The Effects and Antecedents of Conflict in Free and Open Source Software Development

Anna Filippova
Communications and New Media
National University of Singapore
Singapore
annaf@u.nus.edu

Hichang Cho
Communications and New Media
National University of Singapore
Singapore
cnmch@nus.edu.sg

ABSTRACT
Conflict is an important group process, and more so in self-organizing teams with fluid boundaries and high possibility for turnover. We empirically investigate different types of conflict in Free and Open Source Software development teams, their antecedents and impact on developers' sustained participation. Following a survey of 222 FOSS developers, we find conflict to have an overall negative effect on developer retention. Furthermore, different types of conflict have varying impact on outcomes. In particular, only normative conflict levels negatively impact intention to remain in a FOSS project. Both normative and process conflict negatively affect perceptions of team performance, while the co-occurrence of task and process conflict exacerbate negative effects on outcomes. Though we find structural factors like task interdependence and geographical distribution increase overall conflict levels in FOSS teams, participatory decision-making and a transformational leadership style have an ameliorating effect. Implications for theory and practice are discussed.

Author Keywords
Virtual Teams; Free and Open Source Software; Conflict; Peer-Production; Identification; Participation; Performance.

ACM Classification Keywords
K.4.3 [Computers and Society]: Organizational Impacts---Computer-Supported Collaborative Work.

INTRODUCTION
Many sufficiently popular virtual communities have experienced a dramatic and public departure during the course of their group's collective history. The Free and Open Source Software space has seen a number of examples over the years, from the departure of Open Office core team members, to the more recent gendered commit debate that led to the resignation of a Node.JS core contributor [19]. Voices within the community have linked developer departures with a toxic environment created by high levels of conflict [45]. This is because Free and Open Source Software (FOSS) development teams are largely voluntary self-organizing systems with porous boundaries that allow relatively easy entry and exit [12]. Given the tenuous nature of their membership and the fact that developer time is primary currency in FOSS, it is important to investigate factors that affect developer perceptions of the team and their retention.

CSCW '16, February 27-March 02, 2016, San Francisco, CA, USA
© 2016 ACM. ISBN 978-1-4503-3592-8/16/02…$15.00
DOI: http://dx.doi.org/10.1145/2818048.2820018

A healthy body of work has already examined various aspects of participation in virtual communities. We know that the continued involvement of FOSS developers depends on both their individual motivations and team structural factors, such as leadership, ideology and interpersonal relationships among contributors [51]. Motivation, in turn, can be intrinsic, such as the desire to help others, as well as extrinsic, involving rewards like reputation and career advancement [39]. Additionally, developers are more motivated if they identify strongly with the team and subscribe to the project ideology [53].

There are also a growing number of studies on conflict in peer production settings. Prior work has found that in the context of Wikis, there are high levels of conflict in discussions [10], conflict increases with community complexity [30] and negatively impacts the quality of articles produced [3]. In FOSS projects, qualitative studies have shown that conflict is also an important process [17] and manifests in different ways [19]. However, research has yet to examine the effect of conflict on outcomes in FOSS teams [12].

The present research takes the next step in understanding conflict prevalence in FOSS teams by examining the connection between conflict and team outcomes. Using a large survey of the Free and Open Source Software developer population (n=222), this study makes several key contributions to our understanding of peer production dynamics. First, we quantitatively examine the overall prevalence of different types of conflict in FOSS teams to understand the most commonly occurring disagreements among developers. Second, we examine different FOSS team structures to understand where conflict is more prevalent. And finally, we investigate what kinds of conflict affect developers' attitudes towards their team and their sustained participation.

RELATED WORK

Figure 1 presents the theoretical model we employ in this study. Our model follows the classic input-process-output structure and includes both conflict outcomes and antecedents [12]. First we describe our conceptualization of outputs and the four conflict processes proposed in earlier work, and link them together. Then, we examine possible inputs that may affect levels of conflict in FOSS teams.

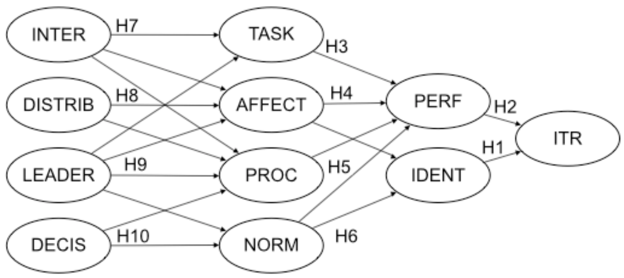


Figure 1 Input-Process-Output Model of Conflict in FOSS Teams

Outcomes

The present study examines sustained developer participation and attitudes towards their project in connection with conflict. We thus consider three outcome variables: developers' intention to remain in the project (ITR), their perceptions of project performance (PERF), and their identification with the team (IDENT).

Continued developer involvement, or the intention to remain in the project, is a significant contributor towards FOSS vitality [11]. While the motivation to join a project depends on a combination of intrinsic and extrinsic factors [39], Fang and Neufeld [18] found that sustained participation rests on the learned experience with the project as well as team identification. Similarly, Bagozzi and Dholakia [4] have found that intention to participate in Linux User Groups depends on the attitudes of developers towards the group, as well as their identification with the group. Taken together, the body of work suggests sustained participation in peer production rests on developers' attitudes towards the project, such as their evaluation of how well the project is doing, as well as the strength of their identification with the project. We thus examine developer perceptions of team performance and their identification with the team in connection with developers' intention to remain in the project.

Team success is most commonly studied within the classical IS framework by DeLone and McLean [14], which measures the quality of software produced and its market success. Crowston and colleagues [11] drew on developer feedback to apply these success measures to the FOSS context. They found success depended on developer involvement, success with users and perceptions of software and process quality. Similarly, Lee and colleagues [31] have found software quality and user satisfaction are significantly related to FOSS use. We build on this conceptualization to investigate developers' perceptions of their project performance as a function of their evaluation of the software quality and success with users [36].

Identification with the team is another important variable to examine in connection with sustained FOSS participation. According to Turner [47], when individuals categorize themselves as part of a group, they perceive other group members to be more similar, and adopt group goals as their own. Identification leads to an emotional attachment with the group, and the desire to maximize group outcomes – that is, the group's collective success becomes reflective of one's own success. Zhu and colleagues [53] showed how identification with a Wiki group not only encourages participation, but also focuses efforts on accomplishing group goals. Similarly, Stewart and Gosain [46] found that in FOSS projects, adherence to team ideology was beneficial to the team in attracting and retaining developers.

Thus we anticipate the following relationship between our outcome variables:

Developers' perception of project performance (PERF) (H1) and identification (IDENT) (H2) will be positively related with intention to remain in the project (ITR).

Conflict

Conflict within groups has been studied across many disciplines, and conceptualized in different ways. Drawing from literature on organizational teams, we define conflict as an "interactive process manifested in incompatibility, disagreement, or dissonance within or between social entities" [43]. Crucially, conflict involves opposing interests that are manifest in some way within the group. That is, for conflict to occur, incompatibilities must not only be present, but also visible and perceived as conflict by one or more group members. Consequently, in this study we investigate developer perceptions of manifested conflict within their team. Following prior work by Arazy and Nov, Van Wendel de Joode, and Filippova and Cho [3,19,49], we draw on literature on traditional virtual teams and distinguish between several different forms of conflict.

Task Conflict

Past research highlights a cognitive dimension to conflict that involves differing opinions about the work that needs to be done [26]. Task conflict in peer production settings involves differences of opinion about what content to add and proposed changes to the project output [3]. In the case of FOSS projects, this means disagreements about feature development, and changes to the code base, such as removing or refactoring code [19]. As many FOSS projects are modular and depend on other projects, task conflict may also arise from differing opinions on the choice of dependencies or between factions of the community that support different software or versions [19].

On one hand, cognitive (or task) conflict is expected to have a positive influence on team performance because it stimulates creativity through diversity of ideas and prevents groupthink [26]. On the other hand, team members may interpret disagreements concerning the work they are doing as a negative assessment of their own abilities, causing stress and dissatisfaction [50]. In practice, task conflict has had mixed effects on group outcomes in traditional virtual teams. A review by De Dreu and Weingart [15] found task conflict to be negatively related to both performance and satisfaction, however a later meta-analysis by De Wit et al. [50] showed no direct effects on either. Similarly, in the peer production context, Arazy and Nov [3] did not find significant effects of task conflict on the quality of Wikipedia articles.

One possible reason for these mixed findings is that task conflict has varying effects on performance depending on the complexity of the task being performed [28]. Specifically, task conflict is beneficial for teams performing more complex tasks as it stimulates creativity and new ideas. However, for teams working on more routine tasks task conflict may be detrimental because it takes focus away from work. As FOSS development is a creative task that involves writing complex software, task conflict may have a more positive impact on team performance. Thus we propose:

H3: Task conflict (TASK) will be positively related with performance perception

Affective Conflict

Affective (or relational) conflict involves emotional disagreements and interpersonal dissonance, thus it is often presented in contrast to task conflict [26]. These displays of emotion and personal incompatibilities are expected to have an overall negative effect on the team because they take focus and resources away from accomplishing the group goals [26,37]. In virtual communities like FOSS projects, affective conflict may manifest as disinhibited interactions between members that involve personal attacks, such as flaming [3,20]. Furthermore, in the absence of situational or contextual cues, FOSS contributors may attribute failure to personal incompetence or inattention, rather than situational factors such as lack of information [19]. This may result in a more pronounced effect of affective conflict.

The direct impact of affective conflict has not been empirically investigated in research on peer production. However, work by Arazy and Nov[3] on the interaction between task and affective conflict suggests affective conflict, when present, has a negative impact on quality of Wikipedia articles. Furthermore, in work on virtual teams in traditional organizations, affective conflict also has a largely negative impact on team outcomes like member satisfaction and performance [50]. Thus we expect affective conflict to be negatively related with perceived performance as it detracts from the task at hand. We also expect affective conflict to negatively impact identification with the team, as high affective conflict levels may make developer efforts seem unappreciated. We propose that:

H4: Affective conflict (AFFECT) will be negatively related with perception of project performance and identification with the team

Process Conflict

Process conflict is a third conceptual conflict dimension that involves disagreements over *how* to perform a task and team responsibilities. Process conflict has appeared less frequently in research due to measurement difficulties in distinguishing it from the first two conflict types [50]. However, Behfar et al. [5] have recently re-examined this dimension and highlighted its distinctiveness in predicting team outcomes. Specifically, disagreements about procedure introduce confusion about the correct course of action, and take attention and resources away from the task at hand. Therefore, in the context of traditional virtual teams, process conflict has had largely negative effects on team performance [50].

FOSS teams evolve and infer procedures based on visible indirect cues and the behavior of others [6]. This lack of explicit coordination may lead to instances of process conflict through misunderstandings about issues such as access to the code base, distribution of work load, as well as how to structure contributions in terms of an expected coding style [19].

The direct effect of process conflict on outcomes has not been examined in peer production settings as well. However, Arazy and Nov [3] have found that in Wikipedia, the presence of process conflict alongside task conflict has a negative effect on article quality. Taken together with earlier work on traditional virtual teams, we can expect high levels of process conflict to negatively impact perception of project performance:

H5: Process conflict (PROC) will be negatively related with perception of project performance

Normative Conflict

Lastly, a fourth conflict dimension has been proposed for more long-term oriented, self-organizing teams that have had time to develop group norms, such as FOSS projects [19,41]. Normative conflict, or conflict about group norms, involves higher order disagreements about group function that do not directly arise out of working together. Instead, conflict emerges from a perceived dissonance between the prescriptive norms of the group, that is, group values and expected behavior, and the descriptive norms, or members' actual behavior [41]. Thus conflict may involve ideological debates and recursive, meta-level, discussions on the state of the community as a whole. Though outlined in qualitative work [19], this form of conflict has not yet been empirically examined in the context of peer production.

Because normative conflict highlights inconsistencies in the goals and direction of the project, it may lead to perceptions that group cohesion is overall lower, and thus reduce identification with the project. Additionally, similar to affective and procedural conflict, normative conflict may take attention away from the activity of writing software while the community debates larger issues, thus negatively impacting perceptions of performance. We propose that:

H6: Normative conflict (NORM) will be negatively related with performance perception and identification with the team

Interactions

Finally, a number of recent studies suggest that these conflict dimensions do not occur in isolation, and influence the effects of other conflict types on outcomes [3,19,50]. Specifically, Arazy and Nov[3] found that the presence of either affective or process conflict alongside task conflict led to reduced Wikipedia article quality. Similarly, in traditional virtual teams, De Wit et al. [50] have found that in the presence of affective conflict, task conflict and performance are less positively associated. Finally, qualitative work by Filippova and Cho [19] suggests that affective and process conflict may interact with task conflict, as well as with each other. As there appear to be a number of potential interactions, and the interaction of normative conflict with other conflict types has not yet been examined, we propose a more general research question:

RQ: How do interactions between task, affective, process and normative conflict influence perceptions of team performance and identification with the team?

Conflict Antecedents

The rich history of inquiry into virtual team conflict identified a number of relevant antecedent factors both at an individual and group level [50]. As in this study we are interested in manifested intragroup conflict, we thus consider group level antecedents in our model. Specifically, we examine four relevant conflict antecedents proposed by earlier work on FOSS conflict [19]: task interdependence, geographical distribution, leadership style and the relative distribution of decision-making, alongside several control variables.

Task Interdependence

FOSS teams vary in their levels of interdependence. As Bolici, Howison and Crowston [6] show, some teams, especially larger projects, tend towards a more modular structure. Instead of explicit coordination, teams use implicit "stigmergic" strategies instead [6]. Traditional virtual team research shows that modular teams experience lower levels of conflict than groups in which members have to rely more on each other to accomplish tasks [52]. In the context of Wikipedia, Kittur and colleagues [29] have also found that contributor density, or the degree to which editors are forced to interact on an article, increases conflict levels. Similarly, qualitative work by Van Wendel de Joode [49] suggests that modular FOSS projects may experience less overall conflict.

This is because task interdependence creates more opportunities for opinions to diverge, which may lead to more task conflict. At the same time, greater interdependence also requires more complex rules, access and authority structures and therefore greater opportunity for process conflict. Finally, interdependent teams may experience the effects of affective conflict more severely than more modular team structures where individuals can more easily avoid confrontation. Thus we propose that:

H7: Interdependence (INTER) will be positively related with the occurrence of task, affective, and process conflict.

Geographical Distribution

Distance can be a mixed bag for virtual teamwork. On one hand, more distributed teams have less shared context, and therefore experience greater overhead when coordinating [24]. This can lead to more misunderstandings about procedure and thus greater process conflict in both traditional [37] and FOSS teams [19]. Furthermore, meeting in person is challenging for highly geographically distributed teams, and continual reliance on computer-mediated communication may result in depersonalization of other team members and more disinhibited communication, leading to greater affective conflict [37]. On the other hand, distance may also make incompatibilities easier to avoid when developers do not physically interact as much with each other, especially when tasks can be performed independently [6]. Thus we propose that geographical distribution will have only a selective effect on conflict occurrence:

H8: Geographically distributed (DISTRIB) teams will experience more procedural and relational conflict

Leadership Style

Prior work distinguishes between two leadership styles – transactive and transformative [27,32,42]. Transactive leaders rely on rewards and punishments to manage and motivate their team, while transformational leaders are charismatic, display strong commitment to ideals, inspire and lead by example [27]. Purvanova and Bono [42] have found that transformational leaders have a greater impact on performance when there is more uncertainty, such as in virtual rather than face-to-face settings. Giuri and colleagues [21] have also found that leadership is an important part of FOSS project success, while Li et al. [32] have shown that transformational leadership is an important factor in motivating FOSS contributions.

Leadership has an important effect on conflict in distributed work settings. Wakefield and colleagues [48] have shown that effective leaders mitigate task conflict by acting as monitor, and process conflict through coordination activities. Additionally, transformational leaders may mitigate affective conflict levels by acting as mediators [49] and normative conflict levels through charisma and inspiring developers to align towards group goals. Hence, we propose that:

H9: Transformational leadership (LEADER) will be negatively associated with all types of conflict

Distribution of Decision-making

FOSS teams vary greatly in their decision-making structure, from benevolent dictatorships like the Linux kernel, to flat, one vote per developer, decision-making styles like Debian. O'Mahony and colleagues [38] have shown governance to be an emergent property of FOSS teams. That is, governance evolves over time and is in a constant state of flux. Campion et al. [7] have shown that a participatory decision making style is an important predictor of organizational team success. Though this factor has not been directly examined alongside conflict, we expect participatory decision making to reduce process conflict levels as team members gain more autonomy and contribute more to team organization. Similarly, we expect less need for normative conflict in more participatory teams because members would have more formal opportunities to vote on and impact the growth and development of the team. Thus we propose:

H10: Participatory decision-making (DECIS) will be negatively related with process and normative conflict

Control variables

Though we focus on group-level antecedents and the manifestation of conflict at the group level in our model, there may be other individual- and group-level factors that influence outcomes. Specifically, we also control for *gender* in this study, because Collier and Bear [10] have suggested that highly contentious peer production environments may lead to lower participation intentions for female contributors. As FOSS teams are ongoing teams, individuals who have participated longer may be more likely to identify strongly with the team and have greater intention to continue contributing to the project [44]. Similarly, developers with more central roles in the project may perceive themselves as more prototypical members of the in-group, and have stronger identification and intention to remain in the project [47]. Thus we also examine *length of contribution* and *role in the project* as control variables. Finally, as larger projects tend to be more complex, they may experience greater negative effects of conflict [1], thus we also control for *project size*. In the next section, we describe our operationalization of the above propositions.

METHODOLOGY

Measures

We adapted most of our measures from existing pre-validated scales. Task, affective, process conflict and intent to remain were adapted from Jehn and colleagues' work [25,26]; identification from Greene [23]; interdependence from Sharma and Yetton [52]; geographical distribution from Chudoba et al. [9]; the short transformational leadership scale from Carless et al. [8]; and participatory decision-making from Campion and colleagues [7]. All items were measured on a 5-point Likert Scale (from Strongly Disagree to Strongly Agree). Our control items were single-item measures, and their operationalization is reported under demographic statistics.

We also measured performance as a feature of developers' satisfaction with their project output, that is, the extent to which the project performance matched developers' expectations within the past year. We used a scale adapted from McDonough and colleagues [36] using insights from Lee et al. and Crowston et al. [11,31] that suggest including both software quality and user success dimensions.

While several measures exist for normative conflict [34,40] they are designed to measure internal and individual level dissonance, rather than the outward group-level conflict manifestations this study is interested in. As no measures exist for normative conflict in this context, we developed our own. In doing so, we reexamined and reframed existing conflict dimensions in the context of an additional variable. Thus, before implementing our final survey, we first conducted two pilot tests.

Pilot Tests

We drew on the qualitative work of Filippova and Cho [19] to design an initial set of normative conflict items. We ran our first pilot study with 7 FOSS developers, inviting qualitative feedback on the extent to which our questions reflected the experience of working in a FOSS team. Next, we ran a larger pilot study to identify measurement issues. We distributed questionnaire links on social media and during the annual 2015 FOSSAsia Conference. We received 58 responses in total, with 31 valid for further analysis after attention and quality checks.

Overall item reliability for all the variables of interest was between 0.74 and 0.92, with all conflict variables greater than 0.8 and well above the traditional 0.7 cut-off. An exploratory factor analysis confirmed a 4-factor structure that explained 78% of the total variance. However, two of the items we developed did not load cleanly on only one factor, thus we removed them. Due to space constraints, we only report the final items used: "NORM1: Contributors to this project disagree about project mission", "NORM2: There are disagreements concerning the values this project upholds", and "NORM3: Contributors disagree about what is expected behavior in the project".

Survey

Sample

For our final survey, we randomly sampled developers on GitHub who are involved in public projects containing 2 or more developers. Although GitHub hosted projects are not necessarily Free and Open Source Software (FOSS), GitHub encourages the use of a FOSS license when creating a new project, especially a public one. Thus, in absence of one central repository for all FOSS projects, GitHub offers a close approximation. In addition, invitations and survey instructions highlighted the target audience to be FOSS contributors.

We used the April 2015 version of the GHTorrent [22] dataset to obtain our list of participants. First we determined project sizes by counting the number of users who had commit access to the repository for each project in the dataset and made a list of projects with 2 or more contributors. Next, we made a list of active contributors to the projects by selecting all project members who had a commit merged in the past year (April 2014 or later). Finally, to ensure diversity of projects, we used stratified random sampling to obtain an equal number of unique active contributors from projects of various sizes: 2-10, 21-30, 31-40, 41-50, 51-100, and more than 100. We chose to oversample members from larger projects because there are considerably more small projects on GitHub than very large ones, thus we used combined the 51-100 and over 100 categories.

Procedure

Sampled contributors were invited to participate by e-mail using the SoSci.de survey tool at the end of April 2015. The invitations contained information about the purpose of the study, affiliation and contact information of the principal investigator, together with a personalized survey link. One reminder e-mail was sent a week later to respondents who have not yet clicked on their link. In exchange for participation, at the end of the survey respondents could nominate a FOSS project to receive a $100 donation and were informed that 3 projects will be chosen at the end of the study. All pilot participants received this option as well. Participants were also given the option to stay informed of the study results. As participants may be members of several projects at the same time, they were instructed to answer the questionnaire from the perspective of the project they contribute most actively to. This instruction was repeated across all questionnaire pages.

Demographics

We collected a total of 520 responses over 3 weeks. Of these, 222 were complete, passed our attention checks and were used for all further analysis. We received responses from 59 countries, with USA the largest group (32.8%) followed by Germany (7.8%), and the UK (6%). Consistent with earlier work, our sample was largely male – only 6.6% identified as "Female", and 3.3% identified their gender in another way, including "Gender Queer", "Feminine Male" and "Jedi". Our sample was balanced in terms of developer tenure and roles within the project: 28.3% have contributed to their project for over 4 years, 30.7% for less than 1 year and the remaining between 1 and 4 years. 31.6% were project leaders or founders, 34.3% core contributors or maintainers and 21.1% were members with commit access. 45.5% of projects were small (between 2 and 10 developers), 16.6% were very large (more than 50 contributors) with the remaining in between.

Model Specification

To verify the validity of our measurement model, we first conducted a confirmatory factor analysis (CFA). The lavaan package in R (version 0.5-15) and full information maximum likelihood (FIML) estimation were used to run the CFA and subsequent structural equation modeling (SEM) analyses on a covariance matrix (n=222). We used FIML over listwise deletion because it performs better across all forms of missing data [35]. In our initial analysis, some indicators had low factor loadings thus we removed all indicators with loadings less than .55. Standardized factor loadings in the final model ranged from .619 to .945.

The final measurement model had a good fit as Table 1 demonstrates. Though the Chi-square p value is typically expected to be above >.05, this test is sensitive to sample size and number of variables [35]. Thus we consider the remaining size independent fit indices instead, which are all within acceptable ranges (CFI and TLI > .90, RMSEA < .05, SRMR <.08) [35]. We also compared this measurement model against a model with all conflict indicators loading on one factor and found that the four-factor solution had a significantly better model fit (Table 2).

Finally, we tested the convergent and discriminant validity of all factors in the measurement model by examining the Average Variance Extracted (AVE) and Square Maximum Correlations (SMC). Table 2 presents the AVE scores for our model variables, with all values above .50 showing that our indicators accounted for more variance in the latent factors than standard error, confirming their convergent validity. Additionally, Table 2 shows that AVE was larger than the squared maximum correlations (SMC) between the variables, confirming discriminant validity of our factors. Table 2 also reports the composite reliability (CR) for each latent factor.

After confirming the validity of our measurement model, we ran a structural equation model (Model 1) to investigate the hypothesized relationships indicated in Figure 1. As Table 1 shows, fit indices were slightly less than acceptable in this original model. As it is possible that our model antecedents may also have direct effects on outcomes, to improve model fit we then added paths between our

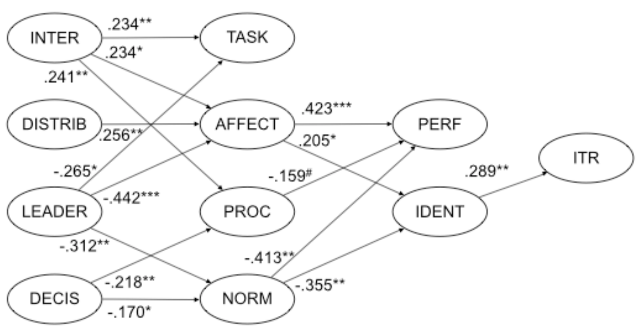
Model (N=222)	**Fit Indices**				
	χ^2/df	CFI	TLI	RMSEA	SRMR
4-Factor	1.45	0.932	0.921	0.045	0.061
1-Factor	2.35	0.788	0.765	0.078	0.076
Model 1	1.70	0.893	0.879	0.056	0.085
Model 2	1.50	0.924	0.914	0.047	0.062

Table 1 Fit indices for measurement model with four conflict factors, a one factor solution and structural model iterations

Factor	Statistics			
	MEAN (SD)	CR	AVE	SMC
AFFECT	2.22 (1.00)	.897	.745	.317
TASK	3.06 (.82)	.818	.602	.303
PROC	1.88 (.82)	.871	.698	.234
NORM	1.88 (.87)	.836	.629	.317
PERF	3.78 (.68)	.792	.559	.162
ITR	4.02 (.95)	. 789	.557	.206
IDENT	3.96 (.80)	.811	.519	.206
INTER	3.01 (.96)	.839	.568	.061
DISTRIB	4.10 (1.17)	.802	.577	.113
LEADER	3.87 (.82)	.908	.665	.190
DECIS	3.88 (.89)	.843	.648	.205

Table 2 Means, Standard Deviations, Reliability and Validity Statistics

antecedents and outcome variables (Model 2). Specifically, as more transformational leadership styles may directly improve perceptions of performance, identification with the team and consequently, intention to remain, we drew a path between leadership and each of our outcome variables. Stronger identification with the team may also be a direct result of greater distribution of decision-making and task interdependence. Therefore, we modeled these paths as well. Finally, we added covariance among the four conflict variables, as they are known to co-occur. Table 1 shows that our Model 2 had a much improved and acceptable fit. Figure 2 summarizes the significant hypothesized relationships from Model 2, omitting paths from antecedents to outcomes and covariance for visual clarity. The omitted path results are discussed in the following section together with our main findings.



\# p< .1 * p<.05 ** p<.01 ***p<.000

Figure 2 Standardized Factor Loadings for Conflict Model 2

RESULTS

Descriptive statistics

Table 2 presents the means and standard deviation for the latent variables used in our model. In general, overall conflict levels reported are fairly low (M=1.88 to M=3.06, with task conflict the most frequently occurring) while identification with the team (M=3.96), performance perception (M=3.78) and intent to remain (M=4.02) in the project are relatively high. Teams surveyed are also highly distributed geographically (M=4.10), which is expected in a FOSS context, and have largely participative decision making styles (M=3.88) and transformational leaders (M=3.87).

Hypothesis testing

Results of our structural equation model (Figure 2) show that identification is significantly and positively related to intention to remain in the project (β=.289, p<.01) (H2), but performance perception is not (H1). Thus, interestingly, developers' perceptions of a project's success do not significantly impact their intention to continue contributing to the project. Normative conflict is strongly and negatively related to identification with the team (β=-.355, p<.01), and perceptions of team performance (β=-.413, p<.01)(H6). Contrary to our expectations in H4, relationship conflict has a significant *positive* effect on perception of FOSS team performance when controlling for the effects of other types of conflict (β=.423, p<.000). Relationship conflict also has a significant positive relationship with identification with the team (β=.205, p<.01). Finally, task conflict does not have any direct effect on either performance perception or team identification (thus H3 is not supported), while process conflict has a marginally significant negative effect on performance perception (β=-.159, p=.069) (in partial support of H5).

With regards to conflict antecedents, team interdependence was significantly and positively related to the occurrence of task (β=.234, p<.01), affective (β=.234, p<.05) and process (β=.241, p<.01) conflict in FOSS teams, as predicted in H7. Geographical team distribution was significantly related only to relationship conflict (β=.256, p<.01) (in partial support of H8). A distributed decision-making style was significantly and negatively connected with the occurrence of procedural (β=-.218, p<.01) and normative (β=-.170, p<.01) conflict only (H10). Finally, a transformational leadership style was significantly negatively related to the occurrence of task (β=-.265, p<.05), affective (β=-.442, p<.000) and normative (β=-.312, p<.01) conflict (in partial support of H9).

There were also significant direct relationships between leadership and performance perception (β=.401, p<.000) as well as intention to remain (β=.291, p<.01). Team identification was also significantly predicted by interdependence (β=.239, p<.05) and distribution of decision making (β=.333, p<.01). Finally, the covariance between all 6 possible pairs of conflict were all significant at a minimum of p<.01.

Interaction Effects

As the four types of conflict may have potential interactions with each other, we ran additional tests to address our research question (RQ). Though it is possible to include interaction terms in an SEM model, this procedure makes effects interpretation more difficult. Additionally, there are 6 potential interactions possible for each outcome, and including these would make our research model overly complex. Thus we chose to perform this analysis post-hoc using hierarchical multiple regressions instead. This also allowed us to include control variables: gender and role in the project were coded as dummy variables, while length of contribution (in years) and team size were treated as continuous, and included in the first step. In the second step, we entered the four conflict variables. In the third step, we entered all 6 possible interaction effects between the 4 conflict types.

We used residual centering to compute the interactions, rather than traditional mean centering, to allow for easier interpretation of the main and interaction effects. Residual centering allows the computation of interaction terms that are orthogonal to their respective main effects (i.e. have no effect), and thus allow main effects to be interpreted together with their interaction effects [33]. To compute the interaction terms, we followed procedure proposed by Little et al. [33].

Regression results showed that among the four conflict types, only normative conflict had a direct and negative main effect (β= -.279, p <.01) on identification with the team. Unlike the SEM model in Figure 2, affective conflict did not have a significant main effect on identification in the regression model, possibly because the regression does not control for the direct effects of antecedents on outcomes the way the SEM model does. Among the control variables, only length of contribution had a significant positive relationship with identification (β = .167, p<.05), while committers (β = -.272, p<.01) and members with no official title (β =-.217, p<.01) reported significantly less identification compared to team leaders. There were no significant interactions between conflict types associated with identification.

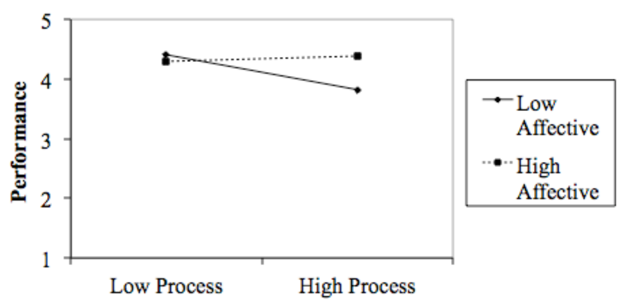


Figure 3 Interacting effects of affective and process conflict on performance perception in FOSS teams

Affective (β = .189, p<.05), process (β = -.155, p<.05) and normative (β = -.251, p<.01) conflict had significant main effects on performance perception. None of the control variables had significant main effects on performance. There was also a significant and positive interaction between affective and process conflict (β= .217, p<.05). Figure 3 details this two-way interaction – at low levels of affective conflict, greater procedural conflict had a negative effect on performance but at high levels there was no difference. There was also a significant and positive interaction between normative and task conflict (β = .224, p<.05). As shown in Figure 4, at low levels of task conflict, normative conflict had a significantly negative effect on performance perception, but at high levels of task conflict there was no such effect.

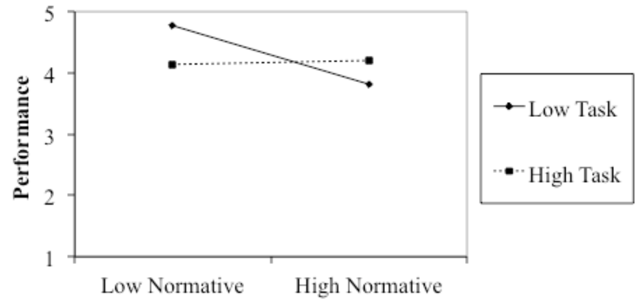


Figure 4 Interacting effects of normative conflict and task conflict on performance perception of FOSS teams

The lack of significant main effects of task conflict on any of the outcome variables measured was surprising. Research has found that task conflict can have a curvilinear relationship with outcomes like performance both in traditional virtual teams [16] as well as peer production communities like Wikipedia [2]. Specifically, low and high levels of task conflict have a negative effect on performance, while moderate levels are beneficial for team performance [2,16]. This may offer a potential explanation of our findings, thus we decided to investigate this relationship further in a post-hoc test. To test for a possible quadratic effect, we entered the square of our original task conflict variable into the third step of the same hierarchical equations predicting performance and identification. The squared term was not significant in either regression, and it did not result in significant changes to the main effects and interactions reported above. Thus, in contrast to earlier work [2,16], the results of our post-hoc test did not show a curvilinear relationship between task conflict and performance perception or identification with the team.

DISCUSSION

The present study examined the diverse effects of conflict on Free and Open Source (FOSS) contributors' sustained participation. To do so, the study aimed to a) quantitatively assess prevalence of different kinds of conflict in FOSS teams; b) understand their differential impact on developers' perceptions of team performance, identification with the team and intention to remain as project contributors; and c) explore structural antecedents to different conflict dimensions experienced in FOSS projects.

This study makes a number of important theoretical and practical contributions. First, this study makes a methodological contribution in developing and validating a set of measurement tools for normative conflict, as well as reformulating conflict into a four-dimensional model.

Second, we contribute to larger theory on conflict in virtual work by explicating the effects of normative conflict on team outcomes, hitherto unstudied. Specifically, we find that levels of normative conflict are significant predictors of both identification with the team, as well as perceptions of team performance. Additionally, normative conflict is the only conflict type to negatively affect developers' intention to remain in the project, acting indirectly by reducing identification with the team. Because normative conflict highlights a dissonance between community best practices and actual team behavior, it may reduce developers' sense of group cohesion, affecting their identification with the group and consequently their intention to continue contributing to the project. Normative conflict may also crystallize subgroups of opinions within the community, and if enough developers perceive this dissonance, it may even lead to a project fork. Our findings suggest the need for further work on the impact of normative conflict on other team variables.

Thirdly, we contribute to literature on peer production by exploring the complex effects of different types of conflict on developers' perception of team performance. For instance, we find that task conflict moderates the negative relationship of normative conflict with performance perception. To recap, at low levels of task conflict, high levels of normative conflict have a greater negative impact on performance. Thus, interestingly, when there is a lack of debate about task related issues, such as adding new features or disagreeing about a proposed change to the code base, normative conflict has the potential to "steal the show" in community discussion and lead to reduced performance. This finding has potentially important implications about managing normative conflict in FOSS teams. Given the overall negative effects of normative conflict on team outcomes, it would be worthwhile to examine whether technical or social mechanisms could be used to divert community attention toward task related issues and thereby improve team function during high levels of normative debate.

Additionally, we find that affective conflict has a surprisingly positive relationship with performance perception and identification with the team. Thus, paradoxically, developers who perceive greater levels of affective conflict also perceive their teams to be the best performing and identify most strongly with the group. Perhaps the presence of affective conflict in voluntary communities like FOSS suggests greater emotional involvement in the project [3], thus indirectly contributing to team success. Affective conflict levels may also be indirect indicators of project activity. As Dabbish et al. [13] found, even potentially negative indicators such as visibility of membership turnover can lead to greater participation intentions, as they demonstrate a certain level of activity within the group. They also found the effect particularly salient in groups with strong team identification. Thus greater team cohesion may mitigate conflict effects, opening up a potentially fruitful avenue for future research.

Finally, like recent prior work on Wikipedia [3] and in contrast to other studies both on peer production and traditional virtual teams [2,15,50], we find no significant main effects of task conflict on either perceptions of team performance or identification with the team. We also investigated the possibility of a more complex curvilinear relationship [2,16], but did not find differing task conflict intensities to have a significantly different impact on individual outcomes. This may be because task conflict can become an expected part of the software development process. As Filippova and Cho [19] allude, FOSS developers expect a certain level of critique when submitting patches. In fact, systems like GitHub and code review or "request for comment" practices are designed to stimulate feedback when contributions are submitted. Thus task conflict may be seen in FOSS projects as a part of the process, rather than an extraneous force on the team. It may be enlightening for future work to examine if the presence of a pro-feedback group norm would moderate the relationship between task conflict and team or individual outcomes.

Fourth, in examining group-level conflict antecedents, we also offer practical recommendations for team design and conflict management. For instance, we find that greater participatory decision-making significantly reduces levels of procedural and normative conflict in FOSS teams, and thereby helps to moderate their negative effects on team performance perception and identification. Hence, teams with more hierarchical structures that experience high levels of disagreements about process or the team direction may consider delegating greater decision-making power to contributors. We also found that a transformational leadership style was significantly related with lower levels of procedural and normative conflict in the team. Thus when selecting new team leadership, either by vote or through appointment by a former leader, FOSS teams should consider not just the individual(s) technical contribution but also their charisma and the extent to which they inspire other developers to follow them in taking the project further. Additionally, our findings suggest that reducing interdependence and increasing task modularity can help to mitigate negative effects of process conflict.

Finally, our findings echo and extend earlier work on conflict in peer production by showing both direct negative effects of process conflict on performance perception, as well as its interaction with affective conflict. Specifically, the presence of greater procedural conflict at lower levels of affective conflict significantly reduces developers' perception of their team's performance. This finding echoes work by Arazy and Nov [3] on Wikipedia who found the co-occurrence of process conflict has a negative impact on article quality, as well as findings in traditional virtual team literature [50].

Taken together, these findings suggest that different kinds of conflict have differing and complex relationships with

FOSS team outcomes, thus reaffirming the need to examine them individually. Additionally, while task conflict occurs most frequently, normative conflict has by far the most negative effects on FOSS team outcomes. Given that earlier qualitative work on normative conflict has anticipated an overall positive impact on group outcomes [19], an interesting future research question is whether and how these effects may change over time. For instance, though contributors may have greater intentions to leave following normative conflict, is this necessarily a bad thing for the team in the long-term? When a FOSS community forks, developers who leave frequently do so on ideological grounds, forming (relatively) more cohesive new communities. Future longitudinal studies of normative conflict effects in FOSS teams may shed light on this dynamic.

LIMITATIONS AND FUTURE WORK

This study is not without limitations. We use self-reported perceptions of outcomes and intentions, rather than measuring outcomes and behavior directly. Specifically, we ask developers the extent to which they are satisfied with the recent performance of their team, their level of identification with the team and intention to remain in the project. As we are interested in predicting developer intentions to remain in the project, measuring individual satisfaction with project performance is not inappropriate. This is because an individual's intention to engage in a behavior is a factor of their subjective evaluation of the activity [1]. In fact, Campion et al. [7] have shown that perceptions of group members about the team, its processes and success are a good predictor of subsequent behavior and actual outcomes. While the link between intentions and subsequent behavior has been established across a number of earlier studies, it would be helpful to validate these initial findings with future work using more objective performance and behavioral measurements.

Furthermore, as with all survey methodology, it is possible that there is a self-selection of respondents who are more identified with their teams and thus are more inclined to participate and present their project in a positive light. Additionally, as Packer [41] suggests, individuals who have lower identification with the team may withdraw from the group when they experience greater normative conflict levels. This group of respondents may not be present in our sample of active contributors. Now that a link has been quantitatively established between normative conflict and developer retention, we therefore recommend for future work to build on the expanded taxonomy presented here and examine conflict manifestations directly in FOSS team communication archives. This can capture both experiences of developers over time who may have left the group, alongside objective behavioral outcome measures derived from artifacts such as commit logs and project downloads.

Finally, identification with the team can also be an antecedent to conflict, rather than an outcome as specified in our model. Team members who identify strongly with their group are more likely to express dissent and initiate normative conflict [41]. Thus identification may have a cyclical relationship with conflict, both encouraging more manifest normative conflict in strongly identifying members, and in doing so potentially reducing the identification of other team members. This is also a fruitful area for future inquiry in a more controlled environment such as ongoing groups with fixed membership structures, or a longitudinal case study of a single FOSS project, tracing the evolution of normative conflict levels and membership over time. Aside from the inclusion of identification as antecedent, there are a number of other relevant factors for future work to consider that are outside of the scope of the present study. For instance, various conflict resolution strategies are known to moderate the effects of conflict on outcomes [43,49]. Also, greater team diversity can stimulate conflict due to differences in background and understanding [25]. We hope to see future work extend and validate the above model with the inclusion of additional factors.

CONCLUSION

The present study investigated the effects of different conflict types on Free and Open Source Software contributors' attitudes towards their project and their sustained participation. We found that normative conflict was the only negative predictor of intention to remain in a FOSS project, suggesting further work is needed to explore normative conflict in virtual work. Additionally, task, affective, process and normative conflict had differing and complex effects on developers attitudes towards their project. The study contributes a theoretical conflict explication to research on peer production, and makes practical suggestions for managing unproductive conflict by employing more transformational leadership and participatory decision-making.

REFERENCES

1. Allen C. Amason and Harry J. Sapienza. 1997. The Effects of Top Management Team Size and interaction Norms on Cognitive and Affective Conflict. *Journal of Management* 23, 4: 495–516.
2. Ofer Arazy, Oded Nov, Raymond Patterson, and Lisa Yeo. 2011. Information Quality in Wikipedia: The Effects of Group Composition and Task Conflict. *Journal of Management Information Systems* 27, 4: 71–98.
3. Ofer Arazy, Lisa Yeo, and Oded Nov. 2013. Stay on the Wikipedia task: When task-related disagreements slip into personal and procedural conflicts. *Journal of the American Society for Information Science and Technology* 64, 8: 1634–1648.
4. Richard P. Bagozzi and Utpal M. Dholakia. 2006. Open Source Software User Communities: A Study of Participation in Linux User Groups. *Management Science* 52, 7: 1099–1115.

5. Kristin J. Behfar, Elizabeth A. Mannix, Randall S. Peterson, and William M. Trochim. 2011. Conflict in Small Groups: The Meaning and Consequences of Process Conflict. *Small Group Research* 42, 2: 127–176.
6. Francesco Bolici, James Howison, and Kevin Crowston. 2009. Coordination without discussion? Socio-technical congruence and Stigmergy in Free and Open Source Software projects. In *Socio-Technical Congruence Workshop in conj Intl Conf on Software Engineering*, 1 – 9.
7. Michael Campion, Ellen Papper, and Gina Medsker. 1996. Relations Between Work Team Characteristics and Effectiveness: a Replication and Extension. *Personnel Psychology* 49, 2: 429–452.
8. Sally A Carless, Alexander J Wearing, and Leon Mann. 2000. A short measure of transformational leadership. *Journal of Business and Psychology* 14, 3: 389–405.
9. Katherine M Chudoba, Eleanor Wynn, Mei Lu, and Mary B. Watson-Manheim. 2005. How virtual are we? Measuring virtuality and understanding its impact in a global organization. *Information Systems Journal* 15, 4: 279–306.
10. Benjamin Collier and Julia Bear. 2012. Conflict, Confidence, or Criticism: An Empirical Examination of the Gender Gap in Wikipedia. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (CSCW '12), 383.
11. Kevin Crowston, James Howison, and Hala Annabi. 2006. Information systems success in free and open source software development: Theory and measures. *Software Process Improvement and Practice* 11, 2: 123–148.
12. Kevin Crowston, Kangning Wei, James Howison, and Andrea Wiggins. 2012. Free/Libre open-source software development. *ACM Computing Surveys* 44, 2: 1–35.
13. Laura Dabbish, Rosta Farzan, Robert Kraut, and Tom Postmes. 2012. Fresh faces in the crowd: Turnover, Identity and Commitment in Online Groups. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (CSCW '12), 245.
14. William H. DeLone and Ephraim R. McLean. 2003. The DeLone and McLean Model of Information Systems Success : A Ten-Year Update. *Journal of Management Information Systems* 19, 4: 9–30.
15. Carsten K W De Dreu and Laurie R Weingart. 2003. Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis. *The Journal of Applied Psychology* 88, 4: 741–749.
16. Carsten K. W. De Dreu. 2006. When Too Little or Too Much Hurts: Evidence for a Curvilinear Relationship Between Task Conflict and Innovation in Teams. *Journal of Management* 32, 1: 83–107.
17. Margaret S Elliott and Walt Scacchi. 2003. Free Software Developers as an Occupational Community : Resolving Conflicts and Fostering Collaboration. In *Proceedings of the 4th International Conference on Supporting Group Work* (GROUP '03), 21–30.
18. Yulin Fang and Derrick Neufeld. 2009. Understanding Sustained Participation in Open Source Software Projects. *Journal of Management Information Systems* 25, 4: 9–50.
19. Anna Filippova and Hichang Cho. 2015. Mudslinging and Manners: Unpacking Conflict in Free and Open Source Software. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '15), 1393–1403.
20. V. Franco, R. Piirto, H. Y. Hu, B. V. Lewenstein, R. Underwood, and N. K. Vidal. 1995. Anatomy of a flame: conflict and community building on the internet. *IEEE Technology and Society Magazine* 14, 2: 12–21.
21. Paola Giuri, Francesco Rullani, and Salvatore Torrisi. 2008. Explaining leadership in virtual teams: The case of open source software. *Information Economics and Policy* 20, 4: 305–315.
22. Georgios Gousios. 2013. The GHTorrent dataset and tool suite. In *Proceedings of the 10th Working Conference on Mining Software Repositories* (MSR '13), 233–236.
23. Steven Greene. 2014. Understanding Party Identification: A Social Identity Approach. *Political Psychology* 20, 2: 393–403.
24. Pamela J Hinds and Diane E Bailey. 2003. Out of Sight, Out of Sync: Understanding Conflict in Distributed Teams. *Organization Science* 14, 6: 615–632.
25. Karen A Jehn, Gregory B Northcraft, and Margaret A Neale. 1999. Why Differences Make a Difference: A Field Study of Diversity, Conflict, and Performance in Workgroups. *Administrative Science Quarterly* 44, 4: 741–763.
26. Karen A. Jehn. 1995. A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative science quarterly* 40, 2: 256–282.
27. Timothy A Judge and Ronald F Piccolo. 2004. Transformational and transactional leadership: a meta-analytic test of their relative validity. *The Journal of applied psychology* 89, 5: 755–768.
28. Atreyi Kankanhalli, Bernard C.Y. Tan, and Kwok-Kee Wei. 2007. Conflict and Performance in Global Virtual

Teams. *Journal of Management Information Systems* 23, 3: 237–274.

29. Aniket Kittur and Robert E Kraut. 2010. Beyond Wikipedia: Coordination and Conflict in Online Production Groups. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (CSCW ’10), 215.
30. Aniket Kittur, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. 2007. He says, she says: conflict and coordination in Wikipedia. In *ACM Conference on Human Factors in Computing Systems* (CHI 2007), 453 – 462.
31. Sang-yong Tom Lee, Hee-woong Kim, and Sumeet Gupta. 2009. Measuring open source software success. *International Journal of Management Science* 37, 426–438.
32. Yan Li, Chuan Hoo Tan, and Hock Hai Teo. 2012. Leadership characteristics and developers’ motivation in open source software development. *Information and Management* 49, 5: 257–267.
33. Todd D. Little, James A. Bovaird, and Keith F. Widaman. 2006. On the Merits of Orthogonalizing Powered and Product Terms: Implications for Modeling Interactions Among Latent Variables. *Structural Equation Modeling* 13, 4: 497–519.
34. Rachel I McDonald, Kelly S Fielding, and Winnifred R Louis. 2012. Energizing and De-Motivating Effects of Norm-Conflict. *Personality and Social Psychology Bulletin* 39, 1: 52–72.
35. Roderick P McDonald and Moon-Ho Ringo Ho. 2002. Principles and practice in reporting structural equation analyses. *Psychological methods* 7, 1: 64–82.
36. Edward F McDonough, Kenneth B Kahn, and Gloria Barczak. 2001. An investigation of the use of global, virtual, and colocated new product development teams. *Journal of Product Innovation Management* 18, 2: 110–120.
37. Mark Mortensen and Pamela J Hinds. 2001. Conflict and Shared Identity in Geographically Distributed Teams. *The International Journal of Conflict Management* 12, 3: 212–238.
38. Siobhán O’Mahony and Fabrizio Ferraro. 2007. The emergence of governance in an open source community. *Academy of Management Journal* 50, 5: 1079–1106.
39. Shaul Oreg and Oded Nov. 2008. Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values. *Computers in Human Behavior* 24, 5: 2055–2073.
40. Dominic J Packer and Alison L Chasteen. 2010. Loyal deviance: testing the normative conflict model of dissent in social groups. *Personality and social psychology bulletin* 36, 1: 5–18.
41. Dominic J Packer. 2008. On being both with us and against us: a normative conflict model of dissent in social groups. *Personality and Social Psychology Review* 12, 1: 50–72.
42. Radostina K. Purvanova and Joyce E. Bono. 2009. Transformational leadership in context: Face-to-face and virtual teams. *Leadership Quarterly* 20, 3: 343–357.
43. M. Afzalur Rahim. 2002. Toward a Theory of Managing Organizational Conflict. *International Journal of Conflict Management* 13, 3: 206–235.
44. C. S. Saunders and M. K. Ahuja. 2006. Are All Distributed Teams the Same? Differentiating Between Temporary and Ongoing Distributed Teams. *Small Group Research* 37, 6: 662–700.
45. Sarah Sharp. 2013. Re: [00/19] 3.10.1-stable review. Retrieved from http://marc.info/?l=linux-kernel&m=137390362508794&w=2
46. Katherine J Stewart and Sanjay Gosain. 2006. The impact of ideology on effectiveness in open source software development teams. *MIS Quarterly* 30, 2: 291–314.
47. John C. Turner, Michael A. Hogg, Penelope J. Oakes, Stephen D. Reicher, and Margaret S. Wetherell. 1987. *Rediscovering the Social Group: A Self-Categorization Theory.* Basil Blackwell, New York.
48. Robin L. Wakefield, Dorothy E. Leidner, and Gary Garrison. 2008. A model of conflict, leadership, and performance in virtual teams. *Information Systems Research* 19, 4: 434–455.
49. R. van Wendel de Joode. 2004. Managing Conflicts in Open Source Communities. *Electronic Markets* 14, 2: 104–113.
50. Frank R. C. de Wit, Lindred L. Greer, and Karen a. Jehn. 2012. The paradox of intragroup conflict: A meta-analysis. *Journal of Applied Psychology* 97, 2: 360–390.
51. Bo Xu, Donald R. Jones, and Bingjia Shao. 2009. Volunteers’ involvement in online community based software development. *Information and Management* 46, 3: 151–158.
52. Philip Yetton and Rajeev Sharma. 2007. The Contingent Effects of Training, Technical Complexity, and Task Interdependence on Successful Information Systems Implementation. *MIS Quarterly* 31, 2: 219–238.
53. Haiyi Zhu, Robert Kraut, and Aniket Kittur. 2012. Organizing without formal organization: Group Identification, Goal Setting and Social Modeling in Directing Online Production. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (CSCW ’12), 935–944.