THIS IS AN UNPUBLISHED DRAFT OF A CHAPTER FROM A FORTHCOMING BOOK. WHEN PUBLISHED, THE CITATION FOR THIS CHAPTER WILL BE: Hill, B. M., & Shaw, A. D. (2019) Studying Populations of Online Communities. In Foucault Welles, B. & González-Bailón, S. (Eds.), *The Handbook of Networked Communication*. New York, New York: Oxford University Press.
BEFORE CITING OR DISTRIBUTING, PLEASE CHECK FOR AN UPDATED VERSION AT: https://mako.cc/academic/hill_shaw-populations_of_communities-DRAFT.pdf

Studying Populations of Online Communities

Benjamin Mako Hill*
makohill@uw.edu

Aaron Shaw*
aaronshaw@northwestern.edu

INTRODUCTION AND BACKGROUND

The discussion sections of empirical research on digital communities are littered with unsatisfying phrases. Networked communication researchers have read and, in many cases written, sentences like the following: our study is only of one web forum and we cannot know how our work generalizes to others; because we have gathered data only from Wikipedia, we only propose that our findings might be antecedents of that community's success; messages in our dataset were collected from one Facebook group and we cannot speak to all the conversation that may have happened elsewhere. These phrases are acknowledgements of real problems. We think we can do better.

The limitations identified in the sentences above each stem, in large part, from the fact that most empirical online community research looks within individual communities. Although there have been literally thousands of papers published about individual peer production projects like Wikipedia and Linux, there has been vanishingly little research that compares peer production projects to each other. Likewise, studies of communication in social net-

*Authors contributed equally and author order was determined randomly.

working sites and discussion groups have nearly always focused on interactions within a single network like Facebook or an individual discussion forum. Such research does not always generalize well. Furthermore, we foreclose many types of research questions by selecting research sites for their size, longevity, or the engagement level of their participants.

In this chapter we argue that by studying groups of communities, we can enhance the quality of online communities research in multiple ways. First, we can mitigate many common threats to the validity of online community research designs. Moreover, by studying groups of online communities, we open the door to answering types of questions that are unanswerable when our datasets begin and end at communities' borders.

A few key points about population-level analysis and online communities will help contextualize the rest of our argument. First, we are not advocating for anything especially radical by arguing that research on networked communication in online communities ought to do more to adopt population-level approaches. Empirical communication research into online communities has existed for several decades[1] and we cite many examples of population-level approaches. Moreover, population-level analysis exemplifies well-established traditions in organization science, which has sought a turn away from models of organizations as self-contained entities to models that treat them as actors in complex social environments; a shift that inspired a proliferation of comparative, population-level approaches to organizational analysis (Scott & Davis, 2006). We draw inspiration from these approaches and seek to describe how they might be applied in the context of research on online communities. We also believe that the empirical contexts of online communities present an extraordinary opportunity to extend these concepts and frameworks.

A second point concerns what exactly we mean when we talk about populations of online communities. In ecology, the term "population" is used to describe all interbreeding organisms that share a geographical area. In demography, the term is used to refer to groups of humans. In organizational sociology, it is used to refer to similar types of organizations that might compete for resources like customers or suppliers. Although others have defined the term more narrowly (e.g. Weber, Fulk, & Monge, 2016), we use the term "population" broadly and inclusively to refer to any groups of online communities whose membership is defined through similarity, competition, or

[1] We do not attempt to provide a comprehensive accounting of this prior work. Interested readers might seek out Kraut and Resnick (2012), which is an excellent book-length synthesis of much empirical online community research.

interaction. As is the case in ecology, demography, and organizational sociology, there are many ways to define populations of online communities and a single online community might belong to a large number and variety of distinct populations.

Most of our discussion will focus on populations of online communities defined through their use of a common technical platform. Examples of populations of this type might include discussion forums that use the same bulletin board software, wikis that use the same hosting service, or software development projects using the same tools to host their code and coordinate their work. For example, SourceForge or Github each provide a common technological platform that hosts millions of different online communities dedicated to creating software projects, LinkedIn and Facebook each act as platforms to host online communities associated with many different offline organizations and interest groups, and Reddit hosts hundreds of thousands of "sub-reddit" communities around different discussion topics. Because communities within this type of population are mediated by a common technology, data is often more readily available and consistent as a single host or platform provider may provide access via a single Application Programming Interface (API) or database. The resulting datasets support the advantages of computational social science (Lazer et al., 2009) and can often facilitate direct comparative analysis.

Of course, there are other ways to define populations. For example, a population of online communities might include all communities hosting conversations about a set of topics or beliefs or serving a single geographic area. Examples might be as narrow as the group of all Star Wars fan culture communities or as broad as all music sharing networks. The analogy to an industry of firms or a denomination of churches maps more or less directly in cases like these. A population might also be described in terms of communities whose membership overlaps or through which messages diffuse. An exciting body of research, including studies of diffusion which we discuss below, has constructed large platform-spanning populations of communities.

In the rest of this chapter, we use examples from recent empirical research to highlight five benefits of studying populations of online communities. First, we argue that studies of populations can lead to increased generalizability. That said, community-spanning research designs not only improve the type of research already done within online communities, but also bring entirely different questions into the realm of answerability. In particular, we highlight the ability to study community-level variables like group size, inter-community behavior like knowledge diffusion, and the ways that communities affect each other through dynamics like competition. We also argue that data from populations of online communities makes it possible to combine many of these benefits with the benefits of intra-community analyses. Finally, we discuss a series of limitations before concluding.

Benefit 1: Generalizability

The first benefit of studying populations of communities is simple and straightforward: studies of a single community – no matter how exhaustive, granular, and expertly designed – may produce findings that hinge on the idiosyncrasies of the community being studied. While this is intuitive and widely-acknowledged, analyses of networked behavior and group processes frequently focus on a single platform or online community (e.g., Facebook, Twitter, Wikipedia) and result in findings that do not apply more broadly. Analyzing data drawn from multiple sites, communities, or platforms can lead to greater generalizability.

Early scholars of online communities gathered and compared evidence from case studies of the discussion-based systems Usenet, text-based role-playing games like "multi-user dungeons" (MUDs), the pioneering community "the WELL," and others (e.g., Kollock & Smith, 1999). Over time, the forces that brought about the rise of computational social science (Lazer et al., 2009), including widespread availability of digital trace data, and reduced costs of computing and storage, have made more direct comparisons across communities possible. However, despite the existence of opportunities for inter-community empirical work, the vast majority of studies in online community research still focus on a single site. Generalizability is hurt because communities are heterogeneous in ways that are frequently poorly understood.

This point is illustrated by Hargittai (2015) who uses data from offline surveys to describe the user bases of large online communities like Facebook and Twitter in traditional demographic terms. Although Hargittai's point is that participants in these communities are far from a cross-section of society, she also suggests that the demographics of online communities' user bases are distinct from each other. Of course, decades of social science have shown that demographic characteristic like age, race, gender, and skills are correlated with many of the attitudes and behaviors that social scientists study. Because online communities are different from the general population in demographic terms, Hargittai suggests that generalizability to society is often unjustified. An implication of her findings is that, to the extent that communities also

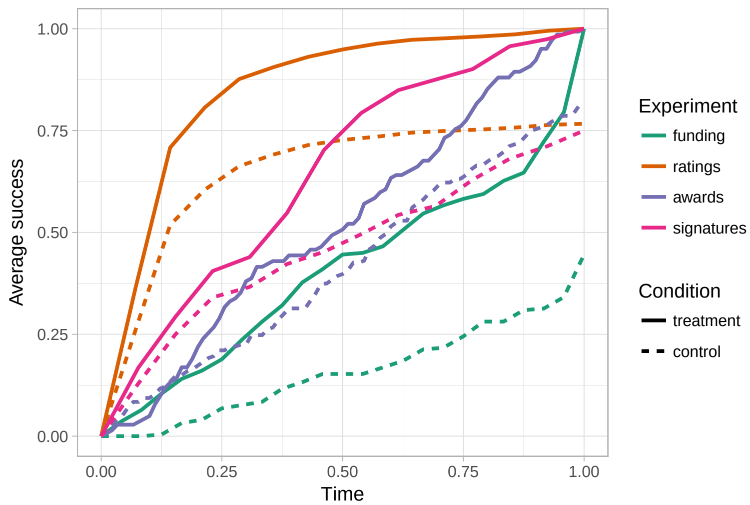

Figure 1: The success-breeds-success effect over time. The curves represent running numbers of donations (blue), positive ratings (red), awards (yellow), and campaign signatures (green) in the experimental conditions (solid lines) and the control conditions (dashed lines). The horizontal axis is normalized so that 0 marks the time of experimental intervention, and 1 marks the end of the observation period. The vertical axis is normalized so that for each system a value of 1 equals the maximum across time and conditions. (Figure and caption adapted from van de Rijt, Kang, Restivo, and Patil (2014)).

differ systematically from each other, generalization between communities will be fraught as well.

Studies across populations of online communities can help increase confidence in generalizability. For example, two studies published by Michael Restivo and Arnout van de Rijt look at the effects of social awards in Wikipedia called "barnstars" (Restivo & van de Rijt, 2012, 2014). In the first of these projects, the authors select very active Wikipedians and award barnstars to random subsets. Award recipients go on to edit the encyclopedia more than users not given an award. They also go on to receive more awards in the future. The authors suggest that such peer-to-peer forms of public recognition may produce a "success-breeds-success" dynamic. In a follow-up paper, van de Rijt, Kang, Restivo, and Patil (2014) compare the evidence of this effect in Wikipedia to very similar field experiments in three different online communities and show that the dynamic they identified in Wikipedia is also present in donations to Kickstarter, positive ratings of products on Epinions.com, and signatures on the e-petition site Change.org. The effects in all four experiments are shown together in Figure 1.

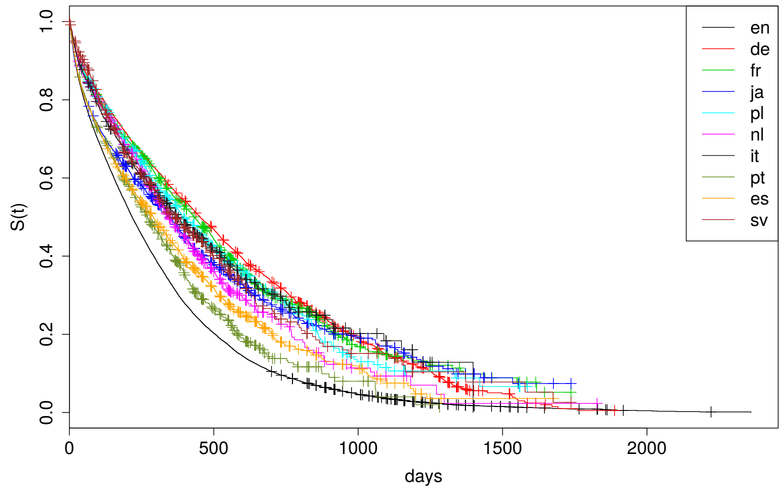

Figure 2: Survival functions that show the likelihood that editors participating in the core group of contributors continue to participate in each project once they cease to be a member of that core group. English Wikipedia is labeled "en" and is represented by the black line closest to the x-axis. (Figure adapted from Ortega (2009) page 125.)

Many of the most studied online communities – even those that appear extremely similar – are unusual in ways that impact findings and limit generalizability. For example, in an extensive comparative analysis of the ten largest language editions of Wikipedia, Ortega (2009) details many instances where the English edition is very different from Wikipedias in other languages. One example is editor retention. English Wikipedia's low editor retention has been an extremely widely studied problem (see Halfaker, Geiger, Morgan, & Riedl, 2013; Halfaker, Kittur, & Riedl, 2011), but most analyses of it can only speculate as to whether editor retention represents an issue endemic to all other large wikis. Ortega shows that the English-language community has extremely low retention among the most committed editors compared to other language Wikipedias (Figure 2). His results suggest that something particular about the English language community should explain at least part of its retention issues.

Generalization unsupported by direct comparison is risky business. This concern is akin to the challenge of scientists using "model organisms" to understand basic biological processes (Fields & Johnston, 2005). For example, much of what scientists know about genetics stems from research conducted with *Drosophila melanogaster*. That said, the question of when we should, and should not, generalize from *Drosophila* to other species is often not clear. Moreover, model organisms – like the most widely studied online communities – are selected precisely because of their extreme or idiosyncratic characteristics. Geneticists study *Drosophila* because it reproduces more quickly than other species. Online community researchers study Wikipedia precisely because it has attracted so many readers, contributions, and contributors and because it has generated articles of exceptional quality. We do not always understand the relationship between these unusual features and the basic social and communicative processes we seek to study. In such situations, a combination of careful theorizing combined with healthy doses of analytic modesty and skepticism become a researcher's most important assets.

Benefit 2: Studying Community-level Variables

Although Wikipedia's enormous size and high quality articles motivate its starring role in online community research, understanding *why* Wikipedia became so big and high quality is remarkably difficult. The question "Why did Wikipedia succeed?" implies treating Wikipedia itself as the unit of analysis and measuring its success (however you define it) at the level of the community. As introductory research methods textbooks explain (e.g., King, Keohane, & Verba, 1994, pg. 139–142), answering this type question requires variation across multiple projects, some of which succeed and some of which fail. Ideally, to study why Wikipedia succeeded, you would compare it to failed Wikipedias or, at least, to projects that were trying to be something very similar to what Wikipedia became (Hill, 2013). This sort of comparison and inference represents an especially compelling opportunity created by studying populations of online communities.

Organization-level variables have a long tradition in organizational research where scholars have used data on variation in firm performance to understand which kinds of firms thrive. For example, a body of work has established and tested a theory of the "liability of newness" that suggests that newer firms are more likely to fail (Schoonhoven, 2015). Other research has examined how the structure of communication affects the performance of teams and work groups engaged in collaboration of various kinds (e.g., Crowston & Howison, 2006; Cummings & Cross, 2003). Because structure exists within groups, studying multiple groups provides variation that makes it possible to understand how structure can affect group-level outcomes. As a result, a large body of research in organizational communication has focused on group, team, or even firm-level variables.

Both the scholarly and popular literature about the power of online communities makes claims about project-level performance and outcomes. As a result, there have been many prior calls to pursue studies across projects and communities (e.g., Benkler, Shaw, & Hill, 2015; Crowston, Wei, Howison, & Wiggins, 2008; Kraut & Resnick, 2012). Schweik and English's (2012) book *Internet Success* provides a compelling example of the power of this approach. Deeply influenced by Ostrom's (1990) work on common pool resource management, Schweik and English frame their analysis of free/libre open source software (FLOSS) projects in terms of questions about the successful provision of information commons. They create a stage model of when FLOSS projects will be abandoned and when they will develop into effective communities. They seek to explain these outcomes using dozens of project-level variables like leadership style and the copyright license of the project. They find that measures of clarity in leadership communication are among the best predictors of successful FLOSS commons and that other potential antecedents of FLOSS success cited in earlier literature, like license choice, do little to explain project outcomes.

Schweik and English's variables are nearly all at the level of the FLOSS project or community. Although previous case studies of leadership and governance reveal quite a bit about how FLOSS communities function, there is a sense in which they are limited to describing *how* these communities have operated. Because the sample of FLOSS projects analyzed by Schweik and English (2012) in *Internet Success* varies in terms of success (as the authors define it) and its theoretical antecedents, their analysis can begin to answer the question of *why* communities succeed. In this way, population-level research can engage with substantively and theoretically important questions in a manner that smaller scale analyses or comparisons struggle to do.

Perhaps the most important organization-level variables in networked communication research are related to technology. However, individual online communities tend to run on a single piece of software that consistently mediates the experience of every participant and the vast majority of research on the effect of technology on online communities uses data in which the technology itself is held constant. In the best cases, researchers are able to take advantage of a technological change or shock within a community to understand the impact of technology (e.g., Geiger & Halfaker, 2013). By looking across numerous communities, researchers can also build samples where technology varies as a way of testing these claims systematically (e.g., Shaw & Benkler, 2012). Such platform-level comparisons exemplify the sorts of analyses that become possible with the inclusion of community-level variables.

Benefit 3: Studying Diffusion Between Communities

The ability to answer questions related to the way that information and practices flow between communities is another benefit of population-level research. For example, we know that a huge proportion of messages on social media systems like Twitter include links to sources outside Twitter and to other online communities (Bruns & Stieglitz, 2012, 3/4). As a result, studies that attempt to characterize processes of information transmission without looking beyond the boundaries of an individual community or platform are necessarily incomplete. Although it can be difficult, research across communities can paint more accurate pictures of the diffusion.

Studies of diffusion have played a central role in communication scholarship since Rogers's (1962) seminal work on the diffusion of innovation. Subsequent studies by Valente (1995) and others have situated these diffusion processes within networks. Most of this research has sought to understand how information flows between individuals. For example, Bennett and Segerberg (2013) used data on message diffusion on Twitter to understand processes related to power, politics, and collective action. Of course, Twitter makes up only a part of Twitter users' media diets and many tweets point to further engagement in other media (Bruns & Stieglitz, 2012, 3/4). This speaks to the importance of research designs that can encompass a larger set of communities and thereby offer a more comprehensive explanation of diffusion processes.

The power of looking across communities is illustrated by a recent study by Graeff, Stempeck, and Zuckerman (2014) that traces the diffusion and development of conversations about the shooting of Trayvon Martin in February 2012. In addition to traditional media sources including television transcripts and newspapers, the analysis uses data from Media Cloud, an enormous dataset of many thousands of online media sources (Benkler, Roberts, Faris, Solow-Niederman, & Etling, 2013); search volume data from Google; messages on Twitter; signature data on the e-petitioning site Change.org; and a unique dataset from bit.ly that provided measures of how often people viewed social media material in a variety of communities including Twitter and Facebook. The paper traces how, after initially being reported only in local news, Martin's killing faded entirely from the media for weeks. A publicist hired by Martin's family and their lawyer then thrust the event back into the local media, bringing national media coverage that led to an online petition on the website Change.org. Then, a staff member at Change.org took notice of the petition and successfully engaged celebrities like Talib Kweli and Spike Lee to tweet messages about it, drawing increased engagement and attention. This

Date	Change.org petition traffic	Traffic referred by social media
12 March	11,141	7,486
13 March	34,345	20,712
14 March	305,672	45,952
15 March	190,354	72,165
16 March	80,268	50,798

Table 1: Total traffic to the Trayvon Martin Change.org petition page and traffic referred to by social media, several days when the petition was active. (Table from Graeff, Stempeck, and Zuckerman (2014).)

increased attention led to further media coverage.

Graeff and colleagues' accounting provides a compelling answer to the question of how Trayvon Martin's killing became one of the most important news stories of 2012 precisely because they follow the story as it diffuses through different media sources and across different online communities, networks, and platforms. By looking across communities, the authors discover that activity in social media communities like Twitter and Facebook was driven by activity in Tumblr and Blogger, and that the activity in these communities was driven by earlier activity in Twitter, which in turn was driven by activity around an e-petition. Table 1 reproduces some of Graeff et al.'s (2014) data summarizing traffic to the Change.org petition as well as traffic referred through social media sites. The multiple data sources combine to illustrate how the controversy developed across many types of media.

Although the dataset on the Trayvon Martin controversy used by Graeff and colleagues remains far from comprehensive, it offers the opportunity to go much further than an analysis of any single source could. The study also builds on previous work that had looked at network gatekeeping within Twitter (Meraz & Papacharissi, 2013) and is able to extend this theory beyond a single communication network. As is always the case with new methodologies, this approach also introduces new limitations. In particular, there are important challenges in terms of source selection, new opportunities for sample bias, and significant technical difficulties caused by the need to collect and integrate heterogeneous datasets.

Other studies consider other types of information transmission processes across community boundaries. A highly-cited study by Leskovec, Backstrom, and Kleinberg (2009) used an enormous dataset of 1.6 million different mainstream media sites and blogs to theorize a set of temporal patterns through which short distinctive phrases diffuse from news sites to blogs. A literature on code reuse in FLOSS has shown how code written for one project can be used in another (e.g., Mockus, 2007) and research in computer science has looked at how images are copied and moved between websites (Seneviratne et al., 2009). These lines of inquiry illustrate how studies of information flow across communities are increasingly possible and can lead to theoretically salient insights. By considering a larger part of the communicative and collaborative landscape, these approaches enable the analysis of diffusion processes that may not exist at all within communities.

Diffusion processes can also involve the transmission of organizational practices. In organization science, sociologists have studied dynamics of organizational "isomorphism" whereby organizations adopt routines and structures used in other organizations (DiMaggio & Powell, 1983; Meyer & Rowan, 1977, e.g.). Although studies of this type of diffusion remain much less common in online communities, they are possible. In one recent study, Zhu, Kraut, and Kittur (2016) considers the diffusion of a model for coordinating through topic-focused task forces within Wikipedia called "Collaboration of the Week." Because they look at both a group-level process and group-level outcomes of productivity across many groups, Zhu and her colleagues' can evaluate the effectiveness of the practice at a more general level. Their work shows how a population-level research design also makes it possible to understand the ways that the task force model of collaboration spread and, together with the Graeff et al. study, illustrates how such research can alter and enhance our understanding of multiple kinds of diffusion processes.

Benefit 4: Studying Ecological Dynamics

Research on diffusion and information transmission underscores how online communities do not exist in a vacuum. By studying individual communities, we tend to study them in isolation from each other and from the external forces that shape or threaten them. This point leads us to another line of inquiry that population-level datasets of online communities can address: the way that online communities interact with their environments and the surrounding ecology of similar communities. We use the terms "environment" and "ecology" here in the same way that scholars in organizational studies have applied them to describe a collection of factors and pressures outside organizations' boundaries.[2] In studies of firms, an environment consists of

[2] See Scott and Davis (2006) for both an excellent general overview of organizational theory and a detailed account of organization scientists' definitions of organizational environ-

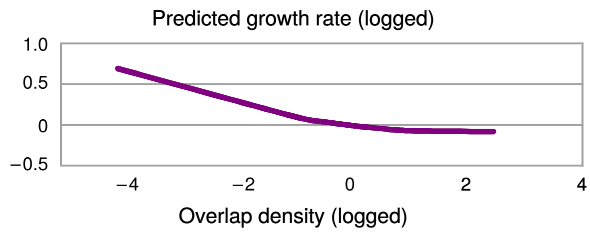

Figure 3: Relationship between member overlap density and membership growth as Usenet groups as predicted by an empirical model. Figure adapted from Wang, Butler, and Ren (2013), page 426.

competitors, suppliers, and customers. For example, one could describe the studies of diffusion discussed above as a type of research on environments. In this section, we highlight the opportunity to learn about online communities by studying how they experience ecological forces through their environments. Questions of this type remain almost entirely unexplored in online community research, but present promising opportunities.

Ecological studies of organizations are well established within organizational research. Seminal work in organizational science borrowed the metaphor of ecologies from biology to assert a set of theories around resource competition among firms (Carroll & Hannan, 2000; Hannan & Freeman, 1993). A central insight drives this body of work: firms succeed or fail because of conditions in their environment like the the number of competing firms. It turns out that the much of the likelihood of a firm's failure can be predicted based on what other firms in the same "niche" are doing. To pick just one example, firms in either very empty or very competitive niches (i.e., brand new sectors versus markets that are crowded with other firms) fail more frequently that firms in moderately competitive markets (Carroll & Hannan, 2000). Although originally focused only on for-profit firms, this work has been extended to non-profit and social movement organizations (McPherson, 1983; Soule & King, 2008).

Several pieces of recent research on online communities adopt this type of ecological approach. One important example is Wang, Butler, and Ren (2013) which uses 64 months of longitudinal data from a stratified random sample of 240 Usenet discussion groups to understand how competition for participants' attention affects communities' growth. For each discussion group, the authors created a measure of membership overlap by identifying participants and then measuring the number of other Usenet discussion groups that

ments.

each user participates in during every month. Wang and colleagues show that groups whose members participate in more groups grow more slowly. The relationships predicted by their model are shown in Figure 3. They use group-level variables like community size to unpack this finding and conclude that larger groups experience more difficulty in growth and are more vulnerable to the deleterious effects of competition.

Population-level datasets make it possible to imagine research like that the Wang et al. study. Additionally, taking an ecological perspective allows them to identify substantively important relationships. Their models of community growth rate improve substantially when they take membership overlap into account, explaining up to 14% of the variance overall and returning significantly better fit than alternative specifications without the overlap measure. Although their strongest controls are group-level factors like total membership, Wang et al.'s paper shows that we can make important predictions about how much an online community will grow by considering activities outside communities' boundaries.

An ecological approach to studying online communities remains rare, but other examples exist. Gu, Konana, Rajagopalan, and Chen (2007) considered competition between online investment communities. Two other recent pieces have extended the findings and approach of Wang et al. (2013): Zhu, Chen, et al. (2014) studying a group of 9,495 IBM discussion groups and Zhu, Kraut, and Kittur (2014) in a community of 5,673 wikis. Remarkably, the latter finds that membership overlap improves the survival rate for new wikis. Future work may confirm or diverge from these findings as well as the earlier findings of organizational research that focused exclusively on firms.

Benefit 5: Studying Multilevel Processes

A final benefit of population-level analyses of online communities is the ability to analyze complex, multilevel social processes. This benefit derives partly from the extraordinary granularity and breadth of online trace data, with which there is often no need to sacrifice micro-level detail in the analysis of meso- and macro-level dynamics. However, while multilevel analysis may have an affinity with exhaustive digital trace data, it need not proceed through data-intensive or computational approaches. Insightful qualitative inquiry can also account for multilevel social processes. We believe that research on populations of online communities present an exceptional opportunity in this regard.

Scholars of organizational communication and group processes have implemented multilevel approaches in a variety of domains. Prior work by Faraj and Johnson (2011) advocates the use of multilevel network analysis to examine complex social processes at the individual-level. Indeed, the rise of massive scale online datasets has coincided with the creation of novel multilevel methods, including the use of multilevel network analysis to study (among other things) multi-team systems (Contractor, 2013; Zaccaro, Marks, & DeChurch, 2012). Howard's (2002) concept of "network ethnography" provides a mixed-methods example of this approach that utilizes exhaustive knowledge of macro- and meso-level network structures both to identify appropriate cases for in-depth analysis as well as to guide the process of interpreting and generalizing findings. To date, relatively little research capitalizes on the opportunities to apply multilevel analysis across many online communities. Several studies from organizational research use multilevel approaches to study individual and group-level relational processes across multiple online communities (Faraj & Johnson, 2011), but such work remains exceptional. We think there should be more.

A study of social capital and relationships in World of Warcraft (WoW) guilds by Williams et al. (2006) provides an excellent example of multilevel work that combines the advantages of large-scale trace data with in-depth qualitative evidence. WoW guilds are arbitrarily sized clans of players which vary in size, scope, formality, and strategic focus. A player might, for instance, want to join up with a "raid-oriented family friendly guild" of marauding orcs who meet on weekends. The authors use an exhaustive observational dataset collected with automated programs connecting to different types of WoW game servers over several months to build a representative sample of individual players and guilds.[3] The authors created a stratified sample of these guilds and a representative sample of their members. They then conducted 48 interviews with individuals from the sample over the game's chat platform to understand how members participate in guilds.

Williams and colleagues' elaborate their account of both group-level and individual-level patterns of activity using quotations contextualized through their knowledge of player and guild characteristics. For example, they show how players articulate group-level identities and find very legitimate forms of social support through participation in guilds. One interviewee explained the experience of playing in a large guild with everything but a citation to

[3]Note that additional details of the data collection are published in a companion article by Ducheneaut, Yee, Nickell, and Moore (2006).

Putnam's (2000) influential argument about the putative decline of civic associations in America:

> It's kind of like a bowling team or a softball league...it's just as social event in here, probably more often then [*sic*] bowling since I talk to these people several nights a week (Williams et al., 2006, pg. 347).

The quote illustrates the existence of large-scale solidarity and associationism in the supposedly alienated world of video games. The authors can generalize from this kind of evidence because they used guild-level data to ensure that their findings are representative of experiences in guilds of varying sizes and characteristics. The findings integrate the richness and depth of insightful interviewing with the analytic leverage of a research design that can generalize across multiple communities.

An important takeaway is that individuals' experiences online often occur within the context of sub-communities like guilds. As a result, understanding whether and how WoW players build sustained social relationships and social capital entails accounting for a complex interplay of both factors at the game, guild, and individual levels. In this way, Williams and colleagues' study illustrates the value of multilevel modeling strategies as a tool for incorporating measures nested across levels of analysis and social interaction in online communities that are, themselves, part of larger communities. Although their study is qualitative in nature, Williams has collaborated with others on quantitative studies using the same dataset in similar ways (e.g., Shen, Monge, & Williams, 2014).

Multilevel analyses can also account for ecological dynamics in communities' environments and look across non-nested communities. For instance, Faraj and Johnson (2011) test claims about the prevalence of patterns of interpersonal exchange using exhaustive micro-level data collected from a random sample of discussion groups. They draw conclusions at the individual and group levels while accounting for community-level variation as well. Studying interpersonal communication processes in multiple online communities by simultaneously analyzing multiple levels of social organization and behavior offers a particularly promising path for future work.

Limitations

Alongside these benefits, population-level research on online communities introduces a number of limitations. Some of these limitations are inherent to research on populations and others derive from particular characteristics of populations of online communities. First, multiple populations of communities can interact in ways that suggest an even higher level of analysis that may bring additional insights. Second, identifying distinct communities poses difficulties in many situations. Third, research in this area exists in the interstices of multiple fields, leading to challenges related to integrating disparate literatures and concepts. Finally, it is important to acknowledge that studying populations of communities can require more effort and can be more skill and resource-intensive than studying a community in isolation.

While studies on populations of communities may provide insights and generalizability that exceed smaller scale comparisons or case studies, it is important to note that a population of communities (as we described it above) may be distinct from still larger sets of populations, known as a metapopulation in ecology research (Hanski, 1999). Along these lines, organizational sociologists have suggested that studying organizations within industries ignores important high-level effects across industries and that it is also important to study populations of industries – literally, populations of populations (e.g. Ruef, 2000). In the context of online communities, examples of populations and corresponding metapopulations might look like all of the communities on a particular platform or hosting service (e.g., Wikia, GitHub, or EdX) and the full set of potentially comparable communities (e.g., all wikis, FLOSS projects, or MOOCs). There may also be other dimensions along which to define metapopulations, such as the sets of communities encompassed by particular language groups, legal regimes, or infrastructural features.

The distinction between populations and metapopulations implies new types of limitations of generalizability analogous to those confronted by single case studies. While a study may include all of the software development communities hosted by GitHub or SourceForge, or even both, there may be underlying biases that make these two groups systematically different from other otherwise similar communities. There may also be ecological forces that reflect competition between the different populations. As usual, unknown bias has no empirical remedy other than an effort to encompass more cases – in this case more populations. Likewise, the identification of ecological forces at a metapopulation level remain inaccessible to studies within a single population. These possibilities suggest future opportunities for research.

A second issue relates to the ways in which the nature and character of online communities resist easy classification or definition. A single community, such as the English language Wikipedia, may contain hundreds or even thousands of distinct sub-communities. In Wikipedia, these include WikiProjects – groups of individuals who work together on topics of shared interest (Morgan, Gilbert, McDonald, & Zachry, 2013). Like guilds in WoW, WikiProjects are online communities nested in English Wikipedia (for example), which is nested within the broader Wikipedia project, which is nested within the even broader Wikimedia movement. Such a fractal structure is not unique to Wikipedia and may represent a broader characteristic of large-scale, open organizations on the Internet.

Given these examples, language and theories developed to analyze more clearly bounded formal organizations such as firms, non-profits, and even more open volunteer projects or social movements may prove inadequate. Even the term "community" is contested and is used in a wide variety of ways by researchers studying online interaction (Bruckman, 2006). This disagreement about what constitutes a "community" makes the definition of a population (or metapopulation) of online communities analytically slippery as well. Research on organizations and collective action has, historically, elided this concern by adopting a very broad set of inclusion criteria (e.g., Marwell & Oliver, 1993). Ultimately though, if every sort of remotely collective endeavor can be also called a "community," the term ceases to hold meaning.

The murky empirical and conceptual boundaries around online communities make defining populations difficult. We advocate a pragmatic, but ultimately somewhat unsatisfying, approach to this issue: researchers should define populations by a combination of objectively observable criteria along with folk and scientific wisdom. Sometimes, as in studies examining all of the communities hosted through a common software platform or company, this may be straightforward and the results may correspond to both scientific intuition as well as the understanding of participants within these populations. Under other circumstances, it may be necessary to choose between different possible analytic boundaries. For instance, in analyses of emergent collaboration networks such as those engaging in breaking news coverage on microblogging platforms, Wikipedia, or Facebook, groups of participants may act in ways that are consistent with researchers' understandings of what it means to be a community even though they do not perceive themselves as such. In these situations, it becomes critical to acknowledge such divergent perspectives and proceed with caution. It may not make sense to address certain theoretically interesting questions with data drawn from unsuitable cases. Just because we can model something as an community does not mean we should. Relatedly, just because a platform describes something as a group (a Facebook group, a guild within a game, a WikiProject) doesn't mean it is the correct unit of analysis for a particular question

A third challenge is that existing research conducting population-level analyses of online communities spans disciplinary boundaries. Our own work has been influenced by organizational studies, social psychology, political sociology, interpersonal communication, and human computer interaction. Such diverse intellectual orientations may create opportunities for academic brokerage, but these opportunities come at a cost. As in finance (Zuckerman, 1999), academic endeavors that fail to fit into existing categorical schema may pay an illegitimacy premium. There are also opportunity costs that come with learning multiple research fields. The institutionalization of novel schema takes time and proceeds unevenly.

We believe the best response to this problem is to draw strategically from the most compelling domains of research. For example, the papers we have highlighted in this essay have drawn on the literature on organizational behavior, work groups, computer science, and teams. We have also turned to the sociological literature focused on comparative analysis of movements. Others have emphasized links to studies of emergent networks and collective intelligence. All of these approaches have borne fruitful results. Of course, this type of pluralistic approach introduces challenges as well. While reading broadly, it is important to ground one's work in established traditions of scholarship. A broad set of influences is not a license to cherry-pick theories to fit data. An open challenge lies in the synthesis and integration of findings across disparate approaches. For the time being, many findings remain mutually contradictory or unintelligible as they speak of similar phenomena in different ways.

A final limitation is purely practical but still important to note. Simply put, studying multiple communities can require more time, more effort, more skills, or more resources than a study of a single community. For one, population-level datasets can be large and unwieldy and building and analyzing them may be outside the abilities of many communication researchers. The dataset used by Graeff et al. (2014) was created by professional engineering staff. Many population-level analyses, like Leskovec et al. (2009), were carried out by computer scientists. When using these large and varied datasets, there are often trade-offs between scope and scale. In the short term, collaborations with computer scientists and engineers is a popular strategy. In the longer term, population-level studies will become more widespread as the technology necessary to complete them becomes more established, easier-to-use, and more reliable. For example, the deployment of standardized APIs and large data releases from large community-hosting platforms has already made some types of population-level studies much more accessible. We expect that this will only increase with time.

Discussion

In summary, we have argued that online communities research can benefit enormously from studying populations of communities. We have tried to show how population-level research designs that span communities can offer increased generalizability and can open the door to new kinds of questions including those that focus on theoretically important community-level variables, processes of diffusion across communities, and the way that communities interact with their environments through ecological competition. Finally, we've argued that digital trace datasets and a variety of analyses that cross multiple levels can allow research to realize many of these benefits without tossing aside the benefits of intra-community studies. By narrating this argument with a series of in-depth examples of exemplary work from communication and beyond, we have also pointed to concrete examples of how this can be done. Although these studies still reflect a small proportion of research on online communities, similar approaches are increasingly common.

Additional benefits from population-level research about online communities may emerge as researchers experiment and explore this type of work. One benefit we find particularly exciting concerns the degree to which analysis across many online communities may provide stronger empirical support for policy-making and design decisions. These decisions can benefit from insights into the effects of specific interventions. This sort of evidentiary basis holds particular relevance for computer-mediated communication systems where code may shape *de facto* institutional arrangements, norms, and behavioral patterns (Lessig, 1999). For example, in our own work we used a population of wikis to explore the effects of a requirement to create an account on subsequent community-level activity. Estimates of an average effects of a widely debated policy decision like this, across many projects, can give designers and community leaders greater confidence in the generalizability of a finding in a way that can inform subsequent policy decisions. Exploration of heterogeneous effects can provide an understanding of how a design might succeed or fail.

In ways that we discussed earlier, many studies of online communities and networked communication seek to understand the mediating effects of specific technologies. Since the design of technical platforms and tools, in most cases, operate at the level of communities, a shift to empirical analysis and theorizing at the level of communities can make this work more directly usable. There are already significant overlaps in the scholarly communities studying networked communication and the designers creating social computing systems. Crafting research capable of speaking across these boundaries will enhance the impact of communication research and will provide communication researchers with greater opportunities to test, extend, and refine their theories in conversation with designers and computer scientists.

Of course, while we see many opportunities in population-level research, the approach is far from a silver bullet. The limitations we have sketched out are real, significant, and a subset of the challenges that confront population-level analysis. Just as we hope that continued growth of population-level studies will help establish benefits of the approach, we also hope it will paint a better picture of the limitations of population-based approaches and support the development of better ones. We believe that an increased attention to populations marks one step toward better online community research and toward a deeper understanding of networked communication.

Acknowledgements

We thank the Community Data Science Collective for helping us develop these ideas. We thank the editors, Brooke Foucault Welles and Sandra Gonzalez Bailon, and reviewers of this chapter, for their helpful feedback. Portions of the project were completed at the Helen R. Whitely Center at the University of Washington's Friday Harbor Laboratories. Financial support for this work came from the National Science Foundation (grants IIS-1617129 and IIS-1617468).

References

Benkler, Y., Roberts, H., Faris, R., Solow-Niederman, A., & Etling, B. (2013). *Social Mobilization and the Networked Public Sphere: Mapping the SOPA-PIPA Debate* (SSRN Scholarly Paper No. ID 2295953). Social Science Research Network. Rochester, NY. Retrieved September 22, 2013, from http://papers.ssrn.com/abstract=2295953

Benkler, Y., Shaw, A., & Hill, B. M. (2015). Peer Production: A Form of Collective Intelligence. In M. Bernstein & T. Malone (Eds.), *The Handbook of Collective Intelligence*. Cambridge, MA: MIT Press.

Bennett, W. L., & Segerberg, A. (2013). *The logic of connective action: Digital media and the personalization of contentious politics*. New York, New York: Cambridge University Press.

Bruckman, A. (2006). A new perspective on "community" and its implications for computer-mediated communication systems. In *CHI '06 extended abstracts on Human factors in computing systems* (pp. 616–621). doi:10.1145/1125451.1125579

Bruns, A., & Stieglitz, S. (2012, July). Quantitative Approaches to Comparing Communication Patterns on Twitter. *Journal of Technology in Human Services*, *30*, 160–185. 00083. doi:10.1080/15228835.2012.744249

Carroll, G. R., & Hannan, M. T. (2000). *The demography of corporations and industries*. Princeton, New Jersey: Princeton University Press.

Contractor, N. (2013, March 28). Some assembly required: Leveraging Web science to understand and enable team assembly. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *371*(1987), 20120385. doi:10.1098/rsta.2012.0385. pmid: 23419854

Crowston, K., & Howison, J. (2006, December). Hierarchy and centralization in free and open source software team communications. *Knowledge, Technology & Policy*, *18*(4), 65–85. doi:10.1007/s12130-006-1004-8

Crowston, K., Wei, K., Howison, J., & Wiggins, A. (2008, March). Free/Libre Open-source Software Development: What We Know and What We Do Not Know. *ACM Comput. Surv.* *44*(2), 7:1–7:35. 00212. doi:10.1145/2089125.2089127

Cummings, J. N., & Cross, R. (2003). Structural properties of work groups and their consequences for performance. *Social Networks*, *25*(3), 197–210. Retrieved June 8, 2015, from http://www.sciencedirect.com/science/article/pii/S0378873302000497

DiMaggio, P. J., & Powell, W. W. (1983, April). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147–160. JSTOR: 2095101

Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2006, January 10). Building an MMO With Mass Appeal A Look at Gameplay in World of Warcraft. *Games and Culture*, *1*(4), 281–317. doi:10.1177/1555412006292613

Faraj, S., & Johnson, S. L. (2011). Network Exchange Patterns in Online Communities. *Organization Science*, *22*(6), 1464–1480. doi:10.1287/orsc.1100.0600

Fields, S., & Johnston, M. (2005, March 25). Whither Model Organism Research? *Science*, *307*(5717), 1885–1886. doi:10.1126/science.1108872. pmid: 15790833

Geiger, R. S., & Halfaker, A. (2013). When the Levee Breaks: Without Bots, What Happens to Wikipedia's Quality Control Processes? In *Proceedings of the 9th International Symposium on Open Collaboration* (6:1–6:6). WikiSym '13. 00010. doi:10.1145/2491055.2491061

Graeff, E., Stempeck, M., & Zuckerman, E. (2014, January). The battle for 'Trayvon Martin': Mapping a media controversy online and off-line. *First Monday*, *19*(2). Retrieved January 27, 2016, from http://firstmonday.org/ojs/index.php/fm/article/view/4947

Gu, B., Konana, P., Rajagopalan, B., & Chen, H.-W. M. (2007). Competition among virtual communities and user valuation: The case of investing-related communities. *Information Systems Research*, *18*(1), 68–85. JSTOR: 23211832

Halfaker, A., Geiger, R. S., Morgan, J. T., & Riedl, J. (2013, May 1). The Rise and Decline of an Open Collaboration System How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist*, *57*(5), 664–688. doi:10.1177/0002764212469365

Halfaker, A., Kittur, A., & Riedl, J. (2011). 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* (pp. 163–172). WikiSym '11. doi:10.1145/2038558.2038585

Hannan, M. T., & Freeman, J. (1993). *Organizational Ecology*. Cambridge, MA: Harvard University Press.

Hanski, I. (1999, March 18). *Metapopulation Ecology*. 03605. Oxford, UK: Oxford University Press.

Hargittai, E. (2015, May 1). Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 63–76. doi:10.1177/0002716215570866

Hill, B. M. (2013). Almost Wikipedia: What eight early online collaborative encyclopedia projects reveal about the mechanisms of collective action. In *Essays on volunteer mobilization in peer production*. PhD Dissertation. Cambridge, Massachusetts: Massachusetts Institute of Technology.

Howard, P. N. (2002, January 12). Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods. *New Media & Society*, *4*(4), 550–574. doi:10.1177/146144402321466813

King, G., Keohane, R. O., & Verba, S. (1994, May 2). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Kollock, P., & Smith, M. (1999, February 10). *Communities in Cyberspace*. London: Routledge.

Kraut, R. E., & Resnick, P. (2012, March 23). *Building Successful Online Communities: Evidence-Based Social Design* (). Cambridge, MA: The MIT Press.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., ... Van Alstyne, M. (2009, February 6). Computational Social Science. *Science*, *323*(5915), 721–723. doi:10.1126/science.1167742

Leskovec, J., Backstrom, L., & Kleinberg, J. (2009). Meme-tracking and the Dynamics of the News Cycle. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 497–506). KDD '09. doi:10.1145/1557019.1557077

Lessig, L. (1999). *Code and Other Laws of Cyberspace*. New York, NY: Basic Books.

Marwell, G., & Oliver, P. (1993). *The critical mass in collective action: A micro-social theory*. Cambridge, Massachusetts: Cambridge University Press.

McPherson, M. (1983). An ecology of affiliation. *American Sociological Review*, *48*(4), 519–532. doi:10.2307/2117719. JSTOR: 2117719

Meraz, S., & Papacharissi, Z. (2013, January 27). Networked Gatekeeping and Networked Framing on #Egypt. *The International Journal of Press/Politics*, 138–166. doi:10.1177/1940161212474472

Meyer, J. W., & Rowan, B. (1977, September). Institutionalized organizations: Formal structure as myth and ceremony. *The American Journal of Sociology*, *83*(2), 340–363. 20516. JSTOR: 2778293

Mockus, A. (2007, May). Large-Scale Code Reuse in Open Source Software. In *First International Workshop on Emerging Trends in FLOSS Research and Development, 2007. FLOSS '07* (pp. 7–7). First international workshop on emerging trends in floss research and development, 2007. floss '07. doi:10.1109/FLOSS.2007.10

Morgan, J. T., Gilbert, M., McDonald, D. W., & Zachry, M. (2013). Project Talk: Coordination Work and Group Membership in WikiProjects. In *Proceedings of the 9th International Symposium on Open Collaboration* (3:1–3:10). WikiSym '13. doi:10.1145/2491055.2491058

Ortega, F. (2009). *Wikipedia: A Quantitative Analysis* (Ph.D. Dissertation, Universidad Rey Juan Carlos, Madrid, Spain). Retrieved June 22, 2009, from http://libresoft.es/Members/jfelipe/phd-thesis

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. New York, NY: Cambridge University Press.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York, NY: Simon and Schuster.

Restivo, M., & van de Rijt, A. (2012). Experimental Study of Informal Rewards in Peer Production. *PLoS ONE*, *7*(3), e34358. doi:10.1371/journal.pone.0034358

Restivo, M., & van de Rijt, A. (2014). No praise without effort: Experimental evidence on how rewards affect Wikipedia's contributor community. *Information, Communication & Society*, *17*(4), 1–12. doi:10.1080/1369118X.2014.888459

Rogers, E. M. (1962). *Diffusion of Innovations*. New York, NY: The Free Press of Glencoe.

Ruef, M. (2000, November 1). The emergence of organizational forms: A community ecology approach. *American Journal of Sociology*, *106*(3), 658–714. doi:10.1086/318963

Schoonhoven, C. B. (2015). Liability of Newness. In *Wiley Encyclopedia of Management*. John Wiley & Sons, Ltd. Retrieved January 30, 2016, from http://onlinelibrary.wiley.com/doi/10.1002/9781118785317.weom030067/abstract

Schweik, C. M., & English, R. C. (2012). *Internet Success: A Study of Open-source Software Commons*. Cambridge, MA: MIT Press.

Scott, W. R., & Davis, G. F. (2006). *Organizations and organizing: Rational, natural and open systems perspectives*. Upper Saddle River, NJ: Pearson Prentice Hall.

Seneviratne, O., Kagal, L., Weitzner, D., Abelson, H., Berners-Lee, T., & Shadbolt, N. (2009). Detecting creative commons license violations on images on the world wide web. In *Proceedings of the 18th Internetional World Wide Web Conference*. Www2009, Madrid, Spain. Retrieved January 31, 2016, from http://dig.csail.mit.edu/2008/Papers/WWW2009/paper.pdf

Shaw, A., & Benkler, Y. (2012). A Tale of Two Blogospheres: Discursive Practices on the Left and Right. *American Behavioral Scientist*, *56*(4), 459–487. doi:10.1177/0002764211433793

Shen, C., Monge, P., & Williams, D. (2014, October 1). The evolution of social ties online: A longitudinal study in a massively multiplayer online game. *Journal of the Association for Information Science and Technology*, *65*(10), 2127–2137. doi:10.1002/asi.23129

Soule, S. A., & King, B. G. (2008, May). Competition and resource partitioning in three social movement industries. *The American Journal of Sociology*, *113*(6), 1568–1610. doi:10.1086/587152. JSTOR: 25145846

Valente, T. W. (1995, January 10). *Network Models of the Diffusion of Innovations*. Cresskill, N.J: Hampton Press.

van de Rijt, A., Kang, S. M., Restivo, M., & Patil, A. (2014, May 13). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, *111*(19), 6934–6939. doi:10.1073/pnas.1316836111. pmid: 24778230

Wang, X., Butler, B. S., & Ren, Y. (2013). The impact of membership overlap on growth: An ecological competition view of online groups. *Organization Science*, *24*(2), 414–431. doi:10.1287/orsc.1120.0756

Weber, M. S., Fulk, J., & Monge, P. (2016, February 11). The Emergence and Evolution of Social Networking Sites as an Organizational Form. *Management Communication Quarterly*, 0893318916629547. doi:10.1177/0893318916629547

Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., & Nickell, E. (2006, January 10). From Tree House to Barracks The Social Life of Guilds in World of Warcraft. *Games and Culture*, *1*(4), 338–361. doi:10.1177/1555412006292616

Zaccaro, S. J., Marks, M. A., & DeChurch, L. (Eds.). (2012, May 22). *Multiteam Systems: An Organization Form for Dynamic and Complex Environments*. Routledge.

Zhu, H., Chen, J., Matthews, T., Pal, A., Badenes, H., & Kraut, R. E. (2014). Selecting an effective niche: An Ecological view of the success of online communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 301–310). CHI '14. doi:10.1145/2556288.2557348

Zhu, H., Kraut, R. E., & Kittur, A. (2014). The impact of membership overlap on the survival of online communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 281–290). CHI '14. doi:10.1145/2556288.2557213

Zhu, H., Kraut, R. E., & Kittur, A. (2016). A Contingency View of Transferring and Adapting Best Practices within Online Communities. In *Proceedings of the Conference on Computer Supported Cooperative Work and social Computing*. Cscw 2016, San Francisco, CA: ACM Press. Retrieved January 29, 2016, from http://haiyizhu.com/wp-content/uploads/2015/09/BestPracticeTransfer.pdf

Zuckerman, E. W. (1999, March). The categorical imperative: Securities analysts and the illegitimacy discount. *The American Journal of Sociology*, *104*(5), 1398–1438. JSTOR: 2990940