hours volunteering with this community over the last three years, at least 120 of which have been spent in collective meetings where I developed an overview of the structure of the organization and the challenges it faced. Over the course of these years of involvement, I have observed, participated, and mentored others in almost every kind of volunteer role. My most regular volunteer position at the Bike Co-op was hosting the sale shift on Saturday mornings, where, along with other volunteers, I would help sell used bikes.

When I began formally collecting data in July 2017, I signed up for several shifts during the summers of 2017 and 2018 to conduct a more thorough participant observation. These included 'open shops' where patrons would bring in bikes to work on, 'volunteer hours' where volunteers would fix up bikes in storage and work on miscellaneous projects around the shop, volunteer orientation sessions, and collective member meetings. I took scratch notes when possible[20], and wrote memos summarizing my observations after I left. In total, I wrote memos after seven different shifts or meetings, based on around 20 hours of observation during the summers of 2017 and 2018.

During my observations, I focused on interactions between volunteers (rather than

[20]Due to the frequently busy and hands-on nature of volunteering at the Bike Co-op, it was not always possible to immediately jot down observations.

those between volunteers and patrons/customers), to understand how skills and access were shared within the volunteer community at the Bike Co-op. In particular, I noted times when newer volunteers needed help or input from regular volunteers. Examples of these include times when volunteers had questions about bike mechanics or the location of specific items, or if they needed help with a task they were not authorized to perform (such as using the cash register). I also noted when and how veteran volunteers delegated tasks to newer volunteers. When I wrote my memos, I would occasionally reference any similar occurrences I had observed in previous shifts, in order to underscore which interactions were more common and associated with usual processes in the Bike Co-op.

Interviews

In addition to my observations, I conducted seven semi-structured interviews of various stakeholders in the bike co-op, including two staff members (who also volunteered their time to the organization), two long time volunteers who had recently been inducted into the collective, and three newer volunteers who had been attending shifts at the Bike Co-op for fewer than two months. I worked with the volunteer co-ordinator of the bike co-op (who was also one of my interviewees) to identify a purposeful sample of potential interview participants. I chose interviewees to ensure I received perspectives from people who represented different roles in the co-op (See Table 8). However, it should be noted that while these roles are generally applicable, they are not the only way that these individuals relate to the organization. For instance, VC and A are staff,

but they also regularly choose to volunteer for the organization outside of their paid activities. My interviews with VC and A touched on their experiences as staff members *as well as* their experiences as volunteers.

Interviews ranged from half an hour to one and a half hours in length, and were mostly conducted in person (one participant was interviewed over video chat). Interviewees were generally not offered compensation for their time[21]. All interviews were recorded and transcribed. Interviewees were assured that their identities would be kept anonymous. In addition to suppressing the names of the organization and the volunteers, I have chosen to refer to all participants using gender neutral pronouns to further occlude their identities.

My interviews focused largely on uncovering staff and volunteer pathways ("How did you come to do what you're doing?") and workflow ("How do you go about doing what you do?"). The relative emphasis of these parts changed somewhat depending on who I interviewed. For instance, my interview with the volunteer co-ordinator focused quite significantly on their workflow, but less so on their pathway into the organization. On the other hand, my interviews with volunteers revealed broadly similar descriptions of workflows during shifts, but exhibited significant variation in the pathways they took within the organization, so I typically spent more time in interviews uncovering those differences.

My analysis of volunteer and staff pathways through the organization was developed through iterative coding of the memo and interview data. Additionally, I used this

[21]On two occasions, I bought my interviewee's coffee since we were conducting the interview at a coffee shop, but they had already agreed to participate before I had offered to purchase their beverage.

ID	Role
VC	Volunteer Coordinator
A	Accountant
CM1	Collective Member
CM2	Collective Member
NV1	New Volunteer
NV2	New Volunteer
NV3	New Volunteer

Table 8: Participant IDs and roles of interview subjects

data to generate consolidated sequence models that illustrated organizational processes, based on the method described by Holzblatt et al[44]. For each of the workflows that my interviewees described, I created abstracted models that represented the process. I then compared these models against one another and found points of agreement and disjuncture among different stakeholders, which informed the analysis I present.

Electronic artifacts

Since I was interested in understanding how the technical infrastructure of the Bike Co-op influenced volunteer management, I also collected a number of electronic artifacts that shed light on how various stakeholders in the Bike Co-op co-ordinated activity with one another. I was given access to the organization's shared Google Drive and Mailchimp account, which contains forms and spreadsheets used to organize workflow and tasks at the Bike Co-op, including shift assignments, volunteer lists, agendas and minutes from meetings, mass email templates and prior newsletters. This shared repository contained information that helped me parse how the volunteer co-ordinator

interacted with volunteers in the organization, and provided a picture of the technical infrastructure that the bike co-op used for volunteer co-ordination.

I drew heavily upon information from my observations and interviews to identify and focus on artifacts regularly used to co-ordinate and schedule volunteer activity at the Bike Co-op. I processed these by noting the *intent* for the creation or use of the artifact, as well as how it was actually used in organizational workflows. I was able to locate these artifacts within the sequence models I described in the previous section to illustrate where each system or artifact fit into larger organizational processes, which informed my subsequent analysis and recommendations.

I also sought to understand how a volunteer-tracking tool (called Trackr[22]) affects the volunteer experience. To inform my understanding of how Trackr works, I had multiple extended conversations with the volunteer who developed the system. This volunteer declined to be recorded or have personal details about them be reported in the study, but agreed to provide information about the architecture of the system. I took notes during our conversations that focused solely on how the system was designed. As with the other artifacts I studied, I drew upon interviews and observations to locate the use of this system within larger organizational processes. I also used data from Trackr to inform my own understanding of which volunteers had been frequently contributing to the Bike Co-op, and used it as a way to identify interview participants.

[22]Name changed to deidentify organization.

Data Analysis and Positionality

I volunteered for the Bike Co-op for over three years, two of which were spent as a collective member. My long-time insider status at the Bike Co-op offered me a wealth of access to this space, and allowed me to draw upon contextual knowledge of this organization that I had developed over the course of several years.

Additionally, understanding the relationship between the Bike Co-op's socio-technical infrastructure and volunteer progression was interesting to me not just as a researcher, but also as longtime member that wished for the organization's success. Indeed, the decision to conduct this project was informed both by my academic interests as well as interest from other collective members and staff in understanding ways they could develop better technical infrastructure for the organization. Thus, I received considerable support from the Bike Co-op's leadership when I proposed conducting this research, and had no problem gaining (or rather, maintaining) my access to all levels of the organization while collecting data.

However, I am also conscious that my insider status has the potential to influence the collection and interpretation of the data in several ways. As a person who is supportive of the Bike Co-op's mission, I am inclined to portray the organization favorably. Additionally, being an insider for a long time caused me to develop my own sense of how the Bike Co-op functioned, in a manner particular to my own trajectory and experience as a volunteer. To ensure that I was not just basing my analysis on my singular perspective of 'how things work', I made sure to interview volunteers and staff who had been involved with the organization for varying lengths of time, and based my

analysis on observing the way that different stakeholders interacted with the technical infrastructure of the Bike Co-op. This process challenged many of my own preconceived notions of how the Bike Co-op functioned. Furthermore, I stepped down from my role as collective member and stopped volunteering in August 2018, which give me more space to analyze the Bike Co-op from the outside. Although stepping down did not make me instantly lose my insider status, it did help me better approach understanding the organization from a more detached analytical perspective, as compared to when I was intimately involved as a stakeholder.

Collecting data while being a visible member of the Bike Co-op presented a few challenges. For instance, I had intended on passively observing a few shifts so that I could take detailed notes on what was happening in the room. This was possible in some cases, like the volunteer orientation session and in collective meetings, where I was not always expected to actively participate. This was much more difficult during shifts like open shop, which could get very busy and were sometimes understaffed. Insisting on being a fly on the wall would have been both difficult and rude, since volunteers were frequently looking for all the help they could get, and knew that I would be able to contribute as an experienced volunteer. In these cases, I took scratch notes whenever I could between tasks, and wrote a more detailed memo after the shift ended.

My leadership position as a collective member of the Bike Co-op also influenced how I conducted my interviews. As staff at the Bike Co-op, the volunteer co-ordinator (VC) and accountant (A) are technically in a reporting relationship to the collective. As I was a collective member at the time I conducted these interviews, I tried to be extra

sensitive to this dynamic while interviewing them. Before beginning the interviews with P1 and P2, I acknowledged our respective positions, and assured them that any information that they shared with me (including and especially about breakdowns or uncertainty in their workflow) would not be used against them personally, or affect their employment status.

It is possible that my status as a long time volunteer at the Bike Co-op also affected my interviews with new volunteers NV1, NV2 and NV3. I was concious of the possibility that being interviewed by a person in a leadership position at the Bike Co-op might lead the new volunteers to present themselves as more enthusiastic or accommodating than they otherwise might be. To mitigate this risk, I chose to interview new volunteers that I had very little previous contact with. Additionally, when I emailed them asking for an interview, I acknowledged that I was a long time volunteer at the Bike Co-op but I mostly emphasized my role as a researcher. While debriefing at the end of these interviews, I talked about my volunteer and collective member role at the Bike Co-op in more detail when asked.

Findings

My findings indicate that the primary design challenge in managing volunteers at the Bike Co-op emerges from the tension between its centralized role formalization process, and decentralized volunteer development process.

New volunteers are brought in from the periphery to the core of the organization

through a gradual process of learning skills associated with the work at the co-op, and being granted increasingly privileged access to tools and other physical artifacts that indicate a form of organizational trust. Examples of these include the cash register and point of sale system, keys, access to the organization's Google Drive, and inclusion within more exclusive mailing lists.

Although access to these artifacts are indicators of volunteer experience and organizational trust, records of which volunteers have what kind of access are not well-maintained, nor are they made clearly visible to all members of the Bike Co-op. These forms of volunteer development are frequently invisible to the centralized volunteer co-ordinator, who is tasked with identifying volunteers who are ready to take on formal leadership roles in the organization.

Task assignment and volunteer promotion at the Bike Co-op can be on some occasions informally determined, other occasions formally specified, and co-ordination between these various facets of the volunteer experience is often informal and unclear. This is where breakdowns in process most often occur, and I argue that this presents the most compelling space to iterate on the design of the Bike Co-op's technical infrastructure.

In the following sections, I describe how the hiring of a volunteer co-ordinator and the subsequent introduction of a new volunteer management system helped formalize the 'shift host' role and centralize the process through which volunteers gained recognition for this role. I contrast this against the way that volunteers are developed informally through mentorship and being granted increasingly privileged access

to shop tools and systems, and proceed to draw on the tensions between these two organizational processes at the Bike Co-op.

Technology adoption and the formalization of roles

In 2016, the Bike Co-op hired a volunteer co-ordinator (VC) to take on the task of managing and scheduling volunteers, and ensuring every shift was adequately staffed. Before the arrival of the VC, the task of recruiting and scheduling volunteers was largely under the purview of the group of collective members (experienced, regular volunteers who were also on the board of the organization). Collective members who volunteered at particular shifts were responsibile for ensuring their shifts were staffed every week, and they frequently did so by emailing or texting volunteers they knew and asking for help when needed.

During this period, they largely used a system called Trackr to manage information about volunteers. Trackr was developed by DEV[23], a long-time volunteer Bike Co-op, and its installation preceded the VC's arrival. It is a bespoke, browser-based system that allows the organization to collect data about volunteer participation. A 'sign-in' computer exists near the shop door with Trackr open in a browser window (see Figure 12). During their first shift at the shop, new volunteers can create a profile, which includes their name, contact information, and a picture of themselves. When signing in for a shift, volunteers search for their name in a search bar in the top left corner of the window, which brings up the profile picture that they used for their

[23]Code assigned to anonymize subject.

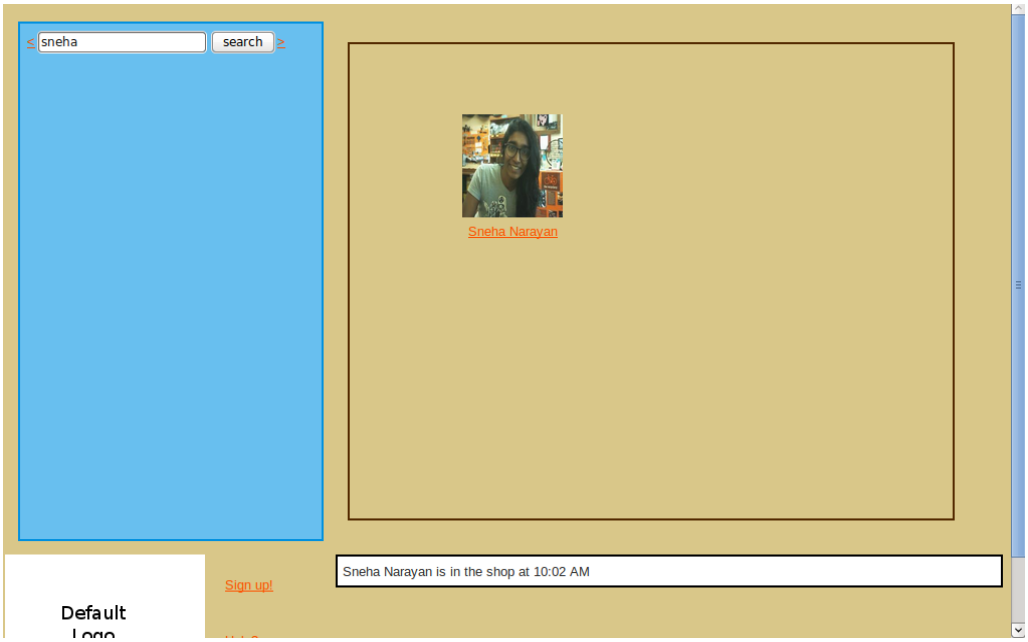


Figure 12: Trackr's interface within a browser

account in the blue area beneath the search bar. The user then drags their picture to a box on the right to mark the start of their shift. At the end of their shift, they move their picture out of the box, and Trackr updates their hours in an internal database.

Every volunteer at the Bike Co-op uses Trackr, albeit briefly, to sign in and out of their shifts, thus creating a record of the hours they have contributed to the organization. Trackr is a system that is largely used and updated by volunteers themselves; volunteers create their own profiles, are responsible for updating information within it, and for remembering to log their shift hours when they enter the Bike Co-op.

Collective members, staff, and other experienced volunteers use Trackr to find contact information for volunteers they wished to reach out to. Additionally, the system is used to generate statistics about the number of people who volunteered at the shop,

or the number of volunteer hours logged in a given time period, typically in the service of grant-writing proposals. Finally, there was a practice of using the data from the system to identify and acknowledge the volunteer who had contributed the most hours in a particular month.

Eventually, the task of volunteer management became sufficiently complex and time intensive that the Bike Coop chose to hire a volunteer co-ordinator (VC) in the summer of 2016, who was tasked with the job of recruiting and orienting new volunteers, and ensuring that volunteers were matched with shifts that needed to be staffed. Thus, the VC was put in a position where they could reshape the volunteer management process at the Bike Co-op.

The VC proceeded to use the tools they knew how to use (i.e. Google Forms, Google Sheets and MailChimp), to structure their workflow according to how what they understood their job to be, which was ensuring that every shift at the Bike Co-op was staffed with volunteers who were well-matched to that role.

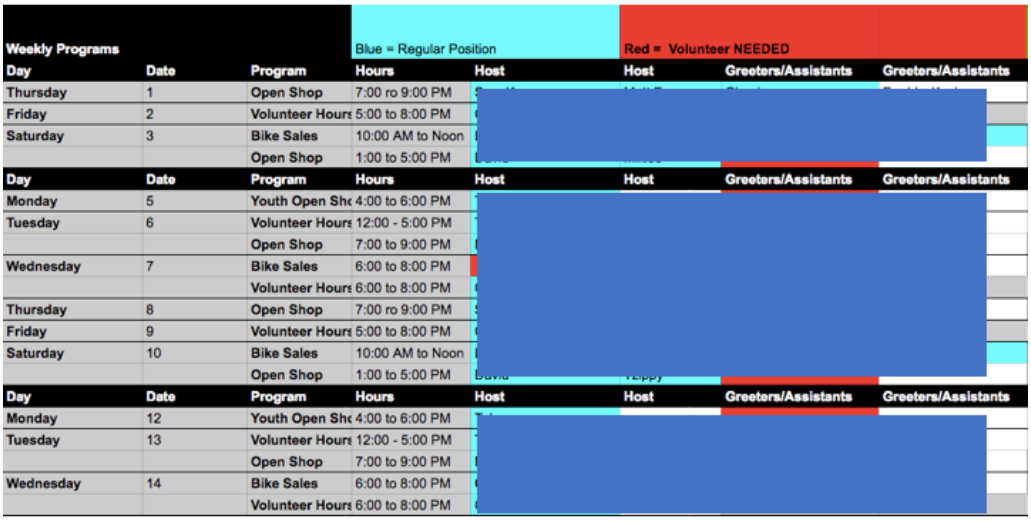
> I've got this process, asking with a simple survey, asking how many times a week you want to volunteer, and which programs are you interested in volunteering in. And I get feedback from those responses and then I construct a calendar for every week in the quarter. And there are holes when things aren't filled up and then I call around and try to remember who has volunteered in the past month or two months and see if I can fill up those spaces. And then it gets changed again because I send that calendar out to all the volunteers, or anyone who's gone through volunteer orientation,

and people sign up for things.

The main tools that the VC uses to co-ordinate volunteers at the Bike Co-op are the MailChimp *volunteer email list* and a Google spreadsheet containing the *weekly program schedule*. After a new volunteer creates their profile on Trackr, they are directed to a webpage that lets them sign up for the Bike Co-op's volunteer mailing list, which records their contact email address in the Bike Co-op's Mailchimp account. Every week, the VC sends out a mass email to the volunteer mailing list, embeds a link to the weekly program schedule, and requests volunteers to indicate their shift preferences either by leaving a comment on the schedule or emailing the VC directly.

This system for managing volunteers was developed almost entirely by the VC alone, and grew out of their understanding of what their duties were in their staff position. In a manner consistent with many other non-profits, they developed 'homebrew databases', using products they were already familiar with or could easily learn to use, such as Google Sheets and Mailchimp.

Although the weekly program schedule is publicably viewable and can be commented on, the VC is the primary user and maintainer of this system. Before sending out the mass email every week, the VC goes through the program schedule, updates it to reflect their knowledge of who is volunteering when, and highlights areas of need. Every month, they create a new tab in the spreadsheet that lists all the shifts and positions available for that month, including any one-off events that might need staffing in addition to the primary programs offered each week. The spreadsheet in its current form now contains dozens of hidden tabs, each one of which contains the schedule for



Weekly Programs			Blue = Regular Position		Red = Volunteer NEEDED		
Day	Date	Program	Hours	Host	Host	Greeters/Assistants	Greeters/Assistants
Thursday	1	Open Shop	7:00 ro 9:00 PM				
Friday	2	Volunteer Hours	5:00 to 8:00 PM				
Saturday	3	Bike Sales	10:00 AM to Noon				
		Open Shop	1:00 to 5:00 PM				
Day	Date	Program	Hours	Host	Host	Greeters/Assistants	Greeters/Assistants
Monday	5	Youth Open Sho	4:00 to 6:00 PM				
Tuesday	6	Volunteer Hours	12:00 - 5:00 PM				
		Open Shop	7:00 to 9:00 PM				
Wednesday	7	Bike Sales	6:00 to 8:00 PM				
		Volunteer Hours	6:00 to 8:00 PM				
Thursday	8	Open Shop	7:00 ro 9:00 PM				
Friday	9	Volunteer Hours	5:00 to 8:00 PM				
Saturday	10	Bike Sales	10:00 AM to Noon				
		Open Shop	1:00 to 5:00 PM				
Day	Date	Program	Hours	Host	Host	Greeters/Assistants	Greeters/Assistants
Monday	12	Youth Open Sho	4:00 to 6:00 PM				
Tuesday	13	Volunteer Hours	12:00 - 5:00 PM				
		Open Shop	7:00 to 9:00 PM				
Wednesday	14	Bike Sales	6:00 to 8:00 PM				
		Volunteer Hours	6:00 to 8:00 PM				

Figure 13: Weekly program schedule. Blue squares obscure names to protect volunteer privacy.

a month in the past.

Of course, the volunteer base at large also uses the weekly program schedule, but to a different extent. Volunteers typically access this spreadsheet while they are outside the Bike Co-op, either when they anticipate signing up for or dropping a shift. Figure 13 shows that some positions are being staffed regularly (indicated by cells with blue backgrounds) - these are usually regular volunteers with weekly shifts. Volunteers with weekly shifts usually do not need to refer to the weekly program schedule regularly, since they already know how they fit into the schedule at the Bike Co-op. While the information on the spreadsheet is visible to most volunteers, the VC is the primary person managing the informaton on it.

On the other hand, every volunteer at the Bike Co-op regularly uses and updates

Trackr, but the VC interacts with this system quite rarely. The main reason for this is that the mailing list and program schedule system they developed is largely sufficient for the task of ensuring that every shift is adequately staffed. Even if a request for volunteers over the mailing list goes unanswered, the weekly program schedule has records of people who have been volunteering recently. In this event, the VC reaches out to recent active volunteers over either email or SMS to ask whether they are available to fill a shift. Eventually, Trackr became largely unnecessary for most of the VC's day to day tasks, although volunteers continue to use it when signing in and out of shifts.

Increased visibility of formal roles

One of the effects of the creation of the weekly program schedule was the formalization of the role of 'shift host'. When the VC scheduled volunteers into shifts, they sought to ensure that there were one or two experienced volunteers on each shift in the schedule. To do so, they included a slots on the schedule that were designated 'host', and encouraged experienced volunteers to take on that slot within a shift for an extended period of time.

CM1 started volunteering occasionally at the Bike Co-op in January of 2015, before the VC was hired. Although collective members co-ordinated to ensure that there were one or two experienced volunteers on every shift, there was no publicized weekly program schedule that named particular people as 'host' when CM1 began attending shifts.

> CM1: I was coming in on different days, like a month apart. One person seemed to be in charge last time, but now he's not here, and I dont even remember who was at Tuesday open shop last time. So if I needed to access the register, I don't know the other volunteers' names to say "hey, this guy needs to check out". But like... the person I ask could also be on their first day, so if I ask the wrong person I don't know if they can jump in and do this as well. So at some level, there's that general structure - one of the benefits of a militaristic organization is that you have symbols on your shoulder that say "oh, he has more symbols, so he probably knows better than me." And then we're here as equals, and then I dont know who to turn to. I think the big issue was like, figuring out who's at the task, and almost having a hierarchy. Like who do I look up to, not so much as an honorary way, but like if I need to get a question answered, who's my go to?

While CM1 did eventually start working a regular shift and was able to identify volunteers with more experience, their observation differed from every other interviewee who started at the Bike Co-op after the VC was hired. Unlike CM1, other volunteers I interviewed immediately described shifts with reference to who the host or hosts were, because the host role was formally defined and reified by the VC, and made visible through their weekly program schedule. The formalization of the host role, made visible through the weekly program schedule, set expectations for a specific point person to help enforce shop policies, and mentor new volunteers.

> CM2: No one is expected to know all of the things about everything, so I make sure the volunteer understands who's the host, that if you have questions, this is the person you can go to. I think that helps us to maintain quality service and make sure policies are being met. But it also helps the volunteer to feel valued and supported, and because things could get a little chaotic sometimes, they don't feel lost in the mix of all the stuff that could be happening in the shop.

In many ways, the creation of the host position was an identification and reification of an intermediate role that existed in the Bike Co-op. Prior to the arrival of the VC, the host role was informally met by one or more experienced volunteers who had the necessary experience (mechanical skills, knowledge of shop policies) and access (posession of keys to the shop, password to open the register, etc) in order to manage a shift. When creating their system of spreadsheets and mailing lists, the VC formalized their understanding of the necessary skills and permissions required of an experienced volunteer to manage the shop by naming the people who embodied this role in each shift as "hosts".

In this sense, we can see how the organizational structure of the Bike Co-op both informed and responded to the creation of this technological system. By naming the host and making their role visible in the weekly program schedule, they also implicitly structured the shift; with the explicit naming of the host role, new volunteers had an identifiable person that they regularly turned to for mentorship, and identified personal goals such as 'becoming a host' for a shift. This process reveals the mutually co-creative

nature of technology adoption and organizational structure in the context of the Bike Co-op.

Volunteer Development through Legitimate Peripheral Participation

Once a new volunteer begins regularly attending shifts at the Bike Co-op, they interact most often with the other volunteers at the shift, rather than the VC. Through repeated and extended interactions with experienced volunteers during their shifts, new volunteers slowly become more integrated into the Bike Co-op through a process of legitimate peripheral participation[57]. During shifts at the Bike Co-op, new volunteers are expected to work with veteran volunteers in performing a combination of technical tasks (such as pumping tires, oiling chains, adjusting brakes) and social tasks (greeting patrons, explaining the shop rules, showing bikes) to contribute to the mission of the Bike Co-op. New volunteers learn these technical and social practices by participating alongside more experienced volunteers as they work together to serve patrons and customers.

Legitimating participation

Part of how the Bike Co-op develops new volunteers is by immediately encouraging them to step up and help wherever they can, without necessarily directing them to perform specific tasks explicitly. In this way, they are asked to participate *alongside*

more experienced volunteers in managing the shift, and are granted autonomy in doing so from the beginning. In doing so, experienced volunteers engage in what Lave and Wenger describe as 'benign community neglect'[55] in order for participants to develop a sense of autonomy surrounding the practices at the co-op.

> NV2: It wasn't so much "here's what to do, do this, go do this", I think it was more like, I just kind of stepped in where I could if [the host] was busy I would help whoever came in the door.

This practice was observed by newcomers like NV2 above, and confirmed as an ideal quality by experienced members as well:

> CM2: One of the things that is necessary [for a new volunteer] is the willingness to jump in and help as needed, and being a volunteer that is comfortable once they're oriented and trained to identify where needs are and where help is needed without a lot of self direction. I think there are definitely some volunteer opportunities out there where you have a set task list and you have a person who's telling you what to do; we don't have the capacity to do that. Sometimes we do, if it's volunteer hours or something like that, but during the busy open shop or when the sale is busy, it's really valuable for the volunteer to understand that we want you to ask questions if you don't know what youre doing, but as long as you feel confident, just jump in, dont wait for me to say "Hey, can you help this person find a bike?" If you see someone walk in the door, just jump in and be like "Lets look for a bike!"

Self-selection and self efficacy scaffold skill building

While the practice of granting the new volunteer autonomy can legitimize their participation and indicate that the organization trusts and relies on the new volunteer, it can also be disorienting to the new volunteer if they feel underprepared. CM2 reflected on their first open shop shift at the co-op after a particularly rushed orientation:

> CM2: It felt intimidating, and a little bit chaotic, not because it was necessarily disorganized, but because we just kind of had this crash course so it was kind of like a mixture of not really having the full grasp of what's going on there, and not being totally competent with some mechanic skills, and also not really feeling the confidence- like who am I? I just came in here, I dont even know anything! So it was interesting for sure.

A significant factor in determinined both shift assignment and promotion at the Bike Co-op was how volunteers perceived their skill set and areas for growth. To be able to participate semi-autonomously in the way that was part of the organizational culture of the Bike Co-op, volunteers describe needing to start with a task they could feel confident about doing, and use that as a starting point from which to build further skills.

NV3, who had volunteered a little with a different bike shop but did not have significant mechanical experience, chose to begin their volunteer tenure by staffing a sale shift, which they felt they had the necessary experience for.

> NV3: I knew right away that I wanted to do sales because I'm not that

> skilled as a mechanic and the time that I was at open shop it just seemed really busy that night, and I know that I'm someone who's kind of shy, and introverted, and just kind of freaked out when there's a bunch of stuff going on like that. So I thought that I would start with the sale because even though I don't know a lot about mechanics, I think I have a good idea of like "okay, who's this person, what do they want to do with their bike?"... so just matching people with their bikes is good for them, and something that I think I felt confident about doing. And then I thought I would do that for a little bit, and start doing more of the mechanic stuff. And then, because those shifts just seem more faster paced, I was a bit more intimidated by it, but now that I'm more familiar with [the Bike Co-op], I signed up on Friday to do the volunteer hours for the fixing the bikes part.

Once NV3 identified an area that they were skilled at, they could use that shift to build familiarity and confidence about their contribution to the Bike Co-op. Once they had that foundation, they were able to start branching out and building new skills at different shifts.

On the other hand, NV1 started volunteering at youth open shop, which is a shift that requires more mechanical skills than the bike sale. However, they came in with experience managing children and youth, which caused them to identify their skill set with this shift (NV1: " I think just being a teacher, I figured it might be an easier transition") even though they had limited mechanical experience. Once they established themselves at youth open shop, they began attending volunteer hours to

gain more mechanical skills, primarily to support their existing work at youth open shop.

> NV1: I'm hoping that if I keep coming to volunteer hours here I'll start picking up more mechanical skills, so I won't feel so dependent on [the shift host]... so just feeling more helpful and a little more independent when Im here during youth open shop.

Access to physical artifacts and tools

As new volunteers continue building the skills and confidence to carry out the tasks for their shifts, experienced volunteers frequently grant them further autonomy by granting privileged access to key tools and artifacts in the shop. The transition from peripheral to core participation in the context of the Bike Co-op is mediated by access to physical and technical artifacts that signify increased institutional trust in the volunteer. These artifacts include things like keys to closets or doors in the shop, the password to access the point-of-sale system, or inclusion in more exclusive mailing lists. New volunteers are granted increased access as they come into contact with more and more of these representative artifacts.

The process of increasing access can start in small and simple ways. For example, the co-op's main space has several tool boards lined along the walls, with many kinds of commonly used bicycle repair tools dangling from them. A locked red cabinet in the corner of the room contains more expensive tools that are only taken out when needed. During an open shop shift I observed, a volunteer asked the shift host where they could

find a chain breaker. The shift host, who was busy helping a patron fix their bike, said "look in the basket on the door of the red cabinet", and handed the volunteer a set of keys. The volunteer retreived the chain breaker from the red cabinet, and returned the keys to the shift host.

Even though the volunteer was not given permanent access to the keys in this instance, this action still served to distinguish them from the patrons in the shop, since keys to the red cabinet are not offered to patrons or customers (unless those individuals are also known volunteers). This form of granting access legitimizes the volunteer, and indicates that they are part of a more exclusive set of people that are allowed to open the red cabinet. This access also opens up new forms of participation for the volunteer - in this case, the volunteer was able to use more restricted tools while helping a patron. Furthermore, the experienced volunteer did not have to pause their task to attend to the red cabinet when they chose to instead grant this access to the newer volunteer, thus widening the circle of trust in the organization.

Many physical and technological artifacts in the shop that have limited access either hold significant monetary value (such as the tools in the red cabinet) or provide access to sensitive information or other valuables (i.e. password for the point of sale system, keys to the safe). Access to these artifacts form a system of permissions that indicate the level to which a volunteer is trusted by the organization to handle valuables responsibly and accurately (*CM1: Obviously if I were on my third day, you wouldn't want me to be working on point of sale*). In addition to monetary value, many of these artifacts are crucial for the successful functioning of the shift. Every shift at the Bike Co-op needs at least one volunteer who has the necessary access and experience to

conduct important tasks like unlocking the shop, accessing the cash in the register, and restocking supplies from the storeroom. All of these tasks involve privileged access to physical and technological artifacts.

Well-developed skills and self-efficacy can go a long way in determining how experienced a volunteer is, yet many interviewees identified these forms of privileged access as a salient mark of official membership in this community.

> CM1: The first few times when I was still rather irregular and didn't have a specific shift or anything I was coming to, it did feel odd that I knew enough about bike mechanics and I could get somebody to change a derailleur and things like that, but the weirdest thing was the patrons would always look up to me in the shop to figure things out like how to pay for stuff, and then I would look up to the actual volunteers.

> NV2: Bicycling and bicycle mechanics was kind of a hobby for me for the last 40 years, so I would say I'm quite experienced about mechanics... but I'm still not really familiar with the [Bike Co-op] concept. I'm... I mostly am, but there are still lots of details I just do not know. For example, how the register works, when people ask me for prices I always have to refer them to someone else.

Experienced volunteers, shift hosts, and staff typically have the authority to grant access to these artifacts to new volunteers when they deem it necessary or appropriate. CM1 recalls when they were first taught how to use the point-of-sale by the host of the shift they were volunteering at early on in their tenure:

> CM1: It was slow in the winter, and [the shift host] was like "I can teach both of you guys how to [use the point-of-sale] because I won't be here next week and you guys need to run the shop". And pretty quickly after that, it was one of the benefits of that little team there... that pretty much from that point somebody could start closing up [the register], somebody could start sweeping, and it was just whoever was there, it didn't have to be a hierarchical situation.

CM1's recollection suggests that access to artifacts like the point of sale system shapes the organizational hierarchy within the Bike Co-op. Although the shift host was still formally seen as and referred to as the shift host, CM1 had been granted access to a similar set of permissions that rendered them capable of performing the shift hosts' tasks in their absence. Once this access was granted, the volunteers on the shift were able to divide and rotate between closing up tasks

Granting increased access to tools and technology can happen outside a particular shift as well. For instance, the VC described a time when a shift host asked for more help with a busy sale shift in the middle of the summer. The host at the time was the only one who knew how to use the point-of-sale system, and was staffing the shift with two relatively new (but regular) volunteers. The VC could not find extra volunteers, so they instead contacted the two newer volunteers on the sale shift, and arranged to meet with them to teach them how to use the point of sale system to help the host out.

The VC recognizes that a lack of access to necessary system permissions is the

reason that promising volunteers might not see themselves as able to host a shift, and sees granting access as a way to build self-efficacy and interest in taking on leadership roles.

> VC: I contact them directly and say 'are you interested in hosting?' and then say 'this is what we need to do to set you up with that, we need to train you on the register, are you willing to come meet with somebody and learn the register, and be the host?... If you're interested in hosting, but you don't feel comfortable on the cash register, that's not a problem, we can teach you.'

Finally, sharing limited access items with new volunteers also helps them obtain 'insider information' about the organization, and helps them feel like a member due to their exclusive knowledge. During a volunteer hours shift I observed, an experienced volunteer told a newer volunteer about the 'treasure chest' - a box of particularly valuable salvaged bike parts that was stored out of reach on a high shelf. The experienced volunteer stood on a chair and brought it down, and together, they spent some time marveling at its contents. This was not necessarily additional access that the newer volunteer needed in order to get a task done, but rather an example of sharing organizational context, and had the effect of bringing the volunteer into an 'in-group'.

Identifying Volunteers for Formal Promotion

In contrast with host assignment at the Bike Co-op, which has been shaped by a centralized technological system, is periodic in nature, and concerns a specific point

person, volunteer development at the Bike Co-op is shaped by access to a variety of physical and technical artifacts, triggered by random events, and enabled by a number of veteran volunteers and staff members. However, despite their differences, both these processes need to work in concert in order for the Bike Co-op to effectively schedule and promote volunteers.

The tension between the scheduling and promotion processes at the Bike Co-op are most visible when the VC tries to schedule shift hosts. Every three months, the VC sends out a Google form to volunteers who have hosted shifts before, and asks which shifts they are willing/able to host the following quarter. If the VC finds enough people among the existing host list to commit to hosting all the slots for the next quarter, the search is complete. If not, the VC tries to identify regular volunteers who have gained sufficient experience within their shifts, and encourage them to step up and lead shifts independently.

Identifying potential new hosts occurs informally; if a collective member happens to converse with the VC about a promising volunteer, or if the VC meets a volunteer who expresses interest in taking on more responsibility, the VC keeps these interactions in mind when making these decisions. Once the VC has identified a new volunteer who can host, their name is then included in the host mailing list, and added to the rotation of volunteers who are able to host.

> If we're talking about identifying and building leadership, ... it's maybe whoever I decided to talk to after another shift, a collective member or a host mentions a new volunteer, and that sort of alerts me, and then I

> say "Oh yeah, this is that new volunteer's third time, I should follow up with them"... actually I did this with XYZ for example, I said that "[Your shift host] said that you could be ready to host stuff really quickly here, it sounds like you want to do that, is that right?"... so it's kind of moments like that. Knowing who's new, communicating with them. I wish it was more systematized, because I know people are falling through the cracks - I'm just relying on my own powers of observation for the most part. "

While this process usually works well enough for ensuring that shifts are staffed, it can result in the VC creating an incomplete list of people who are capable of hosting, or who are interested in hosting. I compared the VC's host list against the volunteer logs from Trackr, and found a number of volunteers who have been regularly attending shifts at the Bike Co-op for over a year, who have possibly gained significant levels of access on their shifts, but do not appear to be in the VC's host lists. The result of this is that the pool of people regularly *solicited* to host shifts does not grow commensurate with the number of people who are *able* to host shifts.

This means that the Bike Co-op is asking the same small pool of volunteers to take on hosting roles every month. While these volunteers may be very committed and good at their jobs, relying on too small a set of experienced volunteers can lead to burnout among people who step up to host shifts very often. Having too small a regular host list also makes the Bike Co-op vulnerable if experienced volunteers need to step down from their duties for any reason.

During discussions with the VC, they explicitly brought up their concern that there

are promising volunteers who might not be on their radar. Since the VC is unable to see which volunteers are experiencing informal promotions within their shifts, they depend on hearing this information through informal channels, and they are aware that they might not hear about every volunteer who has shown commitment and experienced promotions during their shifts.

The VC implicitly understood the importance of access to artifacts that mediates promotion within the Bike Co-op. One of the first concerns that the VC has when assigning a new shift host is whether this person has training and access to the tools they need to effectively work their shift. Understanding which volunteers have access to which artifacts is important, yet difficult to ascertain information from the VC's perspective.

> VC: Well for keys, it's [Bike Co-op] policy that if a host needs a key, we give them a key. At this point, we don't have a good key record keeping system. So that's not good. It's according to the need of the [Bike Co-op].
>
> SN: Do you know who has keys?
>
> VC: No. I mean, I know some, but I couldn't point to a up-to-date list.
>
> SN: Do you know who all knows how to operate the point of sale?
>
> VC: I don't have good data on it.

This frustration is visible through an examination of the documents in the VC's Google Drive. The VC has in fact started multiple spreadsheets that attempt to track which volunteers have what kind of access within the Bike Co-op. In at least two

different files, the VC began lists of which volunteers have point-of-sale access. Both lists are presently incomplete. The VC has made multiple attempts to create more 'homebrew' databases in this manner, but faces a persistent challenge in organizing this information because they often learn about it indirectly.

In addition to the ways that decentralized volunteer development impacts centralized scheduling, decisions around scheduling also have an effect on promotion. The promotion of existing volunteers depends on the presence experienced volunteers with more access granting that access to newer volunteers. Since the VC primarily creates the schedule to ensure the shifts are staffed, volunteers can sometimes be scheduled into shifts that close off their opportunities to advance.

For instance, CM2 had been a volunteer for multiple years, and had become deeply committed to the organization. They were asked by the VC to host a weekly shift staffing bike sales on Wednesday evenings. Since taking on their hosting responsibilities, they became interested in taking more of a leadership role in the Bike Co-op by joining the collective. However, it was not clear how they could go about doing so - since their responsibilities so far were determined by the VC or other hosts/collective members asking for more help, their initial instinct was to just wait until they were noticed enough to be asked to join the collective.

> CM2: I didn't understand what the avenue was for joining [the collective]. I was doing bike sales all summer 2017, but I was the host, and there was no collective member there. And from my understanding, a lot of what happens is that volunteers volunteer alongside a [collective member] host,

> and if the volunteer starts volunteering consistently, then that host says "Hey, you've been volunteering a bit, do you want to join the collective?" I had been involved with the organization more seriously for the past year and a half to two years, but I was never volunteering on a regular basis with a collective member host, I was always the host. So I was just kind of like, well, I don't know... I guess I'll just keep doing what I'm doing and get the courage to ask someone one day.

In this case, the VC had heard informally about a volunteer who had shown interest in taking on additional responsibilities, and asked them to host the sale shift. Once CM2 was asked to take on the shift, they were given a lot of responsibility, but were closed off from further opportunities for advancement because promotion implicitly depends on the presence of more experienced volunteers. This tendency is exacerbated by the shift structure of the Bike Co-op- since regular volunteers participate in weekly shifts, they can contribute to the Bike Co-op for many months before meeting experienced volunteers at other shifts who might pave the way to greater levels of access.

Discussion

Organizations like the Bike Co-op provide an opportunity to study technology use and tool access in cases where a single, comprehensive system for volunteer management was not enforced from the top down. In organizations with little top-down enforcement of tool use, different entities can use and introduce systems and processes based

on their own 'technological frames'[67]. In this case, the volunteer co-ordinator created a scheduling system that was built out of their personal mental model of how the organization operated. Further adoption of this system by other organizational stakeholders led to the formalization of a new role in the organization, as well as an increasingly central role for the VC. Similarly, the nature of work and mentorship during shifts at the Bike Co-op led to a different set of tools and processes that structured promotion. Rather than centering the role of any particular volunteer or collective member in volunteer advancement, access to particular tools such as the cash register or keys became central to building skill and experience at the Bike Co-op. Experienced volunteers structured these processes as they saw fit given their own mental models of how the organization operates.

While the creation and adoption of the VC's technological system led to the formalization and reification of the host role, a fundamental disjuncture in the promotion process at the Bike Co-op grew between the several informal ways in which volunteers are brought closer to the core of the organization, and the process by which leadership positions are identified and made visible.

My findings suggest that the organization needs to develop systems and routines that can help bridge the gap between informal volunteer development in shifts, and formal assignment of roles. Broadly, a way to address this is by making information about which volunteers have acquired privileged access more visible to the VC, and other volunteers.

This can be accomplished by implementing systems that allow for information

about promotion and access to be recorded and updated by volunteers, rather than the VC.The first method centers around building tools that make promotion processes more visible to both the VC, as well as volunteers. Currently, the VC has an incomplete sense of which volunteers are being promoted within their shifts. Pathways for promotion are also not often clear to volunteers - they are frequently granted increased access at the discretion of more experienced volunteers.

A potential way to make promotion processes more visible is by expanding the role of an existing system, i.e. Trackr. Unlike the weekly program schedule that is largely maintained by the VC, and the access to tools, keys and registers that is controlled by veteran volunteers, Trackr is directly accessible by volunteers themselves. Although its sole function right now is to note when volunteers are present in the shop, its role as a system that is accessed by all volunteers can be leveraged more than it currently is. While the VC's centralized weekly program schedule had the effect of making leadership roles more visible to new volunteers, Trackr has the potential of making privileged access more visible to the VC and other organizational stakeholders.

Conclusion

Overall, this dissertation showcases the significant challenges that volunteer-run communities experience in designing technologies to support participation. In Chapter 1, we saw that the Wikimedia Foundation's deployment of an onboarding tutorial did not ultimately elicit more contributions from newcomers (in constrast to expectations from organizational theory that suggested it would do so). In Chapter 2, Wikia's overhaul of discussion software on its wikis did increase the amount of communication that users engaged in across the board, but did not lead to a corresponding increase in contributions (and indeed, negatively impacted contributions from newcomers, again contradicting expectations from prior work in this space). In Chapter 3, the Bike Coop experienced challenges reconciling its formal systems of promotion with the more informal, gradual ways that new volunteers were brought into the organization.

A central theme that runs through my work is that organizations of this form (i.e. volunteer-run organizations with fuzzy boundaries where individuals self-select into tasks) experience the ongoing challenge of reconciling centralizing tendencies with decentralized organizational structure. Every case I've presented here, at some level, engages with this tension. In the case of Chapter 1, a formalized, uniform, and cen-

tralized system for socializing new volunteers did not succeed on Wikipedia, largely because it was at odds with the decentralized, self-selecting way that editors choose to contribute to Wikipedia. In Chapter 2, we see that the introduction of message walls increased the extent to which volunteers conversed with each other, but this intervention perhaps induced many decentralized and disconnected conversations that drew attention away from the central task of synthesizing information to create shared repositories of information. Finally, in Chapter 3, we see that the Bike Co-op struggles with reconciling its more centralized technological systems that formalize promotion decisions with the more decentralized manner that volunteers are developed and granted access to ther organization. While the settings for these studies have differed in both medium and purpose, my work in this dissertation suggests that the push and pull of these different tendencies presents a persistent challenge for volunteer communities of this form. How these communities can successfully navigate this balance, be it through technological systems, organizational workflows, or social norms, remains an open and compelling area of study.

Avenues for Future Work

Each of the studies in this text provide launching points for new avenues of inquiry that can shed more light on the dynamics of volunteer participation in distributed communities, and inform design in these areas.

For instance, a major takeaway of Chapter 1 is that the self-selecting nature of task assignment on Wikipedia makes it difficult to institute systems where new editors are

required to complete tutorials like the Wikipedia Adventure before they begin editing, and expecting users to choose to play the tutorial largely undermined its ability to affect user behavior in our study. However, there are other contexts where such an intervention might work well. Future work could focus on understanding how gamified tutorials like the Wikipedia Adventure shape contribution behavior to Wikipedia when introduced in contexts such as classrooms, or hackathons. Comparing the contribution behavior of groups where such a tutorial was used, to groups when it was not, could help us better understand the effect of such an intervention in contexts where uptake is significantly higher.

There is also potential to extend the analysis in Chapter 2 in several ways. First, we can proceed by exploring the effects of message walls with more granularity. We know that the introduction of message walls increased the number of messages that were sent by new and veteran volunteers. It is worth understanding what kinds of messages in particular faced an increase. For instance, after message walls were instituted, did newcomers ask more questions? Were veteran volunteers more likely to introduce themselves, or offer help? The current study design could be extended to answer questions like this, by, for instance, using techniques from natural language processing to construct new dependent variables that indicated the prevalance of particular kinds of messages. Even if we did not see a corresponding increase to contributions to wikis after message walls were instituted, understanding *how* communication changed (if at all) in these communities could help us understand whether such a technological change would likely produce longer term effects on contribution.

Another way Chapter 2 could be extended is by better understanding and classifying

the kind of variance exhibited by different wikis in the sample. While the aggregate trend of increased communication with no increase in contributions holds across a sample of 275 wikis, there were many wikis within the sample that exhbited different behavior. Understanding what kinds of factors (such as purpose of wiki, or number of administrators, etc) create the conditions for communties to respond to interface changes in different ways could shed light on how things like the purpose and internal organization of these communities shape the way they adopt new technological changes once they are rolled out en masse.

Finally, Chapter 3 presents a contextual inquiry that identifies a compelling design space to make paths to promotion more visible in volunteer communities. Future work in this area could involve using this research to design and deploy new systems that help track and identify promising volunteers within organizations. These systems could then be evaluated to shed more light on the relationship between the visibility and shared awareness of volunteer skills and roles, and volunteer advancement.

Bibliography

[1] Adler, B. T., Chatterjee, K., de Alfaro, L., Faella, M., Pye, I., and Raman, V. Assigning Trust to Wikipedia Content. In *Proceedings of the 4th International Symposium on Wikis* (New York, NY, USA, 2008), WikiSym '08, ACM, pp. 26:1–26:12.

[2] Adler, B. T., and de Alfaro, L. A Content-driven Reputation System for the Wikipedia. In *Proceedings of the 16th International Conference on World Wide Web* (New York, NY, USA, 2007), WWW '07, ACM, pp. 261–270.

[3] Ahuja, M. K., and Galvin, J. E. Socialization in Virtual Groups. *Journal of Management 29*, 2 (2003), 161–185.

[4] Angrist, J. D., and Krueger, A. B. Does compulsory school attendance affect schooling and earnings? Tech. rep., National Bureau of Economic Research, 1990.

[5] Angrist, J. D., and Pischke, J.-S. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ, 2008.

[6] Antin, J., and Cheshire, C. Readers are not free-riders: Reading as a form of participation on Wikipedia. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2010), CSCW '10, ACM, pp. 127–130.

[7] Antin, J., Cheshire, C., and Nov, O. Technology-mediated contributions: Editing behaviors among new Wikipedians. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (New York, NY, USA, 2012), CSCW '12, ACM, pp. 373–382.

[8] Arazy, O., Ortega, F., Nov, O., Yeo, L., and Balila, A. Functional Roles and Career Paths in Wikipedia. In *Proceedings of the 18th ACM Conference*

on Computer Supported Cooperative Work & Social Computing (New York, NY, USA, 2015), CSCW '15, ACM, pp. 1092–1105.

[9] Balkundi, P., and Harrison, D. A. Ties, leaders, and time in teams: Strong inference about network structure's effects on team viability and performance. *Academy of Management Journal 49*, 1 (2006), 49–68.

[10] Bauer, T. N., Bodner, T., Erdogan, B., Truxillo, D. M., and Tucker, J. S. Newcomer adjustment during organizational socialization: a meta-analytic review of antecedents, outcomes, and methods. *Journal of applied psychology 92*, 3 (2007), 707.

[11] Benkler, Y. Coase's penguin, or, Linux and the nature of the firm. *Yale Law Journal 112*, 3 (2002), 369–446.

[12] Benkler, Y. *The wealth of networks: How social production transforms markets and freedom*. Yale University Press, New Haven, CT, 2006.

[13] Benkler, Y. Peer production and cooperation. In *Handbook on the Economics of the Internet*, J. M. Bauer and M. Latzer, Eds. Edward Elgar, Cheltenham, UK, 2016.

[14] Benkler, Y., Shaw, A., and Hill, B. M. Peer production: A form of collective intelligence. In *Handbook of Collective Intelligence*, T. W. Malone and M. S. Bernstein, Eds. MIT Press, Cambridge, MA, 2015, pp. 175–204.

[15] Bryant, S. L., Forte, A., and Bruckman, A. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work* (New York, NY, 2005), GROUP '05, ACM, pp. 1–10.

[16] Burke, M., and Kraut, R. Mopping up: modeling wikipedia promotion decisions. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (2008), ACM, pp. 27–36.

[17] Butler, B., Joyce, E., and Pike, J. Don't Look Now, but We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2008), CHI '08, ACM, pp. 1101–1110.

[18] Cheshire, C., and Antin, J. The social psychological effects of feedback on the production of Internet information pools. *Journal of Computer-Mediated Communication 13*, 3 (2008), 705–727.

[19] Ciampaglia, G. L., and Taraborelli, D. MoodBar: Increasing New User Retention in Wikipedia Through Lightweight Socialization. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (New York, NY, USA, 2015), CSCW '15, ACM, pp. 734–742.

[20] Collier, B., and Bear, J. 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* (New York, NY, USA, 2012), CSCW '12, ACM, pp. 383–392.

[21] Crowston, K., Annabi, H., and Howison, J. Defining Open Source Software Project Success. In *ICIS 2003 Proceedings* (Seattle, Washington, USA, 2003), pp. 327–340.

[22] Cummings, J. N., and Cross, R. Structural properties of work groups and their consequences for performance. *Social Networks 25*, 3 (2003), 197–210.

[23] Deci, E. L., and Ryan, R. M. Intrinsic Motivation. In *The Corsini Encyclopedia of Psychology*. John Wiley & Sons, Inc., New York NY, 1975.

[24] Dee, T. S. Are there civic returns to education? *Journal of Public Economics 88*, 9 (2004), 1697–1720.

[25] Dejean, S., and Jullien, N. Big from the beginning: Assessing online contributors' behavior by their first contribution. *Research Policy 44*, 6 (July 2015), 1226–1239.

[26] Ducheneaut, N. Socialization in an Open Source Software Community: A Socio-Technical Analysis. *Computer Supported Cooperative Work (CSCW) 14*, 4 (2005), 323–368.

[27] Ekstrand, M. D., and Riedl, J. T. rv you're dumb: Identifying Discarded Work in Wiki Article History. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (New York, NY, USA, 2009), WikiSym '09, ACM, pp. 4:1–4:10.

[28] Farzan, R., Kraut, R., Pal, A., and Konstan, J. Socializing Volunteers in an Online Community: A Field Experiment. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2012), CSCW '12, ACM, pp. 325–334.

[29] Fé, E., and Hollingsworth, B. Short- and long-run estimates of the local effects of retirement on health. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 179*, 4 (2016), 1051–1067.

[30] Foote, J., Gergle, D., and Shaw, A. Starting online communities: motivations and goals of wiki founders. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)* (New York, New York, 2017), ACM, pp. 6376–6380.

[31] Frey, B. S., and Jegen, R. Motivation Crowding Theory. *Journal of Economic Surveys 15*, 5 (2001), 589–611.

[32] Geiger, R. S., and Halfaker, A. Using Edit Sessions to Measure Participation in Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2013), CSCW '13, ACM, pp. 861–870.

[33] Geiger, R. S., Halfaker, A., Pinchuk, M., and Walling, S. Defense Mechanism or Socialization Tactic? Improving Wikipedia's Notifications to Rejected Contributors. In *Sixth International AAAI Conference on Weblogs and Social Media* (Dublin, Ireland, May 2012), AAAI Publications, pp. 122–129.

[34] Gerber, A. S., and Green, D. P. *Field Experiments: Design, Analysis, and Interpretation*. WW Norton, New York, 2012.

[35] Guzdial, M., and Tew, A. E. Imagineering Inauthentic Legitimate Peripheral Participation: An Instructional Design Approach for Motivating Computing Education. In *Proceedings of the Second International Workshop on Computing Education Research* (New York, NY, USA, 2006), ICER '06, ACM, pp. 51–58.

[36] Halfaker, A., Geiger, R. S., Morgan, J. T., and Riedl, J. The rise and decline of an open collaboration system: how Wikipedia's reaction to popularity is causing its decline. *American Behavioral Scientist 57*, 5 (May 2013), 664–688.

[37] Halfaker, A., Keyes, O., and Taraborelli, D. Making peripheral participation legitimate: reader engagement experiments in wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work* (New York, NY, USA, 2013), CSCW '13, ACM, pp. 849–860.

[38] Halfaker, A., Kittur, A., Kraut, R., and Riedl, J. A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (New York, NY, USA, 2009), WikiSym '09, ACM, pp. 15:1–15:10.

[39] Halfaker, A., Kittur, A., and Riedl, J. Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (New York, NY, USA, 2011), WikiSym '11, ACM, pp. 163–172.

[40] Hamari, J., Koivisto, J., and Sarsa, H. Does Gamification Work? - A Literature Review of Empirical Studies on Gamification. In *Proceedings of the 2014 47th Hawaii International Conference on System Sciences* (Washington, DC, USA, 2014), HICSS '14, IEEE Computer Society, pp. 3025–3034.

[41] Hanus, M. D., and Fox, J. Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education 80* (Jan. 2015), 152–161.

[42] Haythornthwaite, C. Crowds and communities: Light and heavyweight models of peer production. In *2009 42nd Hawaii International Conference on System Sciences* (2009), pp. 1–10.

[43] Hill, B. M., and Shaw, A. The hidden costs of requiring accounts online: Quasi-experimental evidence from peer production. *Unpublished manuscript Under review*.

[44] Holtzblatt, K., Wendell, J. B., and Wood, S. *Rapid contextual design: a how-to guide to key techniques for user-centered design*. Elsevier, 2004.

[45] Huang, S.-W., Suh, M. M., Hill, B. M., and Hsieh, G. How Activists Are Both Born and Made: An Analysis of Users on Change.Org. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2015), CHI '15, ACM, pp. 211–220.

[46] Jagadish, H. V., Chapman, A., Elkiss, A., Jayapandian, M., Li, Y., Nandi, A., and Yu, C. Making Database Systems Usable. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data* (New York, NY, USA, 2007), SIGMOD '07, ACM, pp. 13–24.

[47] Joseph, D., and Nacu, D. C. Designing Interesting Learning Environments When the Medium isn't enough. *Convergence: The International Journal of Research into New Media Technologies 9*, 2 (June 2003), 84–115.

[48] Kiene, C., Shaw, A., and Hill, B. M. Managing organizational culture in online group mergers. Forthcoming in *Proceedings of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing,* CSCW '18, New York, NY, USA. ACM.), 2018.

[49] Kittur, A., and Kraut, R. E. Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2008), CSCW '08, ACM, pp. 37–46.

[50] Kittur, A., and Kraut, R. E. Beyond Wikipedia: Coordination and Conflict in Online Production Groups. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (New York, New York, 2010), CSCW '10, ACM, pp. 215–224.

[51] Kraut, R. E., and Resnick, P. *Building Successful Online Communities: Evidence-Based Social Design.* The MIT Press, Cambridge, MA, Mar. 2012.

[52] Kraut, R. E., and Streeter, L. A. Coordination in Software Development. *Communications of the ACM 38*, 3 (1995), 69–81.

[53] Kriplean, T., Beschastnikh, I., and McDonald, D. W. Articulations of wikiwork: uncovering valued work in wikipedia through barnstars. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW2008)* (San Diego, CA, USA, 2008), ACM, pp. 47–56.

[54] Lakhani, K., and Wolf, B. Why hackers do what they do: Understanding motivation and effort in free/open source software projects. In *Perspectives on Free and Open Source Software*, J. Feller, B. Fitzgerald, S. A. Hissam, and K. R. Lakhani, Eds. MIT Press, 2005, pp. 3–22.

[55] Lave, J., and Wenger, E. *Situated Learning: Legitimate Peripheral Participation.* Cambridge University Press, Cambridge, UK, Sept. 1991.

[56] Lave, J., and Wenger, E. Legitimate peripheral participation in communities of practice. *Supporting lifelong learning 1* (2002), 111–126.

[57] Lave, J., and Wenger, E. Legitimate peripheral participation in communities of practice. In *Supporting Lifelong Learning: Perspectives on learning*, vol. 1. Routledge, 2002.

[58] Luther, K., Fiesler, C., and Bruckman, A. Redistributing leadership in online creative collaboration. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2013), CSCW '13, ACM, pp. 1007–1022.

[59] Malone, T. W., and Crowston, K. What is Coordination Theory and How Can It Help Design Cooperative Work Systems? In *Proceedings of the 1990 ACM Conference on Computer-supported Cooperative Work* (New York, NY, USA, 1990), CSCW '90, ACM, pp. 357–370.

[60] Morgan, J. T., Bouterse, S., Walls, H., and Stierch, S. Tea and sympathy: crafting positive new user experiences on wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work* (New York, NY, USA, 2013), CSCW '13, ACM, pp. 839–848.

[61] Mugar, G., Østerlund, C., Hassman, K. D., Crowston, K., and Jackson, C. B. 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* (New York, NY, USA, 2014), CSCW '14, ACM, pp. 109–119.

[62] Murnane, R. J., and Willett, J. B. *Methods matter: Improving causal inference in educational and social science research.* Oxford University Press, New York, NY, 2011.

[63] Musicant, D. R., Ren, Y., Johnson, J. A., and Riedl, J. Mentoring in Wikipedia: a clash of cultures. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (New York, NY, USA, 2011), WikiSym '11, ACM, pp. 173–182.

[64] Narayan, S., Orlowitz, J., Morgan, J., Hill, B. M., and Shaw, A. 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* (New York, NY, USA, 2017), CSCW '17, ACM, pp. 1785–1799.

[65] Orlikowski, W. J. The Duality of Technology: Rethinking the Concept of Technology in Organizations. *Organization Science 3*, 3 (1992), 398–427.

[66] Orlikowski, W. J. Improvising Organizational Transformation Over Time: A Situated Change Perspective. *Information Systems Research 7*, 1 (Mar. 1996), 63–92.

[67] Orlikowski, W. J., and Gash, D. C. Technological Frames: Making Sense of Information Technology in Organizations. *ACM Trans. Inf. Syst. 12*, 2 (Apr. 1994), 174–207.

[68] Panciera, K., Halfaker, A., and Terveen, L. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the ACM 2009 international conference on Supporting group work* (New York, NY, USA, 2009), GROUP '09, ACM, pp. 51–60.

[69] Potthast, M., Stein, B., and Gerling, R. Automatic Vandalism Detection in Wikipedia. In *Advances in Information Retrieval*, C. Macdonald, I. Ounis, V. Plachouras, I. Ruthven, and R. W. White, Eds., no. 4956 in Lecture Notes in Computer Science. Springer, Berlin, Germany, Mar. 2008, pp. 663–668. DOI: 10.1007/978-3-540-78646-7_75.

[70] Preece, J., and Shneiderman, B. The reader-to-leader framework: motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction 1*, 1 (2009), 13–32.

[71] Priedhorsky, R., Chen, J., Lam, S. T. K., Panciera, K., Terveen, L., and Riedl, J. Creating, Destroying, and Restoring Value in Wikipedia. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work* (New York, NY, USA, 2007), GROUP '07, ACM, pp. 259–268.

[72] Reagle, J. *Good Faith Collaboration: The Culture of Wikipedia*. MIT Press, Cambridge Mass., 2010.

[73] Ren, Y., Harper, F., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., and Kraut, R. Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *Management Information Systems Quarterly 36*, 3 (2012), 841–864.

[74] Ren, Y., Kraut, R., and Kiesler, S. Applying common identity and bond theory to design of online communities. *Organization Studies 28*, 3 (2007), 377–408.

[75] Restivo, M., and van de Rijt, A. No praise without effort: experimental evidence on how rewards affect Wikipedia's contributor community. *Information, Communication & Society 17*, 4 (2014), 1–12.

[76] Rosenbaum, P. R. *Design of Observational Studies*. Springer series in statistics. Springer, New York, 2010.

[77] Schneider, J., Passant, A., and Breslin, J. G. Understanding and improving Wikipedia article discussion spaces. In *Proceedings of the 2011 ACM Symposium on Applied Computing* (New York, NY, USA, 2011), SAC '11, ACM, pp. 808–813.

[78] Schweik, C. M., and English, R. C. *Internet success: A study of open-source software commons*. MIT Press, Cambridge, MA, 2012.

[79] Shaffer, D. W., and Resnick, M. "Thick" Authenticity: New Media and Authentic Learning. *J. Interact. Learn. Res. 10*, 2 (Dec. 1999), 195–215.

[80] Shaw, A., and Hill, B. M. Laboratories of Oligarchy? How the Iron Law Extends to Peer Production. *Journal of Communication 64*, 2 (Apr. 2014), 215–238.

[81] Shaw, A., Zhang, H., Monroy-Hernández, A., Munson, S., Hill, B. M., Gerber, E., Kinnaird, P., and Minder, P. Computer supported collective action. *interactions 21*, 2 (2014), 74–77.

[82] Suh, B., Convertino, G., Chi, E. H., and Pirolli, P. The Singularity is Not Near: Slowing Growth of Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (New York, NY, USA, 2009), vol. 8 of *WikiSym '09*, ACM, pp. 1–10.

[83] Van Maanen, J., and Schein, E. H. Toward a theory of organizational socialization massachusetts institute of technology. *Research in organizational behavior* (1977), 209–264.

[84] Van Maanen, J., and Schein, Edgar H. Toward a theory of organizational socialization. In *Research in organizational behavior*, B. M. Staw, Ed. JAI Press, Greenwich, CT, 1979, pp. 209–264.

[85] Voida, A., Harmon, E., and Al-Ani, B. Homebrew Databases: Complexities of Everyday Information Management in Nonprofit Organizations. In *Proceedings*

of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2011), CHI '11, ACM, pp. 915–924.

[86] Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., and Smith, M. Finding social roles in Wikipedia. In *Proceedings of the 2011 iConference* (New York, NY, USA, 2011), iConference '11, ACM, pp. 122–129.

[87] Yeow, A., Johnson, S., and Faraj, S. Lurking: Legitimate or Illegitimate Peripheral Participation? *ICIS 2006 Proceedings 62* (2006).

[88] Zhang, X., and Zhu, F. Group size and incentives to contribute: A natural experiment at chinese Wikipedia. *The American Economic Review 101*, 4 (2011), 1601–1615.

[89] Zhu, H., Kraut, R., and Kittur, A. Effectiveness of Shared Leadership in Online Communities. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (New York, NY, USA, 2012), CSCW '12, ACM, pp. 407–416.

[90] Zhu, H., Kraut, R., and Kittur, A. Organizing without formal organization: group identification, goal setting and social modeling in directing online production. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (New York, NY, USA, 2012), CSCW '12, ACM, pp. 935–944.

[91] Zhu, H., Zhang, A., He, J., Kraut, R. E., and Kittur, A. Effects of peer feedback on contribution: a field experiment in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2013), CHI '13, ACM, pp. 2253–2262.

Appendix 1: Chapter 2 Supplement

Summary Statistics

To provide an overview of our dataset, we present summary statistics of our eight dependent variables in Table 9.

Outcome	Min.	Median	Mean	Max.	Std. Dev.
No. of messages (all)	0.00	0.00	5.57	685.00	25.19
No. of communicators (all)	0.00	0.00	1.40	43.00	2.60
No. of contributors (all)	0.00	3.00	5.30	272.00	10.52
No. of contributions (all)	0.00	17.00	80.09	26488.00	470.55
No. of messages (new)	0.00	0.00	0.86	61.00	3.10
No. of communicators (new)	0.00	0.00	0.42	20.00	1.06
No. of contributors (new)	0.00	1.00	2.15	177.00	6.86
No. of contributions (new)	0.00	2.00	8.57	849.00	31.52

Table 9: Summary statistics for our eight dependent variables across all weeks in our study. Each variable describes the amount of activity taking place within a wiki each week.

Retaining Week 0

In our main paper, we presented models where we drop data from the week immediately after the transition (referred to as 'week 0'). We did so because we identified a 'bump' in many of our dependent variables during week 0.

We present additional evidence for this bump in Table 10 by showing the change in medians in the weeks surrounding the transition[24]. In four out of our eight dependent variables, we see that the median value of the measure shows an increase in week 0 compared to the surrounding weeks. Though these increases might appear small in most cases, the distributions of all these variables are quite skewed and include many zeroes, and an increase in the median by even one indicates significant variation.

The biggest evidence for the presence of this bump is seen in Table 11, where we report our findings with data for week 0 included. In M1a, M1b, M3a, and M3b, we see that estimate sizes are significantly higher, in some cases more than double of what they would have been without week 0. We find that M2a and M2b are significant in this specification, although that is no longer the case when we drop week 0. And finally, M4a and M4b undergo a sign change with the presence of week 0.

Although these results portray a much more flattering view of the effects message walls, we chose not to present them as our main finding because in each case, these estimates were largely the result of one specific data point.

[24]We report a similar table that describes the change in means in the main paper.

Outcome	Week -1	Week 0	Week 1
Amt. of communication (all)	1.00	2.00	1.00
No. of communicators (all)	1.00	2.00	1.00
No. of contributions (all)	23.00	37.00	31.00
No. of contributors (all)	3.00	4.00	4.00
Amt. of communication (new)	0.00	0.00	0.00
No. of communicators (new)	0.00	0.00	0.00
No. of contributions (new)	3.00	4.00	3.00
No. of contributors (new)	1.00	1.00	1.00

Table 10: Median outcomes across all wikis for the weeks immediately around the transition.

Outcome	Est.	SE
M1a: No. of messages (all)	1.027 ***	0.10
M1b: No. of communicators (all)	0.64 ***	0.06
M2a: No. of contributors (all)	0.122 ***	0.03
M2b: No. of contributions (all)	0.468 ***	0.07
M3a: No. of. messages (new)	0.724 ***	0.14
M3b: No. of communicators (new)	0.593 ***	0.10
M4a: No. of contributors (new)	0.055	0.05
M4b: No. of contributions (new)	0.097	0.09

Table 11: Estimates for the *msgwall* term in models that include data from the week after the transition to message walls.

Placebo Tests

To ensure that our results were not driven by spurious correlations, we conducted a number of "placebo" tests. To do this, we ran our models on 14 datasets with fictional message wall transition dates. We made these datasets by transposing the real transition by periods of greater than 120 days, thus choosing non-overlapping 16-week periods that bounded 14 different 'fake' transition dates.

We found that about 90% of the models we ran yielded no significant outcomes

around the fictional transition dates. This suggests that the effects we estimate around the true transition dates are likely due to the variation introduced by the message wall feature.

Shorter and Longer Analytic Windows

Another possible threat to our findings concerns the number of observations we choose to make before and after the transition to message walls. It is theoretically possible that our results are driven by the underlying trend in our data rather than time of the transition.

To test our results' sensitivity to this threat, we ran our models on datasets where we changed the analytic windows of our study to 6 weeks before and after the transition, and 12 weeks before and after the transition. Table 12 and Table 13 show the results of these alternate specifications.

Table 12 shows that reducing the analytic window to 6 weeks before and after the transition generally shows similar findings as our original models. We see the same pattern of findings in M1a, M1b, M2a and M2b as we did in the orignal models, with significant, positive estimates for M1a and b and insignificant estimates for M2a and b. For M3 and M4, we see estimate sizes all have the same signs as the original model, yet there is some difference in which models were significant. In this case, M3a is insignificant and M4a is significant- the opposite was true in the original models.

Table 13 shows that increasing the analytic window to 12 weeks before and after

Outcome	Est.	SE
M1a: No. of messages (all)	0.414 ***	0.12
M1b: No. of communicators (all)	0.178 *	0.07
M2a: No. of contributors (all)	-0.03	0.04
M2b: No. of contributions (all)	-0.064	0.08
M3a: No. of. messages (new)	0.256	0.17
M4a: No. of contributors (new)	-0.195 **	0.06
M4b: No. of contributions (new)	-0.418 ***	0.11

Table 12: Estimates for the *msgwall* term in models that include data from six weeks before and after the transition to message walls.

the transition shows the same pattern of significance for all models but one - M4b is no longer significant in this specification.

Outcome	Est.	SE
M1a: No. of messages (all)	0.59 ***	0.09
M1b: No. of communicators (all)	0.309 ***	0.06
M2a: No. of contributors (all)	0.049	0.03
M2b: No. of contributions (all)	-0.002	0.06
M3a: No. of. messages (new)	0.544 ***	0.13
M3b: No. of communicators (new)	0.418 ***	0.10
M4a: No. of contributors (new)	-0.01	0.05
M4b: No. of contributions (new)	-0.07	0.08

Table 13: Estimates for the *msgwall* term in models that include data from twelve weeks before and after the transition to message walls.

Smaller and Larger Time Windows for Binning Data

The analysis we presented was based on measures constructed by binning a week's worth of activity in each data point. It is possible that our findings could be driven by this choice; creating larger bins could over-smooth the data and lose variation, and

whereas smaller bins could introduce much more variation. Both these outcomes could affect the way we estimate the local trend around the transition.

To estimate how sensitive our results were to these threats, we ran our models on data binned with alternate definitions of 'weeks'. In Tables 14 and 15, we report the results of running our models on 4-day long 'weeks' and 10-day long 'weeks' respectively. We find that the size and significance of our estimates are unchanged by binning our data into 4-day weeks. However, increasing the week length to 10 days does indeed cause a loss of variation in our dataset - in this case, M1b, M3a and M3b are no longer significant. Additionally, we find a significant negative estimate for M4a.

Outcome	Est.	SE
M1a: No. of messages (all)	0.577 ***	0.10
M1b: No. of communicators (all)	0.318 ***	0.06
M2a: No. of contributors (all)	0.049	0.03
M2b: No. of contributions (all)	0.045	0.06
M3a: No. of. messages (new)	0.467 **	0.15
M3b: No. of communicators (new)	0.437 ***	0.11
M4a: No. of contributors (new)	-0.053	0.05
M4b: No. of contributions (new)	-0.187 *	0.09

Table 14: Estimates for the *msgwall* term in models where data is binned in four day long 'weeks'.

Dropping Administrators

We were interested in understanding how administrators on Wikia were driving our findings. Given the sparse editor base of many Wikia wikis, it is often the case that the vast majority of activity in each wiki is driven by administrators.

Outcome	Est.	SE
M1a: No. of messages (all)	0.345 **	0.12
M1b: No. of communicators (all)	0.104	0.07
M2a: No. of contributors (all)	-0.026	0.03
M2b: No. of contributions (all)	-0.157	0.08
M3a: No. of. messages (new)	0.118	0.16
M3b: No. of communicators (new)	0.117	0.12
M4a: No. of contributors (new)	-0.136 *	0.06
M4b: No. of contributions (new)	-0.341 ***	0.10

Table 15: Estimates for the *msgwall* term in models where data is binned in ten day long 'weeks'.

This is true of our dataset as well. Table 16 shows compares findings from M2b and and M4b (related to the number of contributions made by different user groups) with a model that estimates the effect of message walls on the contributions made by all non-administrative editors. We find that the estimate for non-administrative editors is virtually the same as the one for newcomers. This demonstrates that administrators produce the vast majority of content in our sample.

Outcome	Est.	SE
M2b: No. of contributions (all)	0.011	0.07
No. of contributions (no admins)	-0.205 **	0.08
M4b: No. of contributions (new)	-0.205 *	0.10

Table 16: Estimates for the *msgwall* term in models that drop admins, in

Varying the Definition of Newcomers

In our study, we categorize an edit as being made by a newcomer if it was made by an account that was a) created in the last three months, and b) had made fewer than 20

edits so far. We sought to understand the extent to which our results were driven by the way we defined newcomers in our study. Thus, we reran the models that pertained to newcomers (i.e. M3a and b, and M4a and b) on datasets where we varied the definition of newcomer by altering either the time since account creation or the number of edits made so far (see Tables 17,18,19,20). We find that in almost every case, variations to the definition of newcomer do not significantly alter our findings. We do see a difference in Table 20, where although M3a and M4a have positive estimates, they are no longer significant.

Outcome	Est.	SE
M3a: No. of. messages (new)	0.424 **	0.16
M3b: No. of communicators (new)	0.308 **	0.12
M4a: No. of contributors (new)	-0.083	0.06
M4b: No. of contributions (new)	-0.196 *	0.10

Table 17: Estimates for the *msgwall* term in models where newcomers have edited for less than two months (and made fewer than 20 edits).

Outcome	Est.	SE
M3a: No. of. messages (new)	0.415 **	0.16
M3b: No. of communicators (new)	0.29 *	0.12
M4a: No. of contributors (new)	-0.09	0.06
M4b: No. of contributions (new)	-0.193 *	0.10

Table 18: Estimates for the *msgwall* term in models where newcomers have edited for less than four months (and made fewer than 20 edits).

Outcome	Est.	SE
M3a: No. of. messages (new)	0.481 **	0.16
M3b: No. of communicators (new)	0.27 *	0.13
M4a: No. of contributors (new)	-0.109	0.06
M4b: No. of contributions (new)	-0.166	0.09

Table 19: Estimates for the *msgwall* term in models where newcomers have made fewer than ten edits (in less than three months).

Outcome	Est.	SE
M3a: No. of. messages (new)	0.238	0.15
M3b: No. of communicators (new)	0.212	0.11
M4a: No. of contributors (new)	-0.081	0.06
M4b: No. of contributions (new)	-0.226 *	0.10

Table 20: Estimates for the *msgwall* term in models where newcomers have made fewer than thirty edits (in less than three months).