The Effects of Diversity in Global, Distributed Collectives: A Study of Open Source Project Success

Abstract

Diversity is a defining characteristic of global collectives facilitated by the Internet. Though substantial evidence suggests that diversity has profound implications for a variety of outcomes including performance, member engagement, and withdrawal behavior, the effects of diversity have been predominantly investigated in the context of organizational workgroups or virtual teams. We use a diversity lens to study the success of non-traditional virtual work groups exemplified by open source software (OSS) projects. Building on the diversity literature, we propose that three types of diversity (separation, variety and disparity) influence two critical outcomes for OSS projects: community engagement and market success. We draw on the OSS literature to further suggest that the effects of diversity on market success are moderated by the application development stage. We instantiate the operational definitions of three forms of diversity to the unique context of open source projects. Using archival data from 357 projects hosted on SourceForge, we find that disparity diversity, reflecting variation in participants' contribution-based reputation, is positively associated with success. The impact of separation diversity, conceptualized as culture and measured as diversity in the spoken language and country of participants, has a negative impact on community engagement but an unexpected positive effect on market success. Variety diversity, reflected in dispersion in project participant roles, positively influences community engagement and market success. The impact of diversity on market success is conditional on the development stage of the project. We discuss how the study's findings advance the literature on antecedents of OSS success, expand our theoretical understanding of diversity, and present the practical implications of the results for managers of distributed collectives.

Key words: Open source software, diversity, global collectives

1. Introduction

> We need to acknowledge that good developers come from many different walks of life, that their motivations differ widely, and that the overall strength of an Open Source project comes from the *diversity* of its users and developers rather than their *homogeneity*. Diversity of interests leads to generality of the code, to development of useful pathways that might otherwise never be trod, and continued support of the project when business falters. Perens 2009

As the Internet spawns virtual communities and facilitates global collaborations, diverse groups are becoming ubiquitous. Escalating group diversity is apparent both inside and outside of formal organizations. With organizations increasingly seeking to draw on the best talent from around the world in order to remain competitive (Gibson and Gibbs 2006, Lee, et al. 2006), evidence of the benefits and consequences of leveraging global, diverse ability is mounting (Van Knippenberg and Schippers 2007). Global alliances include outsourcing software development to offshore destinations (Levina and Vaast 2008), using virtual teams for the execution of complex projects that require diverse skills (Majchrzak and Malhotra 2004), volunteer labor forces responsible for creating strikingly successful products such as Wikipedia, and the increasing prevalence of open source software (OSS) projects.

The OSS development model has fundamentally transformed traditional software development practices. In the OSS model, software is frequently constructed by a virtual community of volunteers, is available for modification by anyone, and is obtainable for free (Scacchi, et al. 2006). OSS applications compete directly with closed and proprietary software, and this competition is altering the strategies and tactics of traditional for-profit software organizations (Jaisingh, et al. 2008). Unsurprisingly, the striking growth and success of the OSS model has spurred a significant body of academic research in the past decade (Chengalur-Smith, et al. 2010, Singh, et al. 2011, Stewart and Gosain 2006).

Despite robust research on OSS, critical questions remain unanswered. In particular, limited work has focused on the role of diversity in open source communities. Why should diversity matter in OSS projects? On the one hand, the open access model underlying OSS development allows participants whose expertise matches the project's needs to contribute with minimal geographic or organizational restrictions; thereby potentially yielding high levels of diversity. Alternatively, to the extent that

individuals seek out others who are like themselves, homophily may result in projects on which participants are more similar than different along a variety of dimensions (McPherson, et al. 2001). Such plausible variation in diversity across OSS projects, coupled with significant prior evidence for the profound impacts of diversity, both positive and negative, on outcomes such as performance, member engagement, withdrawal behaviors, and turnover (e.g., Van Knippenberg and Schippers 2007), suggests that diversity is likely to be an influential factor driving OSS project outcomes (Colazo and Fang 2010, Giuri, et al. 2010, Seidel and Stewart 2011). Indeed, the presence and challenge of diversity in OSS settings has been acknowledged: the Python open source project website noted, "The most important skill Python can teach is the delicate skill of working in a diverse group" (quoted in Ducheneaut 2005: 339).

A deeper understanding of the effects of diversity in OSS projects is important for at least two reasons. First, typical diversity dimensions such as race and gender are not readily observed in computer mediated communications, and other frequently-studied dimensions such as functional background may require different operationalizations due to the unique governance structures in OSS. Thus, the way in which diversity matters may differ between conventional and online settings. Second, findings related to the impact of diversity can be a double edged sword: across research settings, findings related to the impact of diversity are complex and sometimes inconsistent (cf. Homan, et al. 2008, Knippenberg and Ginkel 2010). Equivocal findings suggest that impacts may be context dependent, leading researchers to underscore the importance of contextually relevant moderators (cf. He, et al. 2007, Homan, et al. 2008). Thus it is necessary to understand the specific setting of OSS development to predict the conditions under which different aspects of diversity have greater or lesser effects on relevant outcomes. A contextual factor of particular relevance in OSS projects is development stage. In studies of OSS projects, the impact of other factors on project success has been found to vary across development stages (Setia, et al. 2010, Stewart and Gosain 2006b) and diversity studies in other software development settings also implicate development stage as an important moderating influence (He, et al. 2007).

We study the effects of diversity on two outcomes of significance for volunteer software development organizations: the engagement of community participants as reflected in their contributions

to the effort, and market success indicated by the interest generated by the project in the market (Crowston and Scozzi 2002, Grewal, et al. 2006, Stewart, et al. 2006, Subramaniam, et al. 2009). Because OSS projects fundamentally depend on voluntary contributions to achieve project goals, motivating contributions is crucial for the sustainability of the project. Further, market interest is important because it reflects the number of potential adopters that can be sold consulting services or more sophisticated versions of the application, and can also enhance the reputation of those involved with the project. We develop the theoretical logic explaining how diversity influences the ability of groups to create useful software and affects the key motivations of OSS participants, thereby adding to our understanding of why some projects are more successful than others.

This study makes several contributions to theory and practice. First, though there is a robust literature on the antecedents of OSS success (Grewal, et al. 2006, Singh 2010, Singh, et al. 2011, Stewart, et al. 2006, Subramaniam, et al. 2009), the impacts of diversity on OSS success have generally been overlooked (see Giuri et al. 2010 for an exception). To the degree that diversity varies in voluntary and virtual collectives enabled by technology, and in light of the documented effects of diversity on outcomes in a variety of organizational settings (Joshi and Roh 2009), this is a gap in the literature that needs to be addressed.

Second, though the diversity literature has suggested that context specific moderators are important for understanding the conditions under which diversity matters (Buyl, et al. 2010, Homan, et al. 2008), the effects of contingency variables on the relationship between diversity and outcomes have received minimal theoretical attention. This study develops and tests the logic that explains how application development stage moderates the effects of diversity on market success.

Third, our theoretical development pays special attention to the fact that the nature of diversity that is salient for OSS communities is likely to be different from diversity experienced in collocated non-volunteer contexts. To illustrate, community participants may never see each other so differences that are visually observable recede to the background, whereas differences in communication technology usage patterns become more salient. In isolating aspects of diversity likely to make a difference in OSS

projects, we are able to offer OSS managers practical advice about accentuating the positive and mitigating the negative effects of diversity in their projects. Such guidance is increasingly useful as a greater proportion of organizational work is conducted in work structures that require managers to construct effective groups from a pool of workers that are geographically distributed and fundamentally diverse.

2. Theoretical Background and Prior Research

2.1 The Conceptualization of Diversity in Workgroups

Diversity has been examined in multiple disciplines including sociology, psychology, and organizational behavior from a theoretical and empirical perspective (Van Knippenberg and Schippers 2007). Though a complete review of this work is beyond the scope of this paper, we provide a broad overview. While diversity is defined and conceptualized in a variety of ways across studies, there is agreement that at its essence, it represents "differences" among members of an organizational unit, be it a workgroup, the top management team, or a specific department within a larger organization (Harrison and Klein 2007).

In the past three decades a number of empirical studies have theorized and tested the effects of diversity across a range of settings, workgroups, and tasks (see Horwitz and Horwitz 2007 and Joshi and Roh 2009 for recent meta-analyses of these studies). Despite the fact that diversity exists in many different types of collectives, empirical literature has focused on understanding workgroup diversity – i.e., differences within a clearly identified and bounded collection of individuals with articulated goals and tasks (Van Knippenberg and Schippers 2007). The outcomes for such studies are often metrics related to financial performance (Buyl, et al. 2010), which is not highly relevant for OSS projects. Further, the majority of diversity studies have been conducted in the context of collocated or face-to-face teams, with a handful of recent exceptions examining virtual teams (e.g., Giuri, et al. 2010, Peters and Karren 2009, Staples and Zhao 2006).

Three broad conclusions emerge from the diversity literature. One, diversity may be characterized as a "double-edged sword" (Horwitz and Horwitz 2007, Knippenberg and Ginkel 2010) in that heterogeneity among group members can yield both positive and negative outcomes (Williams and O'Reilly 1998). On the one hand, some types of differences, such as those in functional background, are purported to result in a broader range of perspectives and knowledge becoming available to the group, thereby leading to greater creativity and innovation (Peters and Karren 2009). On the other, differences such as those in cultural norms and beliefs can be a source of conflict among members, yielding dysfunctional outcomes (Kankanhalli, et al. 2006).

A second conclusion that emerges from reviews of the diversity literature is that there is wide variation in the observed effects of diversity. The same diversity dimension may have positive effects in one study, negative effects in another, and no effect in a third (Harrison and Klein 2007). Similarly, the same outcome may be affected differently by different dimensions of diversity. For instance, in a meta-analysis of studies related to the effects of diversity in workgroups, Horwitz and Horwtiz (2007) found that although task-related diversity (acquired characteristics such as expertise and education) was related to team performance, bio-demographic diversity (innate attributes of individuals such as age, gender, and race/ethnicity) was not. This lack of consistency in findings has led researchers to underscore the need for a nuanced approach to studying the impacts of diversity that accounts more carefully for context and conditioning influences (Cannella, et al. 2008, Joshi and Roh 2009).

A final conclusion that can be drawn from the diversity literature is related to the conceptualization of the core construct. Scholars note that there are many attributes and dimensions along which a group can be diverse such as cognitive abilities, personality, age, gender, ethnicity, prior experience, attitudes, etc. (Harrison and Klein 2007, Milliken and Martins 1996). Harrison and Klein (2007) suggest that one explanation for the equivocal findings in diversity research is that the construct has not been precisely defined at a conceptual level. They propose a typology, arguing that diversity encompasses three distinct elements: disparity, separation, and variety, with varied underlying theoretical logics responsible for observed effects.

Disparity diversity reflects situations where some members have more power and resources than others. Findings related to how disparity diversity impacts outcomes are mixed (Anderson and Brown 2010). On one hand, the unequal distribution of resources may have detrimental effects on group outcomes by creating dysfunctional situations such as increased competition among team members, domination and suppression of "voice" by powerful members, and poor communication quality (Pfeffer and Davis-Blake 1992). In contrast to the negative consequences of an unequal power distribution, expectancy theory (Lawler, et al. 1968, Lawler and Suttle 1973) and tournament theory predict disparity positively impacts outcomes (Lazear and Rosen 1981; Rosen 1986). Unequal power distributions can also facilitate the development of hierarchical relationships that speed decision making and facilitate performance (Groysberg, et al. 2011, Overbeck, et al. 2005).

Separation diversity captures differences in position or opinion among unit members, arising from heterogeneity in values, beliefs, and attitudes. To the degree that individuals like to be with "similar others," separation diversity should hinder performance (Locke and Horowitz 1990). Social categorization likewise suggests that homogeneity among group members engenders feelings of an "in-group", thereby facilitating the accomplishment of team tasks (Williams and O'Reilly 1998).

Finally, *variety diversity* indicates the breadth of relevant information that members bring to the group. Greater information breadth leads to higher creativity, more effective idea generation, and broader exploration of diverse perspectives on the task, thereby improving performance (e.g. Jackson, et al. 1995).

Prior work also suggests that the effects of diversity may be conditional on other factors in the workgroup setting. For example, in recent work the impact of variety diversity on outcomes has been argued to be impacted by context specific moderators such as CEO characteristics (Buyl, et al. 2010). Homan et al. also demonstrate diverse teams with higher levels of openness to experience perform better than diverse teams with lower levels of openness to experience (Homan, et al. 2008). Further, as team members spend more time collaborating, the influence of surface level diversity (such as gender and age) tends to decrease (Harrison, et al. 1998, Harrison, et al. 2002). The moderators, CEO characteristics, openness to experience and time, influence how a project is able to manage diversity.

2.2 Diversity in Virtual Contexts

Given our focus on the effects of diversity for global, distributed collectives, prior research on diversity in virtual contexts is relevant. The growth of technology-mediated communication and emergence of virtual teams has motivated researchers to examine the effects of diversity in these settings and, echoing findings for more traditional workgroups, diversity has been found to exhibit equivocal impacts. To illustrate, the literature suggests inconsistent effects for cultural diversity across different contexts. Shachaf (2007) report that cultural diversity can have both positive and negative impacts on global Fortune 500 virtual teams. Lowry et al. (2010) find that in culturally heterogeneous (Chinese and American) technology supported groups, collectivistic members perceive less interpersonal trust than individualistic members, and in their exploratory study of masters level students, Kankanhalli et al. (2006) find that cultural diversity (linguistic and national) is the antecedent of a majority of conflict episodes in the global virtual teams they studied. On the other hand, Alfaro's (2010) results show that national diversity positively impacts team performance in a sample of 247 global software development teams. Likewise, the effects of functional diversity are varied in virtual settings: Kankanhalli et al. (2006) find that functional diversity may generate conflict, while Peters and Karen (2009) report that functional diversity is positively associated with team-rated performance. Thus, as with the wider literature on diversity, it appears context-specific factors must be more closely examined to understand the role of diversity in different types of virtual team settings.

2.3 OSS Project Communities

OSS projects are a type of virtual team comprised of volunteers, and some distinctive characteristics of the context are worth noting. Because the project community's main method of interaction is through technology, diversity dimensions such as race and gender that have implications for group outcomes but require visual observability (Riordan and Shore 1997) may be less salient. Similarly, attributes such as job title that might indicate hierarchical position, authority, or some other form of social order in traditional settings, may not be relevant in a volunteer context, where authority may be acquired in different ways, for example by engagement in and contribution to the community's efforts.

Further, the manner through which individuals garner resources (social and technical) can be distinct in volunteer groups compared to settings where members are paid and there are formal, hierarchical processes that govern the organization. In the latter setting, an individual's wage may signal authority or the possession of resources while wages are not relevant as a differentiator in a volunteer organization. In such contexts reputation in the organization could be a better way to infer individuals' relevant resources, as it is by establishing a positive reputation via contribution that individuals earn decision authority and other privileges.

Third, the OSS community participant functional backgrounds may be distinct from those in more formal software development communities. Prior work indicates that in most OSS communities, two key participant types represent the overwhelming majority: active users and developers (Chengalur-Smith, et al. 2010). Active users include individuals who participate in mailing lists, request features, report bugs, fix bugs, or perform other peripheral tasks that do not involve making substantive code contributions (Ye and Kishida 2003). We include among developers those whom Ye and Kishida classify as project leaders, core members and active developers. Together, these two groups include those whose primary contributions are towards developing the software code (developers) and those who do not have a large direct impact on the code, but contribute in other ways by providing feedback and engaging in discussions (active users).

Finally, OSS community success is distinct from that of other groups. In particular, prior studies on diversity examine success indicators such as financial performance measured by sales (Buyl, et al. 2010), trust (Lowry, et al. 2010) and team rated performance (Peters and Karren 2009). While these outcomes may be relevant for some OSS communities, as discussed, the ability to garner market interest and contributions from the community have been consistently identified as important success dimensions for OSS projects (Crowston and Scozzi 2002, Grewal, et al. 2006, Stewart, et al. 2006).

3. Hypotheses

3.1 Disparity Diversity

Disparity diversity reflects differences in the concentration of valued social assets or resources such as pay, knowledge, status, decision-making authority and reputation among unit members (Harrison and Klein 2007). What effects do such differences have on group outcomes? Noting that disparity diversity in underexplored in the literature, Harrison and Klein synthesize the limited prior work to suggest that disparities are typically associated with negative outcomes such as competition, silence, and suppression of creativity. However, subsequent work notes that effects can be equivocal in that disparity can yield both positive and negative outcomes (Anderson and Brown 2010, Groysberg, Polzer et al. 2011, Halevy, et al. 2011), with its influence varying as a function of the perceived legitimacy of the disparity and the type of outcome considered. For instance, to the extent that paying higher wages to workers who exhibit more effort motivates others to work harder (Becker and Huselid 1992, Ehrenberg and Bognanno 1990), disparity in pay can increase the overall performance and productivity of the group. A belief that harder work leads to higher pay (i.e., a perception of fairness in the distribution system) would lead workers to increase their effort. In contrast, consistent with theories of relative deprivation, if an employee perceives that others are paid more for inappropriate reasons, he may feel discontent, leave the organization, or act out negatively (Harrison and Klein 2007, Pfeffer and Davis-Blake 1992). Thus, disparity is likely to be beneficial for eliciting individual effort and commitment to an organization when the distribution of the key resources is perceived as fair, and negative when it is not (Halevy, et al. 2011).

Eric Raymond, the first president of the Open Source Initiative, characterized the OSS community as a meritocracy in which reputation and authority, valued social assets, are gained by the knowledge that is shared with the community (Raymond 1998). Roberts et al. (2006) found corroborating evidence for this assertion in their study of Apache, a large OSS project, where they observed a meritocratic structure, with progress being based on ability and talent, demonstrated by members' contributions. In such a meritocracy, advancement is predicated on a member's ability to draw attention to his knowledge and talents; in OSS communities, the means by which one does so is by making contributions (von Krogh, Spaeth et al. 2003; Dahlander and O'Mahony 2010). In the OSS context, the

ability to make useful and valued contributions to the core tasks facing the community is heterogeneously distributed among group members. Simply put, some members are capable of providing greater volume of input into the activities of software development than others.

Sharing contributions draws others' attention to the contributor, and not only enhances the reputation of the contributor, it can also create other positive externalities. These include being granted greater authority in the project including the privilege of making changes to the common code repository or, in large projects, becoming a formal member of related governing bodies such as boards of directors (Bird, et al. 2007, Dahlander and O'Mahony 2010, Fang and Neufeld 2009). We conceptualize disparity among participants on the basis of reputation derived from their level of project contribution (e.g., providing bug fixes, answering others' questions, or creating new features). Contributions garner attention that, in turn, enhances the reputation of the contributor.

We expect positive outcomes from this form of disparity based on notions of fairness and legitimacy: a meritocracy is generally considered a legitimate way to distribute resources as those who contribute more reap greater rewards. In a typical OSS project, participants' contributions are visible and transparent, allowing the community to observe the relationship between effort and advancement, and enabling participants to take steps to advance in reputation if they wish. High contribution disparity may elicit greater effort from members by tapping two key OSS participant motivations. The first is attaining and maintaining reputation (Hertel, et al. 2003, Lakhani and von Hippel 2003). Those who contribute extensively can further develop and maintain their reputation based on advice and help distributed to less active contributors. Here, the presence of low contribution members provides an audience for the high contribution members, and a pool of individuals against whom they can differentiate themselves.

As emphasized in prior work, a second critical motivation for OSS project participants is the opportunity to learn (Fang and Neufeld 2009, Roberts, et al. 2006). Disparity in contributions supports participant learning in at least three ways. First, because active contributors openly share application relevant knowledge, they represent a key resource from which those with lesser knowledge, and

correspondingly fewer contributions, can learn. Individuals with fewer resources often prefer to work with those with greater resource endowments (Gould 2002) so that they may reap the benefits of others' endowments. Second, to the degree that a desire to learn motivates greater engagement from the learner (Fang and Neufeld 2009), it may drive low contribution members to increase their level of contribution as well. Finally, we note that high contribution members not only provide low contribution members the opportunity to learn, they also symbolize evidence that increasing efforts can lead to increasing rewards in the project.

Harrison and Klein (2007) note that disparity diversity is not symmetric, i.e., a work group where only 10% of members have a significantly superior endowment of an asset or resource as compared with others is more diverse than one where 90% are highly endowed and 10% are resource poor. In an OSS project with low disparity diversity, all participants are either high or low contributors. In the former case, there are limited opportunities to build reputation as there is no audience to attract attention and impress, in the latter, learning opportunities are limited. By contrast, in an OSS project with high disparity diversity the small number of high reputation participants have access to a larger audience, and lower reputation members have opportunities for learning and advancement because there is "room at the top." To the extent that disparity leads to both learning and reputation-based motivations for participation being fulfilled, we expect:

H1: Contribution-based disparity diversity is positively associated with OSS project community engagement.

We have argued that diversity in contribution levels among community members will create a virtuous cycle of engagement. We expect similar positive outcomes for market success, though the underlying mechanism is different. In order to attain market success with its products, an OSS project must successfully integrate participant contributions into a coherent final product and respond effectively to multiple issues that typically need to be resolved in the artifact construction process. For instance, decisions must be made about the kind of algorithmic solutions employed, the type of programming languages used, and the priority given to additional features. To the degree that technology evolves

rapidly, an application that is able to respond in a timely manner to meet the dynamic needs of users is likely to exhibit superior market performance as compared with products that are slower to react to customer requirements. Rapid response requires efficient decision-making and, as noted in prior research, speedy and efficient decision making in group settings is facilitated by a hierarchical social structure (Cooper and Withey 2009, Groysberg, et al. 2011, Halevy, et al. 2011). The implicit social order legitimizes the authority of those higher in the hierarchy to efficiently resolve disputes among those positioned lower in the hierarchy (Halevy, et al. 2011).

Although OSS projects are generally regarded as non-hierarchical in structure (Gallivan 2001), when the contributions of community participants and their resultant reputations exhibit significant variation, the community gains an emergent social structure, with those who have made more contributions being positioned higher in the structure than those who make fewer contributions (Dahlander and O'Mahony 2010, Seidel and Stewart 2011, Stewart 2005). Because the position in the social structure is based on merit, it is likely to be perceived as fair (Halevy, et al. 2011) and community participants will be more willing to concede decision making and conflict resolution authority to the more active contributors. In contrast, in an OSS community comprised predominantly of participants who contribute at similar levels (i.e., low disparity diversity) it may be less clear who can guide and steer decision making processes efficiently. In this case conflicts can emerge that impede and stifle application development, thereby diminishing the utility of the application for users. Following from this logic we test:

H2: Contribution-based disparity diversity is positively associated with OSS project market success.

3.2 Separation Diversity

Separation diversity "indicates differences in position or opinion among unit members" (Harrison and Klein 2007), especially as they relate to team processes and goals. Such differences have been implicated in a range of negative outcomes, including reduction in group cohesiveness, heightened

conflict, and poor task performance (Harrison and Klein 2007). A recent meta-analysis shows that the negative impacts associated with some separation variables are especially prevalent in high technology settings, and in interdependent teams, both of which characterize OSS communities (Joshi and Roh 2009). A key form of separation diversity that is pervasive in OSS projects by virtue of the global and distributed nature of OSS communities is cultural separation (Ågerfalk and Fitzgerald 2008). Following prior work that recognizes language and country of residence as critical aspects of culture (Alfaro 2010, Anderson and Hiltz 2001, Chen, et al. 2006, Claes 2009, Henderson 2005, Hinds and Mortensen 2005, Staples and Zhao 2006, Vignovic and Thompson 2010, Von Glinow, et al. 2004, Welch and Welch 2008), we focus on diversity in OSS project participants' spoken languages and home countries as salient indicators of cultural separation diversity.

Language yields "separation" among members of a collective because it is through language that those embedded with a specific culture encode and communicate perceptions and associations that disclose a common outlook (Barley, et al. 1988). Cultures have distinctive norms about the use of language: for example Asian cultures tend to prefer indirect expressions of negative feelings whereas New Guinea villagers express anger by shouting abusive comments (Von Glinow, et al. 2004). Language is fundamentally a mental model of the world, framing activity and behavior (Welch and Welch 2008). The way people use language can often surface disparate metaphors for work and related expectations for roles and objectives across national cultures (Gibson and Zellmer-Bruhn 2001). In addition to language, individuals' home country is a second common indicator of culture (Alfaro 2010, Earley and Mosakowski 2000). As has been extensively documented, countries exhibit heterogeneity across a variety of dimensions, including the extent of hierarchy present in the social structure, the level of individualism, and long- versus short-term orientation in decision making (Hofstede 1985). Thus, members of a collective who come from different nations are likely to have distinct values, beliefs, and preferences about interaction and other task-related behaviors.

We expect culture-based separation diversity to influence OSS community engagement by dampening two key motivations for participation: learning and identification. Learning in OSS groups is

facilitated by communication among members, e.g. via the asking and answering of questions or explanations of complex coding decisions (Bechky 2003, Dahlander and O'Mahony 2010). However, diversity in the spoken languages of members from different cultures can impede the effective communication of knowledge (Welch and Welch 2008), thereby inhibiting learning. Further, the technical errors and etiquette breaches made by individuals communicating in a non-native language may reduce others' assessment of their intelligence and desire to work with them (Vignovic and Thompson 2010). In essence, diversity in languages creates barriers to effective communication that limit the learning benefits that might otherwise accrue to participants, thereby dampening their motivation to participate.

In addition to reducing the desire to collaborate, lack of conformance to communication norms in email has been shown to reduce perceived agreeableness and trustworthiness (Vignovic and Thompson 2010). Cultural diversity can trigger social categorization processes (Van Knippenberg and Schippers 2007) that evoke stereotyping, distrust, suspicion, and other negative attitudes and behaviors (Barner-Rasmussen and Björkman 2007, Welch and Welch 2008). Cross-national teams frequently face these types of significant communication challenges that can reduce member identification with the workgroup (Fiol 1991, Gibson and Vermeulen 2003). Identification is an important motivational factor for inducing individuals to engage in interactions and contribute to the workgroup discourse (Bergami and Bagozzi 2000), and it has been specifically isolated as an important motivational factor in the OSS context (Fang and Neufeld 2009, Stewart and Gosain 2006a). We therefore test:

H3: Culture-based separation diversity is negatively associated with OSS community engagement.

Echoing findings regarding the increased difficulty of communication in groups where members come from different cultures and speak different languages, in the context of software development teams in particular, He et al. (2007) find that cultural diversity hampers the development of team cognition – an awareness of "who knows what" and a shared task understanding reflected in common mental models. The lack of a common mental model implies that there will be difficulty in creating cohesion and

congruence in terms of an overall vision for the product. This may reduce overall quality of the final product in two ways: by making it difficult to reconcile disparate perspectives into a coherent codebase and by increasing the likelihood that the community will fracture, leading to forking of the codebase.

It has been suggested that a culturally diverse group may find it difficult to establish common ground in communication (Gibson and Gibbs 2006, Von Glinow, et al. 2004). Dependence on the computer mediated communication provided by the platform and the lack of opportunities for face-to-face communication may exacerbate negative attributions based on different communication styles among participants from different cultures (Vignovic and Thompson 2010) and limit the community's ability to find common ground and resolve potentially conflicting goals for the project (Dreu and Weingart 2003). To the degree that the failure of members to agree on a common vision and course for the project hinders the development of a coherent and useful feature set, the utility of the product will diminish and market interest may be negatively impacted.

Further, limited ability to resolve conflict and multiple ideas about the options for development can lead to attrition; indeed, diversity based conflict has been shown to result in turnover (McCain, et al. 1983). When participants stop contributing, product functionality typically lags evolving user needs. In an OSS environment where participation is voluntary, conflict can result in a phenomenon labeled 'forking' where a subset of the participants starts a different version of the project (Raymond 1998). Over time as different versions become incompatible, the utility of the software declines. Following this logic we test:

H4: Culture-based separation diversity is negatively associated with OSS market success.

3.3 Variety Diversity

Variety diversity is defined as variation in the knowledge and experience of the group: "the number and spread of 'batches' of information content, experience, or unique network ties available across unit members" (Harrison and Klein 2007, p. 1204). Congruent with the logic of theories of information processing and the principles of requisite variety (Harrison and Klein 2007, Van Knippenberg and Schippers 2007), heterogeneity in informational resources is expected to exert a positive influence on

group outcomes. This association is especially true in high technology settings, where combining different sources of knowledge to create innovative products is important for success (Joshi and Roh 2009). Two types of knowledge are useful for communities engaged in technological innovation such as software development, *how-to* knowledge and *awareness* knowledge (Rogers 1995, Tornatzky and Fleischer 1990). The former concerns an understanding of the processes and practices involved in developing software, and includes knowledge about software engineering and programming languages, while the latter relates to the use of the software in a particular context.

How is knowledge variety reflected in OSS communities? Prior work has frequently considered differences in functional background or training as important sources of knowledge variety (e.g. Buyl 2011, Canella 2008). However, while research has shown the importance of integrating different knowledge sets in software development work (Giuri, et al. 2010, He, et al. 2007), OSS projects are not typically organized in functional departments, nor are people employed in specific functional jobs within the project (Seidel and Stewart 2011). Thus functional background is not an appropriate indicator of variety in this setting. A construct that is conceptually similar to functional background in the OSS context is the participant role (Ye and Kishida 2003), which reflects how much and what type of knowledge the individual personally possesses or can access.

As noted earlier, the dominant roles within an OSS project community are those of developer and user, and occupants of these roles bring different knowledge and experience to the project. To the extent that the designation of the developer role is predicated on demonstrated mastery of programming and the use of the development platform tools (Ducheneaut 2005, von Krogh, et al. 2003), developers are likely to possess the how-to knowledge needed to further project goals. However, developers may possess a limited stock of awareness knowledge based only on their own local contexts of use. Users are positioned to bring greater awareness knowledge to the project derived from their more varied usage settings (i.e., use of the software in the different organizational environments in which the users reside, in the different software environments present in those organizations, and for a potentially wider variety of tasks).

When there is a balance between types of participants and the different types of knowledge they

possess, there are opportunities for learning and reputation building. Often, project developers begin their association with a project as active users (von Krogh, et al. 2003). These contributors with less how-to knowledge may be motivated to participate by the opportunity to learn from the more technically savvy members and further develop and refine their technical skills. On the other hand, high ability workers may choose to engage in teams with a wide range of others because of the opportunity to add to their reputation and exercise decision authority among a peer group (Hamilton, et al. 2003). Being able to build their reputations in front of a wider and more diverse audience may potentially enhance developers' career opportunities (Roberts, et al. 2006). Further, the developers may also be motivated to participate by the opportunity to learn about different usage contexts, and thereby extend their knowledge of how they can apply their skills in different contexts. Based on this logic we predict:

H5: Role-based variety diversity is positively associated with OSS community engagement.

Indeed, users possess critical knowledge to guide the development of the software in the direction where it will create the most value in its uses, and the importance of obtaining user input for ensuring success is widely recognized in the literature (Lakhani and von Hippel 2003). A broad repertoire of knowledge available in a community, reflected in a balance between users and developers, should enable the OSS community to construct a product that meets a variety of user needs. In other words, the product should be able to satisfy the instrumental needs of a larger user base, thereby increasing its utility and stimulating higher market interest. Thus, we predict:

H6: Role-based variety diversity is positively associated with OSS project market success.

3.4 The Moderating Effect of Development Stage on Market Success

Given some inconsistency in past work regarding the influences of diversity on group outcomes, several authors have highlighted the need to consider contextually appropriate moderating factors (Homan, et al. 2008). In top management teams, for instance, characteristics of the CEO have been explored as a moderator of the effects of diversity (Buyl, et al. 2011). Time has been identified as a potentially important moderator across a wide range of team settings, including software development teams. Evidence suggests that the effects of diversity in teams change over time (Harrison, et al. 1998,

Harrison, et al. 2002, He, et al. 2007). For example, as team members spend more time working together, the influence of surface level diversity (such as gender and age) tends to decrease (Harrison, et al. 1998, Harrison, et al. 2002).

The importance of time in conditioning the effects of diversity is predicated on its role as a proxy for the extensiveness of interaction within the workgroup. What does the concept of time reflect in the OSS context? We suggest that time is unlikely to accurately signify the amount of interaction among the OSS workgroup participants as it does in more traditional team settings for at least two reasons. First, though OSS projects may exist for a long time, they often experience periods of dormancy with limited or no interaction among participants (Evangelopoulos, et al. 2008). Second, the volitional nature of membership often creates turnover in the project, further reducing the appropriateness of time as an indicator of individuals' experience in working together (Robles and Gonzalez-Barahona 2006). We therefore suggest that in the OSS setting, progress on the team's task, i.e., project development stage, is a more relevant indicator of how much the community has worked together to accomplish project goals compared to time.

Project development stage not only captures the amount of OSS project members' shared history, it also indicates the degree to which members' past contributions have been successfully integrated into the main shared artifact, the software. OSS projects typically move in a sequential manner through planning, alpha, beta, and more mature stages, with the software code being further developed and refined in each stage (Stewart and Gosain, 2006b). Indeed, development stage has been established as a salient characteristic of software engineering groups and open source software communities (Sarker, et al. 2001, Setia, et al. 2010, Stewart and Gosain 2006b, Subramaniam, et al. 2009, Whang 1995) exhiting both direct and moderating effects on outcomes. For example, Subramaniam et al (2009) show that OSS applications in more advanced stages of development are associated with more user and developer interest and project activity (Subramaniam, et al. 2009). Highlighting the conditioning influence of stage, Stewart and Gosain (2006b) demonstrate that stage alters the impacts of trust and ideology on perceived and

actual OSS project effectiveness. Setia et al. (2010) provide evidence that stage alters the relationship between peripheral developer participation and product quality assessment as well as market success.

As shown in Figure 1, development stage is theorized as relevant to understanding the influence of diversity on market success. To the degree that disparity diversity yields an implicit "management structure" for the project and because projects in later stages have more to manage, the benefits of disparity diversity and more efficient decision making should be more apparent for projects in later stages compared to those in earlier stages. That is, as a project progresses to later stages of development, it typically increases in size and complexity, as well as attracting a larger user base, with more varied needs. Consequently, the number of decisions that need to be made increases, as does the complexity of those decisions, amplifying the need for efficient decision-making structures. As contribution-based disparity facilitates those structures, its effects should increase in later stages.

H7a: The positive relationship between contribution-based disparity diversity and OSS project market success will be stronger for projects in later development stages.

For separation diversity, having a more well-defined shared artifact to focus group efforts is likely to reduce the negative impact that separation has on the utility of the software. We argued for a negative effect of separation diversity on market success because such diversity makes it more difficult to develop a cohesive, shared vision for the application. This is likely to be especially evident in early stages when an overall direction for the evolution of the software may not yet be clearly defined, allowing more room for misunderstanding and disagreement. As team members interact to successfully advance the software to more mature stages, they develop team cognition, providing a basis for shared understanding (He et al. 2007). Further, as the software moves to a more mature state, its requirements and functionality become more clearly and explicitly defined (Scacchi 2002). A more advanced artifact creates common ground, acting in essence as a boundary object (Star and Griesemer 1989) that focuses participants' further efforts, helping to ameliorate the negative impact of separation. We therefore hypothesize:

H7b: The negative relationship between culture-based separation diversity and OSS project market success will be stronger for projects in earlier development stages.

For variety diversity, the importance of leveraging different knowledge sets to spur market success may decrease as the development stage of the OSS application advances. Unlike software applications developed using traditional methodologies, OSS projects frequently do not begin with a clear set of requirements outlining the intended functionality of the application to be developed (Scacchi 2002). In the early stages of a project, before the form of the application is clear, the breadth of ideas and knowledge reflected by the OSS community variety diversity can be especially beneficial in arriving at the application feature set. Variety diversity can be used to isolate which features will attract a broad set of users and thus aid in identifying the optimal set of application functions with respect to that OSS community. As the application advances to a stage where its feature set is well defined, the need for innovative ideas is less critical as compared to well-defined activities and processes that will steer the product towards completion and support effective management of ongoing maintenance activities. Consequently, the effects of variety diversity on market success may be less potent as the project advances towards higher development stages and the application, as well as the set of its potential users, becomes more stable. Thus we propose:

H7c: The relationship between role based variety diversity and OSS project market success will be stronger for projects in earlier development stages.

4. Methodology

4.1 Data and Sample Construction

Data for this study are drawn from the SourceForge Research Data Archive, an archival database of OSS projects provided by SourceForge to the University of Notre Dame (Van Antwerp and Madey 2008). The archive includes information about project community participants' patterns of interaction with projects such as their posts and responses in project forums. The archive also provides information about the project, including when the project was registered, the application's development stage, the intended users of the application, and the number of times it has been downloaded.

SourceForge is one of the most popular OSS development platforms[1]. It allows community participants to observe others' project activities, send bug reports and feature requests, post to forums, launch new OSS projects and join existing ones, coordinate and work jointly on specific OSS projects, and integrate the software produced into a larger software application. We sampled projects that registered on Sourceforge between June 1, 2004 and May 30th, 2005.

There were 31,591 projects that registered during this time. However, many of these projects had only one participant (i.e., only the project founder participated). Because we are interested in community level constructs, we excluded projects that had fewer than 5 developers[2], which reduced the sample to 1,351 projects[3]. Finally, many projects, though registered on SourceForge, are never developed on SourceForge, or remain listed on SourceForge after they have effectively been abandoned. We eliminated such projects, only including those that have some activity during a two year observation period. There were 357 projects out of the 1,351 that had activity during the observation period. This final sample contains projects developing a broad range of application types, including projects that develop utilities to aid in software development, management, statistical analysis, education, online sales and entertainment.

4.2 Operationalization of Variables

Level of Analysis. The level of analysis is the OSS project community, which includes both developers and active users focused on a common application. Drawing on prior research that suggests both developers and non-developer users make important contributions to OSS projects (Dahlander and O'Mahony 2010, von Krogh, et al. 2003), and consistent with the theoretical logic underlying the hypotheses, we construct each measure for the entire project community.

Dependent Variables. We used two distinct measures of OSS project success utilized in prior studies: community engagement and market success (Crowston and Scozzi 2002, Grewal, et al. 2006, Stewart, et al. 2006, Subramaniam, et al. 2009). To allow for a time lag between when we observe

[1] According to http://sourceforge.net/about as of February 2009 they hosted more than 230,000 software projects
[2] Below, we discuss the robustness of our findings for different sizes of development groups.
[3] We examine the development group size in July 2006, after all projects have had at least one year to attract developers.

diversity (in year one) and its effects, we use data from *after* the project has been registered on SourceForge for one year to measure community engagement.

Community engagement is operationalized using the number of artifact (e.g. bug report and feature request) closures and responses to forum posts associated with the project made by all community participants during the project's second year on Sourceforge. Closure represents a way to respond to a bug report or feature request. We focus on responses (versus all interactions) because in order to respond to a post or close a bug, the participant has to be aware of both the initial question and also a solution, suggesting he or she is making a valuable contribution to the project. For instance, the responder has to have read a post to respond and understood a bug report to propose a solution.

Because forum posts and artifact closures are qualitatively different types of activities, and because many projects have many more forum posts than artifact closures, we computed a factor score of the total number of closures and the total number of response posts to represent community engagement. Including closures and forum response posts helps ensure that the dependent variable captures the contributions of both users and developers. Table 1 shows that the sum of the number of artifact closed and the number of response posts load together and separately from other measures.

For market success, we use total number of times the application was downloaded from SourceForge as of July 2006. The last project in our sample registered on May 30th 2005 and so observing the number of downloads in July 2006 is a measure of the outcome *after* the first year when the independent variables are measured. The number of downloads is often used as a surrogate for number of units sold when a software product is attainable without cost (e.g. Chandrashekaran, et al. 1999, Crowston and Scozzi 2002, Grewal, et al. 2006, Setia, et al. 2010, Subramaniam, et al. 2009). It is a market-based measure of popularity associated with application use.

Independent Variables. To provide evidence that diversity is an antecedent of the outcomes in this study, we use data from the project's *first* year for the measurement of independent variables. As noted, in the context of OSS projects, attention is garnered, reputation built, and authority achieved based on participants' contributions (Dahlander and O'Mahony 2010, von Krogh, et al. 2003). Drawing on

Dahlander and O'Mahony's (2010) finding that participating in technical conversations significantly impacted OSS participants' progression to project leadership positions, we use a measure that captures the attention a focal participant attracts from all other community participants: addressing artifacts, or responding to posts. We observe the same types of activities for disparity diversity as those that are used to measure community engagement, but any one activity is either used to measure community engagement (if observed in year 2) or contribution based disparity (observed in year 1). A participant's contribution level is measured as the sum of all responses to posts and artifact closures in year one. Unlike the measurement of community engagement, we do not use a factor score to measure an individual participant's contribution because there is no a priori expectation that these two activities would be correlated at the individual level. While responding to posts and closing artifacts both reflect engagement at the community level, many individuals participate only in certain kinds of community activities: i.e., a division of labor often evolves such that some participants focus on code-related activity such as addressing bugs or adding features while others contribute by providing end-user support in discussion forums (Ye and Kishida 2003). Thus while a vibrant community will exhibit high levels of both of these kinds of behaviors, they may not be correlated at the individual level. Following recommendations in prior work for the appropriate index of disparity diversity (Harrison and Klein 2007), we use the coefficient of variation of each participant's contributions to create a community level measure of contribution-based disparity diversity.

We use the language spoken by the participant and the country the participant is associated with in the calculation of the estimate for each project community's cultural separation diversity. We compute Blau's index (1977)[4] of participant language and participant country for each project community. We then use a factor score of the Blau indexes to capture separation diversity. Table 1 shows that the Blau index of language and country load together and separately from other measures.

[4] Blau's index is calculated as $1-\sum p_k^2$ where p is the proportion of unit members in the Kth category. We have 2 categories, developers and active users. Values can range from zero to (K-1)/K. Higher numbers indicate that members are more equally spread across the groups.

Variety diversity, reflected in the degree to which participants are dispersed across roles (developer versus active user) is also operationalized using Blau's (1977) index. Participants are categorized as developers if they are listed as developers or administrators during December, 2005, if they make CVS commits[5] during the first year, or if they are responsible for releases during the first year. The active users for each project are identified as individuals who reported bugs, requested features or support or participated in project-defined forums during the first year, but were not identified as developers using the criteria above.

OSS applications typically progress sequentially through a series of stages; on SourceForge these stages are planning, pre-alpha, alpha, beta, production and mature. To measure development stage, the moderator variable in the research model, we utilize the development stage reported on SourceForge. Although there are six stages, to reduce variance due to reporting errors and form exclusive categories, following the procedure used in prior work we use only three categories: pre-beta, beta, and post-beta (Setia, et al. 2010, Stewart and Gosain 2006b). This categorization captures three crucial phases in development: prior to having a "user ready" version, having a relatively complete product available for testing by users, and having an application considered to be mature. The sample included 126 projects in the pre-beta stage classification (planning, pre-alpha, alpha), 89 in the beta stage and 142 in the post beta stage classification (production and mature).

Control Variables. We controlled for seven additional potential sources of variation in OSS project success. First, we include the total number of participants as a control, as a larger number of participants may indicate more resources that can be leveraged to develop software (Butler 2001), and having more participants, *ceteris paribus*, is likely to lead to a greater number of contributions. Second, we control for whether or not the project uses the Gnu GPL, as prior studies have shown that license choice affects user interest (Stewart, et al. 2006). Third, we control for the type of application because prior work has shown that applications created for use by developers attract more interest (Comino, et al.

[5] The Concurrent Versioning System (CVS) manages the software development process by keeping track of changes: what was changed, when it was changed, who made the change. A commit occurs when a developer uploads altered code.

2007). We also control for the quarter during which the project registered on SourceForge; this captures both the amount of time the project has had to attract user interest and the possibility that there could be external events impacting all projects that register during a given time. Applications vary in the size of the audience they seek to reach and so we control for the scope of the application. We use a factor score based on the number of intended audiences and the number of topics (these are both self-reported on Sourceforge). We also control for the number of other projects our focal project is related to through shared developers because project networks have been shown to drive OSS project success outcomes (Grewal, et al. 2006). Finally, we control for the number of releases because a new release may prompt additional application downloads (the measure of market success).

5. Analysis and Results

5.1 Analysis

Table 2 displays descriptive statistics and summarizes variable intercorrelations. Comparison with descriptive statistics reported in prior studies indicated that our sample appears similar to past work in terms of the number of developers and distribution of licenses (Stewart et al. 2006, Lerner and Tirole 2002, Stewart and Gosain 2006a). Models used to examine the effect of diversity on community engagement are estimated using ordinary least squares regression. The dependent variable in hypotheses related to market success, number of downloads, is a non-negative, skewed count data set; therefore we used negative binomial moderated regression analysis to test these hypotheses (Allison 1999). The negative binomial is a particular case of the Poisson regression. We took several steps to ensure that the data satisfy the underlying assumptions of the statistical techniques. The number of participants in year 1, the number of related groups and the number of releases were left skewed and so we performed a natural logarithmic transformation on them. The model variance inflation factors (all below 3) and the correlations between the independent variables (see Table 2) suggest that multicollinearity is unlikely to be a concern in model estimation. The Durbin Watson value (<3) showed no evidence of serial correlation. Observation of the residuals and Cook and White's test suggest that the errors are normally

distributed.

5.2 Hypothesis Tests

Results are presented in Table 3. Models 1 and 2 predict community engagement , while Models 3-5 estimate the effects of diversity on market success. All models are statistically significant at $p<.01$. Model 1 includes control variables and model 2 includes control variables and main effect variables. The increase in explained variance from model 1 to model 2 is significant ($p<.01$) and so Model 2 is used to test hypotheses. Model 2 provides support for H1 ($\beta =.12$, $p<.05$), H3 ($\beta =-.18$, $p<.01$), and H5 ($\beta =.08$, $p<.1$).

Model 3 includes controls only and model 4 includes controls and main effects. The increase in Likelihood Ratio Chi Square from model 3 to model 4 is significant ($p<.01$). Model 4 provides support for H2 ($\beta=.15$, $p<.1$) and H6 ($\beta =2.49$, $p<.01$). Model 4 contradicts H4 ($\beta =.33$, $p<.01$). Model 5 includes control, main effect and the interaction variables. The increase in Likelihood Ratio Chi Square from model 4 to model 5 is significant ($p<.01$). Model 5 is therefore used to test the interactions. Model 5 suggests that stage moderates the impacts of diversity on market success ($p<.01$).

The significant interactions revealed in model 5 are graphed in figures 2-4. Consistent with hypothesis H7a, figure 2 shows that disparity has a more positive impact on market success in later stages. Surprisingly, figure 3 suggests that separation diversity has a more positive impact on market success in later stages. Finally, figure 4 shows that variety diversity has a more positive impact on market success when the application is in earlier stages of development compared to when the application is in later stages of development, supporting H7c. Results of hypotheses tests are summarized in table 4.

5.2. Robustness Checks: Alternate Model Specifications

To increase confidence in the findings we conducted a series of robustness tests including changing the minimum number of developers used for sample selection (online appendix tables A and B), using alternative measures for disparity diversity (online appendix tables C, D and E), and for community engagement (online appendix table F), as well as changing the observation point for market success (online appendix table G).

While it is important for our research objective to focus on groups with more than one developer (to allow for the presence of diversity), selecting projects based on a cut-off of 5 or more developers is somewhat arbitrary. We examined the models in a sample of projects that had greater than 3 developers (N=486). After eliminating one outlier, the results remain qualitatively the same in terms of the significance and direction of the relationships (see online appendix table A). The research model is also tested in a sample of projects that had greater than 5 developers (N=252). Results were the same with one exception. The impact of disparity diversity on market success was no longer significant. The insignificance may be due to the smaller sample size (N=252 instead of N=357) which reduces the power to detect the effect we see in the larger samples (see online appendix table B).

Online appendix tables C, D and E report the use of alternate measures of contribution-based disparity diversity. Table C shows that if the number of artifacts (bug reports or feature requests) closed is used as the only way to contribute, disparity diversity facilitates community engagement and not market success. In contrast, Table D suggests that when only the number of response posts is used to indicate contribution-based disparity, diversity facilitates market success and not community engagement.

This pattern of results is consistent with the theoretical logic underlying the predicted effects of disparity diversity in hypotheses 1 and 2. Closing artifacts is an internally focused activity, closely related to code development (e.g., closing a bug or adding a feature typically requires code modification), and provides opportunities to learn and build a reputation for technical expertise. Hence it has a positive effect on community engagement as proposed in H1. The discussion forums are more likely to be the channel through which participants make decisions to resolve issues in the community and strategically direct application development. Thus disparity in response posts may be more directly related to the efficiency gains associated with an established decision-making structure, thereby influencing market success, as proposed in H2.

In prior work, CVS commits have been used as a measure of developer contributions (Grewal, et al. 2006). Although our theorizing is at the level of the entire community including active users, we nonetheless checked the robustness of findings using CVS commits to measure participant contribution.

Some of the projects in the sample did not use CVS (i.e., they used another versioning software), limiting the sample for this analysis to 200 projects that had CVS commits in year two. Table E models 2 and 5 show that, consistent with the main analysis, disparity diversity measured using CVS commits positively affects market success and community engagement. Model 3 indicates that stage does not moderate the relationship between disparity diversity (measured using CVS commits) and market success. There are at least two potential explanations. One is that the reduced statistical power in the smaller sample cannot detect the interaction. A second is that the hierarchical social structure that disparity in CVS commits creates facilitates decision-making related to coding issues, which are equally relevant throughout the application lifecycle. In contrast, more strategic decisions that require deliberations in forums are more important in later stages when the project is likely dealing with a wider array of external stakeholders. Hence, hierarchy based on response posts and responding to bugs and features gain importance as project stage advances, but hierarchy based on CVS commits maintains a similar level of impact over time. Table F displays results based on alternate measures of community engagement. The results of the main analysis are consistent when using the number of artifacts opened in year two or the number of artifacts opened plus the number of artifacts closed in year two as the measure of community engagement. As noted, CVS commits are a commonly used measure of OSS project success because they reflect invention that is validated by peer developers on the project (Grewal, et al. 2006, Singh 2010, Singh, et al. 2011). Using CVS commits as a measure of community engagement on the same subsample of 200 projects as above, we find that diversity in the community has no effect. Again, this result may be due to the reduced sample size. Alternately, it could be due to the fact that CVS commits are a very limited measure of engagement. This measure only captures activity by a small subset of participants (those that have attained CVS access) and it only captures a subset of the activities of those individuals (i.e., they may contribute in other ways also).

Because market success (number of downloads) may change quickly over time as users gain or lose interest we tested whether the impacts of diversity on market success was sensitive to the time of measurement. The impact of diversity on the numbers of downloads in July 2006 is the same as the

impact of diversity on downloads in June 2006 and August 2006 (see online appendix table G).

5.3. Discussion

Despite the emergence of the OSS development paradigm as a viable alternative to traditional software development models and the significant human capital vested in this activity, we lack a thorough understanding of how OSS project participant characteristics influence project outcomes. Acknowledging that diversity is likely to be prevalent in varying degrees across open source communities, in this study we leveraged the robust diversity literature to conceptualize diversity as a key antecedent of OSS project outcomes. We drew on the typology developed by Harrison and Klein (2007) and instantiated it in the context of a distributed, volunteer collective. Results from the empirical analysis of data on 357 SourceForge projects generally supported the hypotheses, reinforcing the importance of diversity in driving outcomes for non-traditional work settings. Our theoretical development suggested, and empirical results confirmed, that diversity can have both positive and negative effects on open source project success depending on the type of diversity and type of outcome considered. We also found that application development stage is a contextually relevant moderator that conditions the impact of diversity on market success.

Effects of culture-based separation diversity (H4 and H7b) were contrary to predictions. We found evidence that culture-based separation diversity positively impacts market success, and this effect is stronger in later development stages. To better understand this result, we examined qualitative data (i.e., discussions in project forums) associated with some of the projects that had high separation diversity. The Exponent Content Management System is an exemplar. Examination of this project suggests that having participants who speak different languages and are from different countries facilitates the development of an application usable in differing contexts. For instance, one of the threads that had the most posts was associated with translating the application into other languages, including German and French (http://sourceforge.net/projects/exponent/forums/forum/405066/index/page/34). In other words, greater separation diversity enables the application to be used by a wider range of people, thus market success is enhanced. As shown in figure 3, this effect is most apparent in later development stages. At this point

the application is well-formed and may act as a boundary object facilitating communication across culturally diverse individuals. There is a common core application available to be adapted to use in different contexts. In early stages when there is little functionality, it is more difficult to envision and communicate how the application may be adapted for use in different cultures.

5.4 Limitations

We acknowledge some limitations of the study that also offer useful opportunities for future work. One limitation relates to generalizability. Although SourceForge is the most populous open source development platform, flagship projects such as Linux and Apache are not hosted on it. The OSS projects observed in this study may not be representative of OSS projects that use other platforms (e.g., GForge, at http://gforge.org) or those that do not use a platform that hosts multiple projects. It may be more difficult to attract interest (achieve market success) if the project does not share a platform that draws a large number of users. Thus, results from this study generalize only to communities that use a platform that has a profile and infrastructure similar to SourceForge. However, since SourceForge is one of the largest available platforms, this limitation is minimal. Further, the analysis focused on relatively vibrant communities that had at least 4 developers; further work is needed to assess whether results generalize to projects that have very small development groups.

A second limitation of this study is that the variables in this model are measured using archival data. While the use of archival data is a strength in that the data are objective and free of respondent bias, with archival data it is challenging to map the data perfectly onto the theoretical constructs. Although the proxies used to operationalize the constructs in this study have significant face validity, surveys with measures precisely designed to match the research constructs could be used in addition to the archival data. We note that our findings survive a variety of robustness tests, further increasing our confidence in the quality of the proxy measures. A related limitation is that we identify participants by their userid, and it is possible that multiple participants use the same userid or one participant uses multiple ids. However, this method of identifying participants is commonly used in other studies (e.g. Butler 2001), and the motivation to build reputation reduces the likelihood that participants would disperse their efforts across

multiple ids.

Finally, contribution-based disparity diversity is relevant in other meritocracies in which reputation is based on input provided to further community goals. Measures of reputation based on computer mediated communication are most relevant in situations where there is a public record of contribution. For instance we would expect our results to extend to other online communities where reputation has been measured using forum posts (Reysen, Lloyd et al. 2010). On the other hand, disparity in resources that are not distributed in a manner deemed "fair" may result in opposite effects. To illustrate, in virtual teams composed of members from different organizational subunits, disparity could be based on resources available within a participant's unit, and the sense of deprivation ensuing among those from low resource units could lead to deleterious effects.

6. Contributions, Implications, and Future Research

6.1 Contributions to the Open Source Literature

By focusing on the impact of OSS project diversity this study extends extant literature in three distinct ways. First, although OSS leaders have recognized the influence of diversity in their projects (Ducheneaut 2005, Perens 2009), this is one of the first studies that theoretically develops and empirically tests the influence of distinct types of diversity on two different OSS project outcomes. We highlight specific types of diversity that are relevant for OSS projects (e.g., diversity in roles versus the more commonly used diversity in functional background and diversity in contribution rather than wages) and theoretically differentiate those that have positive compared to negative effects. For example, our logic and results support the idea that the variety of participant knowledge is a strength in OSS projects (as indicated in the earlier quote from Perens), but also show that other kinds of diversity, i.e. cultural separation, may hinder community engagement.

The positive finding regarding role-based variety diversity represents a contribution by providing evidence to support the notion that different kinds of OSS project participants possess dissimilar types of knowledge, and such variety in the community has a positive impact on project success. Although many

developers are also users, it has been widely recognized that participation by non-developers is important (Dahlander and O'Mahony 2010, von Krogh, et al. 2003). Ye and Kishida (2003) distinguished participants according to their role in the project and also acknowledge that projects are different from each other in terms of the number of participants in each role. Prior research stopped short of indicating how different balances of participants of various types can influence the project. Our study is among the first to theorize and provide empirical evidence for the synergistic impact of OSS users and developers, as reflected in role diversity.

Second, we developed and empirically tested how relationships between diversity and OSS market success are moderated by application development stage. Results support the theoretical logic that variety is more critical in early stage projects, highlighting the dynamic nature of the effects of diversity over a project's life. This is especially important because findings indicate that variety may be most beneficial at the stage in a project's life where it may be hardest to attract users (i.e., because the functionality is not yet fully developed). We also found that projects in later stages benefit more from disparity and separation diversity than projects in earlier stages.

Finally, there is the surprising result that cultural separation diversity has a positive effect on market success but a negative effect on community engagement. We offered one plausible explanation based on the increase in the functionality that the software needs to provide. Additional theoretical and empirical research is needed to confirm this explanation or ascertain additional reasons for the observed positive effect.

6.2 Contributions to the Diversity Literature

Diversity research has focused on the effects of differences among unit members in traditional organizations, characterized by paid employees, formal hierarchies, departments, and shared geographical space, with less work examining virtual teams (e.g., Peters and Karren 2009, Staples and Zhao 2006). We applied the theoretical concepts related to diversity developed in this body of work to a setting exemplified by a volunteer labor force, informal hierarchies, geographical dispersion, swift changes in user needs and high levels of user participation. Leveraging these concepts in this context required us to

develop theoretical links to outcomes that are not traditionally a part of diversity research but are critical for OSS projects; viz., community engagement and market success. We make two main contributions to this literature: identifying and connecting relevant diversity and outcome measures for an understudied virtual, volunteer context, and isolating a context specific moderator.

The literature on diversity has long sought to identify attributes that represent consequential differences in a particular setting, i.e. it poses the question, when is a difference salient? Our work contributes to the search for relevant diversity attributes by conceptualizing and testing the relationship between measures that are distinct from those commonly used in diversity research and chosen specifically based on OSS project characteristics. We followed suggestions to consider the specific context in selecting measures that fit well with the underlying interaction dynamics of the setting (Cannella, et al. 2008, Harrison and Klein 2007). Rather than relying on typical socio-demographic attributes such as race and gender that may not be salient in a virtual context, and are also not expected to directly impact contribution to the application's development, we focused on characteristics that are related to interacting with the other OSS project participants about the application development. We found support for the underlying theoretical arguments developed by Harrison and Klein (2007) for differential effects for alternate types of diversity. Measures similar to these may be useful to consider for future researchers that seek to understand the implications of diversity in virtual, globally distributed and volunteer contexts.

Finally, recent work has suggested moderators are likely a useful way to refine our understanding of the impact of diversity on outcomes, but a limited number of moderators have been examined in prior studies (Buyl, et al. 2010, Van Knippenberg and Schippers 2007). By integrating OSS research with diversity research we are able to suggest that stage is an important moderator of the relationship between diversity and success. The finding that the impact of diversity on market success (the number of downloads) is moderated by development stage implies it may be useful to consider other indicators of a group's work history beyond the commonly used "time." For example, research has shown that group efforts often do not flow linearly over time (Gersick 1988, Gersick 1989), thus, taking stage into account

may better explain temporal evolution in the effects of diversity.

6.3 Implications for Practice

As the distinction between OSS development and proprietary software development begins to blur (Dahlander and Wallin 2006, Henkel 2008), traditional software development managers need a deeper understanding of the dynamics underlying the OSS development process. Our findings shed light on the importance of managing the balance among different types of participants, thereby yielding implications regarding the types of individuals that managers should target for OSS project participation. For instance, while many companies have begun paying developers to work on OSS projects (Dahlander and Wallin 2006, Daniel, et al. 2011, Henkel 2008), our findings imply that managers should incentivize participation from users as well, as they play a key role in creating the vibrant setting that leads to success.

A broader implication for managers in traditional work settings relates to a redefinition of the notion of "reputation." The integration of the diversity literature into the OSS setting required the development of granular insights such as understanding the unique process through which reputation is developed in a context where volunteers are prevalent. Understanding that contributions attract attention and build reputation, and that contribution-based disparity across participants can have positive impacts is important as organizational structures become increasingly flexible, and the Internet becomes even more significant as an essential foundation for work practices. As more groups collaborate through this foundation, reputation based on socio-demographic factors or by virtue of occupying the corner office could lose significance (Dubrovsky, et al. 1991). Our results point to the need for careful consideration of how reputation is visibly represented through system design features that allow participants to make their contributions salient and visible to the project community.

6.4 Future Research

OSS projects are at the vanguard of leveraging the Internet for innovation, however there are many other examples of Internet use in innovation and collaboration processes. For example, the Internet is used extensively in the creation and evolution of Wikipedia. Future research could begin to explore

whether the findings in this study extend to contexts such as Wikipedia and academia (Kane and Fichman 2009). The influence of diversity on Wikipedia participation may be distinct from effects observed here because, for example, contributing to a Wikipedia project is likely to be more challenging for people speaking different natural languages than it is for people speaking different languages to contribute to an OSS project.

In addition to exploring the impacts of diversity in other online innovation communities, future research can attempt to identify factors that will ameliorate the negative effects of separation diversity. Often researchers suggest that the impacts of diversity can be mitigated or amplified by recruiting actions that purposively control the level of diversity, but this may be challenging in a volunteer context. Future consideration should be given to how implementing various procedures and practices could alter the negative influence of separation diversity. For instance, routines and guidelines that create project specific cultural expectations could minimize confusion related to national cultural expectations that are left unmet when interacting with participants from other cultures.

Joshi and Roh (2009) suggest that the context in which the team is embedded alters the effects of diversity on team outcomes. While we considered unique aspects of the OSS context generally, individual projects within that context operate within their own specific sub-context. For example, if a project is associated with a for profit software development company, there may be an expectation that some participants will be more active than others and so a disparity measure based on contributions may have less impact. In such projects disparity might be more strongly reflected by organizational affiliation or official role, in which case it is unclear whether disparity will be seen as "legitimate" or not. Likewise, the platform used (i.e. Sourceforge or a company sponsored platform) could affect the relationship between disparity diversity and success, for example based on the extent to which it makes diversity in activity or other characteristics visible.

We identified specific, salient indicators of disparity, separation and variety diversity in the OSS context, but as Harrison and Klein (2007) note, the three types of diversity can be reflected in many different attributes. Future research can extend this work by using alternate methodologies such as

surveys or interviews to directly capture other attributes including opinions and attitudes that reflect, for example, separation diversity, or perceptions of social power that reflect disparity diversity. By revealing distinct aspects of diversity not captured by our archival data such work may provide further insight into why we do not find the expected relationships related to separation diversity.

In addition to exploring factors that may alter the impact of diversity on market success and community engagement, it will be fruitful to explore how diversity influences other outcomes in OSS projects. Identifying the antecedents of developer turnover is especially relevant for OSS project managers and potential application adopters since developers are often volunteers. If an application ceases to be maintained because developers quit, the application is of less value to adopters. Fang and Neufeld (2009) draw upon the legitimate peripheral participation perspective to understand the factors that limit developer turnover in OSS projects, but a diversity lens could be used to complement this perspective. While they identify behaviors of participants that are associated with core developers remaining active, a diversity perspective suggests that characteristics of the developers who contribute less frequently and active users in the group could also help to explain sustained contributions.

This study begins to bridge the gap between the well-developed literature on diversity and one of the most technically advanced organizational structures. We considered unique aspects of OSS development projects and the diversity literature to suggest OSS community characteristics that lead to OSS project success. While the organizational forms that will emerge as a result of the Internet have many elements in common with traditional forms, we hope this research provides a foundation for other work that explores the subtle differences between how constructs influence outcomes in traditional and online environments.

Table 1: Rotated Component Matrix (Notes: N=357)			
	Component		
	Scope	**Separation**	**Community Engagement**
Number of Topics	**0.87**	0.02	0.07
Number of Intended Audiences	**0.86**	0.04	0.08
Blau Index of Languages Spoken	0.13	**0.83**	-0.11
Blau Index of Countries	-0.08	**0.77**	0.25
Number of artifacts (e.g. bugs) closed in year 2	-0.03	-0.06	**0.84**
Number of response posts in year 2	0.20	0.17	**0.68**
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. a. Rotation converged in 4 iterations.			

Table 2: Variable Descriptive Statistics and Intercorrelations (Notes: N=357, $^{+}$p<.1, *p<.05, **p<.01)													
Variable	**Mean**	**Std. Dev.**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1. Community Engagement	0.00	1.00											
2. Market Success	36935.81	252438.50	.01										
3. Disparity Diversity (Contribution)	2.41	1.60	.33**	.27**									
4. Separation Diversity (Culture)	0.00	1.00	.00	.11*	.23**								
5. Variety Diversity (Role)	0.23	0.19	.17**	.05	.22**	.28**							
6. Development Stage	1.04	0.87	.17**	.11*	.16**	.05	.2**						
7. No. of Participants (Year 1)	26.81	50.13	.38**	.24**	.50**	.18**	-.04	.10^{+}					
8. Gnu Public License	7.61	7.51	.09^{+}	.10^{+}	.10^{+}	.17**	.04	.20**	.04				
9. Project Type	1.45	1.50	.09^{+}	-.08	.03	.06	.05	.14**	.07	-.13*			
10. Scope	0.00	1.00	.00	-.03	.05	.00	.01	.31**	.03	.21**	.38**		
11. Number of Related Groups	8.20	7.80	.11*	.06	.23**	.08	-.03	-.12*	.28**	-.14**	.11*	-.11*	
12. Number of Releases	6.59	12.17	.22**	.19**	.17**	.05	.10^{+}	.13*	.07	.02	.02	-.05	.01

Table 3: Regression Results (Notes: N=357, $^{+}$p<.1, *p<.05, **p<.01)

	Community Engagement		Market Success		
Independent Variables	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
No. of Participants (Year 1)	0.42**	0.38**	1.06**	0.60**	0.47**
Gnu Public License	0.09^{+}	0.12*	0.03*	0.02	0.01
Project Type	0.11*	0.12*	-0.07	-0.10	-0.10
SF Registration Quarter 1	-0.07	-0.07	1.07**	0.73*	0.85**
SF Registration Quarter 2	-0.03	-0.03	1.19**	0.49	0.26
SF Registration Quarter 3	-0.05	-0.06	0.23	0.17	0.28
Scope	-0.10^{+}	-0.10^{+}	-0.12	-0.09	-0.05
No. of Related Groups	-0.07	-0.05	0.04	0.01	0.01
No. of Releases (Year 1)	0.00	0.00	0.33**	0.29**	0.27**
Development Stage	0.11*	0.09^{+}	0.33^{+}	0.34*	0.64*
Disparity Diversity (Contribution)		0.12*		0.15^{+}	-0.09
Separation Diversity (Culture)		-0.18**		0.33**	0.01
Variety Diversity (Role)		0.08^{+}		2.49**	5.78**
Disparity Diversity (Contribution) * Stage					0.22**
Separation Diversity (Culture) * Stage					0.37**
Variety Diversity (Role) * Stage					-3.16**
Likelihood Ratio Chi Square/Adjusted R squared	.19**	.22**	149.22**	201.06**	267.37**

Table 4: Results of Hypotheses Tests

Hypotheses	Results
H1: Contribution-based disparity diversity is positively associated with OSS project community engagement.	Supported
H2: Contribution-based disparity diversity is positively associated with OSS project market success.	Supported (Marginal)
H3: Culture-based separation diversity is negatively associated with OSS community engagement.	Supported
H4: Culture-based separation diversity is negatively associated with OSS market success.	Not Supported (Contradicted)
H5: Role based variety diversity is positively associated with OSS community engagement.	Supported
H6: Role based variety diversity is positively associated with OSS project market success.	Supported (Marginal)
H7a: The positive relationship between disparity diversity and OSS project market success will be stronger for projects in later development stages.	Supported
H7b: The negative relationship between separation diversity and OSS project market success will be stronger for projects in earlier development stages.	Not Supported (Contradicted)
H7c: The relationship between role based variety diversity and OSS project market success will be stronger for projects in earlier development stages.	Supported

Figure 1 Research Model

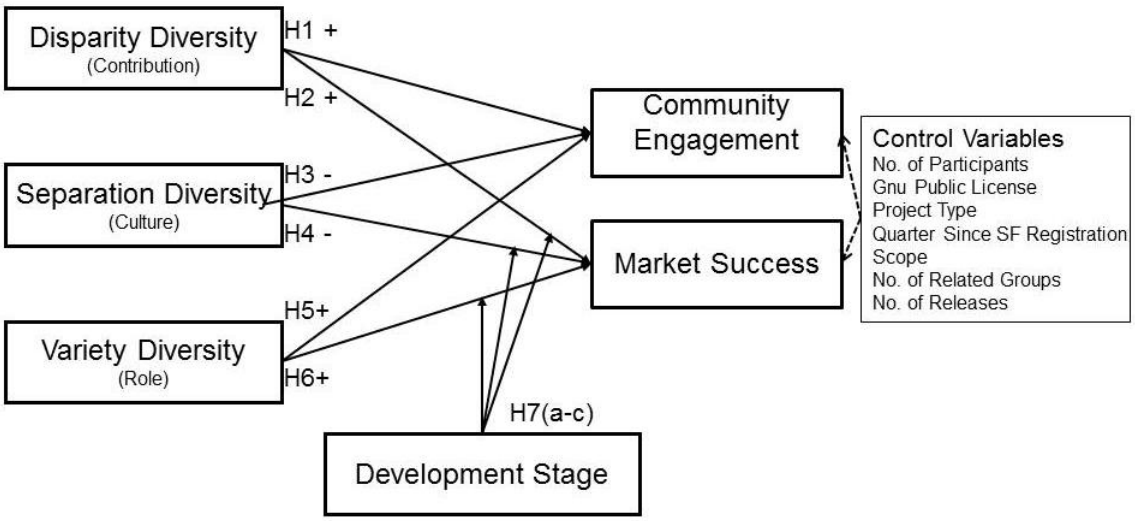


Figure 2: Disparity Diversity * Stage

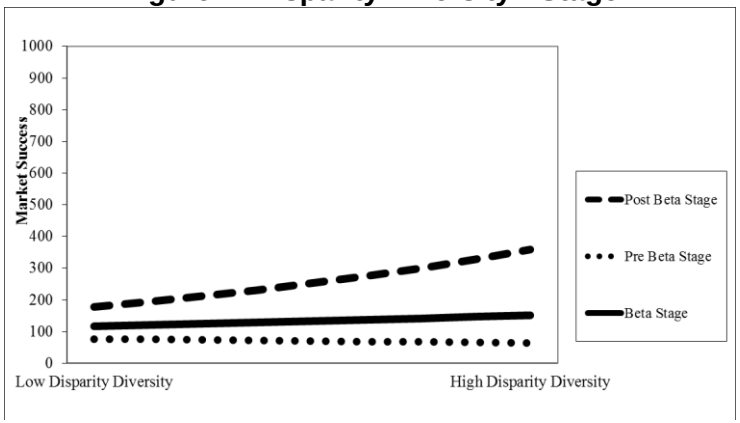


Figure 3: Separation Diversity * Stage

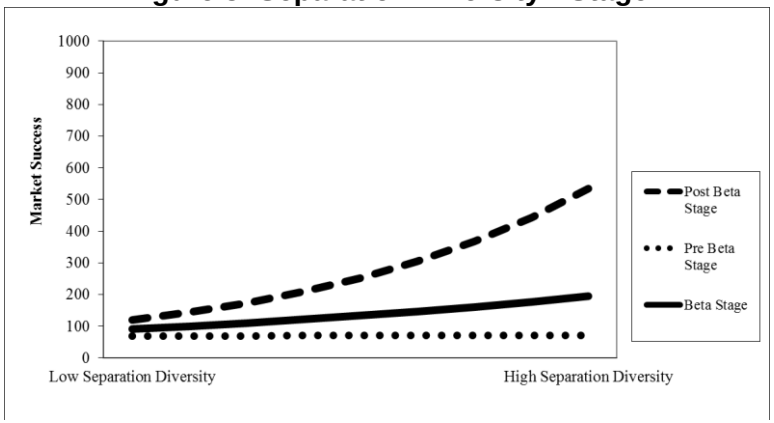
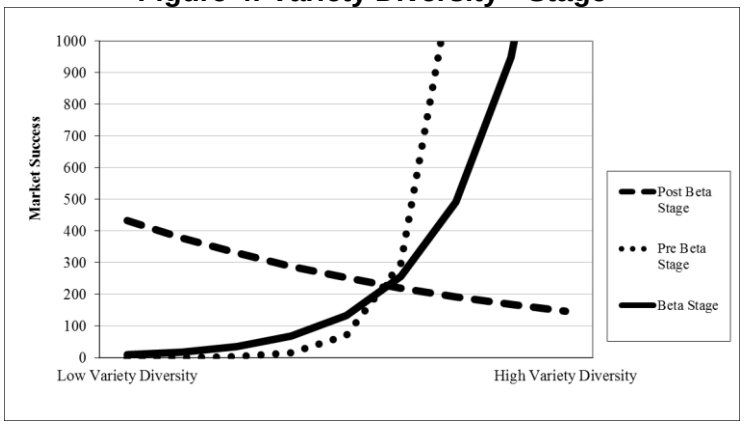


Figure 4: Variety Diversity * Stage



References

Ågerfalk, Pär J. and Brian Fitzgerald, "Outsourcing to an Unknown Workforce: Exploring Opensourcing as a Global Sourcing Strategy," *MIS Quarterly*, 32, 2, (2008), 385-409.

Alfaro, Ivan, "Nationality diversity and performance in global software development teams: The role of temporal disperion and leadership," *Proceedings of the International Conference on Information Systems*, Saint Louis, Missouri, 2010.

Allison, Paul, *Logistic regression using the SAS system: Theory and application. SAS*, SAS Institute and Wiley Cary, N.C., 1999.

Anderson, Cameron and Courtney Brown, "The functions and dysfunctions of hierarchy," *Research in Organizational Behavior*, 30, (2010), 55–89.

Anderson, William and Starr Roxanne Hiltz, "Culturally Heterogeneous vs. Culturally Homogeneous Groups in Distributed Group Support Systems: Effects on Group Process and Consensus " *Proceedings of the Proceedings of the 34th Hawaii International Conference on System Sciences*, Maui, Hawaii, 2001.

Barley, Stephen, Gordon Meyerand Debra Gash, "Cultures of Culture: Academics, Practitioners and the Pragmatics of Normative Control," *Administrative Science Quarterly* 33, 1, (1988), 24-60.

Barner-Rasmussen, Wilhelm and Ingmar Björkman, "Language Fluency, Socialization and Inter-Unit Relationships in Chinese and Finnish Subsidiaries," *Management and Organization Review*, 3, 1, (2007), 105–128.

Bechky, Beth A., "Sharing Meaning Across Occupational Communities: The Transformation of Understanding on a Production Floor," *Organization Science*, 14, 3, (2003), 312-330.

Becker, Brian and Mark Huselid, "The Incentive Effects of Tournament Compensation Systems," *Administrative Science Quarterly* 37, 2, (1992), 336-350.

Bergami, Massimo and Richard P. Bagozzi, "Self-Categorization, Affective Commitment, and Group Self-Esteem as Distinct Aspects of Social Identity in the Organization," *British Journal of Social Psychology*, 39, 4, (2000), 555-577.

Bird, Christian, Alex Gourley, Prem Devanbu, Anand Swaminathanand Grega Hsu, "Open Borders? Immigration in Open Source Projects," Minneapolis, MN (2007).

Blau, Peter, "A Macrosociological Theory of Social Structure," American Journal of Sociology 83, 1, (1977), 26-54.

Butler, Brian, "Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures," *Information Systems Research*, 12, 4, (2001), 346-362.

Buyl, Tine, Christophe Boone, Walter Hendriksand Paul Matthyssens, "Top Management Team Functional Diversity and Firm Performance: The Moderating Role of CEO Characteristics,," *Journal of Management Studies*, 48, 1, (2011), 151-177.

Cannella, Albert, Jong-Hun Parkand Ho-Uk Lee, "Top Management Team Functional Background Diversity and Firm Performance: Examining the Roles of Team Member Colocation and Environmental Uncertainty," *The Academy of Management Journal* 51, 4, (2008), 768 - 784.

Chandrashekaran, Murali, Raj Mehta, Rajesh Chandrashekaranand Rajdeep Grewal., "Market motives, distinctive capabilities, and domestic inertia: A hybrid model of innovation generation," *Journal of Marketing Research*, 36, 1, (1999), 95-112.

Chen, Stephen, Ronald Geluykensand Chong Ju Choi, "The importance of language in global teams: a linguistic perspective," *Management International Review*, 46, (2006), 679.

Chengalur-Smith, InduShoba, Anna Sidorovaand Sherae Daniel, "Sustainability of Free/Libre Open Source Projects: A Longitudinal Study," *Journal of the Association for Information Systems*, 11, (2010), 657-683.

Claes, Marie-Therese, "The linguistic shift in cross-cultural studies: culture in discourse," *European Journal of Cross-Cultural Competence and Management*, 1, 1, (2009), 66-76.

Colazo, Jorge and Yulin Fang, "Following the Sun: Temporal Dispersion and Performance in Open Source Software Project Teams," *Journal of the Association for Information Systems*, 11, (2010), 684-707.

Comino, Stefano, Fabio Manentiand M. Parisi, "From planning to mature: On the success of open source projects," *Research Policy*, 36, 10, (2007), 1575-1586.

Cooper, William H. and Michael J. Withey, "The Strong Situation Hypothesis," *Personality and Social Psychology Review*, 13, 1, (2009), 67-72.

Crowston, Kevin and Barbara Scozzi, "Open Source Software Projects as Virtual Organizations: Competency Rallying for Software Development," *IEE Proceedings Software*, 149, 1, (2002), 3-17.

Dahlander, Linus and Siobhan O'Mahony, "Progressing to the Center: Coordinating Project Work," *Organization Science*, (2010), 1-19.

Dahlander, Linus and Martin W. Wallin, "A man on the inside: Unlocking communities as complementary assets," *Research Policy*, 35, (2006), 1243–1259.

Daniel, Sherae, Likobe Maruping, Marcelo Cataldo and James Herbsleb, "When Cultures Clash: Participation in Open Source Communities and Its Implications For Organizational Commitment" *Proceedings of the International Conference on Information Systems*, Shanghai, China, 2011.

Dreu, Carsten De and Laurie Weingart, "Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis," *Journal of Applied Psychology*, 88, (2003), 741-749.

Dubrovsky, Vitaly, Sara Kieslerand Beheruz Sethna, "The Equalization Phenomenon: Status Effects in Computer-Mediated and Face-to-Face Decision-Making Groups," *Human-Computer Interaction*, 6, 2, (1991), 119 - 146.

Ducheneaut, Nicolas, "Socialization in an Open Source Software Community: A Socio-Technical Analysis " *Computer Supported Cooperative Work*, 14, 4, (2005), 323-368.

Earley, Christopher and Elaine Mosakowski, "Creating Hybrid Team Cultures: An Empirical Test of Transnational Team Functioning " *The Academy of Management*, 43, 1, (2000), 26-49.

Ehrenberg, Ronald G. and Michael L. Bognanno, "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour," *Industrial and Labor Relations Review*, 43, 3, (1990), 74-88.

Evangelopoulos, Nicholas, Anna Sidorova, Stergios Fotopoulos and Indushobha Chengalur-Smith, "Determining Process Death Based on Censored Activity Data " *Communications in Statistics - Simulation and Computation*, 37, 8, (2008), 1647 - 1662.

Fang, Yulin and Derrick Neufeld, "Understanding Sustained Participation in Open Source Software Projects," *Journal of Management Information Systems*, 25, 4, (2009), 9 - 50.

Fiol, C. Marlene, "Managing Culture as a Competitive Resource: An Identity-Based View of Sustainable Competitive Advantage," *Journal of Management*, 17, 1, (1991), 191-211.

Gallivan, M. J., "Striking a balance between trust and control in a virtual organization: a content analysis of open source software case studies," *Information Systems Journal*, 11, 4, (2001), 277-304.

Gersick, Connie J. G., "Time and Transition in Work Teams: Toward a New Model of Group Development," *The Academy of Management Journal* 31, 1, (1988), 9-41.

Gersick, Connie J. G., "Marking Time: Predictable Transitions in Task Groups," *The Academy of Management Journal*, 32, 2, (1989), 274-309.

Gibson, Cristina and Jennifer Gibbs, "Unpacking the concept of virtuality: The effects of geographic dispersion, electronic dependence, dynamic structure, and national diversity on team innovation," *Administrative Science Quarterly*, 51, 3, (2006).

Gibson, Cristina and Freek Vermeulen, "A Healthy Divide: Subgroups as a Stiumulus for Team Learning Behavior," *Administrative Science Quarterly*, 48, (2003), 202-239.

Gibson, Cristina B. and Mary E. Zellmer-Bruhn, "Metaphors and Meaning: An Intercultural Analysis of the Concept of Teamwork," Administrative Science Quarterly, 46, 2, (2001), 274-303.

Giuri, Paola, Matteo Ploner, Francesco Rullaniand Salvatore Torrisi, "Skills, division of labor and performance in collective inventions: Evidence from open source software," *International Journal of Industrial Organization*, 28, 1, (2010), 54-68.

Gould, Roger V., "The Origins of Status Hierarchies: A Formal Theory and Empirical Test," *American Journal of Sociology*, 107, 5, (2002), 1143-1178.

Grewal, Rajdeep, Gary L. Lilienand Girish Mallapragada, "Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems," *Management Science*, 52, 7, (2006), 1043-1056.

Groysberg, Boris, Jeffrey Polzerand Hillary Elfenbein, "Too Many Cooks Spoil the Broth: How High-Status Individuals Decrease Group Effectiveness," *Organization Science*, 22, 3, (2011), 722–737.

Halevy, Nir, Eileen Y. Chou, Adam D. Galinskyand J. Keith Murnighan, "When Hierarchy Wins: Evidence From the National Basketball Association," *Social Psychological and Personality Science*, (2011).

Hamilton, Barton H., Jack A. Nickersonand Hideo Owan, "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *The Journal of Political Economy*, 111, 3, (2003), 465-497.

Harrison, David A., Kenneth H. Priceand Myrtle P. Bell, "Beyond Relational Demography: Time and the Effects of Surface- and Deep-Level Diversity on Work Group Cohesion," *The Academy of Management Journal* 41, 1, (1998), 1998, 96-107.

Harrison, David A., Kenneth H. Price, Joanne H. Gavinand Anna T. Florey, "Time, Teams, and Task Performance: Changing Effects of Surface- and Deep-Level Diversity on Group Functioning," *The Academy of Management Journal* 45, 5, (2002), 1029-1045.

Harrison, David and Katherine Klein, "What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations," *Academy of Management Review*, 32, 4, (2007), 1199-1228.

He, Jun, Brian Butler and William King, "Team Cognition: Development and Evolution in Software Project Teams.," *Journal of Management Information Systems*, 24, 2, (2007), 261-292.

Henderson, Jane, "Language Diversity in International Management Teams," *International Studies of Management and Organization*, 35, 1, (2005), 66-82.

Henkel, Joachim, "Champions of revealing—the role of open source developers in commercial firms," *Industrial and Corporate Change*, 18, 3, (2008), 435–471.

Hertel, Guido, Sven Niednerand Stefanie Herrmann, "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel," *Research Policy*, 32, (2003), 1159-1177.

Hinds, Pamela and Mark Mortensen, " Understanding conflict in geographically distributed teams: The moderating effects of shared identity, shared context, and spontaneous communication," *Organization Science*, 16, 3, (2005), 290-307.

Hofstede, Geert, "The interaction between national and organizational value systems," *Journal of Management Studies*, 22, 4, (1985), 347-357.

Homan, Astrid, John Hollenbeck, Stephen Humphrey, Daan Van Knippenberg, Daniel Ilgenand Gerben Van Kleef, "Facing differences with an open mind: openness to experience, salience of intragroup differences, and performance of diverse work groups," *Academy of Management Journal*, 51, 6, (2008), 1204-1222.

Horwitz, Sujin and Irwin Horwitz, "The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography," *Journal of Management* 33, 6, (2007), 987-1015.

Jackson, Susan E., Karen E. Mayand Kristina Whitney, *Understanding the Dynamics of Diversity in Decision-Making Teams*, San Francisco: Jossey-Bass, 1995.

Jaisingh, Jeevan, Eric See-Toand Kar Tam, "The Impact of Open Source Software on the Strategic Choices of Firms Developing Proprietary Software," *Journal of Management Information Systems* 25, 3, (2008), 241-276.

Joshi, Aparna and Hyuntak Roh, "The Role of Context in Work Team Diversity Research: A Meta-Analytic Review," *The Academy of Management Journal* 52, 3, (2009), 599 - 627.

Kane, Gerald and Robert Fichman, "The Shoemaker's Children: Using Wikis for Information Systems Teaching, Research, and Publication," *MIS Quarterly*, 33, 1, (2009), 1-17.

Kankanhalli, Atreyi, Bernard Tanand Kwok-Kee Wei, "Conflict and Performance in Global Virtual Teams.," *Journal of Management Information Systems*, 23, 3, (2006), 237-274.

Knippenberg, Daan Van and W. van Ginkel, "The categorization-elaboration model of work group diversity: Wielding the double-edged sword. The psychology of social and cultural diversity.," Malden, MA, (2010), 257-280.

Lakhani, Karim and Eric von Hippel, "How open source software works: "free" user-to-user assistance," *Research Policy*, 32, (2003), 923-943.

Lawler, Edward and J. Lloyd Suttle, "Expectancy theory and job behavior " Organizational Behavior and Human Performance, 9, 3, (1973), 482–503.

Lawler, Edward E., Lyman W. Porterand Allen Tennenbaum, "Managers' attitudes toward interaction episodes," *Journal of Applied Psychology*, 52, 6, (1968), 432-439.

Lazear, Edward and Sherwin Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," Journal of Political Economy, 89, 5, (1981), 841-864.

Lee, Daniel, Probir Banerjee, Kai H. Lim, Kuldeep Kumar, Jos van Hillegersbergand Kwok Kee Wei, "Aligning IT components to achieve agility in globally distributed system development," *Communications of the ACM* 49, 10, (2006), 48 - 54.

Lerner, J. and J. Tirole, "Some simple economics of open source," *Journal of Industrial Economics*, 50, 2, (2002), 197-234.

Levina, Natalia and Emmanuelle Vaast, "Innovating or doing as told? Status differences and overlapping boundaries in offshoring collaboration.," *MIS Quarterly*, 32, 2, (2008), 307-332.

Locke, Kennet and Leonard Horowitz, "Satisfactin in Interpersonal Interactions as a Function of Similarity in Level of Dysphoria," *Journal of Personality and Social Psychology*, 58, 5, (1990), 823-831.

Lowry, Paul Benjamin, Dongsong Zhang, Lina Zhouand Xiaolan Fu, "Effects of culture, social presence, and group composition on trust in technology-supported decision-making groups.," *Information Systems Journal*, 20, 3, (2010), 297-315.

Majchrzak, Ann and Arvind Malhotra, "Virtual workspace technology use and knowledge sharingeffectiveness in distributed teams: the influence of a team's transactive memory," *Knowledge Management Knowledge Base*, (2004).

McCain, Bruce E., Charles A. O'Riellyand Jeffrey Pfeffer, "The Effects of Departmental Demography on Turnover: The Case of a University," *Academy of Management Review*, 26, (1983), 626-641.

McPherson, Miller, Lynn Smith-Lovinand James M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 27, (2001), 415-444.

Milliken, Frances and Luis Martins, "Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups," *The Academy of Management Review*, 21, 2, (1996), 402-433.

Overbeck, Jennifer R., Joshua Correlland Bernadette Park, "Internal Status Sorting in Groups: The Problem of too Many Stars," St. Louis, (2005).

Perens, Bruce, "Is Open Source Capitalist or Communist?," Datamation, (2009).

Peters, Linda and Ronald J. Karren, "An Examination of the Roles of Trust and Functional Diversity on Virtual Team Performance Ratings," *Group Organization Management*, 34, (2009), 479-504.

Pfeffer, Jeffrey and Alison Davis-Blake, "Salary Dispersion: Location in the salary distribution, and tur-over among college administrators," *Industrial and Labort Relations Review*, 45, (1992), 753-763.

Raymond, Eric, "Homesteading the Noosphere," (1998).

Reysen, Stephen, Jason Lloyd, Iva Katzarska-Miller, Brett Lemkerand Russell Foss, "Intragroup status and social presence in online fan groups," *Computers in Human Behavior*, 26, 6, (2010), 1314–1317.

Riordan, Christine M. and Lynn McFarlane Shore, "Demographic diversity and employee attitudes: An empirical examination of relational demography within work units," *Journal of Applied Psychology*, 82, 3, (1997), 342-358.

Roberts, Jeffrrey, Il-Horn Hannand Sandra Slaughter, "Understanding the Motivations, Participation, and Performance of Open Source Software Developmers: A Longitudinal Study of the Apache Projects," *Management Science*, 52, 7, (2006), 984-1000.

Robles, Gregorio and Jesus Gonzalez-Barahona, "Contributor Turnover in Libre Software Projects " In *IFIP International Federation for Information Processing*, E. Damiani, Fitzgerald, B., Scacchi, W., Scotto, M., Succi, G. (Ed.), 203, Springer, Boston, 2006a 273-286.

Rogers, E.M., *Diffusion of Innovations*, The Free Press, New York, 1995.

Rosen, Sherwin, "The theory of equalizing differences," In Handbook of Labor Economics, 1986, 641–692.

Sarker, Suprateek, Francis Lauand Sundeep Sahay, "Using an adapted grounded theory approach for inductive theory building about virtual team development," *Database for Advances in Information Systems*, 13, 1, (2001), 38-56.

Scacchi, W., "Understanding the requirements for developing open source software systems," *IEE Proceedings on Software*, 149, 1, (2002), 24-39.

Scacchi, Walt, Joseph Feller, Brian Fitzgerald, Scott Hissamand Karim Lakhani, "Understanding Free/Open Source Software Development Processes," *Software Process Improvement and Practice*, 11, (2006), 95-105.

Seidel, Marc-David and Katherine J. Stewart, "An Initial Description of the C-form," *Research in the Sociology of Organizations*, 33, (2011), 37-72.

Setia, Pankaj, Balaji Rajagopalan, Vallabh Sambamurthyand Roger Calantone, "How Peripheral Developers Contribute to Open-Source Software Development," *Information Systems Research*, (2010).

Shachaf, Pnina, "Cultural diversity and information and communication technology impacts on global virtual teams: An exploratory study," *Information & Management*, 45, 2, (2007), 131-142.

Shah, Sonali, "Motivation, Governance and the Viability of Hybrid Forms in Open Source Software Development," *Management Science*, (2006), 1000–1014.

Silver, Steven, Bernard Cohenand James Crutchfield, "Status Differentiation and Information Exchange in Face-to-face and Computer Mediated Idea Genearation," *Social Psychology Quarterly* 57, 2, (1994), 108-123.

Singh, Param, "The Small World Effect: The Influence of Macro Level Properties of Developer Collaboration Networks on Open Source Project Success," *ACM Transactions of Software Engineering and Methodology*, 20, 2, (2010).

Singh, Param, Young Tanand Vijay Mookerjee, "Network Effects: The Influence of Structural Social Capital on Open Source Project Success," *MIS Quarterly*, 35, 4, (2011).

Staples, D. Sandy and Lina Zhao, "The Effects of Cultural Diversity in Virtual Teams Versus Face-to-Face Teams," *Group Decision & Negotiation*, 15, 4, (2006), 389-406.

Star, Susan and James Griesemer, "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39," *Social Studies of Science* 19, 3, (1989), 387-420.

Steiner, Ivan D., *Group Process and Productivity*, Academic Press Incorporated, 1972.

Stewart, Daniel, "Social Status in an Open-Source Community," *American Sociological Review*, 70, 5, (2005), 823-842.

Stewart, Katherine and Sanjay Gosain, "The Impact of Ideology on Effectiveness in Open Source Software Development Teams," *Management Information Systems Quarterly*, 30, 2, (2006a), 291-314.

Stewart, Katherine and Sanjay Gosain, "The Moderating Role of Development Stage in Free/Open Source Software Project Performance," *Software Process Improvement and Practice*, 11, (2006b), 177-191.

Stewart, Katherine J., Anthony P. Ammeterand Likoebe M. Maruping, "Impacts of License Choice and Organizational Sponsorship on User Interest and Development Activity in Open Source Software Projects," *Information Systems Research*, 17, 2, (2006), 126-145.

Subramaniam, Chandrasekar, Ravi Senand Matthew L. Nelson, "Determinants of open source software project success: A longitudinal study," *Decision Support Systems*, 46, 2, (2009), 576-585.

Tornatzky, L. and M. Fleischer, *The Processes of Technological Innovation*, Lexington Books, Lexington, 1990.

Van Antwerp, Matthew and Greg Madey, "Advances in SourceForge Research Data Archive (SRDA)," *Proceedings of the Fourth International Conference on Open Source Systems*, Milan, Italy, 2008.

Van Knippenberg, D. and M. Schippers, "Work group diversity," *Annual Review of Psychology*, 58, 1, (2007), 515-541.

Vignovic, Jane A and Lori Foster Thompson, "Computer-mediated cross-cultural collaboration: attributing communication errors to the person versus the situation," *Journal of Applied Psychology*, 95, 2, (2010), 265-276.

Von Glinow, Mary Ann , Debra L. Shapiroand Jeanne M. Brett, "Can We Talk, and Should We? Managing Emotional Conflict in Multicultural Teams," *Academy of Management Review* 29, 4, (2004), 578-592.

von Krogh, Georg , Sebastian Spaethand Karim Lakhani, "Community, joining, and specialization in open source software innovation: a case study," *Research Policy*, 32, (2003), 1217-1241.

Welch, Denice E. and Lawrence S. Welch, "The importance of language in international knowledge transfer " *Management International Review* 48, 3, (2008), 339-360.

Whang, Seungjin, "Market provision of custom software: Learning effects and low balling," *Management Science*, 41, (1995), 1343-1352.

Williams, K. and C. O'Reilly, *Demography and diversity in organizations: A review of 40 years of research*, JAI Press, Greenwich, CT, 1998.

Ye, Yunwen and Kouichi Kishida, "Toward an Understanding of the Motivations of Open Source Software Developers," *Proceedings of the Twenty-Fifth International Conference On Software Engineering*, Portland, Oregon, 2003.