DOCTORAL DISSERTATION

Essays on Socialization in Online Groups

submitted to the
David A. Tepper School of Business
in partial fulfillment for the requirements for the degree of

DOCTOR OF PHILOSOPHY
in Industrial Administration (Information Systems)

Bo Reum Choi
Carnegie Mellon University
Tepper School of Business
Pittsburgh, Pennsylvania 15213

DISSERTATION COMMITTEE:

Robert E. Kraut (Chair)
Mark Fichman
Param Vir Singh
Laura Dabbish
John M. Levine

December 2012

Abstract

One challenge of online groups is helping new members adjust to their environment. This adjustment process has been referred to as socialization, or the process by which newcomers make the transition from being organizational outsiders to being insiders. During the adjustment process, new members moved from peripheral to full participation, with their goals, tools, and perceptions of the community changed. However, there has been little empirical research on how socialization is accomplished in online groups and on its effectiveness. The goal of this thesis is to identify what kinds of socialization tactics are used in online groups, how different tactics generate more productive and committed newcomers, and how these change with individuals' tenure and their proactive behaviors involved in the socialization process. The first study relates to the investigation phase of group socialization whereby groups attempt to recruit appropriate people and individuals assess groups. This study suggests and tests how recruitment and assessment techniques affect both assessment quality and turnover from the newcomer's and group's point of view. The second study examines the impact of various socialization practices on socialization outcome once newcomers join the group. This study investigates what types of socialization tactics are used in online groups and which tactic is effective to increases newcomers' commitment to online groups. The third study aims to identify the role of individuals involved in the socialization process. It examines the moderating impact of members' tenure and proactive behaviors on the relationship between the types of socialization tactics and member commitment by using the automated measurement method.

Acknowledgement

My life as a doctoral student at Carnegie Mellon has been a time of fantastic intellectual growth, fostered by a wonderful learning environment. I am lucky to have met and worked with so many talented people. Writing this dissertation would not have been possible without the support of my family, friends, colleagues, and mentors.

At Carnegie Mellon, the first person that deserves my gratitude is my advisor, Bob Kraut. There have been numerous occasions where I remember feeling discouraged and puzzled about the direction of my research, but inevitably, a meeting with Bob would reinvigorate my enthusiasm. He was willing to let me explore my own path, but was always there when I got lost. Bob's steadfast support and intellectual guidance have been invaluable over the years and I feel incredibly privileged to have him as my advisor.

I also thank my committee members, Mark Fichman, John Levine, Laura Dabbish, and Param Vir Singh, who have along the way provided valuable comments and support that have made both this thesis stronger and me a better researcher. Without their sharp insights and feedback, my work would have greatly suffered.

My journey would also not have been possible without the support of colleagues and friends. I would like to thank the Social Computing Lab and fellow Tepper and HCII students and staff for countless, pleasant memories of Pittsburgh. Many thanks in particular to Kira Alexander, Bryon Baliant, Moira Burke, Yeon Woo Chang, Soo Haeng Cho, Ben Collier, Ari Kang, Jungpil Hahn, Elina Hwang, Joon Young Kawk, Byung Cho Kim, Dong Whan Kim, Jinwoo Kim, Seung Hyun Kim, Soo Hyung Kim, Niki Kittur, Tat Koon Koh, Inseong Lee, Juen Lee, Sun Young Park, Lawrence Rapp, Jiwoo Suh, Kyung Chan Sul, Yi-Chia Wang, and Haiyi Zhu.

I have been blessed with a very loving and supportive family. I have learned so much from my parents, Ki Sup Choi and Yeon Hong Kim. I appreciate all of the sacrifices that they have made for me. They are the best parents I could ask for and they, along with my sister Bo Seul Choi, ceaselessly cheered me along the way. I also owe my gratitude to Young Gil Yoo, my father-in-law, Kyung Jae Lee, my mother-in-law, and Ga Hyo Yoo, my sister-in-law. They showed me unwavering trust and support. It is great luck for me to have them as my family.

The last two people I want to thank are my beloved husband and son, Jung Chool Yoo and Jae Yeon Yoo. Jungchool has been a great partner in my life. Even though we often were thousands of miles apart, Jungchool has supported me with great patience. When I need someone to turn to, he was always there. Jaeyeon is a wonderful kid. His bright little smile made me enjoy the life. I need to thank my little boy for being such a bundle of joy and laughter. I hope he always knows how grateful I am for him, and how much I love him.

Table of Contents

Chapter 1

1. Introduction

Online groups provide a place to exchange useful information and social support, meet people, maintain social networks, discuss political and social issues, and entertain. Online groups have been growing rapidly and becoming increasingly important, which produces considerable revenue. For example, the total number of subscribers to World of Warcraft, one of the most popular online games, is 14 million, which is more than the population of New York City. There are more than 5,000 semi-permanent guilds in World of Warcraft. The players in the U.S. and Europe alone produce $800 million revenue per year (World of Warcraft Statistics) through subscription fees. Starting in 2001, English Wikipedia now has more than 4 million articles and 23 million Wikipages, including discussion pages and personal user pages. Fourteen million users participate in Wikipedia, including 650,000 Wikipedians who edited at least 10 times. In addition, more than 2.000 WikiProjects are subgroups in Wikipedia and include collections of editors interested in improving the coverage and quality of articles in a particular domain.

However, success of online groups is limited because they often fail to encourage people to contribute over the long term. For example, 68% of newcomers to Usenet groups never post anything after their first post (Arguello, et al., 2006), and a quarter of the members of guilds in the massive multiplayer game World of Warcraft leave their guilds every month even though they are still playing the game (Ducheneaut, et al., 2007). According to the popular open source portal, SourceForge (http://sourceforge.net/), most OSS projects have ended in failure: 58% do not move beyond the alpha developmental stage, 22% remain in the planning phase, 17% remain in the pre-alpha phase, and some become inactive (Lee, et al., 2009). Inequities in contribution

are extreme, with a small fraction of members doing the vast majority of contribution. In Wikipedia, for example, more than 90% of all edits are made by the top 15% of editors.

Given high turnover rates and lack of contribution by most group members, online groups need to select appropriate members and help them to adjust to their new communication environment. Groups, both conventional and online ones, try to select individuals with a good fit to them (Jones, et al., 2000; Stevens & Campion, 1999) and to train their members effectively for survival (Ahuja & Galvin, 2003; Burke, et al., 2009; Ducheneaut, 2005; Wang, et al., Under Review). Similarly, individuals want to join a proper group and learn their roles and acclimatize to the groups to fulfill their needs. This interactive and dynamic investigation, adjustment, and adaptation process between groups and individuals has been referred to as socialization, or the process by which newcomers make the transition from being outsiders to being insiders (Bauer, et al., 2007; Levine, et al., 2005; Morrison, 1993). Socialization is fundamental to online groups and their members because it helps ensure the continuity of central values and gives the members a framework for responding to events in their groups (Bauer, et al., 1998; Jones, 1986; Van Maanen & Schein, 1979).

While all groups can have both opportunities and problems in assessing, selecting, socializing, and retaining members, some of these opportunities and challenges are particularly frequent in online groups. For example, because online groups do not need to select members based on geographic proximity, they have access to a larger pool of possible members. The ease with which people can manipulate self presentation online, the relatively impoverished information available for person perception, and the relatively weak interpersonal ties between members in many online groups may make it more difficult to recruit appropriate people than in face-to-face groups (Tidwell & Walther, 2002). In addition, unlike conventional offline groups, it

is easy for online group members to leave their groups (Burke, et al., 2009) because online groups comprise voluntary members who are not bound to their groups through an employment contract. This lack of commitment leads many people to leave their groups if they are subjected to enforced training. In addition, lack of an organizational hierarchy makes it difficult for online groups to recruit and train mentors.

In spite of its importance, research on socialization in online groups is still in its early stages. While existing research has accumulated much knowledge on why people contribute to online groups (Bock, et al., 2005; Jeppesen & Frederiksen, 2006; Ma & Agarwal, 2007; Roberts, et al., 2006a; Wasko & Faraj, 2005), little work has explored how to select appropriate individuals before entry and socialize them after entry to increase their commitment to the group. In addition, prior research has emphasized the importance of socialization in online groups in general (Bryant, et al., 2005; Ducheneaut, 2005), and some studies have identified the positive impact of receiving responses on individuals' early contributions to a group (Moon & Sproull, 2008; Wang, et al., Under Review). Still, the impact of various socialization practices on the socialization outcome remains an open question.

Moreover, individuals who play an important role in socialization process have not received much attention in most studies. Socialization literature identified that socialization by supervisors has a great impact on their members' commitment to the groups (Bravo, et al., 2003). Because online groups rarely have a vertical hierarchy and peers with more or less experiences socialize others as a substitute for supervisors, we expect that members' tenure impacts the effectiveness of socialization in online groups. In addition, individuals are not just passively socialized but they proactively shape their own socialization experiences (Miller & Jablin, 1991; Morrison, 1993; Wanberg & Kammeyer-Mueller, 2000). Thus, individuals' proactive behaviors

will moderate the linkage between socialization and the outcomes. However, very little is known about the effect of individuals' tenure and proactive behaviors on the relationship between socialization tactics and outcomes both in online and offline groups.

Given the scarcity of research on socialization in online groups, research on specific socialization practices and individuals' roles that lead to successful relationships between online groups and their members will have important theoretical and practical implications. Thus, the goal of this thesis is to identify what kinds of socialization tactics are used in online groups, how different tactics generate more productive and committed newcomers, and how these change with individuals' tenure and their proactive behaviors involved in the socialization process. Socialization research in traditional organizations is classified into two discrete but related categories (Klein & Heuser, 2008). The first category examines the general approach organizations adopt in structuring the socialization experiences of newcomers (Bauer, et al., 2007; Jones, 1986; Saks, et al., 2007; Van Maanen & Schein, 1979). The second category focuses on the role of specific socialization activities in facilitating newcomer adjustment (Holladay, et al., 2006; Klein & Weaver, 2000; Wesson & Gogus, 2005). Relatively few studies were located for this category. This thesis focuses on the second approach and aims to conduct exploratory and theory-building research to better understand what actually goes on in online groups to select and socialize their members and how these practices work.

This thesis comprises three empirical studies, found in Chapters 2–4. The first study in Chapter 2 relates to the phase whereby groups attempt to recruit appropriate people and individuals assess groups before individuals join their groups. This study aims to explain how the modality of these recruitment and assessment techniques affects both the assessment quality and turnover from the individual's and group's point of view. The research context of this study is

guilds in World of Warcraft, semi-permanent groups in one of the most popular massive multiplayer online games, where groups and individuals actively recruit and select each other.

The second study in Chapter 3 relates to the phase whereby groups try to shape the individuals to make them contribute more to the accomplishment of group goals and individuals attempt to learn about the group and assimilate themselves into it after individuals join their groups. This study explores what types of socialization tactics are used in online groups after newcomers join a group and examines which tactic is effective in increasing newcomers' commitment to online groups. The research context of this study is WikiProjects, subgroups in Wikipedia, where the process and the outcomes of socialization are observable.

Lastly, the third study in Chapter 4 is the follow-up study of the second study. This study extends the previous study by examining the impacts of socialization tactics for experienced members as well as newcomers. This study also aims to contribute to a developing research literature that reveals the important role of individuals in the socialization process in online groups. It tests the moderating impact of members' tenure and proactive behaviors on the relationship between the types of socialization tactics and member commitment by using the automated measurement method (e.g., (Ashford & Black, 1996; Bauer & Green, 1998; Griffin, et al., 2000; Morrison, 1993).

Chapter 2

Matching individuals and virtual groups : The role of assessment

2.1. Introduction

Virtual groups expand an organization's ability to link together resources separated in space and time (Timmerman & Scott, 2006). A virtual group is defined as a group of people who interact on interdependent tasks with a shared purpose across space and time using technology (Lipnack & Stamps, 2000). Virtual groups have much less face-to-face contact than conventional groups and are able to collaborate using computer and communication technologies.

One challenge for virtual groups is selecting appropriate group members. Groups, both conventional and virtual ones, are more successful if they select appropriate members (Jones, et al., 2000; Stevens & Campion, 1999) . Similarly, individuals are more satisfied, have less turnover and work better if they join appropriate groups (Mitchell, et al., 2001). Theoretical and empirical research suggests that individuals and organizations are most effective when their values, needs, and interests are matched. This match is called person-organization (P-O) fit and it influences on commitment, satisfaction, and retention (Chatman, 1991; Meglino, et al., 1989; O'Reilly, et al., 1991), organizational performance (Govindarajan, 1989; Meglino, et al., 1989), and individual health (Moos, 1987). The attraction-selection-attrition (ASA) model implies that fit lead to organization homogeneity (Schneider, 1987). According to this model, people are

more attracted to careers/tasks/majors as a function of their own interests and personality (Holland, 1985), and organizations composed of others similar to themselves (attraction). Groups prefer and select new people who are similar to current members (selection). People who do not fit into the group either leave on their own or are encouraged to leave (attrition), thereby producing a more homogenous group.

While all groups can have both opportunities and problems in recruiting, assessing, selecting, and retaining members, some of these opportunities and challenges are particularly frequent in virtual groups. For example, because virtual groups do not need to select members based on geographic proximity, they have access to a larger pool when recruiting and selecting members. On the other hand, the ease with which people can manipulate how they present themselves online, the relatively impoverished information available for person perception and the relatively weak interpersonal ties between members in many virtual groups may make it more problematic to recruit appropriate people than in face-to-face groups. Given the importance of membership selection, there has been surprising little research investigating the recruiting, selection and attrition cycle, starting with assessment methods and ending with outcomes. Investigation of membership selection in virtual groups is even rarer. Some studies investigate how groups are formed and how individuals make decisions about which groups to join in online and distributed groups. Individuals are more likely to join an open source software development project when they have strong collaborative ties with its initiator (Hahn, et al., 2008). Task skills seem to be the most important in selection decisions for distributed groups due to the lack of physical proximity and visibility, whereas personal characteristics such as gender and race are more important for face-to-face group selection (D'Souza & Colarelli, 2010).

However, these studies do not identify which specific techniques for assessing virtual groups and individuals will help both sides achieve a better fit between the virtual group and new group members. Research evaluating virtual group selection has focused on web-recruiting using *credentials* (D. G. Allen, et al., 2007; Bartram, 2001; Dineen, et al., 2002), but has ignored other methods of selection, such as *brief interactions* between the virtual groups and new recruits (e.g. interviews via Internet), *probationary periods* which give recruits and the group experience with each other, and *referrals* from existing members. Moreover, groups and individuals often do not use the available methods optimally (Rynes, et al., 2002) and there is no apparent consensus about which source or method is most effective (Zottoli & Wanous, 2000). Thus, the field needs to develop and test theory about the assessment methods that reliably and validly measure diverse attributes of individuals and groups, and how the knowledge gained through these assessment methods leads to high-quality outcomes, such as fit with the virtual group and member retention.

The usefulness of these techniques will probably depend upon the attributes the group or individual is attempting to assess. Here we rely on Polanyi's classic distinction between tacit and explicit knowledge (Polanyi, 1961), as elaborated in research on learning, groups and organizations (e.g., Lam, 2000; Reber, 1989; Von Krogh, et al., 2000). Tacit knowledge is intuitive and unarticulated, difficult to formalize and communicate, while explicit knowledge can be codified and abstracted. Whether knowledge is tacit or explicit depends, in part, on inherit qualities of a thing being described, but also upon standard techniques used to measure it. Thus, in everyday life, accurate impressions of people are often based on hard-to-articulate, intuitive judgments formed very quickly on the basis of little information (Ambady & Rosenthal, 1992). However, explicit personality tests (e.g., John & Srivastava, 1999) and behavioral measurements

(Pentland & Pentland, 2008) can codify the basis of some of these intuitive judgments and make them explicit.

In the case of groups selecting members and prospective members choosing groups, some relevant attributes are tacit; they are hard to identify and quantify (Ancori, et al., 2000; Cowan & Foray, 1997). Oftentimes social assessments involve tacit information. For example, they involve understanding a group's culture and values or an individual's habits, personality, and social skills. Subtle task assessment may also involve tacit information. For example, understanding an individual's reliability or the way they approach problems involves assessing tacit information. This type of information is difficult to assess effectively in a short period of time. On the other hand, some of the relevant information can be made explicit, like a candidate's employment history, years of schooling or class rank. It can be articulated, codified, and easily transmitted (Ancori, et al., 2000; Cowan & Foray, 1997). Explicit information is often associated with task-related attributes such as experience and skill, and can be expressed formally, often in quantitative summaries. It does not require personal contact between the assessor and target and is easy to verify.

Both tacit and explicit information are needed to achieve a good fit between virtual groups and potential members. However, tacit information may be even harder to measure in virtual groups than in face-to-face ones because it is mostly transmitted through in-person contact (Ahuja & Galvin, 2003; Finholt, et al., 1991). Thus, assessment methods based on direct and long-term observations may be better suited to evaluate tacit information about virtual groups and individuals. On the other hand, most explicit information is publicly accessible and easy to validate in virtual environments (DeSanctis & Monge, 2000). Thus, shorter-term and indirect assessment methods can be efficient and effective when gathering explicit information.

Choosing an appropriate technique presents a strategic opportunity for enhancing competitive advantage because virtual groups that better attract, select, and retain the proper people should outperform those that do not (Barney & Wright, 1998).

In this study, we propose and test a model showing how the modality of assessment affects assessment quality and turnover from both the individual and virtual group's point of view. We expect that different attributes of individuals and groups require different recruitment and assessment techniques, and the assessment, in turn, will have differential impacts on fit and turnover. More specifically, we expect that tacit and difficult-to-assess characteristics require groups and individuals to employ assessment techniques such as brief interactions, probationary periods, and referrals. On the other hand, because explicit information is easy to communicate and validate, short-term and indirect assessment methods can be efficient and effective for gathering this type of information. When assessors have knowledge about assessees' tacit and explicit attributes, this knowledge should improve the fit between members and groups. Finally, we expect that good fit of group members in the virtual environment may result in higher retention. To reduce the reader's confusion, we note that throughout this chapter, we treat members and virtual groups in the roles of both assessor and assessee or target of the assessment. That is, we are interested in how each side, new members and virtual groups, assesses the other, with the goal of each side making a good decision that its their needs.

The research context of this chapter is semi-permanent virtual groups (guilds) in in the online game, World of Warcraft (WoW). WoW is the most popular virtual world with over 12 million paying subscribers and over five thousands guilds (Blizzard Entertainment, 2010). In WoW, most players join and collaborate in guilds, where their avatars perform activities together with a common purpose. Guilds also serve as a broader social environment by providing a

support network in terms of materials, advice, a pool of collaborators for quests or raids and a source of friendships. Because a guild is a distributed virtual group, we expect that the results of this study can contribute not only to our understanding of groups in online games but other types of virtual groups as well.

The remainder of this chapter is organized as follows: The next section presents the hypotheses on assessment methods, knowledge, fit, and retention. The subsequent section describes the methods. The remaining sections show the results using structural equation models and discuss research implications and suggest future research directions.

2.2. Assessment Methods and Knowledge

Assessment Methods	**Examples**	**Characteristic of information**	**Time needed to make assessment**	**Amount of Interaction**
Credentials	Applications, Recruiting announcement	Explicit	Short	No interaction
Brief interactions	Interviews, Short play	Both Explicit and Tacit	Short	Limited interactions
Probationary period	Group experience	Both Explicit and Tacit	Long	Direct interactions over an extended time
Referrals	Acquaintances of current group members	Both Explicit and Tacit	Long	Extensive trustful historical interaction

Table 1. Characteristics of assessment methods

Assessment techniques lie on a continuum in terms of the amount and validity of information they provide for various types of assessments. The preliminary observations of groups in WoW and interviews with leaders of virtual groups and new group members revealed that virtual groups used selection techniques similar to those used in traditional and face-to-face organizations (Table 1). For example, some virtual groups require individuals submit an application and report credentials such as skills, abilities and experiences via the Internet. Individuals also can see the credentials of virtual groups on their websites or in public websites that track the achievements of groups and individuals (e.g. WoW Armory (http://www.wowarmory.com) or *Census Plus* (http://www.warcraftrealms.com/censusplus.php)). Group members and individuals directly interact by chatting via the Internet (e.g. chatting rooms, private messengers) or by jointly participating in a specific task. In addition, some virtual groups have a mandatory probationary period for new members before approving them as permanent members. Finally, virtual groups recruit and select new members who are acquaintances of current group members.

2.2.1. Credentials

To assess individuals' task-related attributes, such as prior experience, skill, and abilities, groups often use documented credentials, such as resumes, applications and other documentation of individuals' relevant history. Credentials represent explicit information. Some credentials, like employment history, can also be independently verified. Similarly, open source project groups and developers' performance and contribution are publicly accessible. In WoW, public information sources such as the WoW Armory or *Census Plus* provide frequently updated data

about each individual player and guild. Thus, it takes little time or effort for groups to evaluate prospective group members' credentials because the information about them is already codified and accessible.

However, some documents, like statements of purpose, can be written for self-presentation, creating uncertainty about the validity of what is being reported. In addition, credentials generally give the assessor only summary information, and not the raw data from which these summaries were built. This means assessors may see what targets have done, but lose subtle information about how they did it. Thus, while credentials are generally useful for conducting assessments of explicit characteristics, they may be less useful for tacit ones (Hansen & Haas, 2001; Winter, 1987).

Individuals also use credentials to assess groups. They have access to a group's credentials, such as the history of the group, composition of group members, and the previous performance or contribution, again reported in a public website. As the Internet has become an important search tool for job and WoW applicants alike, it is easier for candidates to evaluate this information (D. G. Allen, et al., 2007; Anderson, 2003; Dineen, et al., 2002). Group information is codified as well, providing explicit information about some aspects of groups, but not much tacit information about other aspects of the group, like their culture. Groups may describe their cultures or values, but this information is can easily be manipulated for self-presentational reasons and is difficult to verify without direct experience (Haas & Hansen, 2007).

H1. The more assessors use credentials as an assessment method, the more they are able to learn about the target's explicit attributes (a), but they will not learn about the target's tacit attributes (b).

2.2.2. Brief Interactions

Brief interactions are a second technique that virtual groups and individuals can use to assess a target. Group members and individuals directly interact by communicating with each other (e.g. chatting) or by jointly participating in a specific task. By doing so, they pick up subtle information about intelligence, personality, task style and values, among other dimensions. In studies of person perception, judgments made of others from "thin slices" of their behavior (e.g., a several second video clip) are moderately accurate (Ambady & Rosenthal, 1992; Borkenau, et al., 2004; Curhan & Pentland, 2007). In conventional, face-to-face groups, brief interactions have been found to have some validity and some bias (Baron & Kreps, 1999; Huffcutt, et al., 2001; Posthuma, et al., 2002). Through direct interactions such as in-person interviews and short tests, both groups and individuals can verify the target's explicit attributes such as abilities and skills as well as tacit attributes such as communication and interpersonal skills (Baron & Kreps, 1999; Huffcutt, et al., 2001). However, information provided by brief interaction also can be easily manipulated because of the short duration of the interaction (Posthuma, et al., 2002). Brief interactions rarely provide the assessor with behavioral samples of people engaged in authentic, contextually-valid tasks, although in principle they can.

In virtual groups, brief interactions may be effective in obtaining both tacit and explicit information, and may be even more valuable than they are in conventional groups. First, virtual environments filter many factors that can bias assessors during brief interactions in face-to face conventional groups including gender, race, age, and clothing of people or groups they are assessing (Burgoon, et al., 1996; Floyd & Guerrero, 2006). In addition, chats and short tests in virtual groups are very similar to authentic group tasks, since communication among virtual

group members is mostly done in online chats. As with short tests, brief interactions often involve playing virtual games with some group members for a short period of time using authentic tasks (e.g., killing monsters and completing an in-game quest together), which is very similar to those groups usually perform. Thus, we expect that assessors can learn both explicit and tacit information about targets based on direct observation and interaction with them.

H2. The more assessors use brief interactions as an assessment method, the more they are able to learn about the target's explicit (a) and tacit (b) attributes.

2.2.3. Probationary Periods

At the other end of the spectrum from brief interactions, virtual groups may accept individuals as their members only after a probationary period. During the probationary period, groups can observe if individuals are reliable and trustworthy. Similarly, individuals can join their groups after some probationary period. They decide if they will join the group based on the group experience.

Probationary periods provide a large amount of difficult-to-manipulate data to assess characteristics of a group or an individual (Baron & Kreps, 1999). In many groups, probationary periods last up to several months, with interactions between the individual and group taking place in the work context and involving authentic tasks. In universities, probationary periods for new faculty can last for 6 to 8 years. In virtual groups, probationary periods last up to several weeks, relatively shorter than that of conventional groups.

Compared to brief interactions, characteristics revealed in a probationary period are relatively difficult to manipulate for personal advantage because the interaction takes place over an extended time period (Loh, 1994). In addition, the experience includes both direct exchanges between the newcomer and the group as well as indirect observations (Bull & PieroTedeschi, 1989; Sadanand, et al., 1989), such as observations about the ways in which experienced group members react to the newcomer. This makes probationary periods useful for assessing not only explicit information, but also tacit information. Because candidates know that they can be accurately assessed during a probationary period or because it is too difficult to successfully dissemble over a sustained period, the prospect of probationary periods discourages them from making false claims. Extensive interactions during probationary periods also enable candidates to learn about the group's explicit and tacit information, such as work load or group members' cooperativeness.

Some tacit information is lost in electronic media (Finholt, et al., 1991). However, electronic media can become effective for the exchange of tacit information as individuals gain skill through experience (Carlson & Zmud, 1999; Walther, 1995). That is, when individuals have had experience with each other and with the media, they can interpret the remaining cues more effectively (Walther, 2002; Yoo & Alavi, 1999). Thus, probationary periods are useful for candidates learn the group culture and whether group members are reliable. For groups, they can learn if candidates are reliable and fit into the group culture.

H3. The more assessors use probationary periods as an assessment method, the more they can learn about the target's explicit (a) and tacit (b) attributes.

2.2.4. Referrals

Referrals from existing group members are the fourth assessment technique we consider. Referrals from existing typically have detailed and long-term information about both a potential new recruit and the group. Therefore, current members have a strong understanding of the culture of the group and can provide detailed information to the candidate that would otherwise be hard to learn. Moreover, they are motivated to represent the group and newcomer accurately to each other (Granovetter, 1995; Schwab, et al., 1987; Weller, et al., 2009). They simultaneously want to protect their own reputations and to preserve the quality of the group of which they are a member, so they will refrain from recruiting an inappropriate member (Baron & Kreps, 1999; Breaugh & Starke, 2000). For these reasons they should provide the group with reliable explicit and tacit information about the potential recruit.

Individuals often form virtual groups from among their acquaintances (Williams, et al., 2006). For example, Hahn et al.(2008) found that individuals are more likely to join an open source software development project when they have strong collaborative ties with its initiator. By using referrals, virtual groups and individuals obtain social cues to assess others' underlying quality and the risks involved in working with them. This information about their skills and capabilities comes from trustworthy sources. Thus, referrals are useful to learn about both tacit and explicit information in the virtual group context.

H4. The more assessors use referrals as an assessment method, the more they can learn about the target's explicit (a) and tacit (b) attributes.

2.3. Knowledge and Fit

Both recruits and groups are better off if groups recruit new members whose knowledge and skills fit the group's needs and whose values fit the group culture (Kristof, 1996). To achieve fit, groups and individuals must know about each other. Assessors' knowledge about targets' attributes should help them achieve better fit in virtual groups as it does in face-to-face groups. Although it may be easier to assess fit for explicit attributions than for tacit ones, each kind of information should lead to better fit. For example, if groups know about applicants' skills and abilities in areas relevant to their mission, they can find better task fit. In virtual groups, the task-related information of skills and abilities often comes in codified, explicit form. By knowing their personalities, they can better assess interpersonal fit.

H5. The more assessors learn about target's explicit attributes, the better fit to the target they can achieve.

H6. The more assessors learn about target's tacit attributes, the better fit to the target they can achieve.

2.4. Fit and Retention

Fit, that is compatibility between groups and individuals, has a strong impact on satisfaction, commitment, stress, loyalty and retention (Auh & Johnson, 2003; Kristof, 1996). Individual-group fit becomes critical when selecting employees for long-term employment (Bowen, et al., 1991). The attraction-selection-attrition model implies that fit leads to organization homogeneity (Schneider, 1987). According to this model, people are more attracted to careers/tasks/majors, and organizations composed of others similar to themselves (*attraction*).

Groups prefer and select new people who are similar to current members (*selection*). People who do not fit into the group either leave on their own or are encouraged to leave (*attrition*), thereby producing a more homogenous group. For example, task-oriented groups are likely to select applicants with task-oriented motives; applicants whose orientation towards the task is similar to that of the groups will stay longer. If applicants who are not task-oriented enter into a group that is highly focused on pursuing task goals, then these applicants may try to leave the group early and the group may not want to retain them. Similarly, applicants who value friendliness and reaching agreement with other members are likely to remain longer in a group that also values social experiences; these applicants also tend to be retained by the groups. If applicants with high social orientation enter into a group with low social orientation, they will be likely to drop out early or the group may ask poorly fitted applicants to leave.

H7. Individuals and groups with greater fit will have greater retention.

Figure 1. shows the overall research model.

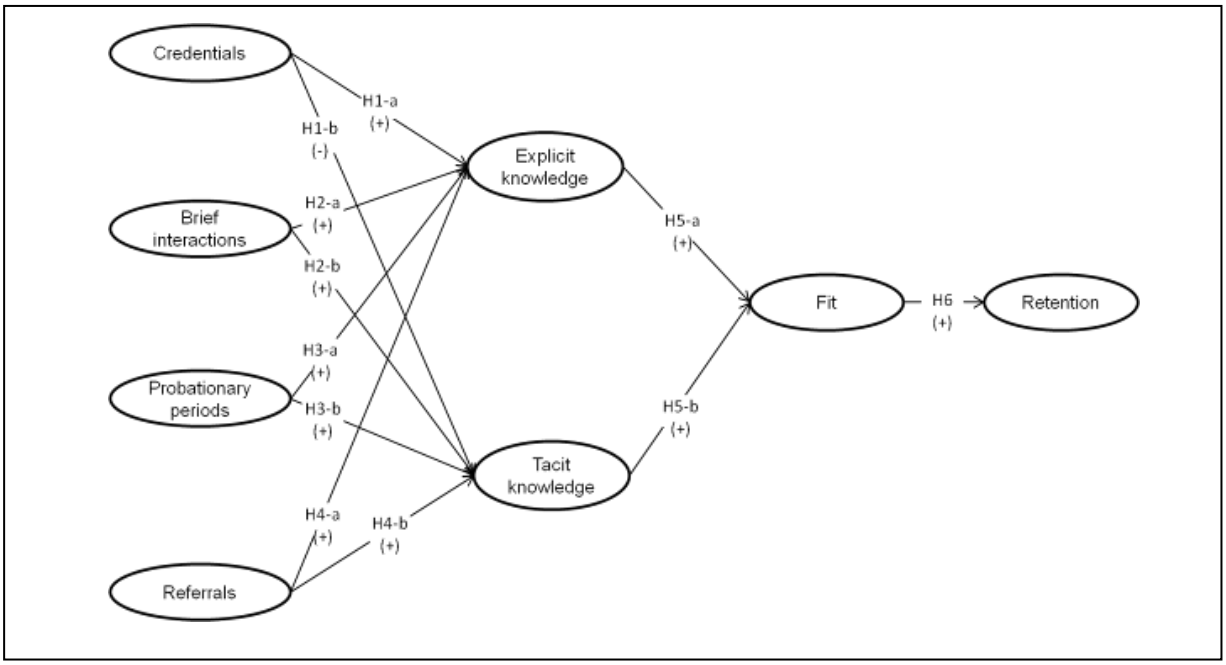

Figure 1. Research Model

2.5. Method

The ideas are tested in the context of guilds in the multi-player game, World of Warcraft (WoW). WoW players, who typically work in groups, perform a sequence of game-defined tasks, such as exploring a region of the game world or killing a specific monster, for which they gain in-game money, skills, loot and status. In performing these tasks, they have many opportunities for social interaction with strangers, opponents, and members of their own group.

We recruited two types of respondents for an online survey. The first are guild leaders, such as guild managers or recruiters; they were informants about the guild's procedures and values and provided the group perspective. The second were new members who had just joined an existing guild; they provided the individual perspective. The survey asked guilds and new members about the methods they used to learn about the other, their explicit and tacit knowledge about the other, and their perceived fit to the other. Lastly, in order to examine how fit influences retention, archival data was used to identify whether players remained in the guilds they joined.

2.5.1. Sample

Survey respondents were recruited from multiple sources: through the WoW's official bulletin board, players' community sites, and direct contact with players in the game. Leaders of guilds with more than forty guild members were randomly recruited and they were asked to fill out a questionnaire about their guilds' motivations and goals and their assessment methods. The guild leader response rate was 28%. The guild leaders were also asked to identify their guilds' newest member. Then the newest members of the participating guilds were contacted and asked

to fill out a questionnaire regarding their motivations and goals. The response rate for new members was 51%. Six hundred and eight guild leaders and 311 new members participated in the online survey. The survey was completed during the summer in 2008. *WoW Armory* and *Census Plus*, two public information sources, were used to measure the number of days players were members of their guilds. The data in "Census Plus" is collected by a Warcraft Census Plus user interface modification, which runs the game's */who* command — a service that lists online character identities and transfers this information to a server website for processing twice a day. This information is visible to the general public. New players' self-reports were compared with this list and excluded from the dataset those who did not appear in the public records. After this deletion, data from 608 guild leaders were retained and members from 273 guilds were verified.

2.5.2. Measure

Assessment Methods: We observed and interviewed guilds and players to examine the various assessment methods they use. As results, four assessment methods– credentials, brief interactions, probationary periods, and referrals were identified. Table 2 shows the questions used to measure groups' and individuals' usage of each method. Some virtual groups require that individuals submit their applications and report credentials such as skills, abilities and experiences. Individuals also can see the credentials of virtual groups in their websites or in public websites. Brief interactions were defined as group members and individuals chatting online or jointly participating in a specific task. In addition, some virtual groups have a mandatory probationary period for new members. Finally, virtual groups recruit and select new members who are referred by current group members.

How much did you learn about the newest member through the following methods?[1] How much did you learn about the guild through the following methods?[2]		Loading	Cronbach's α
Credentials			
Group[1]	Through his/her application in game.	0.94	0.98
	Through his/her character's information on the official game site.	0.99	
	Through his/her application on the guild website.	0.98	
Individual[2]	I heard/read a recruiting announcement in game.	0.99	0.98
	I read a recruitment statement of the guild on the official game site.	0.98	
	I have visited the guild website.	0.97	
Brief-interactions			
Group[1]	Through chatting with him or her once or twice in a game	0.98	0.96
	Through in-game short play with him/her	0.95	
Individual[2]	I chatted with a guild member in game once or twice in a game.	0.93	0.90
	I have played shortly with a guild member.	0.89	
Probationary Periods			
Group[1]	Through his/her probationary period	0.75	0.84
	Through the observation of his/her play more than a month	0.98	
Individual[2]	I began as a probation member.	0.79	0.87
	I was on probation more than a month	0.99	
Referrals			
Group[1]	Through the current members who are friends of him/her.	0.62	0.68
	Through the current members who have known him/her in game.	0.84	
Individual[2]	I talk to friends who are members of the guild.	0.97	0.97
	I talk to the member(s) of the guild that I have known in game.	0.98	

Table 2. Assessment Methods
([1] Guild leader version of the question, [2] The new member version of the question)

Explicit knowledge: Responses described how knowledgeable they were about a target's level, skill, and within-game profession on 5-point Likert scales (Table 3). These are the abilities that player characters incrementally earn in order to gather, make, or enhance items that can be used in WoW gameplay. These factors can be readily represented by information about a player displayed in the virtual game. For example, a player's level is represented by a number from 0-70. Character's skills and within-game professions are easily visible and hard to misrepresent.

Tacit knowledge: Personality of individuals and group members often plays a significant role in personnel selection (Hough & Furnham, 2003; Schmitt, et al., 2003). Responses described how knowledgeable they were about a target's personality on three dimensions (extraversion, agreeableness, and conscientiousness) of the Big Five personality traits (John, et al., 1991), because they are often used in personnel selection (Hogan & Holland, 2003; Hurtz & Donovan, 2000) and are most relevant in the WoW context. Conscientiousness refers to being disciplined and organized, extraversion refers to being out-going and talkative and agreeableness refers to being helpful and cooperative.

How knowledgeable were you about the newest member before he joined? [1] How knowledgeable were you about guild before you joined it? [2]		Estimate	Cronbach's α
Explicit knowledge			
Guild[1]	Level of the newest member	0.98	0.97
	Skill of the newest member	0.96	
	Profession of the newest member	0.93	
Individual[2]	Levels of guild members	0.99	0.96
	Skills of guild members	0.97	
	Professions of guild members	0.97	
Tacit knowledge			
Guild[1]	Conscientious (organized and disciplined) of the newest member	0.90	0.96
	Agreeableness (being helpful and cooperative) of the newest member	0.98	
	Extroversion (being out-going and talkativeness) of the guild members	0.95	
Individual[2]	Conscientious (organized and disciplined) of the guild members	0.92	0.96
	Agreeableness (being helpful and cooperative) of the guild members	0.98	
	Extroversion (being out-going and talkativeness) of the guild members	0.94	

Table 3. Knowledge
([1] Guild leader version of the question, [2] The new member version of the question)

Fit: Following past research fit was defined as congruence between the person and job (Cable & Judge, 1996; Chatman, 1989). Responses reported on the degree of fit they saw between the

entities they were representing (i.e., the group or themselves) and the target. They reported on general fit, quality of the match, and fit in terms of values[1] (Table 4).

Please answer the following questions about your newest member[1]/current guild[2].		Estimate	Cronbach's α
Guild[1]	There is a good fit between him/her and my guild.	0.98	0.96
	The match is very good between him/her and my guild.	0.95	
	I find that his/ her values and this guild's values are very similar.	0.90	
Individual[2]	There is a good fit between my guild and me.	0.98	0.97
	The match is very good between me and my guild.	0.96	
	I find that his/ her values and this guild's values are very similar	0.93	

Table 4. Fit

([1] Guild leader version of the question, [2] The new member version of the question)

Retention: To measure retention, archival data (*WoW Armory* and *Census Plus*) were used. Eight weeks after each new player participated in the survey, the retention whether the player remained a guild member (1) or quit (0) was determined.

2.6. Results

2.6.1. Measurement Model

The data was analyzed by using structural equation models (SEM) with Mplus (version 5.21) (Muthén & Muthén, 1998-2010). For the measurement model, a confirmatory factor analysis was conducted and computed the Cronbach's α for multi-item constructs. The measurement results are presented in Tables 2-1, 2-2, and 2-3. The descriptive statistics, correlation of the variables, and reliability estimates are presented in Table 5 (for guilds) and Table 6 (for individuals).

[1] I get substantively similar results when I use the alternative measure (the absolute difference between values for self and target) and use the self-report one in this chapter for ease of presentation.

		Mean	S.D.	1	2	3	4	5	6	7	8
1	Credentials	3.02	1.49	**0.98**							
2	Brief Interactions	3.74	1.46	0.02	**0.96**						
3	Probationary Periods	3.75	1.37	0.00	0.58**	**0.84**					
4	Referrals	1.35	0.76	0.07	-0.04	-0.04	**0.68**				
5	Explicit Knowledge	4.23	1.26	0.13*	0.14*	0.21**	-0.15*	**0.97**			
6	Tacit Knowledge	3.37	1.30	-0.08	0.32**	0.29**	0.09	0.26**	**0.96**		
7	Fit	4.21	0.56	0.02	0.23**	0.26**	-0.01	0.16**	0.32**	**0.96**	
8	Retention	0.42	0.50	-0.04	0.14*	0.14*	0.05	0.02	0.24**	0.69**	-

Table 5. Descriptive statistics, correlations, and reliability for guilds
(Note: The boldface figures on the diagonal are the Cronbach's α, N=273)

		Mean	S.D.	1	2	3	4	5	6	7	8
1	Credentials	1.50	1.10	**0.98**							
2	Brief Interactions	2.74	1.17	0.01	**0.90**						
3	Probationary Periods	1.46	0.86	0.07	0.10	**0.87**					
4	Referrals	2.64	1.71	-0.02	0.46**	0.06	**0.97**				
5	Explicit Knowledge	3.45	1.55	0.00	0.12	-0.10	0.074	**0.96**			
6	Tacit Knowledge	3.47	1.21	-0.04	0.08	0.10	0.06	0.17**	**0.96**		
7	Fit	4.22	0.56	-0.02	0.07	0.01	0.03	0.03	0.39**	**0.97**	
8	Retention	0.42	0.50	0.07	0.03	0.06	-0.05	-0.10	0.28**	0.69**	-

Table 6. Descriptive statistics, correlations, and reliability for individuals
(Note: The boldface figures on the diagonal are the Cronbach's α, N=273)

The guild measurement model fit the data well (CFI = 0.99.; RMSEA = 0.03; SRMR = 0.03) and so did the individual measurement model (CFI = 0.99.; RMSEA = 0.02; SRMR = 0.03). All the items loaded on their respective constructs ($p < .01$). The estimates obtained from Cronbach's α indicated that all items have reasonable reliability. High convergent validity was demonstrated by the finding that items reflecting the same construct had higher correlation with each other than with items reflecting different constructs.

Discriminant validity was tested according to the procedure in Bagozzi et al. (1991). For each pair of predefined scales, we compared a restricted model where the correlation was fixed to unity, indicating that the two scales were measuring the same construct, with an unrestricted model, where the correlation was freely estimated. In all pair-wise comparisons, the unrestricted models had better fit than the constrained model (for the chi-squared difference tests, all $p < .05$), affirming the discriminant validity of the constructs.

2.6.2. Structural Model

The research hypotheses were tested using SEM. Since the retention is categorical data, the Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator (Muthén, 1984) was used.

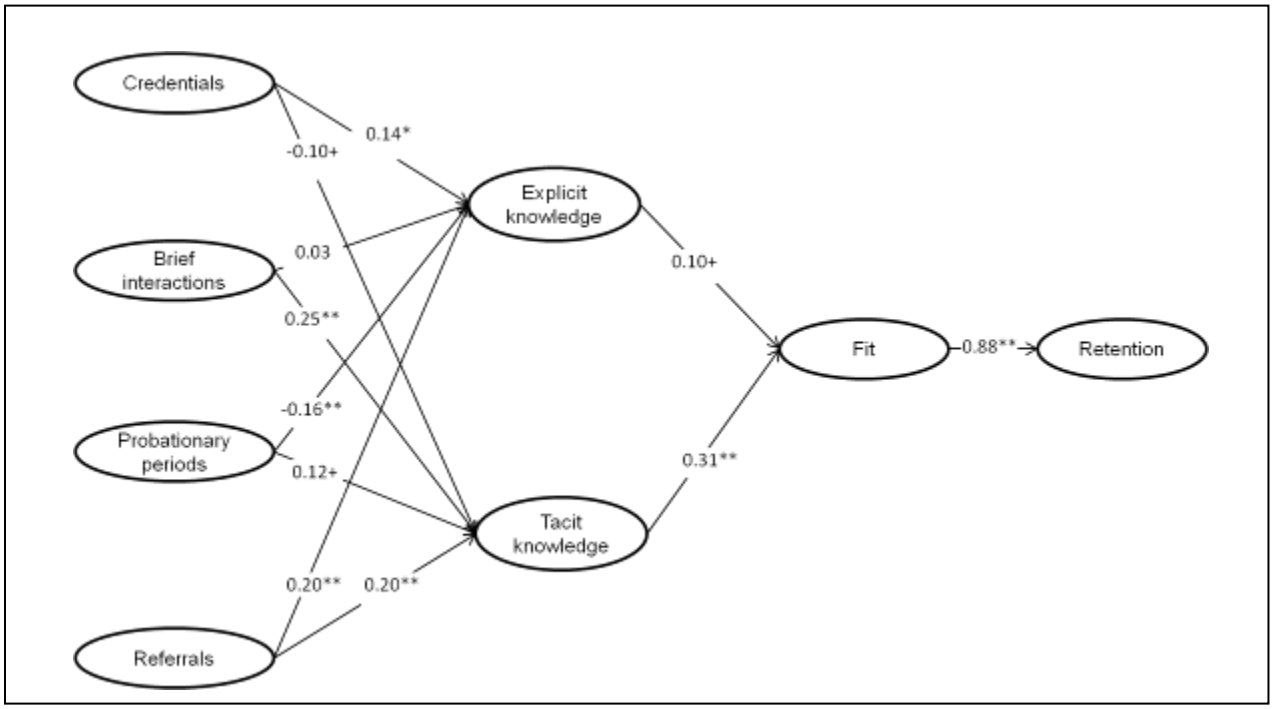

Figure 2. Structural Model Results for Guilds

(Note: Assessors are guilds and targets are new members. † p < .10, * p < .05, ** p < .01 (N=273))

2.6.2.1. Guilds' assessment of individual players

The fit for the structural model testing how guilds assess and retain new members shows good fit (CFI = .96, RMSEA = .07, WRMR = .83, N=273) while the test of model fit is significant ($\chi 2 = 233.61$, d.f. = 12, $p < .01$). Results are summarized in Figure 2. The more guild leaders used credentials as an assessment method, the more knowledgeable they perceived themselves to be about new members' explicit attributes supporting H1-a ($\beta = .14$, $p < .05$), but less knowledgeable they were about their tacit attributes providing marginal support of H1-b ($\beta = -.10$, $p < .10$). Guild leaders' usage of brief interactions did not improve their knowledge about new members' explicit attributes, providing no support for H2-a. Consistent with H2-b, the more leaders used brief interactions ($\beta = .25$. $p < .01$) as an assessment method, the more

knowledgeable they were about new members' tacit attributes. Surprisingly and inconsistent with H3-a, the more leaders used probationary periods as an assessment method, the less knowledgeable they reported being about the target's explicit attributes (β = -.16. $p < .01$). Consistent with H3-b, the more leaders used probationary periods (β = .12. $p < .10$) as an assessment method, the more knowledgeable they were about new members' tacit attributes. Consistent with H4-a and H4-b, the more leaders used referrals as an assessment method, the more knowledgeable they were about new members' explicit attributes (β = .20, $p < .01$) and tacit attributes (β = .20, $p < .01$). H5 was weakly supported. The more knowledgeable leaders were about an individual's explicit attributes, the better was the fit between the guild and the new members (β = .10, $p < .10$). Consistent with H6, the more leaders were knowledgeable about the target's tacit attributes, the better was the fit (β = .31, $p < .01$) between the guild and the new members. Finally, consistent with H7, new members were more likely to remain in their guilds when they were better fitted with the guild (β = .88, $p < .01$).

2.6.2.2. Individual players' assessment of guilds

The indices for the structural model testing how new members assess and remain in guilds shows good fit (CFI = .97, RMSEA = .05, WRMR = .68) while the test of model fit is significant ($\chi 2$ = 216.36, d.f. = 13, $p < 0.1$). Consistent with H5-b and H6, the more knowledgeable new members are about the guilds tacit attributes, the better was their fit (β = .40, $p < .01$) and new members were more likely to remain in their guilds when they fit the guild (β = .88, $p < .01$). However, new members' usage of different selection methods had little impact on the knowledge they had about either explicit or tacit attributes of the guilds they joined. Thus, in the new member model, there was little support for H1 thru H4. The degree to which new members were actually choosing among different guilds may explain these null results.

2.6.2.3. Post-hoc analysis and results

Several researchers have examined the extent to which applicants perceive viable alternatives in employment opportunities (Bauer, et al., 1998). More available opportunities are thought to have a negative effect on attraction to any one specific opportunity. If individuals have more choices, they might want to learn more about an organizations to make a better decision about whether it is the appropriate one for them to join (Chapman, et al., 2005). In this case, because individuals might need to employ assessment techniques to learn about the organization, the impact of assessment methods on gaining knowledge might be more apparent. In contrast, if individuals have no choice, because they are considering only a single organization, they might have less need for using assessment techniques to gain knowledge about it.

In the survey, the new members were asked how many guilds they considered before joining their current guild. The individuals were divided into two groups based on the number of guild choices they had: (a) individuals who had no choice, because considered only one option (N=171), and (b) individuals who had more choice, because they considered more than one option (N=102). A structural model for each group was estimated. For new members with multiple choices the fit was good (CFI = 0.95, RMSEA = 0.07, WRMR = 0.68) while the test of model fit is significant ($\chi 2 = 107.12$, d.f. = 13, $p < 0.1$). For those who applied to only a single guild and had little choice CFI = 0.99, RMSEA = 0.02, WRMR = 0.53), the structural model also fit well while the test of model fit is significant ($\chi 2 = 86.81$, d.f. = 14, $p < 0.1$).

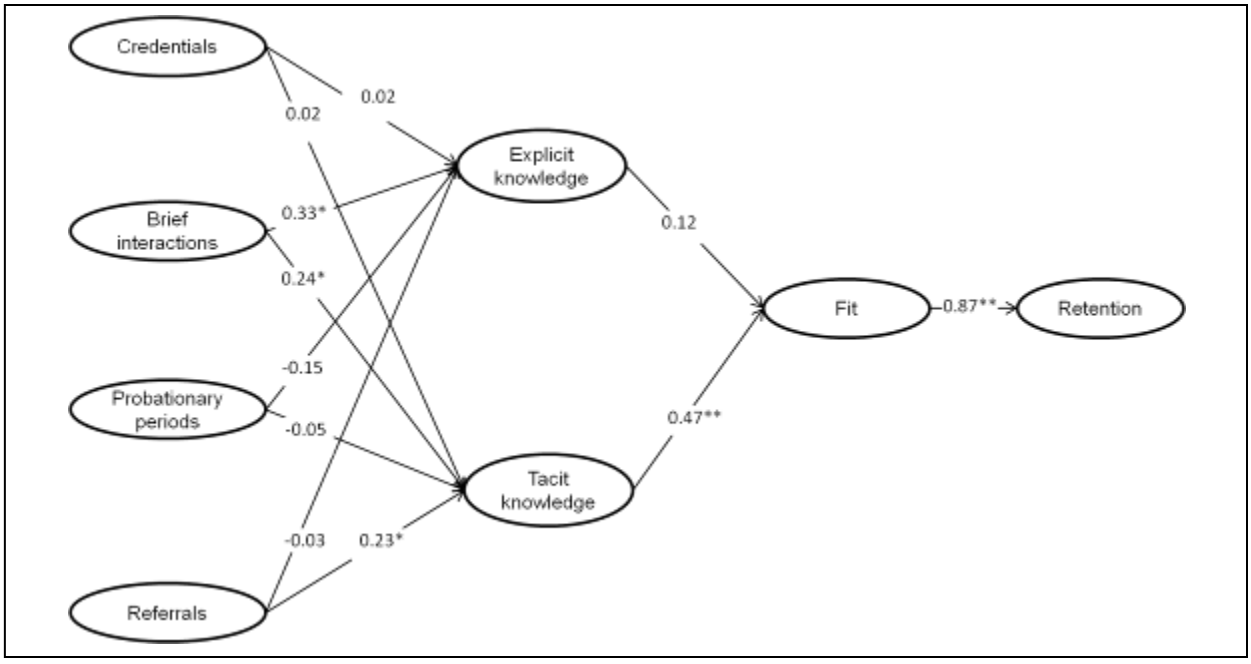

Figure 3. Structural Model Results for Individuals with multiple choices

(Note: Assessors are new members with multiple choices and targets are guilds.* p < .05, ** p< .01, (N=102))

For new members who had applied to multiple guilds, the results are summarized in Figure 3. Brief interactions had a positive impact on both explicit (β = 0.33, $p < .05$) and tacit (β = 0.24, $p < .05$) knowledge, consistent with H2-a and H2-b, and referrals had a positive impact on tacit knowledge (β = 0.23, $p < .05$), supporting H4-b. Knowledge of tacit information had a significant impact on fit (β = 0.47, $p < .01$) and new members were more likely to remain in their guilds when they fit the guild (β = .87, $p < .01$), supporting H5-b and H6. In contrast, for new members who applied to only a single guild, knowledge of tacit information had a significant impact on fit (β = 0.37, $p < .01$) and new members were more likely to remain in their guilds when they fit the guild (β = .87, $p < .01$), supporting H5-b and H6.

2.7. Discussion

Overall the results were consistent with the argument that when recruiting members, groups should employ diagnostic methods that are 1) based on data collected over long periods of time, hence increasing validity, 2) resistant to bias formed by self-presentation, and 3) capable of transmitting tacit information about social style and personality when assessing tacit and difficult-to-assess characteristics. Table 6 shows the summary of the results. In this study, the more guilds used brief interactions, probationary periods, and referrals as assessment methods, the more knowledgeable they were about prospective members' tacit attributes. In contrast, credentials and referrals seem sufficient as assessment methods when assessing recruits explicit attributes. A similar pattern occurs when individuals are assessing which group to join, but only among people who are actually choosing among groups. For individuals, the more they used brief interactions as assessment methods, the more knowledgeable they were about both the guild's explicit and tacit attributes. More use of referrals also increases their knowledge of the guild's tacit attributes. For both individuals and group, having more knowledge about a target's tacit attributes predicted better fit. Finally, we found that players stayed in groups longer when they had a better. Just as in conventional organizations, good fit resulted in higher retention of new members in the virtual environment.

Outcomes	Hypothesis		Assessors	
			Guilds	Individuals with choices
Explicit knowledge	H1-a	The more assessors use credentials as a assessment method, the more they are able to learn about the target's explicit attributes	Supported	Not supported
	H2-a	The more assessors use brief interactions as an assessment method, the more they are able to learn about the target's explicit attributes.	Not supported	Supported
	H3-a	The more assessors use probationary periods as an assessment method, the more they can learn about the target's explicit attributes.	Rejected	Rejected
	H4-a	The more assessors use referrals as an assessment method, the more they can learn about the target's explicit attributes.	Supported	Not supported
Implicit knowledge	H1-b	The more assessors use credentials as an assessment method, the less they can learn about target's tacit attributes.	Supported	Not supported
	H2-b	The more assessors use brief interactions as an assessment method, the more they are able to learn about the target's tacit attributes.	Supported	Supported
	H3-b	The more assessors use probationary periods as an assessment method, the more they can learn about the target's tacit attributes.	Supported	Not supported
	H4-b	The more assessors use referrals as an assessment method, the more they can learn the target's tacit attributes.	Supported	Supported
Fit	H5	The more assessors learn about target's tacit attributes, the better fit to the target they can achieve.	Supported	Supported
	H6	The more assessors learn about other's tacit attributes, the better fit to the target they can achieve.	Supported	Supported
Retention	H7	Individuals and groups with the greater fit will have greater retention.	Supported	Supported

Table 6. Summary of results

The biggest effects to concentrate on in this study are those for brief interactions and referrals improving tactic knowledge. Brief interactions provide direct behavior samples and

people gain subtle information about intelligence, personality, task style and values through "thin slices" of those behaviors. On the other hand, referrals are effective based on indirect but trustful experiences with assesses. The utility of brief interactions and references is not an inherent quality of the thing being assessed. Rather they are useful for attributes where good credentials (abstract, quantitative summaries) are not available. WoW currently provides good abstract, quantitative summaries for skill, but not for personality and values. However, they could provide credentials for personality and values. For example, using peer ratings of friendliness suggests that one could convert tacit attributes into explicit attributes by appropriate measurement and knowledge engineering. The development of Matsushita automatic home bread-making machine is a classic example of converting the tacit into the explicit - a software developer solved the kneading problem with the machine by observing a chief baker's distinctive way of stretching the dough (tacit knowledge) and adding special ribs inside the machine (explicit knowledge) (Nonaka & Takeuchi, 1995). The results show that in the absence of these measurements, brief interaction with direct experience or references work surprisingly well.

Although most results were consistent with the hypotheses, some were unexpected. First, we did not find that brief interactions helped guild leaders gain explicit knowledge of new members. Guild leaders used brief interactions frequently, but variation in use did not predict knowledge about explicit characteristics of new members, although it did predict knowledge about tacit characteristics. One of the reasons can be that they use these methods unreliably without structure, leading to low validity. A structured brief interaction is a standardized method of comparing candidates and typically used when an employer wants to assess and compare candidates impartially and consistently (Huffcutt, et al., 2001). If the position requires specific skills and experience, the employer will draft questions focusing exactly on the abilities the

group is seeking. While unstructured interview is effective in finding out more about a candidate's personality and how a candidate reacts in a pressure situation, structured interview allows the company to find out exactly what it needs to help quickly determine by weeding out candidates that do not have the necessary qualifications (Blackman, 2002). Most of brief interactions in WoW are conducted in a casual manner without structure, and this might lead to poor assessments of the explicit features of new members.

In this domain, credentials may be sufficient to learn about levels and skills. In contrast, new members seemed to gain both explicit and tacit information about guilds from brief interactions. Although some researchers have noted that brief interactions may not be suitable for evaluating tacit knowledge (Huffcutt, et al., 2001), they were useful for guilds and players. In this setting, playing with a pickup organization for a short period of time in a game reveals some authentic guild tasks. While pickup groups are generally smaller and shorter lived than a guild, the players in them perform similar activities (e.g., killing monsters, completing in-game quest together or simply hanging out together between tasks). Completing in-game tasks in a pickup group takes only a few minutes whereas it would usually take several days to complete a project in a real organization. Thus, collaborating in a pickup group is analogous to completing a project with another member in a conventional organization and it may be long enough to assess organizations' and individuals' tacit attributes. In addition, although gender, race, age, and clothing are often factors that cause bias in short interactions in off-line organizations, these attributes may not cause bias in virtual groups because computer-mediated communication often filters out the information about targets' real race, age, and clothing (Burgoon, et al., 1996; Floyd & Guerrero, 2006).

A second unexpected finding was that groups reported knowing less about individuals' explicit attributes when they used probationary periods as an assessment method, inconsistent with hypothesis H3-a. While it is plausible that assessor would get more information and therefore make better assessments from interactions lasting weeks than from interactions lasting minutes, the results contrasting the benefits of probationary periods with brief interactions does not confirm this intuition. However, this result is consistent with research on impression formation about personality and intelligence, which shows diminishing returns with the amount of exposure an assessor has of a target (Ambady, et al., 2006; Borkenau, et al., 2004; Carney, et al., 2007). In addition, prior research suggests that assessors tend to doubt their knowledge about targets' explicit characteristics as they learn more about targets (Ichino & Muehlheusser, 2008). Perhaps as assessors experience more variability in behavior diagnostic of targets' explicit attributes with repeated observations, they become less confident about what they know about them over time. It is possible that during these probationary periods, assessors are not using them as well as they should. For example, if an individual is on probation and no one in the group is tracking their experience with him/her, the group cannot be aware of his/her explicit features that keep changing over time.

Third, the hypotheses about the impact of credentials (H1-a and H1-b) and probationary periods (H3-a and H3-b) on the explicit and tacit knowledge for individuals evaluating guilds were not supported. The hypothesis about the impact of referrals on the explicit knowledge for individuals evaluating guilds (H4-a) was not supported as well. We assumed that impacts of different assessment methods are the same for groups and individuals. In fact, many guilds did not describe themselves well in their credentials whereas individual players usually provide detailed information in their answers to the questions posed by the guild. Similarly, a

probationary period or knowing at least one member of current guilds is often required by guilds to evaluate individuals even if individual players do not want to spend time as a probationary member or do not know any of current members. These might be the reasons that individuals do not know as much about guilds as guilds know about individuals by using these methods. Although groups and individuals engage in a two-sided search and assessment process (Saks & Ashforth, 1997; Schwab, et al., 1987) and selection decisions are jointly determined (Kozlowski & Klein, 2000), the impact of using assessment methods can be different depending on how the methods are executed.

Fourth, we found a marginal positive impact of explicit knowledge on fit from a group point of view, but we did not find a significant result for individuals. One of the reasons might be that it is not that important for individual players to know about guild members' codified skills and level. For them, it might be more important to know how the guild members execute their abilities. For example, even if guild members have high abilities, it is possible that they are unpleasant to play with, are not trustworthy, or do not share their rewards fairly. Individuals need the tacit information about whether members are enjoyable to talk with, helpful and trustworthy, which would help predict a good fit to the group. For individuals who have high abilities and professions and good skills, however, explicit knowledge of guild members can be important. They want to play with players who have a certain explicit level, skills, and professions representing that the group members are able to complete complex tasks with or help them while doing a task. For example, a player with a 10 level can give little help to a player with an 80 level while they try to kill a level 80 monster.

2.8. Implications

The results of this study contribute to the existing literature in several ways. First, this is one of the first studies to examine the entire process by which modality of recruitment and assessment affects both assessment quality, fit and turnover rates in virtual groups. Even in the literature on selection in conventional groups, little research has examined the pathways that produce this effect or have examined multiple assessment methods and mediators simultaneously (e.g., Hom, et al., 1999; Meglino, et al., 1988). In addition, there is no apparent consensus as to which recruitment source or method is most effective in virtual groups (Zottoli & Wanous, 2000). The results of this study should help to develop a fuller understanding about which recruiting and assessment techniques can reliably and validly measure different types of attributes of individuals and groups. Furthermore, this study should help us identify how knowledge of particular attributes can lead to a high-quality outcome. These insights provide greater understanding about the relationship between assessment method and outcomes.

This study's use of empirical and longitudinal data both from groups and individuals remedies methodological problems found in most prior research in the selection and assessment research area. Most prior research on group assessment used of cross-sectional surveys based on self-reported simplified, single item measures (Ployhart, 2006). Another limitation of assessment research is that most looks at selection from a single view point--either from the individual recruit's or the group's perspective--even though selection decisions are jointly determined (Kozlowski & Klein, 2000). This research makes an important contribution to the literature by providing a systematic approach for measuring the effect assessment methods have on retention, using more precise measurements, and taking data both from individuals and groups.

From an applied point of view, the results of this study are relevant to both distributed virtual groups and groups in the real world that use computer-mediated communication (CMC) to interact with each other. As more organizations increase their global operations, the formation of distributed virtual groups becomes a cost effective way to deal with new workplace demands. Virtual groups can apply the results of this study according to their goals and characteristics. The results of this study suggest that using referrals would be the best way to find people for all virtual groups. If using referrals is not possible, and groups are interested in learning about an individual's tacit information, this information might be identifiable via a brief interaction or probationary period rather than credentials. For other virtual groups having an interest in identifying explicit information rather than tacit information using credentials might be more beneficial

Designers of virtual groups in virtual games can develop tools to help individual players and guilds find appropriate matches. For example, they can develop a peer-review reputation system that encourages players to evaluate each other on several dimensions such as friendliness, skill in playing the game, and reliability following interaction with each other, and afterwards displays this information publicly. Using this system, players can see a potential partner's reputation in terms of explicit or tacit attributes before deciding to play with that character. This system would help players make more accurate decisions about interaction, improve the group experience for virtual group members or virtual world users, and thus enhance online collaboration.

2.9. Limitations and future directions

We extracted three personality measurements—extraversion, agreeableness, and conscientiousness—relevant to the WoW context in order to measure the knowledge of tacit attributes. In principle, however, it should be possible to develop improved diagnostic instruments for the tacit dimensions that organizations care about. For example, practitioners can use the short form Big Five personality measurements to measure tacit characteristics of employees. Future studies are needed to investigate specific assessment methods in terms of their validity for assessing particular types of traits, knowledge, skills and abilities.

Another limitation of the current model is that it does not take the actual knowledge of individuals and organizations into account. We measured the perceived knowledge of explicit and tacit attributes, but did not measure the extent to which this perceived knowledge about targets was correct. Even if explicit attributes such as levels, skills and professions can be verified in wow armory and census where information about guilds and individual players is archived, we did not measure whether assessors' judgments of these attributes were accurate. The relationship between knowledge and outcomes are likely to depend upon the accuracy of knowledge about targets. For example, Vandenberg and Scarpello (1990) found a positive relationship between accuracy of pre-entry information and the degree to which a new job's rewards matched entrants' needs. The analysis of this study could be improved by comparing actual knowledge to perceived knowledge.

We examined the impact of assessment techniques on retention only during the initial eight weeks after newcomers joined guilds. However, later retention of new members can be influenced by socialization tactics that may differ depending on how individuals were selected

(Kammeyer-Mueller & Wanberg, 2003). A potential extension of this chapter is to study groups' and individuals' behavior changes in later stages.

Overall, this research contributes to a richer understanding of which recruiting and assessment techniques can reliably and validly measure different types of attributes of individuals and groups. It is one of the few studies that have examined the entire process by which modality of recruitment and assessment methods affects both assessment quality, fit and turnover rates in groups. This study is a real step forward in systematically studying the attraction-selection-attrition model in any kind of group, be it virtual or a more traditional face to face group. The results of this study can help any organization, virtual or real, develop a better, more valid group recruiting and selection process; one that makes the best use of different methods for eliciting information about tacit and explicit attributes of individuals and groups.

Chapter 3

Socialization in Online Groups

3.1. Introduction

Online groups have been rapidly growing and increasingly important. Online groups provide place to exchange useful information and social support, meet people, maintain social networks, discuss political and social issues, and entertain. Firms have investigated ways to use online groups for help with marketing, sales, customer service, and product design. Among various online groups, they are particularly interested in online production groups in recognition of their value-creating potential. These groups of individuals are distributed worldwide and voluntarily collaborating online to produce informational goods and services ranging from software to free encyclopedias (Benkler, 2006). For example, the Open Source Software (OSS) groups have received enormous attention in the last several years because they are a fundamentally new way to develop software with large numbers of volunteers. Wikipedia has also been consistently ranked among the top ten visited websites on the Internet by Alexa.com and studies have shown its quality to be comparable to traditional encyclopedias.

However, success of online groups is limited because they often fail to encourage people to contribute over the long term. For example, 68% of newcomers to Usenet groups never post anything after their first post (Arguello, et al., 2006), and a quarter of the members of guilds in the massive multiplayer online game World of Warcraft leave their guilds every month even though they are still playing the game (Ducheneaut, et al., 2007). According to the popular open

source portal, SourceForge (http://sourceforge.net/), most OSS projects have also ended in failure: 58% do not move beyond the alpha developmental stage, 22% remain in the planning phase, 17% remain in the pre-alpha phase, and some become inactive (Lee, et al., 2009). Inequities in contribution are extreme, with a small fraction of members doing the vast majority of contribution. In Wikipedia, for example, more than 90% of all edits are made by the top 15% of editors.

Given these high turnover rates and lack of contribution by most group members, online groups need to help newcomers to adjust to their new communication environment. This adjustment and adaptation process has been referred to as socialization, or the process by which newcomers make the transition from being outsiders to being insiders (Levine, et al., 2005; Morrison, 1993). Socialization is fundamental to online groups and their members because it helps ensure the continuity of central values and gives the members a framework for responding to events in their groups (Bauer, et al., 1998; Jones, 1986; Van Maanen & Schein, 1979). Recent research on socialization in online groups highlights the importance of socialization to the overall function of groups (Ahuja & Galvin, 2003; Burke, et al., 2009; Ducheneaut, 2005; Wang, et al., Under Review). In one example, Ducheneaut (2005), using ethnographic analysis, found that successful newcomers engage in identity construction over time and forge alliances with other group members.

In spite of its importance, research on socialization in online groups is still in its early stages. Only a handful of studies have examined the impact of socialization in online groups. Some of this research has investigated how newcomers to online groups change throughout the socialization process. For example, Ducheneaut (2005) found that newcomers in OSS projects start finding bugs and providing source code to solve those bugs. They become developers who

integrate their modules to the core code and their influence gets higher. Bryant et al. (2005) conducted telephone interviews with committed Wikipedia editors about their experiences and found that, as these editors moved from peripheral to full participation, their goals, tools, and perceptions of the community changed (e.g., newcomers did not perceive a sense of community within Wikipedia, while experienced users did). However, these studies did not identify how to socialize newcomers and make them progress from peripheral to full participation.

In general, traditional work organizations apply two kinds of tactics. First are institutional socialization tactics that provide newcomers with formal socialization experiences set by organizations, and secondly, there are individual socialization tactics that provide newcomers with informal and a random set of training experiences (Jones, 1986; Van Maanen & Schein, 1979). Unlike with work organizations, socialization mainly occurs through informal, peer-to-peer interaction among the members in non-employment situations, such as the Red Cross or book clubs. This is due to a lack of managerial effort and bureaucratic control that can provide institutional socialization. In online groups, such informal interaction among peers is even more important because members there voluntarily work from at distance at different times with no face-to-face interaction.

Some studies have already identified the positive impact of the number of informal communications on individuals' contributions to their online group (Moon & Sproull, 2008; Wang, et al., Under Review). However, the impact of the various content of these communications on socialization outcome remains to be still studied. Thus, empirical research that examines in more detail how informal peer-to-peer interaction leads to successful socialization outcomes has important theoretical and practical implications.

The remainder of this paper is organized as follows: The next sections review prior research on socialization in conventional work organizations (offline), volunteer organizations (offline), and online groups. The subsequent sections report two studies conducted in Wikipedia, one examining the socialization tactics used and the second their effectiveness. The final two sections discuss our results and suggest future research directions.

3.2. Socialization in conventional work organizations

Van Maanen and Schein's (1979) and Jones' (1986) typology of socialization tactics is the standard model for effective socialization in conventional work organizations. Jones classified socialization tactics into two categories (institutionalized and individual socialization tactics) based on the results of Van Maanen and Schein's study (Jones, 1986; Van Maanen & Schein, 1979). Institutional socialization tactics require more organizational efforts to organize and manage to utilize them than individualized ones that provide opportunities to learn about the group individually. Institutional socialization tactics include five tactics -- collective(newcomers receive training and other socialization experiences as a part of a group of other newcomers), formal (newcomers are segregated from others) , sequential (the organization puts newcomers through a coherent sequence of training and job experiences that build on each other), fixed (newcomers are provided with a clear timetable for training experiences), serial (newcomers are provided with experienced mentors who help them learn their jobs), investiture (the organization acknowledges and builds upon the newcomers' existing skills and abilities). Individualized socialization tactics include five tactics: individual (newcomers accumulate unique experiences separate from other newcomers), informal (there is little separation between newcomers and

existing members), random (the organization provides a random set of training experiences), variable (newcomers have little idea about when training will occur), disjunctive (newcomers are not provided any mentors), and divestiture (the organization demands that newcomers change). Prior research has found that institutionalized socialization leads to more positive outcomes for both individuals and organizations than individualized socialization in traditional work organizations by reducing newcomers' anxiety and uncertainty (Bauer, et al., 2007; Saks, et al., 2007).

3.3. Socialization in non-employment organizations

Socialization occurs in many non-employment groups as well, ranging from volunteer organizations like churches and the Red Cross to informal social groups like book clubs. Non-employment groups are less structured than employment groups due to a general lack of organizational hierarchy in the former. They also do not have as much ability to systemize socialization tactics that require organizational effort. This limited bureaucratic control and a lack of general managerial effort make it difficult for non-employment groups difficult to provide effective institutional socialization. Instead of providing institutional socialization, socialization in non-employment groups occurs through an informal, peer-to-peer interaction among group members.

Pearce (1993) found that the level of interpersonal influence was significantly higher among volunteers than among paid workers. More frequent informal contact with other volunteers helps build relationships, aids communication, and clarifies the expectations of volunteers. The frequency of peer-to-peer communication with other volunteers also indicates

the extent to which the volunteer role is a shared community object for each individual (Piliavin & Callero, 1991). Similarly, Grube and Piliavin (2000) found that social relationships positively relate to volunteer work hours. Farmer and Feber (2001) also found that more frequent informal, peer-to-peer communication with other volunteers was positively associated with the level of contributions.

Prior research on mentoring also indicates that peer-to-peer interaction is an important component of socialization in off-line groups (T. D. Allen, et al., 2007; Kram, 1985; Ragins & Cotton, 1999). Mentoring refers to a relational process where a more experienced individual, usually one that is more senior, contributes to the professional development of a newcomer by providing psychosocial support (e.g., counseling and friendship) and career-oriented support (e.g., coaching and sponsorship) (Burke, et al., 1993; Kram, 1985). Mentoring relationships can vary in both their form and context (Noe, 1988). Informal mentoring is not managed, structured, or formally recognized by an organization. It is based on spontaneous relationships that develop and continue without actual external involvement from the organization. In contrast, formal mentoring refers to those programs that are managed and sanctioned by the organization.

Informal mentorships often will grow out of informal relationships and colleague interactions between senior and junior organizational members. These relationships may be based on work or non-work topics. Mentors often select newcomers with whom they can identify easily and with whom they can develop a relationship and to whom they can devote attention. In contrast, formal mentorships are typically not based on initial informal relationships or interactions between two organizational members. The match between mentor and newcomers may instead range from a random assignment to a committee assignment to actual mentor selection based on newcomer files. Compared with informal mentors, formal mentors may not

view their mentees as particularly worthy of special attention and support. Further, a longer adjustment period may be required for formal mentors and mentees for them to get to know one another.

Despite the potential benefits of formal mentoring, the research suggests that informal mentoring based on informal peer-to-peer communications is more effective than formal mentoring (Chao, et al., 1992). For example, Noe (1988) found that mentees involved in informal mentorships reported higher levels of intrinsic job satisfaction and better organizational socialization than those in formal mentorships. This difference may occur because formal mentoring relationships develop based on an organization's attempts to match the mentor and mentee's shared interests or the mentee's development needs, rather than mutual identification and attraction between mentor and mentee (Ragins & Cotton, 1999).

Informal mentorships often arise because of a desire on the part of the mentor to help the mentee and a willingness on the part of the mentee to be open to advice and assistance from the mentor. Formal mentorships, on the other hand, entail a degree of pressure; the mentor and the mentee may be required to participate in the mentorship program as a function of their positions. This pressure could decrease a mentor's motivation to help the mentee and also decrease the mentee's willingness to be fully open to assistance from the mentor.

3.4. Socialization in online groups

As in most volunteer groups, socialization in online groups mainly occurs through informal peer-to-peer communication among group members. Our casual observation indicates that institutional tactics are used in some online groups, but are not heavily used. For example,

organizational teaching and couching newcomers is not common in OSS project groups, and these groups hide their social and material network. Thus, newcomers are required to uncover this network on their own by communicating with peers and spend time getting information about who is related to what part of the code to become an important developer of their groups (Ducheneaut, 2005). In online games, most players have to learn how to play the game by informally interacting with other players. Some online games, like the football-manager game hattrick.org, socialize newcomers in cohorts and provide incentives for new players to gain formal training through what is called "Hattrack University." Wikipedia allows new editors to make changes to articles with no training at all, even though the site does have voluminous policies that prescribe how one should behave as an editor and also institutions like welcoming committees that have the potential to provide newcomers with structured guidance. Our own initial observations also found that institutional tactics are rare in WikiProjects. Some mature WikiProjects present their goals and guidelines in their main pages but they do not check whether newcomers read and learn them before the newcomers work on project-related pages. Instead, most newcomers in WikiProjects edit something first and then receive feedback on their work afterwards from other members. Research in online groups has shown that the number of messages or feedbacks that newcomers receive increases their continued contribution to the group (Moon & Sproull, 2008; Wang, et al., Under Review).

However, these studies only investigated the impact of the mere existence of a social relationship between the group and the newcomers without examining the specific content of these messages or the types of feedback or any research that examines the impact of that informal communication content on member commitment is still rare. In online production groups, the content of socialization is particularly important. The goal of online production

groups is not conversation, but rather the production of more and better informational products, such as encyclopedia and open source software. For example, Wikipedia clearly indicates that it is not a type of social media where people communicate with each other as in Facebook; it is a place to work. Thus, merely testing the impact of the number of responses on members' contributions is not enough to understand the impact of socialization in this instance without also considering the content of socialization messages in online production groups.

Ahuja and Galvin (2003) examined the frequency of content of information exchange (normative, regulative, and cognitive) within an e-mail-based network of virtual groups, but they did not test the impact of these messages on member commitment. Given the lack of studies that have investigated the practices that online groups use to socialize their new members and the impact of those tactics on member commitment, further empirical studies examining the impact of socialization tactics used by online groups on member commitment would be useful.

To address this gap, we conduct exploratory and theory-building research to better understand what actually goes on in online groups and their members in the early days and how these practices work. Our two research questions are: 1) what kinds of socialization practices are used in online groups? and 2) what kinds of socialization practices enhance newcomers' commitment to these groups? This paper focuses on members' contribution and retention as two indicators of member commitment to an online group. People who commit to their group are more likely to contribute more, and therefore, the success of an online group does rely on the voluntary contributions from their members (Ma & Agarwal, 2007; Roberts, et al., 2006b; Wasko & Faraj, 2005). Maintaining membership in an online (or offline) group is also a fundamental component of commitment to that group (Ashforth & Saks, 1996; Wang, et al., 2012). According to Levine and Moreland (1994), and their group socialization model, the

outcome of the evaluation process between members and groups determines member commitment to the group, which in turn affects the likelihood that members will remain in it and expend effort to achieve its collective goals.

In summary, informal, peer-to-peer communication is also important in off-line employment organizations, may be more important in volunteer organizations, and even more important in online groups. The contribution of this research is the descriptive detail it delivers for how informal peer-to-peer interaction improves peoples' ability to work in a group or organization and contribute to it with a positive feeling.

3.5. Study platform

To answer our research questions, we chose Wikipedia as the research site. Wikipedia is a large, task-focused community whose goal is to produce a free online encyclopedia. Wikipedia is a highly popular website, with over 17 million registered editors as of August, 2012, and over 4 million content pages ("Wikipedia Statistics,"). Most newcomers on Wikipedia start reading articles as unregistered invisible users. When they become interested in noting mistakes or omissions, and correcting them, they then edit articles. Even as they contribute to articles, new users tend to make only minor changes. At some point, they decide to create an account on Wikipedia, which allows them to create a watch list, track their own contributions, and maintain a consistent identity on the site. When a user registers a username, a user page with a user talk page on which they can provide their biographical information and have other members post messages is automatically created. As new users become experienced members of Wikipedia,

they will make major changes to improve the quality of articles and take managerial roles as well to grow the community overall.

When editors become interested in editing particular topics and put their efforts toward specific articles, they can join a Wikiproject, which are collections of editors interested in improving the coverage and quality of articles in a particular domain. For example, members in the Georgia WikiProject create, assess, and improve pages related to the history, geography, culture, and other attributes of the US state of Georgia. Each WikiProject has a dedicated page that exists in a namespace that is separate from regular article content (Figure 1). Editors can join a project simply by adding their name to the member list on this dedicated page (Figure 2), although some projects move members who have not been active enough to an inactive list.

Figure 1. The main page of WikiProject Georgia

Members [edit]

Please feel free to add yourself here, and to indicate any areas of particular interest.

1. Gladtohelp 16:41, 2 October 2007 (UTC)Gladtohelp - Looking forward to working on this Georgia project. I've already been making various updates.
2. PTO 20:42, 24 January 2007 (UTC) - Just here until this project gets on its feet.

11. Mjrmtg 16:07, 23 February 2007 (UTC) I've been adding Infobox County templates to Georgia Counties, started with southern counties as I'm from Lowndes County and added surrounding counties then started doing alphabetical. ~~Recently, today, notice that the infoboxes are showing two maps of Georgia instead of one showing the county in Georgia and the second being Georgia in the US. I don't know why they aren's showing the USA map correctly.~~ Infoboxes are displaying maps correctly now.

Figure 2. Example of member list

Most WikiProject members are registered users and do have some editing experience on other Wikipedia articles. Only 20 % of Wikipedia editors, however, join a WikiProject as soon as they join Wikipedia. WikiProjects also provide mechanisms for members to self-identify and acknowledge each other. Members can also place project banners on their user pages to identify their online personas with the project. Identification with a project seems to influence their behavior. After editors join a Wikiproject, they direct more of their work to articles within the scope of an individual project (Kittur, et al., 2009). In addition, because WikiProjects are social groups with a smaller numbers of members compared to Wikipedia editors with millions of editors, they will bond together and take care of their newcomers more. WikiProject members also respect other members' opinions and talk in a proper manner compared to Wikipedia where some users use harsh words to punish other users when they violate norms or there is conflict about edits.

WikiProjects provide a good setting to study the effects of socialization in online volunteer groups for several reasons. First, WikiProjects are particularly interesting to group

researchers because they incorporate many characteristics of traditional work groups even though they are online. For instance, WikiProject members set goals, develop task criteria, maintain diverse collaborative processes, keep track of work that needs to be done, discuss issues of interest using a forum, develop project-specific norms, and reward each other for good performance. In addition, like a "real-world" work group, the success of a Wikiproject depends on the editors' ability to function as a cohesive group working toward a common goal. In summary, WikiProjects have characteristics in common with other online (and offline) work groups and therefore, we believe our findings can be generalized. Second, we can examine the socialization processes executed by WikiProject members and measure members' contribution to the group and membership changes over time. WikiProjects manage their membership using a member list on which members can sign or remove their names. They also provide rich historical data including all communications among members because users are registered users with user talk pages compared to Wikipedia with unregistered and registered users. Thus, WikiProjects provide a lens through which to examine the socialization processes executed by WikiProject members. We conducted two studies in WikiProjects to answer our research questions about the types of socialization practices used online and their effects.

3.6. Step 1

The goal of this step was to identify the socialization tactics used in WikiProjects.

3.6.1. Data collection

We randomly selected 22 WikiProjects from the 50 projects focusing on US states. Because all of the state projects have similar content (e.g., cities, government, geography, history,

culture), restricting the sample to US states helps to control many variables associated with topics.

Socialization occurred primarily through the interpersonal exchanges existing project members had with the newcomers on project and user talk pages. In order to welcome and socialize new members, experienced members usually begin by recognizing the addition of a new member's username to the member list on the project's main page and then communicating with him or her. This socialization occurs on project talk pages, members' personal user talk pages, and project related article talk pages. In general, on project talk pages, users discuss how to develop criteria and maintain the diverse collaborative processes of the project. Each user has a personal page and a related talk page, which could be used to discuss various subjects ranging from personal issues to article conflicts. Article talk pages are used to discuss and build consensus on changes to the article page. In this research, we restricted our focus to project talk pages and user talk pages, because socialization of WikiProject members mainly occurs on these two kinds of pages. Many anonymous users, who have neither specific user names nor user pages, and many editors who are not members of projects post frequently on the article talk pages, and so these pages are not used for socialization. Article talk pages are commonly used to discuss article content itself rather than to socialize editors.

We identified all new members of each project from the beginning of the project to September, 2007. After identifying the date they joined the project, we visited project talk pages and each member's user talk page. We collected all communication a new member received from already existing project members during the month after joining. Some users were blocked by Wikipedia because of their vandalistic edits, and their user talk pages were removed. We

excluded those users. The sample consists of 579 newcomers and 1150 socialization messages they received from existing project members.

3.6.2. Analysis

In this study we used grounded theory methods to examine the socialization tactics used by WikiProjects (Glaser & Strauss, 1967). We examined each message that newcomers received in a month after their joining the project. Each message could include multiple socialization tactics. We continued analyzing the project data until we reached the point of theoretical saturation, which occurred when the researchers no longer identified additional types of socialization tactics (Glaser & Strauss, 1967). After examining messages in three projects, five distinct categories emerged. Two independent coders coded the messages received by newcomers into the five categories using a standardized codebook. Each message could be associated with multiple coding categories. We coded each distinct socialization tactic received by each newcomer, resulting in total 1150 socialization tactics used by twelve WikiProjects. The overall inter-rater reliability was high (r = 0.94), and the lowest reliability for any category was 0.85.

3.6.3. Results

Coding of the interactions between existing members and newcomers indicates that Wikiprojects uses five types of socialization messages: Constructive criticism of work, positive feedback on work, requests to work on a particular task, welcome messages, and personal comments.

We also found two formats for messages: Standardized and personalized. Standardized messages were formal and used templates to ensure that all newcomers received the same communication, while personalized messages were informal, and content was tailored to particular newcomers (Jones, 1986). Sending standardized messages required a group's collective efforts. That is, a group created a template that clearly represented that group's intention and make a decision about who was in charge of sending it, who would be the receivers and when that message should be sent. They also had to monitor whether the messages are utilized properly. In this sense, standardized messages exhibited certain characteristics of institutional socialization tactics.

On the other hand, personalized messages were more likely to utilize individualized socialization tactics. Personalized messages let newcomers learn about the group individually by including a review of newcomers' recent work, thanking them for particular contributions, suggesting improvements to these contributions, requesting them to do a specific task, and mentioning off-topic interests by using more friendly and informal expressions. Newcomers can develop their own ways to contribute to the group by receiving these personalized messages.

Of the five types of socialization messages, we found that task requests and welcome messages had different formats, namely, both standardized and personalized variants. The remaining tactics had only personalized forms. We describe these socialization tactics in more detail below. Examples are shown in Table 1.

Constructive criticism was sent to newcomers when they did something wrong based on newcomers' recent work. Senders suggest improvements to these contributions in constructive criticism messages and they help newcomers learn how to edit project-related articles correctly

and how to discuss in a proper manner, both in Wikipedia and the WikiProject they joined. For example, in Table 1, an existing project member asked a newcomer not to violate group norms.

Old-timers often gave ***positive feedback*** on a newcomers' work. Existing project members either praised new members' edits on project-related pages or gave them an award.

Task requests were sent to newcomers when the group asked the newcomers (and potentially existing members) to do a certain job. These too came in standardized and personalized variants. Standardized task requests often occurred in the form of a template-driven formal newsletter asking all project members to do a job or class of work. Personalized task requests were personal messages sent from an old-timer to a newcomer to do a specific task.

Welcome messages were sent to newcomers by old-timers shortly after the newcomer had registered to be a member of the WikiProject. The messages signaled that the group was interested in him or her and wanted to develop a positive relationship. Welcome messages came in two variants: standardized ones, in which old-timers pasted a formal project-specific welcome template on newcomers' personal pages and personalized ones in which the old-timer added a personal message to the standard welcome.

Personal comments were friendly and supportive, aimed at maintaining close social relationships supporting group cohesion. These communications are about the sender's interests or off-topic content which was not related to WikiProject tasks.

Constructive criticism

I have seen you adding the links to UofO's special collections, and it looks like good portion of your material for articles come from that source. Just be careful about copyright infringement and Plagiarism, as I am sure you don't want your articles to be deleted for either reason. I find it best to re-order information and integrate multiple sources, change phrases etc. to avoid making it look like anything I add is a copyright violation, and then I always reference the item to avoid plagiarism.
Positive feedback
Thanks for all your work on the article. I saw that and was delighted, well done sir. I give you this award for your great contributions to and your even greater patience with a limited mind. May God bless you.
Standardized task request
Geetings, WPOR member, we are starting a weekly collaboration project where we will announce two articles that are currently stubs that we hope to work together to improve. No pressure to help, but if you would like to, just stop by one of the articles and see if you can find information to expand the article with, copy edit what is there, help with formatting, or add some images. This week's articles are: Alis volat proprils and Fusitriton orgonensis.
Personalized task request
Hey, one of these days do you think you could take some pictures at Mission Mill? I'd like to spruce up the article but it really needs some photos. Thanks!

Standardized welcome message
Welcome to WikiProject Oregon! If you'd like, you can add the WP Oregon userbox to your user page using this code: {{User WikiProject Oregon}}. Check out the ongoing and archived discussions at WT:ORE and be sure to add the page to your Watchlist. If you are new to Wikipedia, it's a good idea to browse through the core principles of Wikipedia as well. The project home page at WP:ORE has many useful links to get you started. The recent changes and recent discussions links will display recent edits on articles within the project's scope. Welcome!
Personalized welcome messages
Welcome to WikiProject Alabama! I saw your name posted on the members list and wanted to welcome you. I'm from the other end of the state, Daphine, near Mobile. I was raised near Tuscaloosa and yes, I am a 'Bama fan. Seems like you Auburn guys have got us beat around here. Anyway we are glad to have you. If I can help at all let me know or any of the other folks around the project. There's a lot of work to be done! Welcome!
Personal comments

I was there from 1995-1999 (was it that long ago?). As an Econ major, I didn't take many Anthropology or Art classes… although Ann Nicgorski was my World Views professors. :) I loved my experience at Willamette. I joined WP:WPOR, so maybe I'll see you around in some articles.

Table 1. Sample messages sent by WikiProject editors and their corresponding socialization tactics

In summary, we identified five socialization tactics in Study 1. We also found that two tactics (task requests and welcome messages) had both standardized and personalized variants. The remaining tactics (providing constructive criticism, positive feedback, and personal comments) were used only in a personal manner. We found that personalized socialization tactics were used more often than standardized ones in WikiProjects. Of the 780 socialization identified in the communications, 87% were personalized whereas only 13% were standardized. It means that socialization in most online communities is mostly based on personalized interpersonal communication rather than standardized one.

3.7. Step 2

In Step 2, we investigated the impact of the five socialization tactics on newcomer commitment to the project. First, we examined the impact of a number of messages that newcomers received, regardless of their content. Second, we investigated the impact of each socialization tactic on newcomers' subsequent contribution and retention. Lastly, we classified socialization tactics into standardized and personalized tactics and examined their differential impact on newcomers' subsequent contributions and their retention.

3.7.1. Data Collection

To measure newcomers' commitment to the project, we collected data reflecting newcomer's contributions and retention (i.e., how long new members remain active in the group). Wikipedia provides a data dump that lists all edits made to every page in Wikipedia up to September, 2007. WikiProjects designate their scope by placing a project-specific template on the articles that they oversee. MySQL database queries allowed us to identify each page a newcomer edited both inside and outside a project before and after joining the project.

3.7.1.1. Independent variables

Number of messages. We counted the number of messages a newcomers received from old-timers during the month after the date that they joined the project. For example, if someone joined on Apr 5, we used socialization messages from Apr 5 to May 4 to predict the number of edits the newcomer made from May 5th onwards. As noted earlier, old-timers were members who had joined the project prior to a newcomer's joining date.

Each type of socialization message. We used count data for all five socialization messages that we found in step 1.

3.7.1.2. Control variables

Projects. We included project as a dummy variable to control for unknown differences among the 22 WikiProjects.

Number of edits on Wikipedia pages before joining the project. We counted the number of edits made by each newcomer to Wikipedia pages before he or she joined the project.

Duration. The number of months between the editor's join date and their last edit in project related pages.

3.7.1.3. Dependent variables

Member commitment is one of the most important indicators of socialization success (Wang, et al., Under Review). We selected member contribution and retention to use to measure that commitment. Paid workers contribute to their organizations mainly because of momentary rewards they receive. Even if they are not committed to their groups, they continue to work to meet performance standards tied to an employment contract. To the contrary, volunteer groups and organizations often endure marginal contributions because coercive and performance assessments are rarely conducted in volunteer groups (Farmer & Fedor, 2001; Pearce, 1993). Thus, volunteer contribution levels are highly variable depending on their commitment levels. Similarly, it is easier for volunteers to leave their groups because they are not bound to any employment contract. Volunteers stay only if they are truly committed to their groups. Presumably, volunteers will contribute as long as they stay, and even small effort is better than none at all. Thus, retention is the indicator of volunteer commitment.

Contribution. We counted the number of edits made by each newcomer to project-related pages after he/she joined the project. We measured editors' contributions towards Wikipedia articles through their revision count (i.e., number of edits). Edits are a direct measure of editors' effort, indicating the number of changes they made to articles during a period of time. Each edit indicates a set of editing actions, for example adding, changing, deleting or reverting text, references or illustrations, or communicating with other editors. The dependent measure was the log transformed edits from the month he/she joined the project until the last month he/she edited

in the project-related pages. Because the logarithm of zero is undefined, we added one before computing the logarithm.

Retention. The time between newcomers' join date and last edits or end of the dataset in the case of truncation represents their survival time. The failure (or drop out) event in the survival analysis is defined as a person's last date of edit to the project-related articles. Because people who edited close to the end of data collect might still be participating, we considered those who last edited within the month of the end of data collection as right censored. Thus, we take into account the truncated nature of time-to-event data (i.e., at the time of data collection, some people who will eventually leave the group have not yet left).

3.7.2. Analysis

We examined the influence of socialization tactics on the commitment of newcomers using two analytic approaches. First, we examined the influence of socialization messages on newcomers' subsequent edits. The analysis used negative binomial regression as implemented in Stata (nbreg) because the dependent variable, number of edits, is over-dispersed count data. Then, we examined the impact of socialization messages on newcomers' retention. When the dependent variable involves time and possible censoring, survival analysis is the appropriate technique to use (Singer & Willett, 2003). Survival analysis is often used to study the time between entry to a study and the experience of an event because standard regression procedures produce biased estimates without taking into account the truncated nature of time-to-event data. We conducted survival analysis using Cox proportional hazards model.

3.7.3. Results

Table 2 reports the descriptive statistics and correlations for the variables before standardization to be entered into the regression and survival models.

		Mean	Median	S. D.	1	2	3	4	5	6	7	8	9
1	N of edits after joining a WikiProject	320.87	15	1451.52	1.00								
2	Duration in months	5.98	4	5.76	υ.21**	1.00							
3	N of edits before joining a WikiProject	1401.3	86	4087.23	0.13	0.15**	1.00						
4	N of messages	1.99	0	3.63	0.46**	0.12**	-0.08*	1.00					
5	Constructive criticism	0.42	0	1.08	0.37**	0.03	-0.06	0.68**	1.00				
6	Positive feedback	0.45	0	1.12	0.28**	0.06	-0.08	0.80**	0.34**	1.00			
7	Task requests	0.68	0	1.56	0.44**	0.16**	-0.05	0.88**	0.48**	0.58**	1.00		
8	Welcome messages	0.24	0	0.52	0.12**	-0.02	-0.05	0.48**	0.28**	0.35**	0.28**	1.00	
9	Personal comments	0.2	0	0.64	0.34**	0.17**	-0.06	0.57**	0.18**	0.42**	0.44**	0.68**	1.00

Table 2: Descriptive statistics and correlations (* p < .05, ** p< .01)

Tables 3, 4, and 5 show the impacts of socialization messages on newcomers' commitment to their groups. In each table, the first column shows the impact of newcomers' communication to existing project members in their first month on the newcomers' subsequent number of edits. Effects are reported as Incidence Rate Ratios (IRR), the ratio where an independent variable increases by a unit to change the dependent variable. (e.g., because the number of messages is measured by the log base 2 scale, the IRR for the number of messages of 2.35 indicates there were 135% ((2.35-1)*100%) more in-project edits for every doubling of messages that a newcomer received. The second column shows the impact of socialization messages on how long new members remain active in the WikiProject. The hazard ratio in this survival analysis is the effect of an explanatory variable on the hazard or risk of an event. For example, a hazard ratio of 0.80 for the number of edits that newcomers made before joining the project means that there is a 80% smaller hazard for every doubling of messages that newcomers received. That is, each doubling of the number of edits that newcomers made before joining the project led to a 20% increase in the time that newcomers stayed in the project.

	Amount of edits		Retention	
	IRR	Std. Err.	Haz.Ratio	Std. Err.
Duration	1.18**	0.03		
N of edits before joining (log2)	1.16**	0.03	0.80**	0.02
N of messages (log2)	2.35**	0.17	0.71	0.71

Table 3. The impacts of the number of messages on newcomers' amount of edits and retention (**: $p < 0.01$)

In Table 4, we found that each doubling of duration led to an 18% increase in newcomer edits the first month after joining the project. Each doubling of the number of edits that newcomers made before joining the project led to a 16% increase in their edits in the first month

after joining the project. Each doubling of messages increased newcomers' edits by 135%. The next column indicates that each doubling of the number of edits that newcomers made before joining the project produced a 20% increase in the time that a newcomer stayed in the project in the first month after joining the project. Each doubling of messages also increased the time a newcomer stayed in the project by 29%.

	Amount of edits		Retention	
	IRR	Std. Err.	Haz.Ratio	Std. Err.
Duration	1.19**	0.03		
N of edits before joining (log2)	1.16**	0.03	0.82**	0.02
Constructive criticism (log2)	1.48**	0.21	0.74	0.15
Positive feedback (log2)	1.84**	0.30	1.06	0.26
Task requests (log2)	1.24	0.19	0.91	0.15
Welcoming (log2)	1.85**	0.43	1.09	0.30
Personal comments (log2)	2.12**	0.47	0.27*	0.15

Table 4. The impacts of five socialization messages on newcomers' amount of edits and retention (**: $p < 0.01$, *: $p < 0.05$)

Table 4 decomposes the total messages into five socialization tactics. As shown in the first column, doubling the constructive criticism, positive feedback, welcome messages, and personal comments increased newcomers' edits by 48%, 84%, 85%, and 112%, , respectively, whereas we did not find a significant impact on task requests. We suspect that this finding might be because two variants of task requests (standardized and personalized) will have quite different impacts. The second column indicates that doubling the personal comments resulted in a 73% increase in the time that a newcomer stayed in the project, while we did not find any significant impact of other socialization tactics.

	Amount of edits		Retention	
	IRR	Std. Err.	Haz.Ratio	Std. Err.
Duration	1.18**	0.03		
N of edits before joining (log2)	1.17**	0.03	0.80**	0.02
Constructive criticism (log2)	1.61**	0.25	0.68	0.15
Positive feedback (log2)	1.67**	0.28	1.01	0.24
Task requests (log2)				
Standardized (log2)	0.47	0.23	1.76*	0.43
Personalized (log2)	1.42*	0.23	0.64*	0.13
Welcoming (log2)				
Standardized (log2)	1.48	0.53	1.33	0.43
Personalized (log2)	2.00*	0.56	0.49	0.26
Personal comments (log2)	1.97**	0.44	0.32*	0.17

Table 5. The impacts of five socialization messages with decomposing task requests and welcome messages into standardized and personalized messages (**: $p < 0.01$, *: $p < 0.05$)

Table 5 decomposes the task requests and welcome messages into standardized and personalized socialization tactics. As shown in the first column, we found that doubling the number of personalized task requests produced a 42% increase in newcomers' amount of edits and we did not find any significant impact of standardized task requests. In fact, the impact of two formats task requests was significantly different ($\chi 2$ (1) = 4.01, $p < 0.05$). Figure 3 shows these different impacts of standardized and personalized task requests on newcomers' contributions and shows that newcomers will increase their edits as they receive more personalized task requests and decrease their edits as they receive more standardized ones. Similarly, doubling the number of personalized welcomes resulted in a 100% *increase* in newcomer edits, whereas we did not find any significant impact for standardized welcome messages. However, the impact of the two formats of welcome messages was not significantly different ($\chi 2$ (1) = 0.49, $p > 0.05$).

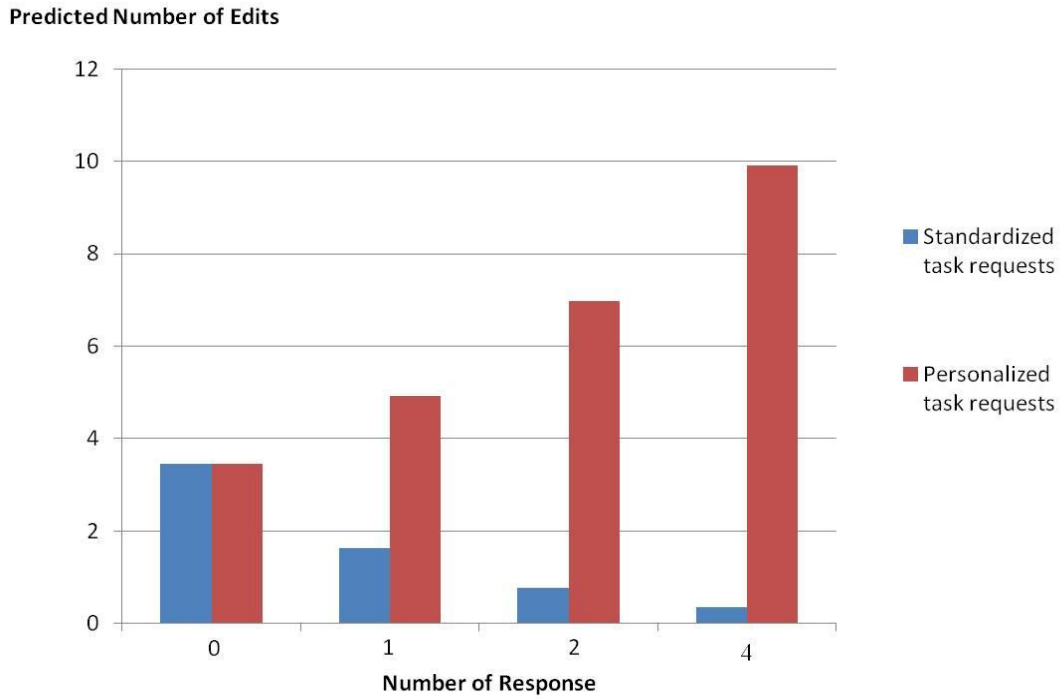

Figure 3. The impact of personalized and standardized task requests on newcomers' amount of edits

The second column indicates that doubling the number of personalized task requests resulted in a 36% *increase*, whereas doubling the number of standardized task requests resulted in a 76% *decrease* for the time a newcomer stayed in the project. In fact, the impact of personalized and standardized task requests was significantly different ($\chi 2$ (1) = 7.68, $p < 0.01$).

Figure 4 illustrates these results graphically, showing three survival curves. The top curve shows survival upon receiving two personalized task request messages, while the middle curve shows survival upon receiving one. The bottom curve shows survival with no personalized task request message. We did not find any significant impact for either standardized or personalized welcome messages for the time a newcomer stayed on the project. The impact of the two formats of welcome messages was also not significantly different ($\chi 2$ (1) = 2.69, $p > 0.05$).

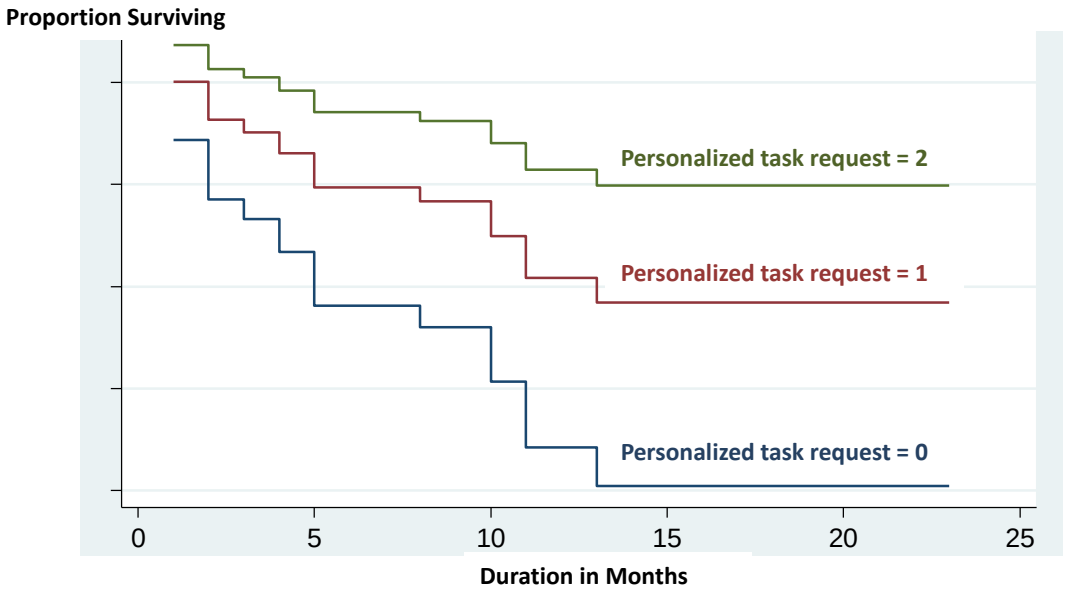

Figure 4. Survival graph for newcomers who received personalized task requests

3.8. Discussion

At the descriptive level, we identified five socialization tactics that are generally overlooked in any research on socialization in conventional organizations. Of these five tactics, we found that two (task requests and welcome messages) had both standardized and personalized variants. The remaining tactics (constructive criticism, positive feedback, and personal comments) were used only in a personal manner. The analysis showed that personalized socialization tactics were used much more frequently than standardized tactics for these WikiProjects. We suspect that because of limited bureaucracy control, there are not many projects that facilitate formal templates that require organizational effort. Instead, WikiProjects use a personalized format more often to socialize their newcomers.

The variables identified in our research appeared to influence newcomers' commitment to online groups. All the control variables significantly predicted newcomers' subsequent editing on a project. The longer a newcomer remains in a group, the more that individual contributes to the group, and the more newcomers edited in the project before joining it, the more contributions they made after joining. This result may simply reflect general consistency (i.e., those who have contributed will continue to contribute in the future). The result may also reflect the group's greater willingness to maintain a relationship with individuals who did participate actively in the past (Wang, et al., Under Review).

Because the groups were more willing to engage individuals who have already shown commitment (Moreland & Levine, 1988, 2001), individuals might feel obliged to reciprocate by contributing more after being welcomed into a group. Consistent with prior research (Duck, et al., 1991; Golder & Donath, 2004; Kraut, et al., Under Review; Williams, et al., 2000), the number of messages that newcomers received was directly associated with the amount of edits that these newcomers subsequently did. Independent of the content of the messages, their mere existence may signal that a social relationship exists between the group and the newcomer, and hence may affect an individual's commitment to the group.

Positive feedback increases the newcomer's amount of edits by increasing newcomer motivation to contribute to a group more often. Interestingly enough, constructive criticism also will increase the newcomer's amount of edits. In our observation, constructive criticism utilizes intimidation and reprimands to decrease undesired behaviors from targets mostly in a positive manner and provides direction to targets and also specifies their roles and responsibilities. Thus, newcomers who received constructive criticisms from old-timers have an opportunity to learn how better to contribute to their groups by correcting what they did wrong when following

directions. This effort might lead to a newcomer's greater contribution to their groups. Welcome messages and personal comments also increase the amount of edits from newcomers. As in the previous research, we identified that maintaining a close social relationship with newcomers by communicating with them in a friendly and supportive manner increased their contributions to their online group.

However, we did not find a significant increase in the amount of newcomer edits when he/she received task requests from an old-timer. This outcome might be because of the different impacts of standardized and personalized messages. When we differentiated task requests into two formats, we found that personalized task requests increased newcomers' contributions and retention, whereas standardized task requests decreased retention. Crampton, Hodge, and Mishra (1998) identified that when employees feel threatened, insecure, and under stress, ultimately those employees are found to rely more on informal communication. Similarly, newcomers in Wikiprojects are uncertain what they have to do in the initial stage. They need more time to investigate and negotiate what the groups need them to do and what they can do. Standardized task requests are often one-sided communication initiated by the groups and focus on reporting group strategy and needs without considering what newcomers want. Newcomers might think that such bureaucratic control is too much control, feel pressure, and be overwhelmed by receiving these messages, especially when they are voluntarily working for the groups (Hager & Brudney, 2004). Ultimately, this initial pressure may scare the newcomers away, so ultimately they leave the group.

Newcomers may feel less pressure when they receive more personalized messages because old-timers request tasks that are mostly based on what the newcomers already have done or make suggestions for what they might be interested in doing. Newcomers can then develop

their own way to contribute to their groups following this kind of guidance. During this process, newcomers will also have an opportunity to bridge the gap between what the groups want them to do and what they personally want to do. Consequently, they gradually assimilate into the group and also remain in the group longer, as they feel little initial stress or pressure.

In addition, when we differentiated welcome messages into two variants, we found that personalized welcoming increased newcomers' contributions, whereas standardized welcome messages did not significantly increase it. While newcomers may think that standardized welcome messages with boilerplate greeting are sent automatically by the system, they will feel more welcomed and accepted by the group when receiving personalized welcoming messages that review a newcomer's personal characteristics, such as recent work, and offer assistance. This kind of messaging may also increase the newcomers' subsequent contributions to the group.

In summary, standardized tactics can encourage newcomers to passively accept the requirements of tasks or roles, while personalized tactics may provide newcomers with opportunities to develop differentiated reactions to common situations with a unique set of learning experiences and thus adopt innovative orientations toward roles (1986). Given that people participate in different locations and are relatively anonymous in online groups, newcomers may prefer to learn about a group individually and develop their own ways to contribute to it. We can conclude then that in online groups, standardized messages that have some of the same characteristics as institutional ones are less effective in increasing newcomers' contribution to online groups than are more personalized messages.

Lastly, even though we found that only personal comments significantly increase both member contribution and retention, the pattern of coefficients for the amount of edits and survival was very similar. The only difference was with positive feedback. The failure to see the

same significant result may be because of the lower power of the survival analysis. The other reason can be that the role of personal comments does enhance newcomer emotional and social well-being. In addition to a delivery of learning, individuals also need to receive social support and feel a sense of belonging and identity (Brass, 1984; Ibarra, 1992; Podolny & Baron, 1997). Personalized comments provide friendship and acceptance, and newcomers feel emotional well-being. For example, in a recent study in an online health support group, Wang et al. (2012) found that emotional support, which is similar to personal comments, was positively associated, whereas informational support was negatively associated with how long members remained in the group. They speculated that emotional support enhanced members' relationships with one another or the group as a whole which in turn increased their feelings of commitment to the group, whereas informational support satisfied only members' short-term information needs.

3.9. Implications

There are several important theoretical and practical implications from this research. First, this study contributes to the growing stream of studies examining how to increase newcomers' commitment to their online groups by identifying socialization practices, which has been a relatively neglected area of research on online groups. This research filled the gap in current research by identifying which types and formats of socialization tactics are effective to socialize newcomers in online groups.

Our paper also provides empirical evidence for the impact of various socialization tactics by conducting a field research in a real online community. Despite the presence of an extensive social science literature examining factors that contribute to the successful socialization of

newcomers, researchers do not yet have a good sense of how groups actually execute many of the tactics mostly described in Jones's self-report questionnaire (1986). This is an understandable limitation: The methodological difficulties involved in collecting and evaluating behavioral data from naturally occurring groups in the field are profound and are challenging even in the laboratory. With an accessible archive of text-based and conversational data, we empirically tested some of the above ideas about socialization behaviors in groups and showed different impacts of various types and formats of socialization practices on newcomers' commitment to their online groups. In summary, the contribution of this research is the descriptive detail it delivers for how informal peer-to-peer interaction improves peoples' ability to work in a group or organization by using empirical data.

Our findings also have important practical implications. Our research suggests that practitioners should consider how to encourage effective socialization tactics and prevent ineffective tactics from members. Our research shows that the two socialization formats—standardized and personalized—are not equally beneficial. In particular, personalized socialization practices had positive effects on newcomers' commitment to their groups, suggesting that interfaces and mechanisms that make it easier for editors to provide more tailored task requests, connect with, and express their interests for newcomers may have the greatest benefits. On the other hand, standardized socialization practices can harm newcomers' sense making process in early stage and thus it should be prohibited.

Although these results were obtained in the context of projects within Wikipedia, we believe that the basic idea of utilize socialization tactics may generalize to other kinds of online communities and offline organizations. For example, these ideas may work well in virtual teams where cohort socialization tactics cannot be utilized due to members' different working times

and locations. We recommend practitioners pay close attention to providing social opportunities such as informal communication channels to interact with newcomers.

3.10. Limitations and Future Directions

Although there are approximately forty thousand members of over two thousand Wikiprojects, the use of hand coding limited this research to approximately 600 individuals in 22 projects. Automated coding of messages will enable us to go beyond these small samples. In addition, while Wikipedia is one of the most public and successful of the online production communities, our results may not generalize to other groups with different forms and governance structures. Future research on socialization in other groups is therefore necessary.

The paper primarily examined socialization tactics delivered via interpersonal communication between experienced members of a group and newcomers, because truly institutionalized socialization tactics like formal mentoring programs or group training were so rarely observed. To the extent, however, that the use of these institutionalized practices vary among WikiProjects, our research methods could not identity their impact, because of the small sample of 22 projects and the use of the project a control variable.

We examined groups' socialization behavior only during the initial months after newcomers joined projects. Socialization tactics may differ depending on how long newcomers have been in the group. An important extension of this paper would be to study newcomers' behavior changes on the later stage. Similarly, the data examined here are limited to socialization tactics used by the group and its experienced members. Because socialization is the process by which groups and prospective members find one another and negotiate a mutually beneficial

relationship, future studies need to investigate prospective members' information seeking as well (Levine & Moreland, 1994).

Although Wikipedia has an enormous amount of archival data, these data are observational, and the receipt of a socialization message is not a true experimental treatment. The treatment here, as with most events in real world, is endogenous in the sense that it is caused by other factors inside the system. In our data, the messages a recipient gets are partially a message to the recipient's previous behaviors. For example, old-timers invited and talk nicely to those they most want to keep. People who get positive feedback and negative feedback are different initially. Evidence of good contribution and commitment may cause the communication. Not controlling for confounding factors that influence both the treatment and the outcome can lead to biased estimation of the treatment effects. To ameliorate this endogeneity problem, we need to use propensity score matching (PSM) to approximate randomization. PSM builds experimental and control groups by balancing the groups on potential confounding factors. PSM can effectively reduce the bias caused by these conditioning factors (D'Agostino, 1998; Rosenbaum & Rubin, 1983). Although we cannot use PSM procedures because of a lack of data to find a good match in this study, we need to use them to control for endogenity in the future study. For example, we can build balanced experimental and control groups based on prior levels of activity.

Chapter 4

Effectiveness of Socialization in Online Groups: The Moderating Impact of Member Tenure and Proactive Behavior on Member Commitment to Online Groups

4.1. Introduction

Despite the rapid growth of online groups, their success has been limited because they often fail to encourage people to contribute over the long term. For example, according to the popular open source portal, SourceForge (http://sourceforge.net/), most Open Source Software (OSS) projects have ended in failure: 58% do not move beyond the alpha developmental stage, 22% remain in the planning phase, 17% remain in the pre-alpha phase, and many become inactive (Lee, et al., 2009). Online groups need to find a way to help their members adjust to their new communication environment. This process is known as socialization, the process by which newcomers make the transition from being outsiders to being insiders (Bauer, et al., 2007; Levine, et al., 2005; Morrison, 1993).

However, the research on socialization in online groups is still in its early stages. Only a handful of studies have examined the impact of socialization in online groups. Some studies have identified the positive impact on diverse outcomes of receiving messages from other members. For example, research on Usenet newsgroups suggests that the number of messages that newcomers receive is associated with their continued contributions and willingness to stay in the

group (Burke, et al., 2009; Wang, et al., Under Review), also known as retention . Research on online technical groups has shown that receiving feedback from the group increased the speed of individuals contributions (Lampe & Johnston, 2005) and how long they continue contributing(Moon & Sproull, 2008).

However, different types of communication can lead different outcomes. For example, Ahuja and Galvin (2003) found that newcomers actively engaged in discussions regarding cognitive information, but not for normative and regulative information. Choi (2012) examined the impact of five types of socialization messages (constructive criticism, positive feedback, task requests, welcoming, and personal comments) on member commitment in Wikipedia. Her results showed that task requests had a negative impact while the other types of messages had positive impacts on member commitment. However, this study is suffered from small samples.

It is likely that the status of both the senders and receivers of these messages will influence their impact, but the role of status has not received much attention in studies of online groups. The literature on socialization in offline groups has, however, identified that socialization by supervisors has a great impact on members' commitment (Bravo, et al., 2003). Since online groups rarely have a vertical hierarchy and peers socialize each other instead, we expect that members with more group experience will substitute for supervisors. That is, they can socialize other members more effectively because of their better knowledge about the group and their communication skills based on their own diverse experience. They provide more detailed explanations on how to correct mistakes as well as useful discussion about rules and guidelines that all need to know (Hovland & Weiss, 1952).

In addition, existing theory in traditional organizations suggests that early socialization experience in a members' career has a greater impact than later experiences (Klein, et al., 2006. Compared to experienced members who already know their roles and tasks and have their own networks, newcomers can reduce their uncertainty and build their connection to the group by receiving messages from others who are familiar with the group. In online groups, it is important to know if socialization is more effective for newcomers because online groups need to decide which message should be sent to whom to achieve effective socialization with limited labor. In addition, it is possible that different types of socialization can have different impacts on the outcomes of newcomers' socialization. However, it has not been tested for which types of specific socialization practices are more effective to newcomers than for experienced members in online groups.

The effectiveness of socialization will in part depend on members' active participation in the process. Individuals are not just passively socialized; they proactively shape their own socialization experiences (Miller & Jablin, 1991; Morrison, 1993; Wanberg & Kammeyer-Mueller, 2000). In particular, individuals seek information to build relationships with other members and gain feedback about their role in the group (Griffin, et al., 2000; Wanberg & Kammeyer-Mueller, 2000).

The impact of socialization on socialization outcomes can vary depending on receiver's information seeking behaviors (Gruman, et al., 2006; Kim, et al., 2005). In online groups, it is important to know whether socialization has more impact on individuals who sought information or not to accomplish more effective socialization. Although there are many studies on the information seeking behavior of newcomers (Ashford & Black, 1996; Bauer & Green, 1998; Griffin, et al., 2000; Morrison, 1993)), we know very little about the effect of individuals'

information seeking behaviors on the relationship between the actual socialization tactics used and their outcomes in both online and offline groups. Moreover, it is not entirely clear whether these interactions between socialization tactics and information seeking behaviors are positively or negatively related to socialization outcomes.

In summary, individuals' tenure and proactivity can offer important, complementary perspectives on what goes on during socialization process in online groups. The goal of the research reported here is to bring these complementary perspectives together in a single study. That is, this study investigates the moderating impact of individuals' tenure and their proactive behaviors on the relationships between different types of socialization message and the commitment of individuals in those online groups. This paper focuses on members' contribution as an indicator of member commitment to an online group. People who commit to their group are likely to contribute more; therefore, the success of an online group does rely on the voluntary contributions from their members (Ma & Agarwal, 2007; Roberts, et al., 2006b; Wasko & Faraj, 2005). To overcome the limitations of previous studies done with small samples, this study develops automated measurement models as measuring types for socialization messages with a large amount of data that includes all communications between members in 1,180 Wikiprojects.

The remainder of this paper is organized as follows: The next section reviews the prior research on socialization in online groups and builds our hypotheses. The subsequent sections report on the measurement and methods conducted in WikiProjects. The final two sections discuss both results and limitations and also implications for application and future research.

4.2. Theory and Hypothesis Development

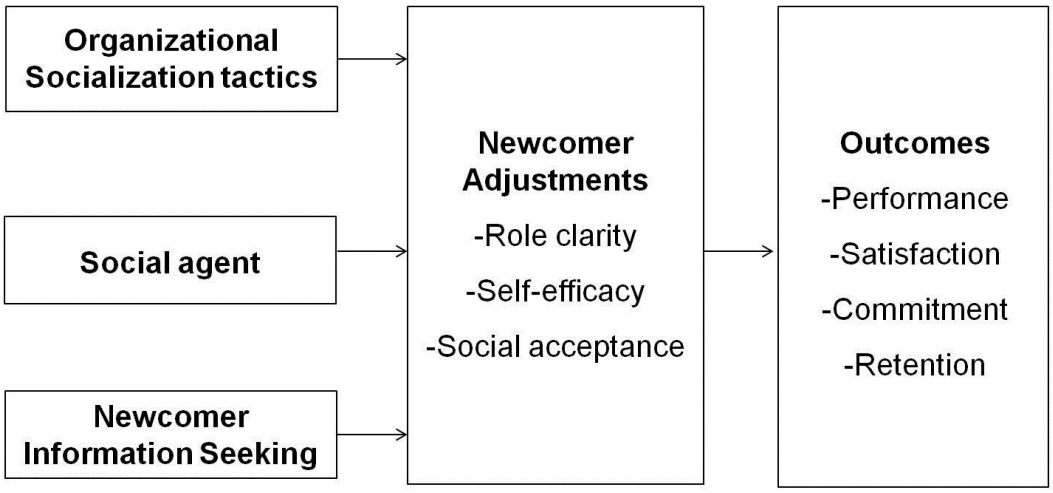

Figure 1. Model of Newcomer Adjustment During Socialization
(Bauer, et al., 2007; Klein & Heuser, 2008)

Figure 1 illustrates the socialization model that Bauer, et al. (2007) have proposed in a recent meta-analysis. The model treats role clarity, self-efficacy, and social acceptance as the three key indicators of newcomer adjustment. Researchers have frequently used these indicators as a sign of newcomer successful adjustment (Bauer, et al., 1998). Role clarity refers to understanding the tasks a job entails and thus being able to perform and understand task priorities and time allocation. Self-efficacy refers to learning the tasks of a new job and gaining confidence in the role. Social acceptance refers to the feeling of being liked and accepted by peers (Feldman, 1981). Newcomer adjustment has been associated with better outcomes, including job satisfaction, organizational commitment, job performance, intention to remain, and reduced turnover (Bauer, et al., 1998). Organizational socialization tactics, social agents, and newcomer information seeking are the proposed antecedents of positive newcomer adjustment.

The relationships between newcomer adjustment and their outcomes in Figure 1 have been studied to varying degrees in traditional work groups, and such adjustment is often found to be positively related to socialization outcomes. However, the nature of these relationships is unclear in online groups because of there has been little empirical research. Based on Bauer, et al. (2007)'s model, we propose and test other hypotheses to examine the relationship between the three antecedents of newcomer adjustment as well as their impact on socialization outcomes in online groups. Among the various possible socialization outcomes, we focus on commitment because that outcome is especially important in online groups (Bateman, et al., 2006; Wang, et al., Under Review).

4.2.1. Socialization tactics in offline and online groups

Traditional work oriented-organizations apply a variety of socialization tactics and practices to socialize newcomers, in general, using institutional and individual socialization tactics (Jones, 1986; Van Maanen & Schein, 1979). Institutional socialization tactics provide newcomers formal socialization experiences as part of a group of other newcomers in a coherent sequence, using a clear timetable and experienced mentors. In contrast, individual socialization tactics are not structured, nor formally recognized by the organization and provide more informal socialization experiences based on spontaneous relationships that occur without external involvement from the organization.

In non-employment organizations, ranging from churches and the Red Cross to informal social groups like book clubs, highly institutional or formal socialization techniques are not used much due to lack of organizational hierarchy and lack of managerial effort. Instead, socialization

in non-employment groups has mainly occurred thorough informal, peer-to-peer interaction among members. Pearce (1993) found that the level of interpersonal influence was significantly higher among volunteers than among paid workers. More frequent informal contact with other volunteers helps builds relationships, aids communication, and clarifies the expectations of volunteers. Other researchers also found that having more frequent informal contact, peer-to-peer communication, or informal mentoring with other volunteers was positively associated with their level of contributions (Farmer & Fedor, 2001).

As in volunteer groups, socialization in online groups is mainly based on informal communication among members. Research in online groups has shown that the number of messages or amount of feedback that newcomers receive increases their continued contribution to the group (Moon & Sproull, 2008; Wang, et al., Under Review). However, these studies only investigated the impact of the mere existence amount of communication between the group and the newcomers without examining its specific content or the types of feedback. Choi's (2012) study is one of the few to examine the specific socialization tactics that online groups use and their impact on newcomer commitment to a group. By examining the messages exchanged among members in 22 WikiProjects, this research identified five socialization tactics: Constructive criticism, positive feedback, task requests, welcome messages, and personal comments. ***Constructive criticism*** refers to suggesting improvements to what individuals did wrong (e.g. violation of group norms). It helps individuals learn how to perform a task correctly and how to communicate with others in a more civil manner. ***Positive feedback*** refers to praise given to individual's efforts or giving members an award. ***Task requests*** refer to asking individuals to do a certain job or task. ***Welcome*** messages refer to greeting members after they join the group. These messages signal that the group is interested in the individuals and wants to

develop a positive relationship. ***Personal comments*** aim at maintaining close social relationships to support group cohesion. These messages discuss off-topic content not related to specific group tasks. Of these, constructive criticism, positive feedback, and task request are task-oriented messages, while welcome messages and personal comments are socially-oriented messages.

Choi, et al.(2012) tested the impact of these messages on newcomer commitment to their groups only in the initial stage using small samples like most studies on socialization in online groups (Hahn, et al., 2008; Major, et al., 1995). However, these messages can also influence experienced members' commitment to their groups. Because the majority of group members are experienced members and their commitment to a group is vital (Ahuja & Galvin, 2003; Levine & Moreland, 1994; Levine, et al., 2005) the impact of socialization in experienced members should be examined. Thus, we hypothesize that both newcomers and experienced members will contribute more to a group when they receive the types of socialization messages examined by Choi (2012).

We expect that task-oriented messages (constructive criticism, positive feedback, and task request), will increase member contribution because these message reduce role ambiguity and increase self-efficacy. By receiving them, members learn what they are expected to do or not do often very specifically. The messages reduce uncertainty about roles and increase member knowledge of their groups. For example, the suggestions for improvements that constructive criticism provides will help receivers learn how to perform better. Positive feedback also increases receivers' competence in completing tasks and reaching goals by giving praise to members for their work. Task requests assign specific tasks to receivers and these tasks make them be clear about their roles. Socially-oriented messages (welcome messages and personal comments) should also increase contribution by helping receivers to feel they are liked by their

groups and maintain a strong attachment to the group as well as establish effective relationships with peers. In summary, we expect that members will contribute more to a group after they receive each type of socialization message, described previously.

H1: Each type of socialization message has a unique impact on individual commitment to an online group beyond receiving non-socialization messages.

4.2.2. The moderating role of people in the socialization process in online groups

Individuals or groups can help newcomers to online groups make sense of and develop an identity in their new environment (Ashforth, 2001; Sluss & Ashforth, 2007). These agents provide information, feedback, resources, social support, and a sense of social validation (Cooper-Thomas & Anderson, 2006). Several types of socializing agents have been identified in the literature and include supervisors, coworkers, team members, colleagues from other departments, mentors, and even other individuals outside of the organization. The literature on socialization in traditional organizations has identified that socialization by supervisors has a major impact on newcomers' commitment to the groups (Bravo, et al., 2003). The research on both offline and online groups has shown that individuals with high status tend to have greater influence than those with low status (Berger, et al., 1977; Weisband, et al., 1995) because the former are more likely to be perceived as credible and competent than the latter (Collins & Stukas, 2006).

Compared to traditional offline groups and organizations, in online groups, however, there are only a small number of formal leaders who socialize newcomers in because of their lack of vertical hierarchy (Forte, et al., 2008; Zhu, et al., 2012). Instead, regular members socialize each other as peers in online groups. Among regular members, more experienced group members are likely to have better knowledge of the group norms based on their own diverse experience. They provide more detailed explanations on how to correct mistakes as well as useful discussion about rules and guidelines that all need to know (Hovland & Weiss, 1952). Thus, members who receive messages from more experienced members compared to less experienced ones will likely better understand their roles and the tasks they need to perform and learn the tasks of the new job and thus gain confidence in their roles. In other words, communication with more experienced members will lead new members to enhance their self-efficacy and role clarity, which in turn increases their commitment to the group (Bauer, et al., 2007; Feldman, 1981). As a result, communication with more experienced members will tend to be more valuable than with members without less experience.

Moreover, experienced members are often perceived as central members of online groups and may induce a sense of relationship and identification with the group, which in turns increases the positive valence of contributing to the group (Zhu, et al., 2012). That is, messages from experienced members increase the receivers' sense of social acceptance. Thus, interacting with more experienced members of a group can be a positive experience for an individual. We expect that receiving socialization messages from more experienced members will lead to a stronger relationship with the full group. In other words, experienced members will be more powerful in terms of influencing members' contribution than will new members in online groups.

H2: Socialization messages will have a greater impact on individual contribution to an online group when senders are experienced members and not newcomers.

While social interactions should generally affect the formation and maintenance of relationships within an online group, individuals will be especially sensitive to socialization messages during their early experiences with the group. Group socialization theory suggests that evaluation and engagement processes are qualitatively different after people transition from being newcomers to being experienced members (Moreland & Levine, 1988, 2001). Uncertainty reduction theory also suggests that newcomers desire to increase the predictability of the interactions between themselves and others within the new environment (Berger & Calabrese, 1975).

Similarly, newcomers in online groups actively explore and evaluate their fit with the group through their interactions with it (Ahuja & Galvin, 2003; Burke, et al., 2009; Wang, et al., Under Review). Therefore, socialization messages are more informative for newcomers who just started interacting with a group than those messages would be for those with more experience with the group. Because individuals gain direct experience and knowledge about a group by actively engaging with it, their level of prior participation serves as an indication of how much they know about the group and the extent to which a relationship with it already exists or is still developed. In other words, compared to experienced members who already know their roles and tasks and have their own networks, newcomers can reduce their uncertainty and build their connection to the group by receiving messages from others who are familiar with the group.

Therefore, the impact of socialization messages should to be higher for newcomers when they initially interact with a group than when they have more experience with it.

H3: Socialization messages will have a greater impact on individual contribution to an online group when receivers are newcomers than when they are experienced members.

4.2.3. The moderating role of information seeking in the socialization process in online groups

Proactive behaviors refer to any behavior through which a newcomer takes the initiative to improve existing circumstances or create new ones, such as information seeking, positive framing, and relationship building (Crant, 2000). Of the three, information seeking is the most present proactive behavior that helps individuals learn about their abilities, better understand their work environment and tasks, and adjust their behavior to improve socialization and career success (Ashforth & Saks, 1996; Morrison, 1993; Ostroff & Kozlowski, 1992; Wanberg & Kammeyer-Mueller, 2000). Individuals who engage in information- seeking acquire more necessary information, and this information allows them to make better sense of their surroundings (Gruman, et al., 2006). Socialization studies on online groups also found that information- seeking behaviors are associated with positive outcomes (Burke, et al., 2009).

Even though many studies have investigated the impact of information- seeking on socialization outcomes, both for offline and online groups, not much research as yet has been done on the moderating effect of information- seeking on the relationship between socialization and group outcomes (Gruman, et al., 2006; Kim, et al., 2005). Kim et al. (2005) found that information- seeking did not moderate the effect of institutional socialization tactics on the

Person-Organization fit and noted that information seeking had no moderating effect because gathering information may show some employees how well they fit, but show other employees that they do not fit at all. Thus, the act of seeking organizational information can lead to positive outcomes for some and negative outcomes for others, depending on each individual's values and needs, thus producing a null effect across the full range of employees.

In contrast, Gruman (2006) found that the relationship between institutional socialization tactics and outcomes was much stronger for newcomers who engaged in less information-seeking because information- seeking behaviors substituted for the socialization tactics. The main reason that firms use institutionalized tactics is to remove the uncertainty of a new environment by offering information that can guide new employee behavior (Bauer, et al., 1998; Jones, 1986; Van Maanen & Schein, 1979). To the extent that some employees gain this information about norms, expectations, and standards themselves, these individuals preempt a key element of the institutionalized socialization process. In other words, if newcomers obtain information and social support through their own efforts, then the beneficial effects of any socialization tactics by others will decrease.

Similarly, proactive members of online groups can access other sources to gain more knowledge and fulfill their needs. For example, they can read more of the guidelines, observe the behaviors of others, or obtain information from people not in the group. Thus, the benefits of receiving unsolicited feedback from other members can diminish. On the other hand, individuals who do not ask any questions, but receive feedback on their efforts or task requests could reduce their uncertainties and be clearer about their roles in these online groups. Individuals who receive friendly messages before they ask any questions will think that the group cares about them and they belong to that group.

H4: Socialization message has will have a greater impact on individual commitment to an online group when receivers have not sought information earlier.

4.3. Methodology

4.3.1. Study Platform

Wikipedia is a large, task-focused community whose goal is to produce a free online encyclopedia. Wikipedia is a highly popular website, with over 17 million registered editors as of August, 2012, and over 4 million content pages ("Wikipedia Statistics,"). Most newcomers on Wikipedia start reading articles as unregistered invisible users. When they become interested in noting mistakes or omissions, and correcting them, they then edit articles. When they first contribute to articles, new users tend to make only minor changes (Bryant, et al., 2005). At some point, they decide to create an account on Wikipedia, which allows them to create a watch list, track their own contributions, and maintain a consistent identity on the site. When a user registers a username, they automatically create a user page on which they can provide their biographical information and a user talk page on which other members can post messages. As new users become experienced members of Wikipedia, they will make major changes to improve the quality of articles and take managerial roles as well to grow the community overall (Bryant, et al., 2005).

Some editors who become interested in editing particular topics and put their efforts toward specific articles, they join a Wikiproject, which are collections of editors interested in improving the coverage and quality of articles in a particular domain. For example, members in the Georgia WikiProject create, assess, and improve pages related to the history, geography,

culture, and other attributes of the US state of Georgia. Each WikiProject has a dedicated page that exists in a namespace that is separate from regular article content. Editors can join a project simply by adding their name to the member list on this dedicated page, although some projects move members who have not been active enough to an inactive list.

Most WikiProject members are registered users and have editing experience on other Wikipedia articles as well. Eighty percent of people who join WikiProjects had prior experience in Wikipedia before joining. WikiProjects provide mechanisms for members to self-identify and acknowledge each other. Members can place project banners on their user pages to show identification with the project. Identification with a project seems to influence their behavior. After editors join a Wikiproject, they direct more of their work to articles within the scope of the project (Kittur, et al., 2009). In addition, because WikiProjects are social groups with a smaller numbers of members compared to Wikipedia as a whole, they will bond together and take care of their newcomers more. WikiProject members tend to respect other members' opinions and talk in a more civil manner compared to Wikipedia where some users use harsh words to punish other users when they violate norms or there is conflict between them on edits.

WikiProjects provide a good setting to study the effects of socialization in online volunteer groups for several reasons. First, WikiProjects are particularly interesting to group researchers because they incorporate many characteristics of traditional work groups even though they are online. For instance, WikiProject members set goals, develop task criteria, maintain diverse collaborative processes, keep track of work that needs to be done, discuss issues of interest using a forum, develop project-specific norms, and reward each other for good performance. In addition, like a "real-world" work group, the success of a Wikiproject depends on the editors' ability to function as a coordinated group working toward a common goal. In

summary, WikiProjects have many characteristics in common with other online (and offline) work groups and therefore, we believe our findings can be generalized.

Second, we can examine the socialization processes executed by WikiProject members and measure members' contribution to the group and membership changes over time. We can tract WikiProject memberships, by examining the member list on which members can sign or remove their names. The WikiMedia foundation provides rich historical data, including all communications among members. In contrast withWikipedia, which has both unregistered and registered users, WikiProject members are all registered users with user talk pages Thus, WikiProjects provide a lens through which to examine the socialization processes executed by WikiProject members.

4.3.2. Measurement of socialization behaviors

We measured socialization behaviors as exchanged in communication among WikiProject members by examining the messages members left on each others' user talk pages. We examined the five message types found in Choi (2012)'s study. However, we combined the welcoming message category and the personal message category into a social message category because they were conceptually and empirically correlated ($\gamma = 0.68$). In addition to four types of messages, we also measured information seeking. In summary, we classified messages into five categories of socialization behaviors: constructive criticism, positive feedback, task request, social messages and information seeking.

This research used a machine learning approach to classify messages into these five types of socialization communication. Specifically, we trained statistical models on a small set of

human-coded data and validated them using a separate set of human-coded data. Then we applied the model to a larger data set that had not been human coded. A machine learning approach has three main components - training sets, representation of messages for machine learners (feature sets), and training algorithms, which we explain in more detail in the following sections.

4.3.2.1. Creating the Human-Coded Dataset

To construct a hand-coded dataset for training the machine learning models, we randomly selected 999 messages and employed Amazon Mechanical Turk (MTurk) workers to categorize each message. Amazon Mechanical Turk is an online marketplace for crowd sourcing. Snow, et al. (2008) and Yichia, et al. (2012) have shown that the combined judgments of a small number (about 5) naïve annotators on MTurk will lead to classfications of texts that are very similar to those of experts. Snow, et al. (2008) showed high agreement between MTurk annotations and existing gold standard labels provided by expert labelers for affect recognition, word similarity, recognizing textual entailment, event temporal ordering, and word sense disambiguation. Yichia, et al. (2012) also showed that MTurk workers' judgments of emotional and information support were similar to those of highly trained, expert coders. We posted the definition of each message type with examples, and workers selected all categories that applied as shown in Appendix A. We first identified MTurk workers familiar with Wikipedia by posting a knowledge test that aimed to measure Wikipedia experience and knowledge(e.g. how many articles have you worked on? What does NPOV mean in Wikipedia?). Then, we only used those answers that passed the test. We posted these messages until each message was judged by five people who passed the knowledge tests. To assess the reliability of the workers' ratings, we calculated intra-class correlation coefficients for each message type (Koch, 1982). Intra-class correlation is often used

to assess the consistency of quantitative measurements, when different judges are rating different objects. The intra-class correlation for each socialization message is reported in Table 1.

	Positive Feedback	Constructive Criticism	Task Request	Social	Information Seeking
ICC	0.80	0.73	0.73	0.78	0.80

Table 1. Intra-class correlation (ICC) for each type of message

4.3.2.2. Learning Algorithms

Our goal was to build classify messages into the five types of socialization messages from a set of features relevant to Wikipedia. Since the five categories were not exclusive, we formulated the classification task as five binary decision problems. We conducted experiments on the training set with varied learning algorithms implemented in Weka (Witten & Frank, 2005), including decision trees, Adaboost on decision trees, Naïve Bayes, and linear Support Vector Machines (SVM). We chose to use linear SVM because it worked consistently well on all five categories.

We used features composed of words and phrase patterns frequently used in Wikipedia to express different intents. These features included 21 domain features from Zhu, et al.'s (2012) and 20 more domain features added here, for a total of 41 features (see Appendix B.). Here are some examples of the features.

- Negative jargon (neg_jargon): Frequency of negative Wikipedia-specific jargons such "vandalism" and "nonfree"
- Barnstar: Frequency of the words presenting awards such as "barnstar", "reward" and "medal".

- Advice: Frequency of the words used to give advice such as “citation”, “reference”, and “guidelines”
- Suggestion: Frequency of suggestion phrases such as “check out” and “may be interested”
- Question terms (qterms): Frequency of question words such as “who is”, “where is”, and “what is”
- Seekinfo: Frequency of information seeking phrases such as “any idea” and“share any information”

4.3.2.3. Classification Results

The results of ten-fold cross-validation of the trained model are represented by Kappa as agreement between machine and human judges (Table 2). The Kappa for all categories was substantial or excellent. We also applied our models on another 447 annotated messages to assess whether the models overfit the training set. The results on this extra test set were quite good, with an average kappa of 0.67. The SVM classifiers calculate the weighted sum of feature counts plus intercept for each message, then determines whether the message belongs to the category depending on whether the sum is larger than 0 or not.

	Constructive Criticism	Positive Feedback	Task Request	Social	Information Seeking	Number
Kappa (Train)	0.68	0.74	0.67	0.78	0.67	999
Kappa (Test)	0.61	0.75	0.61	0.70	0.65	477

Table 2. Agreement between classification and human judges in the training and test data. (Results often-fold cross-validation using the SVM algorithm)

The most important features predicting each of the five types of socialization behavior are listed below. Table 3 reports those features with the top ten large weights. Appendix C. shows the weight for each feature in five categories.

	Constructive Criticism		Positive Feedback		Task Request		Social		Information Seeking	
	Weight	Feature	Weight	Feature	Weight	Feature	Weight	Feature	Weight	Feature
1	5.99	Neg_jargon	6.03	Acknowledge1	4.86	Suggestion	9.99	Greetings	5.80	Qterms
2	-4.23	Smiley	3.90	Welcome	3.47	Collaboration	8.00	Acknowledge1	2.61	Seekinfo
3	-3.94	Acknowledge1	3.59	Barnstar	-2.31	Rfa	-4.67	Neg_jargon	-1.87	Strongpos
4	3.07	Length	3.50	Congrats	2.28	Invitation	4.50	Smiley	1.30	<modal+you>
5	3.00	Negation	-2.48	<modal+you>	2.20	Rollcall	4.33	Barnstar	-1.01	Newsletter
6	-2.39	Apology	-1.50	Collaboration	1.73	<modal+you>	4.00	acknowledge2	-1.00	Rollcall
7	-2.35	Barnstar	1.46	Neg_jargon	-1.67	Neg_jargon	4.00	Apology	0.90	Questionmark
8	-2.33	Dyk (Did you Know)	-1.18	Newsletter	-1.45	Negation	4.00	Welcome	0.56	<modal+I>
9	-2.33	Wikiproject	1.12	Length	1.20	Acknowledge2	3.00	Invitation	0.50	Negation
10	1.95	Advice	-1.04	Suggestion	-1.17	Advice	2.50	Congrats	0.39	Congrats

Table 3. Features with Top 10 large weights in the SVM Models

4.4. Analysis

The goal of this study is to identify the effect of receiving different types of socialization messages and interaction with members' tenure and information seeking to predict subsequent contribution. We will compare the amount of subsequent contribution of those who received socialization messages (treated group) to those who do not receive messages (control group).

4.4.1. Data preparation

The data were longitudinal, following the same editors across different weeks. For the analysis, we first defined whether an editor was active in a given week (the focal week) in terms of whether the editor made any edits during a five-week period (centered on the focal week, two weeks before and two weeks after the focal week). Then we did an editor-week level analysis, restricted to the weeks in which the editor was active. As a result, our data comprised 29,095 unique editors and 6,563,411 editor-week observations within 1180 WikiProjects.

4.4.1.1. Dependent Variables

• Commitment. We selected the members' subsequent contribution to measure their commitment. Unlike paid workers who contribute to their organizations mainly because of momentary rewards they receive and employment contracts they are tied to, volunteers contribute to their groups only when they are committed to their groups (Farmer & Fedor, 2001). Thus, volunteers' contribution levels are highly variable depending in part on their commitment levels. We measured editors' contributions towards Wikiproject articles through their number of edits. Each edit indicates a set of editing actions, for example adding, changing, deleting or reverting text, references or illustrations, or communicating with other editors. The dependent measure was the log transformed number of edits in the week after the focal week. Because the logarithm of zero is undefined, we added one before computing the logarithm.

4.4.1.2. Independent Variables

Our major independent variables measure whether an editor received no communication during the focal week, non-socialization messages or one of the four types of socialization messages.

• Non-socialization Messages. This dummy variable indicates whether the editor received any messages but the t socialization messages (e.g. constructive criticisms, positive feedback, task requests, social messages) during the focal week. One indicates that the editor received at least one non-socialization message, while zero indicates that the editor received no messages.

• Constructive criticism. This dummy variable indicates whether in the focal week the editor received any message categorized as constructive criticism. One indicates that the editor received at least one message with constructive criticism, and zero indicates that the editor received no constructive criticism. The following three variables are similar.

• Positive Feedback. This dummy variable indicates whether the editor received any message categorized as positive feedback during the focal week or not.

• Task request. This dummy variable indicates whether the editor received any message categorized as task request during the focal week or not.

• Social Message. This dummy variable indicates whether the editor received any message categorized as social exchange during the focal week or not.

4.4.1.3. Moderating variables

• Sender's Wikipedia Tenure. We measured a message sender's tenure in Wikipedia: how long the sender has been a member of Wikipedia when he/she send the message. We used tenure

in Wikipedia rather than amount of prior participation as a measure of experience because it captures the experience that lurkers gain by observing the interactions of other editors (Chen, et al., 2010; Preece, et al., 2004). We also chose Wikipedia tenure over WikiProject tenure because experience in Wikipedia as a whole transfers readily to projects (Chen, et al., 2010). For example, a lot of the rules, norms and culture in projects are inherited from the norms in Wikipedia itself. Thus, editors who learn how Wikipedia rules work can apply those rules within any WikiProject.

• Receiver's Wikipedia Tenure. This variable indicates a message receiver's tenure in Wikipedia: how many weeks an editor has been a member of Wikipedia when he/she received the message during the focal week.

• Information seeking. This dummy variable indicates whether the editor sent any message categorized as information seeking during the focal week.

4.4.2. Propensity score matching

In our data, the messages a receiver receives are partially a response to the receiver's previous behaviors. For example, the number of edits an editor made in a previous week may cause others to send them messages in the next week. Similarly, editors who produce good edits may cause others to send them positive feedback whereas those who produce poor edits may cause others to send them constructive criticism in a subsequent week. Not controlling for confounding factors that influence both the treatment and the outcome can lead to biased estimation of the treatment effects.

To ameliorate the endogeneity problem, we use propensity score matching (PSM) to approximate randomization. PSM builds experimental and control groups by balancing the groups on potential confounding factors. PSM can effectively reduce the bias caused by these

conditioning factors (D'Agostino, 1998; Rosenbaum & Rubin, 1983). In an analogy to a true experiment, we used PSM to compare the changes in editing behavior of those who received a type of socialization message (treated group) to those who do not receive that type of socialization message (control group). Since we are interested in four types of socialization messages, we applied PSM four times, one for each type.

Propensity score matching (PSM) involved three steps. First was to estimate the propensity score (i.e., the probability of receiving a type of socialization message from others) from a set of conditioning variables. The variables we used to predict receiving a message were the editors' prior activities (e.g., number of edits in previous week, number of messages received in previous week, tenure in Wikipedia). The rationale was that these factors might both cause other editors to communicate with them and also be correlated with subsequent edits in effort. Therefore, we chose ten of the editors' previous activities listed below as conditioning variables. The ten predictors are listed below.

- Edits (t-1). The log of the number of edits done by the editor in the week before the focal week.
- MsgReceived (t-1). The log of the number of messages the editor received in the week before the focal week.
- MsgSent (t-1). The log of the number of messages the editor sent in the week before the focal week.
- MsgReceived (<t-1). The log of the total number of messages the editor received any time prior to the week before the focal week.
- MsgSent (<t-1). The log of the aggregate number of messages the editor sent any time prior to the week before the focal week.
- Tenure. The number of weeks between the editor's first edit and the focal week.

- Constructive criticism (t-1). The log of the number of constructive criticism the editor sent in the week before the focal week.
- Positive feedback (t-1). The log of the number of positive feedbacks the editor sent in the week before the focal week.
- Task request (t-1). The log of the number of task requests the editor sent in the week before the focal week.
- Social Message (t-1). The log of the number of social messages the editor sent in the week before the focal week.

In the second step, we matched each editor who received a type of socialization message in a focal week with another editor who did not receive that type of socialization message, but who had the most similar propensity score based on the ten conditioning variables. Propensity scores allow researchers to control for many variables simultaneously by matching on a single scalar variable. The variables with higher correlation with the treatment (also having higher risk to introduce bias) will be balanced better than the variables with lower correlation with the treatment. At the end of the second step, we were able to check whether the treatment group and control group were well matched in terms of the conditioning variables we were interested in.

Variable	% reduction in the bias after matching			
	Constructive Criticism	Positive Feedback	Task Request	Social Messages
Edits(t-1)	99.70	95.50	93.00	95.20
MsgReceived(t-1)	98.20	95.80	99.50	99.30
MsgReceived(<t-1)	97.00	17.30	24.60	37.70
MsgSent(t-1)	95.40	87.70	98.30	99.30
MsgSent(<t-1)	92.00	77.20	-69.90	53.80
Tenure	94.60	80.60	90.20	89.20
Constructive criticism(t-1)	95.20	98.90	88.80	88.20
Positive feedback(t-1)	84.60	91.30	92.50	90.20
Social message(t-1)	90.20	94.40	90.90	90.40
Task request(t-1)	90.60	94.50	91.40	90.80

Table 4. The percentage reduction in the bias after matching

Table 4 reports the percentage reduction in the bias after matching showing that the treatment group and the control group are balanced. While editors who received or failed to receive a type of message differed substantially on all the conditioning variables before matching, they had similar means after matching. Appendix D. shows means, percentage of bias and the percentage reduction in the bias after matching for each message type.

In the third step, we ran fixed effects regression analyses to estimate the effect of receiving one type of socialization messages on the treated groups and matched controls. Fixed effects allowed each pair to have different intercept (pre-existing difference among pairs were embodied in the intercept); independent variables—the event of receiving messages and types of messages—determined the scope, which was the same for all pairs (indicating an average effect of receiving certain type of messages).

4.5. Results

Table 5.1-5.4 shows the descriptive statistics and correlations for each socialization message.

	Variable	Obs	Mean	S.D.	Min	Max	1	2	3	4	5	6
1	Edits in project-related Page(Log)	20338	1.720	2.310	0	9.785	1.000					
2	Other Messages	20338	0.409	0.492	0	1	-0.080	1.000				
3	Constrcutive Criticism	20338	0.500	0.500	0	1	0.172	-0.832	1.000			
4	Receiver's Wikipedia Tenure (Weeks)	20338	52.162	41.377	0	305	-0.054	0.054	-0.007	1.000		
5	Sender's Wikipedia Tenure (Weeks)	20338	38.931	59.802	0	202	0.136	-0.542	0.651	0.008	1.000	
6	Information Seeking	20338	0.079	0.270	0	1	0.116	-0.133	0.184	-0.042	0.142	1.000

Table 5.1. Descriptive statistics and correlations for constructive criticism

	Variable	Obs	Mean	S.D.	Min	Max	1	2	3	4	5	6
1	Edits in project-related Page(Log)	10578	1.647	2.294	0	10.091	1.000					
2	Other Messages	10578	0.395	0.489	0	1	-0.072	1.000				
3	Positive Feedback	10578	0.500	0.500	0	1	0.177	-0.809	1.000			
4	Receiver's Wikipedia Tenure (Weeks)	10578	47.480	40.291	0	278	-0.008	0.056	0.035	1.000		
5	Sender's Wikipedia Tenure (Weeks)	10578	38.251	59.282	0	204	0.142	-0.522	0.645	0.016	1.000	
6	Information Seeking	10578	0.074	0.262	0	1	0.115	-0.060	0.117	-0.010	0.105	1.000

Table 5.2. Descriptive statistics and correlations for positive feedback

	Variable	Obs	Mean	S.D.	Min	Max	1	2	3	4	5	6
1	Edits in project-related Page(Log)	19248	1.601	2.238	0	10.091	1.000					
2	Other Messages	19248	0.399	0.490	0	1	-0.043	1.000				
3	Task Request	19248	0.500	0.500	0	1	0.144	-0.814	1.000			
4	Receiver's Wikipedia Tenure (Weeks)	19248	49.542	40.496	0	278	-0.048	0.018	0.017	1.000		
5	Sender's Wikipedia Tenure (Weeks)	19248	39.022	59.822	0	200	0.115	-0.531	0.652	0.014	1.000	
6	Information Seeking	19248	0.067	0.250	0	1	0.130	-0.074	0.125	-0.027	0.098	1.000

Table 5.3. Descriptive statistics and correlations for task request

	Variable	Obs	Mean	S.D.	Min	Max	1	2	3	4	5	6
1	Edits in project-related Page(Log)	19744	1.578	2.248	0	10.091	1.000					
2	Other Messages	19744	0.395	0.489	0	1	-0.052	1.000				
3	Social message	19744	0.500	0.500	0	1	0.152	-0.808	1.000			
4	Receiver's Wikipedia Tenure (Weeks)	19744	49.526	40.479	0	294	-0.030	0.020	0.016	1.000		
5	Sender's Wikipedia Tenure (Weeks)	19744	38.389	59.200	0	199	0.119	-0.524	0.649	0.011	1.000	
6	Information Seeking	19744	0.068	0.252	0	1	0.130	-0.063	0.117	-0.021	0.096	1.000

Table 5.4. Descriptive statistics and correlations for social message

Table 6 shows the impact of receiving a socialization message on an individual's contribution to his/her group. Each column shows the result of the impact of each socialization message based on four samples from propensity score matching.

Variables	Socialization message							
Edits in project-related Page(Log)	Constructive Criticism		Positive Feedback		Task Request		Social Message	
	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.
Other Messages	0.685**	0.080	0.509**	0.105	0.625**	0.076	0.444**	0.074
Socialization Message	1.355**	0.072	1.214**	0.093	1.142**	0.068	1.033**	0.065
Intercept	0.762**	0.069	0.838**	0.088	0.782**	0.064	0.886**	0.062
Number of Observation	20338		10578		19248		19744	
Number of Group	10169		5289		9624		9872	

Table 6. The impact of socialization messages on member contribution
(** $p < 0.01$, * $p < 0.05$)

Each coefficient represents the log base 2 of the expected number of edits the editor will produce when increasing the independent variable by one unit, when other variables in the model are held constant at zero. For example, the intercept for constructive criticism indicates that individuals who received no messages can be expected to make 1.696 (2^.0.762) edits to the focal article. Individuals who received other messages can be expected to make an additional 1.031 (2^(0.762+0.685)- (2^.0.762)) edits and those who received constructive criticism can be expected to make an additional 2.642((2^(0.762+1.355)-)- (2^.0.762))edits compared to those who received no messages.

Hypothesis 1 expected that each socialization message would increase individuals' subsequent edits. The result shows that editors who received non-socialization messages in a focal week subsequently edited more than did those who did not ($\beta = 0.685$, $p < 0.01$; $\beta = 0.509$, $p < 0.01$; $\beta = 0.444$, $p < 0.01$; $\beta = 0.625$, $p < 0.01$). Receiving each type of socialization message

led to additional increases in subsequent edits (constructive criticism: $\beta = 1.355$, $p < 0.01$; positive feedback: $\beta = 1.214$, $p < 0.01$; task request: $\beta = 1.142$, $p < 0.01$; social message: $\beta = 1.033$, $p < 0.01$), thus supporting H1. Figure 2 shows the impact of each socialization message on member contribution.

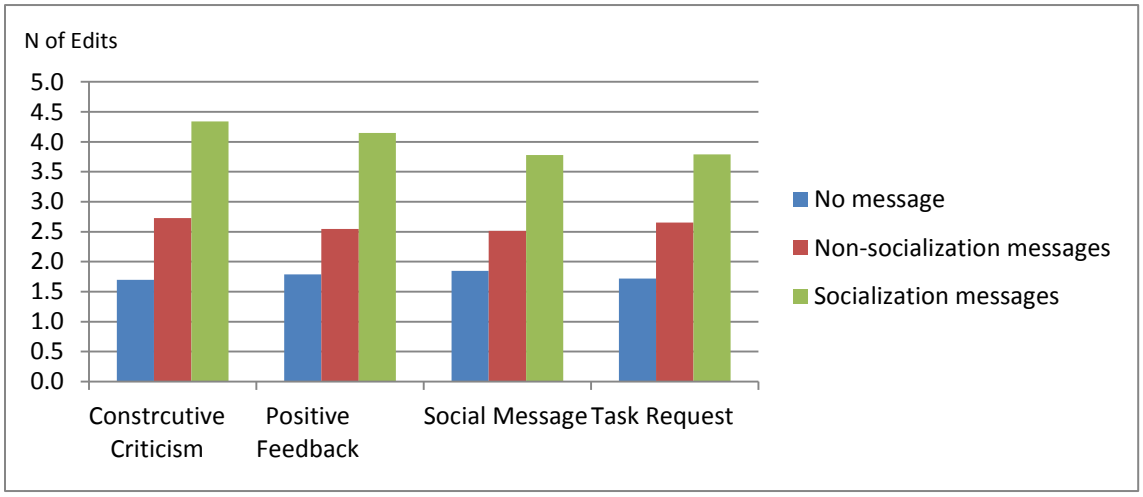

Figure 2. The impact of different types of socialization message on member contribution

Variables	Socialization Message							
Edits in project-related Page(Log)	Constructive Criticism		Positive Feedback		Task Request		Social Message	
	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.
Non-socialization Messages	0.691**	0.080	0.507**	0.105	0.626**	0.076	0.449**	0.074
Socialization Message	1.239**	0.080	1.048**	0.106	1.033**	0.077	0.896**	0.074
Sender's Wikipedia Tenure	0.002**	0.000	0.002**	0.001	0.001**	0.000	0.002**	0.000
Intercept	0.757**	0.069	0.840**	0.088	0.781**	0.064	0.882**	0.062
Number of Observation	20338		10578		19248		19744	
Number of Group	10169		5289		9624		9872	

Table 7. The moderating impact of sender's tenure on the relationship between different types of socialization message and member contribution (** $p < 0.01$, * $p < 0.05$)

Table 7 shows the moderating impact of sender's tenure on the relationship between each socialization message and member contribution. It shows receiving socialization messages from more experienced member compared to less experienced ones led to an increase in subsequent edits (constructive criticism: $\beta = 0.002$, $p < 0.01$; positive feedback: $\beta = 0.002$, $p < 0.01$; task request: $\beta = 0.001$, $p < 0.01$; social messages: $\beta = 0.002$, $p < 0.01$), supporting H2 (Table 6).

Figure 3 shows the graph of this result. To show the difference in number of edits between individuals who received each socialization message from newcomers and experienced members, newcomers are defined as members whose tenure is zero and experienced members are defined as members with two standard deviations more weeks of tenure in Wikipedia. . For example, the standard deviation of sender's tenure for constructive criticism is 59.8. Thus, the experienced members are individuals whose tenure is 119.6 (59.8*2) weeks. The first graph of Figure 3 shows that individuals who received constructive criticism from experienced members made 0.719((2^(0.757+1.239+0.002*119.604)) - (2^(0.757+1.239))) more edits than those who received it from newcomers.

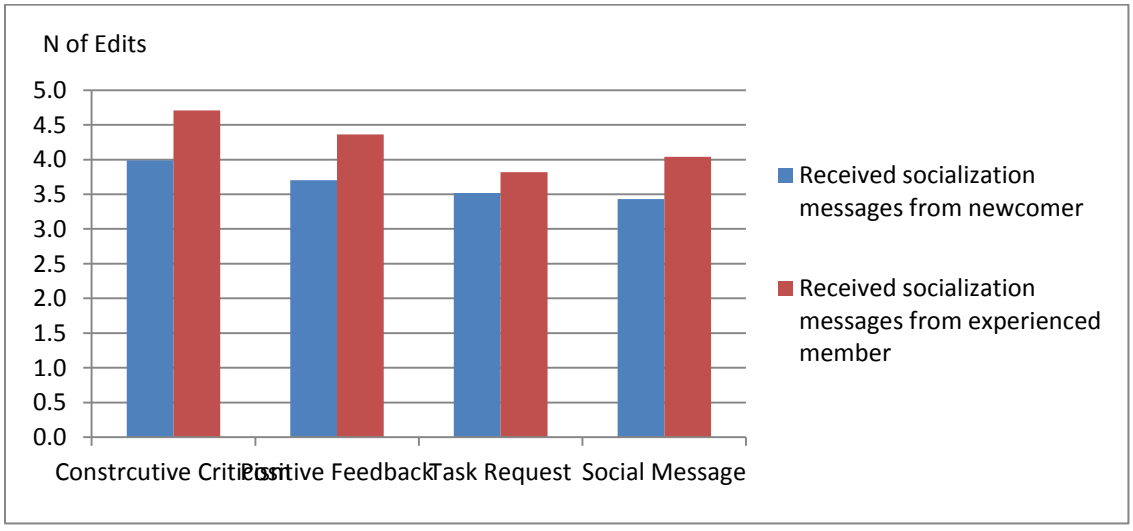

Figure 3. The moderating impact of sender's tenure on the relationship between different types of socialization message and member contribution

Variables	Socialization message							
Edits in project-related Page(Log)	Constructive Criticism		Positive Feedback		Task Request		Social Message	
	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.
Non-socialization Messages	0.672**	0.080	0.489**	0.107	0.631**	0.076	0.433**	0.074
Socialization Message	1.432**	0.081	1.221**	0.102	1.054**	0.077	1.070**	0.075
Receiver's Wikipedia Tenure	0.000	0.001	0.002	0.001	0.000	0.001	0.002	0.001
Socialization Message *Receiver's Wikipedia Tenure	-0.002*	0.001	-0.001	0.001	0.002*	0.001	-0.001	0.001
Intercept	0.759**	0.082	0.758**	0.103	0.788**	0.081	0.804**	0.079
Number of Observation	20338		10578		19248		19744	
Number of Group	10169		5289		9624		9872	

Table 8. The moderating impact of receiver's tenure on the relationship between each socialization message and member contribution (** $p < 0.01$, * $p < 0.05$)

Table 8 shows the moderating impact of receiver's tenure on the relationship between each socialization message and member contribution. Even though members contribute more to their project as they are more experienced, the impact was not significant ($\beta = 0.000$, $p > 0.05$; $\beta = 0.002$, $p > 0.05$; $\beta = 0.000$, $p > 0.05$; $\beta = 0.002$, $p > 0.05$). The negative coefficient for the interaction term between receiver's tenure and constructive criticism ($\beta = -0.002$, $p < 0.05$) indicates that the impact of receiving constructive criticism was larger when receivers were newcomers rather than experienced members, thus supporting H3. The impact of receiving positive feedback and social messages did not differ between newcomers and experienced members ($\beta = -0.001$, $p > 0.05$; $\beta = -0.001$, $p > 0.05$). In contrast, receiving a task request had a larger impact on individuals' contributions when receivers were experienced members rather than newcomers ($\beta = 0.002$, $p < 0.05$), opposite to H3.

Figure 4 shows the graph of this result. Experienced members who received constructive criticism made 0.495 (2^(0.759+1.432+0.002*82.754)- 2^(0.759+1.432)) less edits than newcomers who received constructive criticism made. Experienced members who received task

request made 0.426 (2^(0.788+1.054+0.002*98.992)- 2^(0.788+1.054)) more edits than newcomers who received task request.

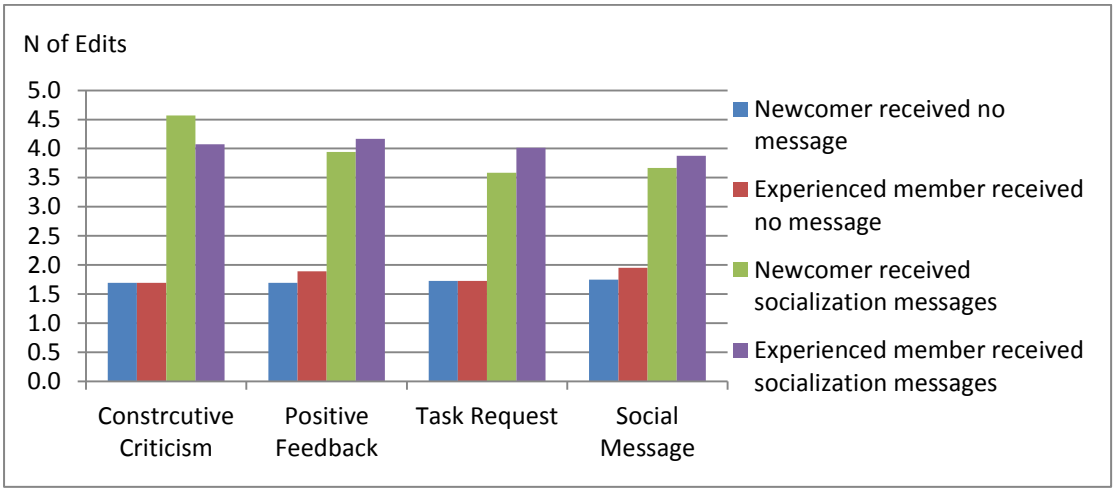

Figure 4. The moderating impact of receiver's tenure on the relationship between each socialization message and member contribution

Edits in project-related Page(Log)	Constructive Criticism		Positive Feedback		Task Request		Social Message	
	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.	Coef.	S.D.
Non-socialization Messages	0.644**	0.080	0.451**	0.105	0.583**	0.076	0.406**	0.074
Socialization Message	1.341**	0.072	1.205**	0.093	1.132**	0.068	1.023**	0.066
Information Seeking	1.322**	0.182	1.332**	0.210	1.293**	0.165	0.999**	0.157
Socialization Message * Information Seeking	-1.168**	0.200	-1.128**	0.240	-1.069**	0.190	-0.807**	0.183
Intercept	0.756**	0.069	0.826**	0.088	0.769**	0.064	0.878**	0.062
Number of Observation	20338		10578		19248		19744	
Number of Group	10169		5289		9624		9872	

Table 9. The moderating impact of information seeking on the relationship between different types of socialization message and member contribution (** $p < 0.01$, * $p < 0.05$)

Lastly, Table 9 shows the moderating impact of information seeking on the relationship between each socialization message and member contribution. It shows information seeking behavior was associated with an increase in subsequent edits ($\beta = 1.322$, $p < 0.01$; $\beta = 1.332$, $p <$

0.01; β = 1.293, p < 0.01; β = 0.999, p < 0.01), whereas the negative coefficient for the interaction term between information seeking and each socialization message type (β = -1.168, p < 0.01; β = -1.128, p < 0.01; β = -1.069, p < 0.01; β = -0.807, p < 0.01) suggests that unsolicited socialization messages had more impact than ones following information seeking, thus supporting H4.

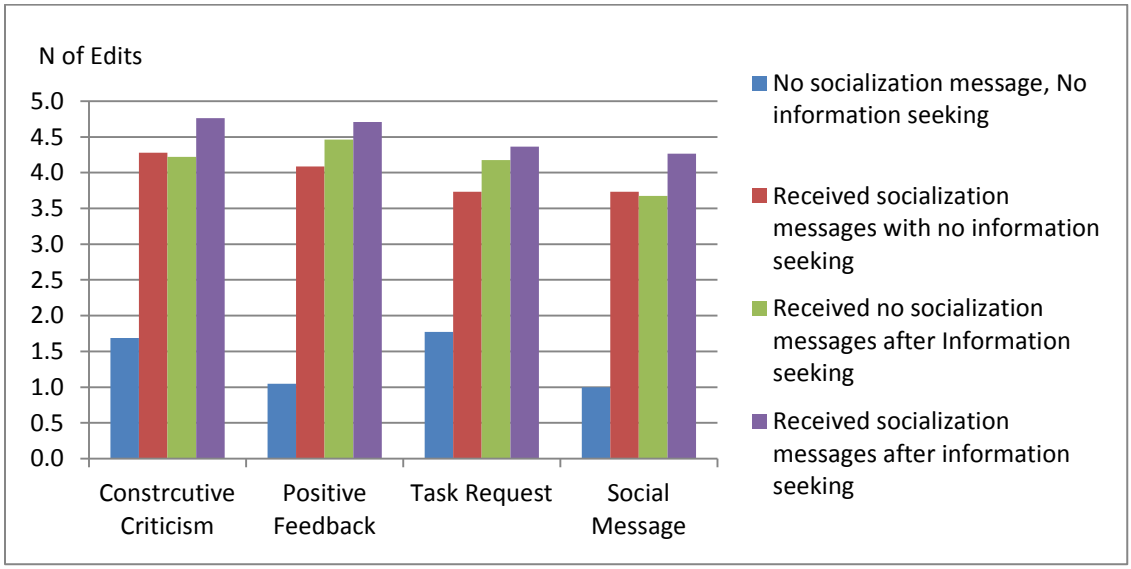

Figure 5. The moderating impact of information seeking on the relationship between each socialization message and member contribution

Figure 5 shows the graph of this result. For example, the intercept for constructive criticism indicates that individuals who received no messages can be expected to make 1.689 (2^.0.756) edits to the focal article. Individuals who received unsolicited constructive criticism can be expected to make 2.589(2^(0.756+1.341)-2^0.756) more edits and those who sought information can be expected to make 2.553(2^(0.756+1.322)- 2^0.756)) more edits who compared to those with no messages. Individuals who received constructive criticism after

seeking information can be expected to make 3.071 (2^(0.756+1.341+1.351-1.168)- 2^0.756)) more edits compared to those who with no messages.

4.6. Discussion

Our results show that receiving any types of socialization messages – constructive criticism, positive feedback, task requests, or social message -- was associated with increases in the amount that individuals subsequently edited. It is interesting to contrast the results in the current study to the results from Zhu, et al. (2012) who found that some type of feedback lead editors to edit less. Zhu, et al. (2012) identified the impact of shared leadership, i.e., an interactive influence process among group members that led them to achieve group goals, and its impact on member contributions on Wikipedia. As in the current study, they used an automated tool to measure the four types of shared leadership: Providing positive feedback which is similar to positive feedback in Choi (2012)'s study; negative feedback which is similar to constructive criticism in Choi (2012)'s study; social which is similar to welcome messages and personal comments in Choi (2012)'s study; and directing. Results for positive feedback and social message were similar. However, negative feedback in their study, which is conceptually similar to constructive criticism in the current study, decreased the number of edits.

To understand this discrepancy, we compared MTurk workers' judgments of constructive criticism to two experts' judgments of negative feedback for a set of 500 messages used in both studies. The results showed that the two categories are not highly correlated ($\gamma = 0.18$). In fact, constructive criticism in the current study includes both negative feedback and advice on how to correct the errors that a receiver generated, while aversive leadership from Zhu et al. (2012) only

included negative feedback. Thus, while these two categories are conceptually similar, humans judge them quite differently. This might be because the messages that WikiProject members exchange are different from those that Wikipedians do. Because WikiProjects are social groups with a smaller numbers of members compared to Wikipedia as a whole, they will bond together and take care of their newcomers more. Thus, WikiProject members are gentler in criticism than Wikipedians at large and provide advices about how to improve tasks.

Our results also show that all the socialization messages had a larger impact on the individual's relationship to a group when senders were experienced members rather than when they were newcomers, mainly because the content of the messages sent by old-timers and newcomers is different. For example, below are constructive criticisms sent by an experienced member and by a newcomer.

Constructive criticism sent by an experienced member:

"A good place to start is Wikipedia:WikiProject Schools/Article guidelines. I'll go through what I deleted and why. First, all of the additions were unsourced. They need a reliable source. The rival school definitely needs a reliable source. This sort of thing is often changed by high school students, it becomes a magnet for vandalism and original research. The list of schools in the conference is better listed at Oregon School Activities Association than on the school page. It isn't related to the school except perphirally. … This shouldn't discourage you. With reliable sources (such as newspaper articles), there are plenty of things that can be added. When was the school building built? What led to the school being built? Can you take a picture of the school? (this doesn't even require sources) Are there any notable alumni? (they must have an article written about them first) Has the school won any sports championships? (again, needs a citation) Let me know if I can help further; I have a brief list of sources for Oregon schools."

Constructive criticism sent by newcomers:

"I notice you recently uploaded this image. Please be more careful when checking the copyright of images in future."

As shown in these examples, experienced members provide more detailed explanations on how to correct mistakes as well as useful citations about rules and guidelines that others need to know through constructive criticism. Newcomers cannot provide the same high quality messages to others members. Thus, members who receive messages from experienced members are likely to more clearly understand their roles and the tasks they need to perform, learn the tasks of their new jobs, and gain greater confidence in their roles within the group.

In addition, receiving positive feedback from experienced members can be perceived as an official endorsement to contribute because experienced members are viewed as more central members of the group. Individuals may also feel better connected to other members when they receive social messages from experienced members. Lastly, when experienced members request tasks of new members, those members want to respond to receive their positive attention and make a good social connection with them.

The role of the recipients' tenure is more complicated. Newcomers edited more after receiving constructive criticism while experienced editors more contribute after receiving task requests. One reason for these effects of a task request could be that newcomers lacked enough knowledge about the group to be able to successfully accomplish tasks they were assigned. The task request could be different from what they expected or were interested in doing. The newcomers were uncertain what they had to do and how to contribute to the group at the initial stage. They need more time to investigate what the group needs them to do and negotiate what

they want to contribute and also to realize what they can do most effectively and with full confidence.

In addition, it is possible that newcomers might not know how to complete a certain task. Newcomers feel pressure and they can be overwhelmed by too many task request messages when they do not yet know how to perform a task (Hager & Brudney, 2004). On the other hand, experienced members may feel less pressure when receiving task requests because they do have the knowledge about how to respond to a request. Instead of feeling pressure, experienced members may be encouraged to contribute by a specific task request and believe they are authorized to help others or complete the assigned task.

Lastly, we found that individuals who sought out information increased their contribution in subsequent weeks. However, socialization messages had less impact among members who sought information compared to those who did not. These findings are similar to those reported by Gruman, et al. (2006) who showed that new employee information- seeking activities can replace organizational socialization tactics. It is possible that WikiProject members perceive socialization less valuable because they already received the answer they needed from other sources. Experienced members, in particular, already have built their own network with other members in Wikipedia. Thus, they can obtain information from other people not in the project. On the other hand, individuals who have not sought information, but received socialization messages from their groups, could learn more about the group through those messages, think that the group cares about them, and thus commit more to their group. Thus, even though both individuals who received socialization messages and those who sought information contributed more to the group than those who did not, the impact of receiving socialization messages was greater for those who did not ask any questions.

4.7. Implications

4.7.1. Theoretical Implications

Socialization is fundamental to online groups because it helps ensure the continuity of central values, and it gives their members a framework for responding to events in their group environment (Bauer, et al., 1998; Jones, 1986; Van Maanen & Schein, 1979). Despite the importance of understanding how socialization helps online group members to commit to their groups, very few studies have examined this relationship. Given the scarcity of research on socialization tactics and member commitment in online groups, one important result from this study was extending work by Choi (2012) showing the general linkage between socialization tactics and newcomer commitment.

Notwithstanding the general positive linkage between socialization tactics and member commitment, perhaps the most important implication of our findings is that members' characteristics play an important role in socialization process. First, this study identifies the moderating effect of member tenure on the linkage between four types of socialization messages regarding member commitment. We found that the impact of the four messages types depended on the recipients'' experiences: Constructive criticism was more valuable to newcomers, while task requests are more beneficial to experienced members. In addition, the results suggest that receiving socialization messages from experienced members is more advantageous.

Second, we found the moderating effect of individual information seeking on the relationship between socialization tactics and members in online groups. Our finding of less impact of socialization on individuals who seek information is consistent with the findings of Guman et al.(2006)'s research in traditional offline organizations, although additional research is

necessary. Given the scarcity of research on the role of the moderating effect of information seeking on the relationship between socialization and those outcomes, this research will contribute to the role of proactive behaviors in socialization research both in online and offline groups or organizations.

In general, these research findings do contribute to a developing literature that reveals the important role of individuals in the socialization process for online groups. These results are also important to help develop and refine socialization theory about how members adapt to their online groups and thus increase their commitment to their group.

4.7.2. Practical Implications

In this study, we demonstrated the possibilities of going beyond small samples by using automated coding of socialization behaviors. Without automated coding of behavior, research on socialization in online groups is restricted to small samples. For example, Choi (2012) hand coded communication for approximately 600 individuals in 22 Wikipedia projects. In contrast the current study was comprised of 29,095 unique editors in 1180 WikiProjects. We can also apply this automated coding method to other research. For example, the machine learning technique used in this study can be used to provide automated intervention, such as suggesting wording changes, offering tips, or providing successful examples of socialization messages to other members, each of which could help online groups socialize their members more effectively.

Secondly, practitioners should consider how to encourage socialization behaviors. Our research shows that the four socialization types—constructive criticism, positive feedback, task

requests, and social messages—are not equally beneficial to newcomers. In particular, constructive criticism seems especially beneficial to newcomers, while task requests are more beneficial to experienced members. Thus, the results of this study suggest that interfaces and mechanisms that make it easier for editors to connect with, reward, and express their appreciation to newcomers and ask for a specific task to be undertaken to experienced members may produce additional benefits.

In addition, the results suggest that receiving socialization messages from experienced members is not always more beneficial, as receiving task requests and social messages from newcomers had the same effect on member commitment as receiving those messages from experienced members. Thus, for effective resource or labor allocation, it would be better for experienced members to focus on sending constructive criticism and positive feedback.

Our results also suggest that for groups sending socialization messages before individuals seek information, given the current system where individuals are having a hard time finding people to answer questions, tools and interfaces can make it simpler to find appropriate people to ask questions and thus help individuals to acquired more benefits from information-seeking behaviors. For example, providing tables or graphs that show the people who are able to answer other members' questions on specific topics may help newcomers seek more information with less effort.

4.8. Limitations and Future Directions

This study has several limitations and definitely suggests interesting questions for further investigation. First, in our data, the messages an editor gets are partially a response to that

editor's previous behaviors. For example, a person who made lots of contribution in a previous week may cause others to send messages the following week and also influence this person's subsequent behavior. Even though we used sophisticated matching procedures in the form of propensity score matching to control for endogeniety, we cannot draw a full causal relationship even though we did observe the correlation between the event of receiving leadership messages and recipient behavioral change. Thus, experiments that can randomly assign different types of socialization messages to WikiProject members are important.

The data examined in this study are limited to public behavior. The results presented here show that socialization messages have an impact on members' contribution to online groups. However, private communication through emails or chatting and their reading behavior remain invisible. Since private interactions may operate differently from public, active ones, these invisible behaviors could be a source of noise that obscured our results. The existence of other forms of interactions suggests interesting questions for future research**.** Is invisible socialization through private communication more effective for newcomers or experienced members in terms of commitment to their groups? Does information seeking in private channels positively or negatively moderate the impact between socialization and its outcomes? In considering these types of questions, future research can move toward more complete and coherent explanations of the range of behavior seen in online groups.

We only investigated the effects of socialization for the amount of member contribution in this study. However, there can be other measurements that more directly estimate the extent that the socialization experiences influence the psychological states identified by Bauer or by Moreland and Levine, such as role clarity, self-efficacy , social acceptance, commitment or the amount newcomers know about their group. In addition, more study to investigate the effects of

socialization on the quality of their contribution is still needed. Quality can be measured through, various measurements such as Wikipedia Assessment Ratings based on the peer review process (Kittur, et al., 2009), the number of views that a given word has (Priedhorsky, et al., 2007), or how long the edits a user does lasts (Panciera, et al., 2009). We hope to come up with additional socialization strategies to maximize the amount of high quality work.

5. References

Ahuja, M. K., & Galvin, J. E. (2003). Socialization in Virtual Groups. *Journal of Management, 29*(2), 161-185.

Allen, D. G., Mahto, R. V., & Otondo, R. F. (2007). Web-based recruitment: Effects of information, organizational brand, and attitudes toward a Web site on applicant attraction. *Journal of Applied Psychology, 92*, 1696-1708.

Allen, T. D., Eby, L. T., O'Brien, K. E., & Lentz, E. (2007). The state of mentoring research: A qualitative review of current research methods and future research implications. *Journal of Vocational Behavior, 73*, 343-357.

Ambady, N., Krabbenhoft, M. A., & Hogan, D. (2006). The 30-sec sale: Using thin-slice judgments to evaluate sales effectiveness. *Journal of Consumer Psychology, 16*(1), 4-13.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological bulletin, 111*(2), 256.

Ancori, B., Bureth, A., & Cohendet, P. (2000). The economics of knowledge: the debate about codification and tacit knowledge. *Industrial and Corporate Change, 9*(2), 255-287.

Anderson, N. (2003). Applicant and recruiter reactions to new technology in selection: A critical review and agenda for future research. . *International Journal of Selection and Assessment, 11*, 121-136.

Arguello, J., Butler, B. S., Joyce, L., Kraut, R., Ling, K. S., Rosé, C. P., et al. (2006). *Talk to me: Foundations for successful individual-group interactions in online communities.* Paper presented at the CHI 2006, New York, NY.

Ashford, S. J., & Black, S. J. (1996). Proactivity during organizational entry: The role of desire for control. *Journal of Applied Psychology, 81*, 199-214.

Ashforth, B. E. (2001). *Role transitions in organizational life: An identity-based perspective*. Mahwah, NJ: Erlbaum.

Ashforth, B. E., & Saks, A. M. (1996). Socialization tactics: longitudinal effects on newcomer adjustment. *Academy of Management Journal, 39*, 149-178.

Auh, S., & Johnson, M. D. (2003). Compatibility effects in evaluations of satisfaction and loyalty. *Journal of Economic Psychology, 26*(1), 35-57.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing Construct Validity in Organizational Research. *Administrative Science Quarterly, 36*.

Barney, J. B., & Wright, P. M. (1998). On becoming a strategic partner: The role of human resources in gaining competitive advantage. *Human Resource Management, 37*(1), 31-46.

Baron, J. N., & Kreps, D. M. (1999). *Strategic Human Resources: Frameworks for General Managers*. New York: John Wiley & Sons.

Bartram, D. (2001). Internet recruitment and selection: Kissing frogs to find princes. *International Journal of Selection and Assessment, 8*, 261-274.

Bateman, P. J., Butler, B. S., & Gray, P. H. (2006). *Community Commitment: How Affect, Obligation, and Necessity Drive Online Behaviors.* Paper presented at the Twenty Seventh International Conference on Information Systems, Milwaukee, WI.

Bauer, T. N., & Green, S. G. (1998). Testing the combined effects of newcomer information seeking and manager behavior on socialization. *Journal of Applied Psychology, 83*, 72-83.

Bauer, T. N., Maertz, C. P., Dolen, M. R., & Campion, M. A. (1998). Longitudinal assessment of applicant reactions to employment testing and test outcome feedback. *Journal of Applied Psychology, 83*, 892-903.

Bauer, T. N., T., B., Tucker, J. S., Erdogan, B., & Truxillo, D. M. (2007). Newcomer adjustment during organizational socialization: a meta-analytic review of antecedents, outcomes, and methods. *Journal of applied psychology, 92*(3), 707 -721.

Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven, CT: Yale University Press.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 99-112.

Berger, J., Fisek, M. H., Norman, R. Z., & Zelditch, M., Jr. (1977). *Status characteristics and social in-teraction: An expectation states approach*. New York: Elsevier.

Blackman, M. C. (2002). Personality Judgment and the Utility of the Unstructured Employment Interview. *Basic and Applied Social Psychology, 24*(3), 241-250.

Blizzard Entertainment. (2010, Oct 26). Word of Warcraft subscriber base reaches 12 million worldwide. http://us.blizzard.com/en-us/company/press/pressreleases.html?101007.

Bock, G.-W., Zmud, R. W., Kim, Y.-G., & Lee, J.-N. (2005). Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organiza-tional Climate. *MIS Quarterly, 29*(1), 87-111.

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and social Psychology, 86*(4), 599.

Bowen, D. E., Ledford, G. E., & Nathan, B. R. (1991). Hiring for the organization, not the job. *Academy of Management Executive, 5*(4), 35-51.

Brass, D. J. (1984). Being in the right place: A structural analysis of individual influence in an organization. *Administrative Science Quarterly,, 29*, 518-539.

Bravo, M. J., Peiro, J. M., Rodriguez, I., & Whitely, W. T. (2003). Social antecedents of the role stress and career-enhancing strategies of newcomers to organizations: A longitudinal study. *Work & Stress, 17*, 195-217.

Breaugh, J. A., & Starke, M. (2000). Research on employee recruitment: so many studies, so many remaining questions. *Journal of Management, 26*(3), 405-434.

Bryant, S. L., Forte, A., & Bruckman, A. (2005). *Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia.* Paper presented at the GROUP 2005, Sanibel Island, Florida, USA.

Bull, C., & PieroTedeschi (1989). Optimal Probation for New Hires. *Journal of Institutional and Theoretical Economics, 145*(4), 627-642.

Burgoon, J. K., Buller, D. B., & Woodall, W. G. (1996). *Nonverbal communication: The unspoken dialogue* (2nd ed.). New York: McGraw-Hill.

Burke, M., Kraut, R., & Joyce, E. (2009). Membership Claims and Requests: Conversation-Level Newcomer Socialization Strategies in Online Groups. *Small Group Research, 41*(1), 4-40.

Burke, R. J., McKeen, C. A., & McKenna, C. (1993). Correlates of mentoring in organizations: The mentor's perspective. *Psychological Reports, 72*, 883-896.

Cable, D. M., & Judge, T. A. (1996). Person-organization fit, job choicedecisions, and organizational entry. *Organizational Behavior and Human Decision Processes, 67*(3), 294-311.

Carlson, J. R., & Zmud, R. W. (1999). Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal, 42*(3), 153-170.

Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality, 41*(5), 1054-1072.

Chao, G. T., Walz, P. M., & Gardner, P. D. (1992). Formal and Informal Mentorships: A Comparison on Mentoring Fuctions and Contrasts with Nonmentored Counterparts. *Personnel Psychology, 45*, 619-636.

Chapman, D., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., & Jones, D. A. (2005). Applicant attraction to organizations and job choice: A meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology, 90*, 928-944.

Chatman, J. (1991). Matching people and organizations: Selection and socialization in public accounting firms. *Administrative Science Quarterly, 36*, 459-484.

Chatman, J. A. (1989). Improving interactional organization research: A model of person-organization fit *Academy of Management Review, 14*, 333-349.

Chen, J., Ren, Y., & Riedl, J. (2010). *The effects of diversity on group productivity and member withdrawal in online volunteer groups.* Paper presented at the CHI '10 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA.

Choi, B. R. (2012). *Essay on Socialization in Online groups.* Carnegie Mellon Uniersity, Pittsburgh.

Collins, D. R., & Stukas, A. A. (2006). The Effects of Feedback Self-Consistency, Therapist Status, and Attitude Toward Therapy on Reaction to Personality Feedback. *Journal of Social Psychology, 146*(4), 463-483.

Cooper-Thomas, H. D., & Anderson, N. (2006). Invited manuscript: Organizational socialization: A new theoretical model and recommendations for future research and HRM practices in organizations. *Journal of Managerial Psychology, 21*(5), 492-516.

Cowan, R., & Foray, D. (1997). The economics of codification and the diffusion of knowledge. *Industrial and Corporate Change, 6*(3), 595-622.

Crampton, S. M., Hodge, J. W., & Mishra, J. M. (1998). The informal communication network: Factors influencing grapevine activity. *Public Personnel Management, 27*(4), 569-584.

Crant, M. C. (2000). Proactive behavior in organizations. *Journal of Management, 26*(3), 435-462.

Curhan, J., & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first five minutes. *Journal of Applied Psychology, 92*, 802-811.

D'Agostino, R. B., Jr. (1998). Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine, 17*, 2265-2281.

D'Souza, G. C., & Colarelli, S. M. (2010). Team member selection decisions for virtual versus face-to-face teams. *Computer in Human Behavior, 26*, 630-635.

DeSanctis, G., & Monge, P. (2000). Introduction to the special issue: Communication processes for virtual organizations. *Organization Science, 10*(6), 693-703.

Dineen, B. R., Ash, S. R., & Noe, R. A. (2002). A Web of applicant attraction: Person-organization fit in the context of Web-based recruitment. *Journal of Applied Psychology, 87*, 723-734.

Ducheneaut, N. (2005). Socialization in an open source software community: A socio-technical analysis. *Computer Supported Cooperative Work, 14*(4), 323-368.

Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2007). *The Life and Death of Online Gaming Communities: A Look at Guilds in World of Warcraft.* Paper presented at the Proceedings of the CHI 2007 Conference on Human Factors in Computing Systems, San Jose, CA.

Duck, S. W., Rutt, D. J., Hurst, M. H., & Strejc, H. (1991). Some Evident Truths About Conversations in Everyday Relationships. *Human Communication Research, 18*(2), 228-267.

Farmer, S. M., & Fedor, D. B. (2001). Changing the focus on volunteering: an investigation of volunteers' multiple contributions to a charitable organization. *Journal of Management, 27*, 191-211.

Feldman, D. C. (1981). The multiple socialization of organization members. *Academy of Management Review, 6*, 309-318.

Finholt, T., Sproull, L., & Kiesler, S. (1991). Communication and performance in ad hoc task groups. In J. Galegher & R. E. Kraut (Eds.), *Intellectual teamwork: Social and technological foundations of cooperative work* (pp. 291-325). Hillsdale, NJ: Lawrence Erlbaum.

Floyd, K., & Guerrero, L. K. (2006). *Nonverbal communication in close relationships*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Forte, A., Larco, V., & Bruckman, A. (2008). Decentralization in Wikipedia Governance. *Journal of Management Information Systems, 26*(1), 49-72.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory.* . Chicago, IL: Aldine Transaction.

Golder, S. A., & Donath, J. (2004). Social roles in electronic communities. *Internet Research, 5*, 19-22.

Govindarajan, V. (1989). Implementing competitive strategies at the business unit level: Implications of matching managers to strategies. *Strategic Management Journal, 27*(25-41).

Granovetter, M. (1995). *Getting a job: A study on contacts and careers* (2nd ed.). Chicago: University of Chicago Press.

Griffin, A. E. C., Colella, A., & Goparaju, S. (2000). Newcomer and organizational socialization tactics: an interactionist perspective. *Human Resource Management Review, 10*(4), 453-474.

Grube, J. A., & Piliavin, J. A. (2000). Role Identity, Organizational Experiences, and Volunteer Performance. *Personality and Social Psychology Bulletin, 26*, 1108-1119.

Gruman, J. A., Saks, A. M., & Zweig, D. I. (2006). Organizational socialization tactics and newcomer proactive behaviors: An integrative study. *Journal of Vocational Behavior, 69*, 90-104.

Haas, M. R., & Hansen, M. T. (2007). Different knowledge, different benefits: toward a productivity perspective on knowledge sharing in organizations. *Strategic management Journal, 28*, 1133-1153.

Hager, M. A., & Brudney, J. L. (2004). *Volunteer Management Practices and Retention of Volunteers*: Urban Institute.

Hahn, J., Moon, J. Y., & Zhang, C. (2008). Emergence of new project teams from open source software developer networks: Impact of prior collaboration ties. *Information Systems Research, 19*(3), 369-391.

Hansen, M. T., & Haas, M. R. (2001). Competing for attention in knowledge markets: electronic document dissemination in a management consulting company. *Administrative Science Quarterly 46*(1), 1-28.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. . *Journal of Applied Psychology, 88*, 100-112.

Holladay, C. L., Donnelly, T. M., Murray, S., & Halverson, S. K. (2006). *On-boarding employees: A model examining manager behavior, socialization, and commitment.* Paper presented at the Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

Holland, J. L. M. v. c. (1985). *A theory of careers*. Englewood Cliffs, NJ: Prentice-Hall.

Hom, P. W., Griffeth, R. W., Palich, L. E., & Bracker, J. S. (1999). Revisiting met expectations as a reason why realistic job previews work. *Personnel Psychology, 52*, 97-112.

Hough, L. M., & Furnham, A. (Eds.). (2003). *Use of personality variables in work settings* (Vol. 12). Hoboken, NY: Wiley.

Hovland, C. I., & Weiss, W. (1952). The influence of source credibility on communication effective-ness. *Public Opinion Quarterly., 15*, 355-350.

Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology, 86*, 897-913.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*, 869-879.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly, 37*, 422-447.

Ichino, A., & Muehlheusser, G. (2008). How often should you open the door? Optimal monitoring to screen heterogeneous agents. *Journal of Economic Behavior & Organization, 67*(3-4), 820-831.

Jeppesen, L. B., & Frederiksen, L. (2006). Why Do Users Contribute to Firm-Hosted User Communities? The Case of. Computer-Controlled Music Instruments. *Organization Science, 17*, 45-63.

John, O., Donahue, E., & Kentle, R. (1991). *The Big Five Inventory*. Berkeley, CA.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. E. John (Eds.), *Handbook of personality: Theory and research (2nd ed.)* (pp. 102-138). New York, NY: Guilford Press.

Jones, G. R. (1986). Socialization tactics, self-efficacy, and newcomers' adjustments to organizations. *Academy of Management Journal, 29*, 262-279.

Jones, R. G., Stevens, M. J., & Fischer, D. L. (Eds.). (2000). *Selection in team contexts* San Francisco, CA: Jossey-Bass.

Kammeyer-Mueller, J. D., & Wanberg, C. R. (2003). Unwrapping theorganizational entry process: Disentangling multiple antecedents and their pathways to adjustment. *Journal of Applied Psychology, 88*, 779-794.

Kim, T., Cable, D. M., & Kim, S. (2005). Socialization tactics, employee proactivity, and person-organization fit. *Journal of Applied Psychology, 90*(2), 232-241.
Kittur, A., Pendleton, B., & Kraut, R. E. (2009). *Herding the cats: The influence of groups in coordinating peer production.* Paper presented at the the 5th international symposium on wikis and open collaboration.
Klein, H. J., & Heuser, A. E. (2008). The learning of socialization content: A framework for researching orientating practices. In J. Martocchio, H. Liao & A. Joshi (Eds.), *Research in Personnel and Human Resources Management* (Vol. 27, pp. 279-336): Emerald Group.
Klein, H. J., & Weaver, N. A. (2000). The effectiveness of an organizational-level orientation training program in the socialization of new hires. . *Personnel Psychology, 53*(1), 47-66.
Koch, G. G. (1982). Intraclass correlation coefficient. In S. Kotz & N. L. Johnson. (Eds.), *Encyclopedia of Statistical Sciences* (Vol. 4). New York: John Wiley & Sons.
Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (Vol. 29). San Francisco, CA: Jossey-Bass.
Kram, K. E. (1985). *Mentoring at work: Developmental relationships in organizational life*. Glenview, IL.
Kraut, R., Wang, X., Butler, B., Joyce, E., & Burke, M. (Under Review). Beyond Information: Developing the Relationship between the Individual and the Group in Online Communities.
Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology, 49*(1), 1-49.
Lam, A. (2000). Tacit knowledge, organizational learning and societal institutions: an integrated framework. *Organization studies, 21*(3), 487-513.
Lampe, C., & Johnston, E. (2005). *Follow the (slash) dot: effects of feedback on new members in an online community.* Paper presented at the ACM SIGGROUP Conference on Supporting Group Work.
Lee, S.-Y. T., Kim, H.-W., & Gupta, S. (2009). Measuring open source software success. *Omega, 37*, 426-438.
Levine, J. M., & Moreland, R. L. (1994). Group socialization: Theory and research. In W. Stroebe & M. Hewstone (Eds.), *The European Review of Social Psychology* (Vol. 5, pp. 305-336). Chichester: John Wiley & Sons.
Levine, J. M., Moreland, R. L., & Hausmann, L. R. M. (2005). *Managing Group Composition: Inclusive and Exclusive Role Transitions. The social psychology of inclusion and exclusion.* New York, NY: Psychology Press.
Lipnack, J., & Stamps, J. (2000). *Virtual teams: People working across boundaries with technology*. New York, NY: John Wiley & Sons Inc.
Loh, E. S. (1994). Employment Probation as a Sorting Mechanism. *Industrial and Labor Relations Review 47*(3), 471-486.
Ma, M., & Agarwal, R. (2007). Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities. *Information Systems Research, 18*(1), 42-67.
Major, D. A., Kozlowski, S. W. J., Chao, G. T., & Gardner, P. D. (1995). A longitudinal investigation of newcomer expectations, early socialization outcomes, and the

moderating effects of role development factors. *Journal of Applied Psychology, 80*(3), 418-431.
Meglino, B., Ravlin, E., & Adkins, C. (1989). A work values approach to corporate culture: A field test of the value congruence process and its relationship to individual outcomes. *Journal of Applied Psychology, 74*, 424-432.
Meglino, B. M., DeNisi, A. S., Youngblood, S. A., & Williams, K. J. (1988). Effects of realistic job previews: A comparison using an enhancement and a reduction preview. *Journal of Applied Psychology, 73*, 259-266.
Miller, V. D., & Jablin, F. M. (1991). Information seeking during organizational entry: Influences, tactics, and a model of the process. *Academy of Management Review, 16*, 92-120.
Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C., & Erez, M. (2001). Why people stay: Using job embeddedness to predict voluntary turnover. . *Academy of Management Journal, 44*, 1102-1121.
Moon, J. Y., & Sproull, L. S. (2008). The Role of Feedback in Managing the Internet-based Volunteer Work-force. *Information Systems Research, 19*(4), 494-515.
Moos, R. H. (1987). Person-environment congruence in work, school, and health care settings. *Journal of Vocational Behavior, 31*, 231-247.
Moreland, R. L., & Levine, J. M. (1988). *Group dynamics over time: Development and socialization in small groups* (Vol. 91). Thousand Oaks, CA: Sage Publications.
Moreland, R. L., & Levine, J. M. (2001). *Socialization in organizations and work groups*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, .
Morrison, E. W. (1993). Longitudinal study of the effects of information seeking on newcomer socialization. *Journal of Applied Psychology, 78*(2), 173-183.
Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*, 115-132.
Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus User's Guide* (Fifth ed.). Los Angeles, CA: Muthén & Muthén.
Noe, R. A. (1988). An investigation of the determinants of successful assigned mentoring relationships. *Personnel Psychology, 41*, 457-479.
Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York, NY: Oxford University Press.
O'Reilly, C., Chatman, J., & Caldwell, D. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal, 34*, 487-516.
Ostroff, C., & Kozlowski, S. W. J. (1992). Organizational socialization as a learning process: The role of information acquisition. *Personnel Psychology, 45*(4), 849-874.
Panciera, K., Halfaker, A., & Terveen, L. (2009). *Wikipedians Are Born, Not Made: A Study of Power Editors on Wikipedia.* Paper presented at the Group '09.
Pearce, J. L. (1993). *Volunteers: The Organizational Behavior of Unpaid Workers*. New York, NY: Routledge.
Pentland, A., & Pentland, S. (2008). *Honest signals: how they shape our world*: The MIT Press.
Piliavin, J. A., & Callero, P. L. (1991). *Giving blood: The development of an altruistic identity*. Baltimore, MD: Johns Hopkins.
Ployhart, R. E. (2006). Staffing in the 21st century: New challenges and strategic opportunities. *Journal of Management, 32*, 868-897.

Podolny, J. M., & Baron, J. N. (1997). Resources and relationships: Social networks and mobility in the workplace. *American Sociological Review, 62*, 673-693.
Polanyi, M. (1961). Knowing and being. *Mind*, 458-470.
Posthuma, R. A., Morgeson, F. P., & Campion, M. A. (2002). Beyond employment interview validity: a comprehensive narrative review of recent research and trends over time. *Personnel Psychology, 55*(1), 1-81.
Preece, J., Nonnecke, B., & Andrews, D. (2004). The top 5 reasons for lurking: Improving community experiences for everyone. *Computers in Human Behavior, 20*(2), 201-223.
Priedhorsky, R., Chen, J., Lam, S. T. K., Panciera, K., Terveen, L., & Riedl, J. (2007). *Creating, Destroying, and Restoring Value in Wikipedia.* Paper presented at the GROUP '07.
Ragins, B. R., & Cotton, J. L. (1999). Mentor functions and outcomes: A comparison of men and women in formal and informal mentoring relationships. . *Journal of Applied Psychology, 84*(4), 529-550.
Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of experimental psychology: general, 118*(3), 219.
Roberts, J., Hann, I.-H., & Slaughter, S. (2006a). Understanding the Motivations, Participation and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects. *Management Science, 52*(7), 984-999.
Roberts, J., Hann, I.-H., & Slaughter, S. (2006b). Understanding the Motivations, Participation and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects. *Management Science, 52*(7), 984 - 999.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. . *Biometrika, 70*, 41-55.
Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management, 41*, 149-174.
Sadanand, A., Sadanand, V., & Marks, D. (1989). Probationary Contracts in Agencies with Bilateral Aymmetric Information. *Canadian Journal of Economics, 22*(3), 643-661.
Saks, A. M., & Ashforth, B. E. (1997). Organizational socialization: Making sense of past and present as a prologue for the future. *Journal of Vocational Behavior, 51*, 234-279.
Saks, A. M., Uggerslev, K. L., & Fassina, N. E. (2007). Socialization tactics and newcomer adjustment: A meta-analytic review and test of a model. *Journal of Vocational Behavior*(70), 413-446.
Schmitt, N., Cortina, J. M., Ingerick, M. J., & Wiechmann, D. (Eds.). (2003). *Personnel selection and employee performance* (Vol. 12). Hoboken, NY: Wiley.
Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*, 437-453.
Schwab, D. P., Rynes, S. L., & Aldag, R. J. (1987). Theories and research on job search and choice. *Research in Personnel and Human Resources Management, 5*, 129-166.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis*. New York, NY: Oxford University Press.
Sluss, D. M., & Ashforth, B. E. (2007). Relationship identity and identification: Defining ourselves through work relationships. *Academy of Management Review, 32*, 9-32.
Snow, R., O'Connor, B., Jurafsky, D., & Ng, A. Y. (2008). *Cheap and fast---but is it good? Evaluating non-expert annotations for natural language tasks.* . Paper presented at the in Proceedings of the Conference on Empirical Methods in Natural Language Processing.

Stevens, M. J., & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management, 25*, 207-228.
Tidwell, L. C., & Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research, 28*(3), 317-348.
Timmerman, C. E., & Scott, C. R. (2006). Virtually working: Communicative and structural predictors of media use and key outcomes in virtual work teams. *Communication Monographs, 73*, 108-136.
Van Maanen, J., & Schein, E. H. (1979). *Toward a theory of organizational socialization* (Vol. 1). Greenwich, CT: JAI Press.
Vandenberg, R. J., & Scarpello, V. (1990). The matching model: An examination of the processes underlying realistic job previews. *Journal of Applied Psychology, 75*, 60-67.
Von Krogh, G., Ichijō, K., Nonaka, I., & Nonaka, I. (2000). *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*: Oxford University Press, USA.
Walther, J. B. (1995). Relational aspects of computer-mediated communication: Experimental observations over time. *Organization Science, 6*(2), 168-185.
Walther, J. B. (2002). Time effects in computer-mediated groups: Past, present, and future. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 235-257). Cambridge, MA, US: MIT Press.
Wanberg, C. R., & Kammeyer-Mueller, J. D. (2000). Predictors and outcomes of proactivity in the socialization process. *Journal of Applied Psychology, 85*(3), 373-385.
Wang, X., Kraut, R., Butler, B., Joyce, E., & Burke, M. (Under Review). The Role of Membership Claim and Grant Signals in Individual-Community Relationships *Management Information Systems Quarterly*.
Wang, Y.-C., Kraut, R. E., & Levine, J. M. (2012). *To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups.* Paper presented at the in Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW) Seattle, Washington.
Wasko, M. M., & Faraj, S. (2005). Why Should I Share?: Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly, 29*(1), 35-57.
Weisband, S. P., Schneider, S. K., & Connolly, T. (1995). Computer-Mediated Communication and Social Information. *Academy of Management Journal, 38*(4), 1124-1151.
Weller, I., Holtom, B. C., Matiaske, W., & Mellewigt, T. (2009). Level and time effects of recruitment sources on early voluntary turnover. *Journal of Applied Psychology, 94*(5), 1145-1162.
Wesson, M. J., & Gogus, C. I. (2005). Shaking hands with a computer: An examination of two methods of newcomer orientation. *Journal of Applied Psychology, 90*, 1018-1026.
Wikipedia Statistics. Retrieved August 8 2012, from http://en.wikipedia.org/wiki/Wikipedia:Statistics
Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., & Nickell, E. (2006). From tree house to barracks: The social life of guilds in World of Warcraft. *Games and Culture, 1*(4), 338-361.
Williams, K. D., Cheung, C. K., & Choi, W. (2000). Cyberostracism: effects of being ignored over the Internet. *Journal of Personality & Social Psychology, 79*(5), 748-762.

Winter, S. (1987). Knowledge and competence as strategic assets. In D. J. Teece (Ed.), *The Competitive Challenge* (pp. 159-184). Cambridge, MA: Ballinger.

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (Second ed.). San Francisco: Elsevier.

Yoo, Y., & Alavi, M. (1999). Media and group cohesion: Relative influences on social presence, task participation, and group consensus. *MIS Quarterly, 25*(3), 371-390.

Zhu, H., Kraut, R. E., & Kittur, A. (2012). *Effectiveness of shared leadership in online communities.* Paper presented at the in Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW), Seattle, Washington.

Zottoli, M. A., & Wanous, J. P. (2000). Recruitment source research: Current status and future directions. . *Human Resource Management Review, 10*, 353-382.

6. Appendix A.

Decide the types of messages that Wikipedia users exchange

NOTE!

- **This task is for Wikipedia users who are native English speakers. Only serious responses will be paid. Thank you for your understanding.**

Instructions

- Carnegie Mellon University's Social Computing Group is interested in what types of messages Wikipedia users exchange with each other.
- The goal of this task to decide the category for each message. You will be reading messages posted to Wikipedia user pages. You can help us by indicating which categories each message are involved. **Please select all categories that apply.**
- You need to complete each question to get paid.

Definition

- Each message can contain content following five categories: 1) **Give Advice**, 2) **Give Positive Feedback,** 3) **Assign/Suggest Tasks**, 4)**Talk Socially** and 5)**Seek Information**. Each message can have multiple types of content. Example of each category follows**.**

Definition of types of messages

- **"Give Constructive Criticism"** means giving advices or suggesting improvements on what individuals did wrong.
- **"Give Positive Feedback"** means providing praise individuals' edits or gave them an award
- **"Request Tasks"** means asking people to do a certain job or task
- **"Talk Socially"** means sending friendly messages including welcoming or off-topic content which was not related to Wikipedia task.
- **"Seek Information"** means asking for information, possible ways of action, opinion, evaluation, orientation, and confirmation. request

Content	Example
Give Constructive Criticism/ Advice	You did wrong. You should follow this format/policy. Here is the links.
Give Positive Feedback	It looks great! I'll give you a Barnstar.
Assign/Suggest/Request Tasks	Would you upload the picture? Would you help me doing this?
Talk Socially	Welcome. Happy birthday. I am doing good.
Seek Information	Can you tell me how to do it? What do you think?

Qualification questions

<Wikipedia Experiences>

1. How long have you been a registered user of Wikipedia?

1) None

2) Less than 2 weeks

3) 2 weeks - less than 1 month

4) 1 month - less than 3 months

5) 3 months - less than 6 months

6) 6months – less than 1 year

7) More than 1 year

2. How long have you communicated with other users in Wikipedia?

1) None

2) Less than 2 weeks

3) 2 weeks - less than 1 month

4) 1 month - less than 3 months

5) 3 months - less than 6 months

6) 6months – less than 1 year

7) More than 1 year

3. How many times have you communicated with other users in Wikipedia?

1) None 2) 1-3 3) 3-5 4) 5-10 5) 11-20 6) 21-40 7) More than 40

4. How many articles have you worked on?

1) None 2) 1-3 3) 3-5 4) 5-10 5) 11-20 6) 21-40 7) More than 40

<Wikipedia Knowledge Test>

1. What does NPOV mean in Wikipedia?

1) Non-Profit Organizations and Visionaries

2) Non-Profit Organizations and Values

3) New point of view

4) Neutral point of view

2. How can you sign your posts in talk pages?

1) Using ~~~~ 2) Using **** 3) Using ---- 4) Using ++++

3. What does a Barnstar mean in Wikipedia?

1) Reward 2) Registerd user 3) Bot 4) Administrator

Main Task

Please carefully read the message:

==========

Message

==========

Do the following categories describe the message? Select all that apply.

Give Constructive Criticism	○	Yes	○	No
Give Positive Feedback	○	Yes	○	No
Request Tasks	○	Yes	○	No
Talk Socially	○	Yes	○	No
Seek Information	○	Yes	○	No

7. Appendix B.

Feature	Description	Example
<I+modal>	Frequency of sentences starting with a pronoun “I” immediately followed by a modal word (e.g., should, might, must) or vice versa.	I can, I could, I will, I should, I do, I did,
<If+you>	Frequency of phrase “if you”	if you
<Modal+I>	Frequency of sentences starting with a modal word (e.g., should, might, must) immediately followed by a pronoun “I”	can I, could I, should I, do I, did I
<Modal+You>	Frequency of sentences starting with a modal word (e.g., should, might, must) immediately followed by a pronoun “you”	you can, you could, you shall, you may, you might, you must, you should, you will, you would
<Please+verb>	Frequency of “please”	Please, Plz
<You+Modal>	Frequency of sentences starting with a pronoun “You” immediately followed by a modal word (e.g., should, might, must) or vice versa	can you, could you, shall you, may you, might you, must you, should you, will you, would you
Acknowledge	Frequency of phrase patterns of “thank”	thank, thanks, thankx, thx
Acknowledge 1	Frequency of phrase patterns of “appreciate”	appreciate,appreciated
Advice	Frequency of following words used to give advice	citation, citations, reference, references, guideline, guidelines, policiy, sanbox, instruction, instructions
Appology	Variants of the word of “appology”	appology, appologize, apology, apologies, apologises, sorry
Barnstar	Frequency of the following words: “barnstar”, “reward” and “medal”.	barnstar, reward, medal, exceptional newcomer
Causative/subjunctive verbs	Frequency of causative/subjunctive verbs	ask, hope, propose, recommend,wish.
Collaboration	Variants of the word of “collaboration”	collaboration, collaborations, cotw, collaborate, collaborates
Congrats	Variant forms of the word “congratulation”.	congrats, congratulation, congratulations
DYK	Frequency of Wikipedia-specific jargons meaning "did you know"	did you know, DYK, DYKs
Greetings	Greeting words/phrases, such as hello, and happy new year.	hey, ha,cheers, regards, new year, merry christmas, see you, good day, nice day, good morning, happy birthday, good luck
Invitation	Variants of the word of “invitation”	invitation, invite, invited, join, joining

Length	Number of word tokens in a message	
Need		Need, Needs, Needed, Necessary,
Negation	Frequency of negation words and phrases (e.g., not, shouldn't, doesn't).	not, cannot, can't, doesn't, didn't, shouldn't, should, won't, mustn't, haven't, hasn't, wouldn't, isn't, aren't, wasn't, weren't, couldn't
Negative Conjunction	Frequency of negative conjunctions	however, but
Negative jargon	Frequency of negative Wikipedia-specific jargons such "vandalism" and "nonfree"	spam,revert,reverted, block, blocked,remove, removed, delete, deleted, vandalism,vandalise,vandalised, vandalized, vandalize, violate, violated, nonfree, copyright, disputed, dispute, noneutral, fair use, deleted, removed, deletion
Newsletter	Frequency of the word of "newsletter"	Newsletter
Nomiation	Frequency of Wikipedia-specific jargons meaning "nomination"	nominate, nominated, nomination
Others	Frequency of "he, him, his, she, her".	he/she
Question Mark	Frequency of the question mark "?".	?
Question terms	Frequency of questions words	who is,who are,who was,who were,where is,where was,where were,where are,how is,how was,what is,whats,what are,what was,what were,what does,what do,what did,is there,are there,was there,were there,is it,is this,was it,are these,was this,were these,have you,havent you,have i,havent i, does it,does this,do these,doesnt it,doesnt this,dont these,wouldnt you,shouldnt it,shouldnt this,shouldnt there,would it,whould this,should it,should this,should these,how about,what about,are you,were you,is he,was he,am i,was i,are they,were they,where do,where did,where does,when will,when would,when do,when did,when does,how did,how do,how does,who does,who do,who did
Review	Variants of the word of "review"	review, reviews, comment, comments, opinion, opinions
RFA	Frequency of Wikipedia-specific jargons meaning "request for adminship"	request for adminship, rfa
Roll call	Frequency of Wikipedia-specific jargon to separate the active members of the project from the former/inactive members	Rollcall, recalling
Seekinfo	Frequency of information seeking phrases (e.g. share any information)	question, wonder, wondering, wonddered, wonders, could you explain, can you explain, like to know, share any information, any idea, no ieda, how to

Smiley	Textual expressions such as :), ;).	), :P, :b, :-), :-P, :-b, ;), ;P, ;b, ;-), ;-P, ;-b, ^_^, =), =], smile, smiles, smiling, balloon, balloons
Strong negative-polarity words	Frequency of strong negative-polarity words based on a subjectivity lexicon	(Wilson T., Wiebe J., and Hoffmann P 2009)
Strong positive adjectives	Frequency of strong positive adjectives used in praise, such as “excellent”, “great”, “impressive”, etc.	incredible, wellwritten, great, excellent, successful, outstanding, impressive, best, highest, outstanding, featured, greatest, awesome, fantastic, nice, beautiful, good
Strong positive-polarity words	Frequency of strong positive-polarity words based on a subjectivity lexicon	(Wilson T., Wiebe J., and Hoffmann P 2009)
Suggestion	Frequency of suggestion phrases (e.g., check it, check out)	check it, check out, may want, might want, may be interested, might be interested, will be interested, would be interested, be willing to, do something, have time, have the time, get the time, got the time, drop by, dropping by, stop by, stopping by, swing by, why don't you, take a look, have a look, get a chance, get the chance, do you mind, would you mind, request, visit, offer, update, updating
Weak negative-polarity words	Frequency of weak negative-polarity words based on a subjectivity lexicon	(Wilson T., Wiebe J., and Hoffmann P 2009)
Weak positive-polarity words	Frequency of weak positive-polarity words based on a subjectivity lexicon	(Wilson T., Wiebe J., and Hoffmann P 2009)
Welcome	Variants of the word of “welcome”	welcome, welcomed
Wikiproject	Frequency of the word of “Wikiproject”	wikiproject, wikiprojects

8. Appendix C.

	Weight				
	Constructive Criticism	Positive Feedback	Task requests	Social Messages	Information Seeking
<I+modal>	-0.29	0.32	-0.19	0.00	0.14
<If+you>	0.55	-0.01	0.36	0.00	0.02
<Modal+I>	0.00	-0.70	0.06	0.00	0.56
<Modal+You>	-1.65	-2.48	1.73	0.00	1.30
<Please+verb>	-0.36	-0.51	1.11	0.00	-0.41
<You+Modal>	0.68	-0.01	0.64	0.00	-0.21
Acknowledge	-3.94	6.03	-0.11	8.00	-0.80
Acknowledge 1	-2.18	0.78	1.20	4.00	-0.49
Advice	1.95	0.43	-1.17	0.00	0.05
Appology	-2.39	-1.01	-1.00	4.00	-0.50
Barnstar	-2.35	3.59	-0.94	4.33	0.17
Causative/subjunctive verbs	-1.46	-0.03	0.54	0.00	-0.07
Collaboration	-1.91	-1.50	3.47	2.00	-0.01
Congrats	-1.41	3.50	0.00	2.50	0.39
DYK	-2.33	0.82	0.67	0.00	0.00
Greetings	-2.22	-0.92	0.17	9.99	-0.14
Invitation	-2.18	-1.04	2.28	3.00	-0.67
Length	3.07	1.12	0.06	0.00	-0.38
Need	0.05	-0.09	0.78	0.00	-0.50
Negation	3.00	-0.53	-1.45	0.00	0.50
Negative Conjunction	1.27	0.26	0.25	0.00	-0.84
Negative jargon	5.99	1.46	-1.67	-4.67	0.04
Newsletter	-2.09	-1.18	-0.55	-0.25	-1.01
Nomination	-0.57	-0.14	0.79	0.00	0.00
Others	1.47	-0.45	0.31	0.00	-0.27
Question Mark	-1.22	-0.80	0.00	0.00	0.90
Question terms	0.09	-0.89	-0.87	0.00	5.80
Review	0.34	-0.35	1.00	0.00	-0.68
RFA	-0.05	0.32	-2.31	0.00	0.00
Roll call	-1.11	0.00	2.20	0.00	-1.00
Seekinfo	-1.06	-0.53	-0.27	0.00	2.61
Smiley	-4.23	-0.14	-0.67	4.50	0.00
Strong negative-polarity words	2.16	0.11	-0.51	0.00	-0.82
Strong positive adjectives	-1.68	0.73	-0.55	0.00	-0.42
Strong positive-polarity words	-0.05	0.57	-0.06	0.00	-1.87
Suggestion	-1.58	-1.04	4.86	0.00	-0.40
Weak negative-polarity words	2.06	-0.39	-0.22	0.00	-0.51
Weak positive-polarity words	1.69	0.54	-0.22	0.01	-0.48
Welcome	-2.26	3.90	-0.67	4.00	-0.49
Wikiproject	-2.33	0.28	0.79	0.00	-0.60

9. Appendix D.

The percentage of bias is the mean difference as a percentage of the average standard deviation: where for each covariate and are the sample means in the treatment groups (editors who received a type of messages in the given week) and control groups (editors who did not receive a type of messages in the given week), respectively, and are the corresponding sample variances (Rosenbaum & Rubin, 1983).

N of obs	Full Matched	Treat Treat		Control Control	
Variable	Sample	Treat Mean	Control Mean	%bias	% reduct bias
Edits(t-1)	Full	4.21	3.14	54.80	
	Matched	4.21	4.20	0.20	99.70
MsgReceived(t-1)	Full	1.22	0.70	53.20	
	Matched	1.22	1.23	-0.90	98.20
MsgReceived(<t-1)	Full	3.58	3.28	17.10	
	Matched	3.58	3.57	0.50	97.00
MsgSent(t-1)	Full	1.20	0.70	41.40	
	Matched	1.20	1.22	-1.90	95.40
MsgSent(<t-1)	Full	3.35	3.10	11.20	
	Matched	3.35	3.37	-0.90	92.00
Tenure	Full	52.86	64.26	-26.00	
	Matched	52.86	53.47	-1.40	94.60
Positive(t-1)	Full	0.09	0.00	47.00	
	Matched	0.09	0.10	-7.20	84.60
Constructive criticism(t-1)	Full	0.24	0.00	69.80	
	Matched	0.24	0.23	3.40	95.20
Social(t-1)	Full	0.17	0.00	62.00	
	Matched	0.17	0.19	-6.10	90.20
Task request(t-1)	Full	0.16	0.00	61.80	
	Matched	0.16	0.18	-5.80	90.60

Table 1. Comparison between treatment editors who received *constructive criticism* in the focal week (treat) and control editors (control) before and after propensity score matching (full versus matched).

N of obs	Full Matched	Treat Treat		Control Control	
Variable	Sample	Treat Mean	Control Mean	%bias	% reduct bias
Edits(t-1)	Full	4.11	3.14	47.70	
	Matched	4.11	4.06	2.10	95.50
MsgReceived(t-1)	Full	1.20	0.71	50.40	
	Matched	1.20	1.23	-2.10	95.80
MsgReceived(<t-1)	Full	3.40	3.29	6.20	
	Matched	3.40	3.30	5.10	17.30
MsgSent(t-1)	Full	1.22	0.70	41.70	
	Matched	1.22	1.29	-5.10	87.70
MsgSent(<t-1)	Full	3.20	3.10	4.10	
	Matched	3.20	3.17	0.90	77.20
Tenure	Full	49.88	64.24	-33.00	
	Matched	49.88	47.08	6.40	80.60
Positive(t-1)	Full	0.11	0.00	51.10	
	Matched	0.11	0.10	4.40	91.30
Constructive criticism(t-1)	Full	0.20	0.00	61.30	
	Matched	0.20	0.20	-0.70	98.90
Social(t-1)	Full	0.18	0.00	62.90	
	Matched	0.18	0.17	3.50	94.40
Task request(t-1)	Full	0.18	0.00	63.00	
	Matched	0.18	0.17	3.50	94.50

Table 2. Comparison between treatment editors who received *positive feedback* in the focal week (treat) and control editors (control) before and after propensity score matching (full versus matched).

N of obs	Full Matched	Treat Treat		Control Control	
Variable	Sample	Treat Mean	Control Mean	%bias	% reduct bias
Edits(t-1)	Full	3.94	3.14	39.40	
	Matched	3.94	3.90	1.90	95.20
MsgReceived(t-1)	Full	1.13	0.70	43.90	
	Matched	1.13	1.14	-0.30	99.30
MsgReceived(<t-1)	Full	3.38	3.29	5.40	
	Matched	3.38	3.32	3.40	37.70
MsgSent(t-1)	Full	1.12	0.70	34.60	
	Matched	1.12	1.12	-0.20	99.30
MsgSent(<t-1)	Full	3.14	3.10	1.60	
	Matched	3.14	3.12	0.70	53.80
Tenure	Full	51.25	64.27	-30.00	
	Matched	51.25	49.84	3.20	89.20
Positive(t-1)	Full	0.09	0.00	47.80	
	Matched	0.09	0.08	4.70	90.20
Constructive criticism(t-1)	Full	0.19	0.00	60.10	
	Matched	0.19	0.21	-7.10	88.20
Social(t-1)	Full	0.17	0.00	62.00	
	Matched	0.17	0.16	6.00	90.40
Task request(t-1)	Full	0.17	0.00	62.00	
	Matched	0.17	0.15	5.70	90.80

Table 3. Comparison between treatment editors who received *task requests* in the focal week (treat) and control editors (control) before and after propensity score matching (full versus matched).

N of obs	Full Matched	Treat Treat		Control Control	
Variable	Sample	Treat Mean	Control Mean	%bias	% reduct bias
Edits(t-1)	Full	3.93	3.14	38.80	
	Matched	3.93	3.88	2.70	93.00
MsgReceived(t-1)	Full	1.13	0.70	43.50	
	Matched	1.13	1.13	0.20	99.50
MsgReceived(<t-1)	Full	3.37	3.29	4.90	
	Matched	3.37	3.31	3.70	24.60
MsgSent(t-1)	Full	1.11	0.70	34.10	
	Matched	1.11	1.11	0.60	98.30
MsgSent(<t-1)	Full	3.12	3.10	1.00	
	Matched	3.12	3.08	1.80	-69.90
Tenure	Full	51.17	64.27	-30.20	
	Matched	51.17	49.89	3.00	90.20
Positive(t-1)	Full	0.09	0.00	47.60	
	Matched	0.09	0.08	3.60	92.50
Constructive criticism(t-1)	Full	0.19	0.00	59.90	
	Matched	0.19	0.21	-6.70	88.80
Social(t-1)	Full	0.17	0.00	61.70	
	Matched	0.17	0.16	5.60	90.90
Task request(t-1)	Full	0.17	0.00	61.60	
	Matched	0.17	0.15	5.30	91.40

Table 4. Comparison between treatment editors who received *social messages* in the focal week (treat) and control editors (control) before and after propensity score matching (full versus matched).