2010

PROMOTING GOOD MANAGEMENT: GOVERNANCE, PROMOTION, AND LEADERSHIP IN OPEN COLLABORATION COMMUNITIES

Benjamin Collier
Carnegie Mellon University, bcollier@cmu.edu

Moira Burke
Carnegie Mellon University, moira@cmu.edu

Niki Kittur
Carnegie Mellon University, nkittur@cs.cmu.edu

Robert Kraut
Carnegie Mellon University, robert.kraut@cmu.edu

PROMOTING GOOD MANAGEMENT: GOVERNANCE, PROMOTION, AND LEADERSHIP IN OPEN COLLABORATION COMMUNITIES

Research-in-Progress

Benjamin Collier
Tepper School of Business
Carnegie Mellon University
bcollier@cmu.edu

Moira Burke
Human Computer Interaction
Carnegie Mellon University
moira@cmu.edu

Niki Kittur
Human Computer Interaction
Carnegie Mellon University
nkittur@cs.cmu.edu

Robert Kraut
Human Computer Interaction
Carnegie Mellon University
robert.kraut@cmu.edu

Abstract

While we do have an idea of what the leadership of open collaboration communities is responsible for at a high level, we have little knowledge of what these leaders actually do. Similarly, in an online context we understand pieces of the process managers and leaders go through to be elected, but their change in behavior once they become a manager remains unexplored. What behaviors do these leaders engage in that is different from a typical contributor to the community? How can we empirically distinguish which behaviors are characteristic of a successful leader vs. an unsuccessful leader? We examine leadership promotion and performance in Wikipedia, a prominent open collaboration community. This study extends previous work on governance and leadership by developing and validating a more complete measurement model of leadership performance in an open collaboration community, and proposing a testable model of leadership promotion and performance within this context.

Introduction

Research in open collaboration communities has examined how these communities motivate contributors (Hertel et al., 2003; Lakhani et al., 2005; Roberts et al., 2006; Shah, 2006), establish joining processes (von Krogh et al., 2003), develop social status differences (Stewart, 2005), develop communication ties (Crowston & Howison, 2005; Grewel et al., 2006), and define success for their community (Crowston et al., 2006). Much less attention in the literature has been given to the understanding of governance and leadership in open collaboration communities. Previous research examined the technical contributions, communications patterns, and network structure within open collaboration communities and developed empirical models to predict whether a candidate for promotion to leadership received the promotion (Fleming & Waquespack, 2007; O'Mahony & Ferraro, 2007).

While we do have a vague idea of what the leadership is responsible for at a high level, we have little knowledge of what these leaders actually do. The classical understanding was that managers in traditional organizations plan, organize, coordinate, and control (Drucker, 1954; Fayol, 1916). However, careful observation by Mintzberg later revealed a very different pattern of activities among managers (Mintzberg,1971; Mintzberg; 1990). Similarly, in an online context we understand something of the process managers and leaders go through to be elected, but their change in behaviors once they become a manager remains unexplored. Why do these relatively flat communities need a governing body? What behaviors do these leaders engage in that is different from a typical contributor to the community? How can we empirically distinguish which behaviors are characteristic of a successful leader vs. an unsuccessful leader?

This study examines leadership promotion and performance in Wikipedia, a prominent open collaboration community. Previous studies have shown that as Wikipedia grows, the need for coordination, communication, policy work, and governance increases in relation to other work (Butler et al., 2008; Forte & Bruckman, 2008; Kittur et al., 2007; Suh et al., 2009). An elected group of users within Wikipedia known as administrators contribute heavily to coordination, mediation, and policy issues and will be the primary focus of this study (see Bryant et al., 2005 for a qualitative analysis of Wikipedia administrators). This study extends previous work on governance and leadership by (1) developing and validating a more complete measurement model of leadership performance in an open collaboration community, (2) proposing a testable conceptual model of leadership promotion and performance within this context.

Within Wikipedia we have a history of every article a user has edited, communication between users, and a complete history of all of the user's contributions to Wikipedia policy making and community building. In the first part of our study, we develop a list of categorical behaviors and empirical measures for those behaviors that are characteristic of administrators. The measurement model is validated through a survey of 91 Wikipedia administrators and voters in the promotion process. Using these behaviors as a guide, we then propose a testable model of promotion and administrator performance.

Promotion and Managerial Performance in Open Collaboration Communities

Promotions serve two roles in an organization: selection and incentive (Milgroms and Roberts, 1992; Gibbons and Waldman, 2006). First, promotions are part of a reward and incentive system to induce employees at one level of an organization to work hard in the organization's interest so that they can be promoted to positions of more power and responsibility (Lazear and Rosen, 1981). Second, they are part of a selection process in which people are vetted and assigned to the roles where they can best contribute to the organization's performance (O'Flaherty and Siow, 1992).

Whether promotion functions as an incentive or a selection mechanism is likely to determine the types of evaluation criteria managers use in making promotion decisions. To the extent that promotions serve as incentives, managers should look for evidence that candidates for promotion have performed well in the past, independent of the degree to which prior job performance predicts performance in a subsequent one. Because the promotions are being used to reward past behavior and encourage effort and work quality that the organization judges as valuable, managers should base promotions on the type of behavior the organization wants to encourage, whether or not this behavior would serve the organization or candidate when the candidate ascends to a new position. For example, in a production environment like a factory, managers should promote the most productive workers to supervisor positions. In this sense, the criteria that managers use to make promotion decisions should be backward-looking; they should be looking for evidence of good performance in the prior job.

To the extent that promotions serve as a selection device to identify appropriate candidates for higher-level, more difficult or more valuable organizational roles, then managers should use prior job performance primarily as evidence of future performance in these new roles. In this view, the promotion view is a prediction task, and the promotion criteria that managers use should be prospective or forward-looking as they try to use prior job performance to predict performance in the new job.

In the context of open collaboration communities, to the extent that promotion to an administrator position is a "badge of honor", or an incentive for contributors to work hard, we would expect to see the tasks being performed by administrators and the qualities valued in candidates for promotion to be similar to the tasks performed by contributors. On the other hand, if the administrator position is a selection mechanism for obtaining a contributor who performs a different kind of work, we would expect to see differences in the kinds of tasks administrators perform, and the qualities valued in potential candidates.

Research in open collaboration communities has only begun to address issues of governance and leadership. Shah (2006) examines both gated source and open source communities, and finds that individual motivations are different between the two governance systems: contributors to open collaboration communities' motivations within evolve with experience. Their research suggests something about the governance model and organizational purpose impacts contributors in this setting differently than in traditional organizational designs. Following this research we can then ask if the managers and leaders governing these organizations behave differently or require a different set of experiences than in traditional organizations.

Fleming and Waguespack (2007) examined leadership in open innovation communities using a sample from the Internet Engineering Task Force and found that while technical contributions are important to leadership success, social connections-in terms of both brokerage and boundary spanning across subgroups of the task force-were critical to leadership success. From this we gather that social relationships and interaction with community members are important parts of being in leadership in an open innovation community-above and beyond one's technical or production capabilities. O'Mahony and Ferraro (2007) similarly found that early on potential leaders ran on a platform of technical accomplishment, and later moved to use a higher proportion of organization building language. Additionally, they found that both their technical contributions and their online communication were predictive in their success in gaining promotion to the leadership team.

While both of these studies predict both production-oriented activities and interpersonal activities to increase a candidates likelihood of being promoted, once a candidate is promoted it is unclear whether this balance of activities continues to be important for managers. The role of manager in open collaboration communities likely requires a different set of requirements, and depending on prior experiences, some recently promoted managers may have more or less difficulty in adjusting to the requirements of the new role (Nicholson, 1984; Ashforth & Saks, 1995).

A Wikipedia Primer

Wikipedia is a large, open source, online encyclopedia written and edited by volunteers. Although Wikipedia is written and edited by volunteers and is not supervised by a professional staff, evidence suggests that the quality in Wikipedia is comparable to that of the Encyclopedia Britannica (Giles, 2005). Wikipedia as an organization is substantially less hierarchical than a conventional production company. By contrast, Encyclopedia Britannica has a CEO and other corporate officers, an editor-in-chief, a scholarly advisory board, and a staff of 19 full-time editors who vet and supervise the work of over 4000 contributors. While any one may alter an article page within Wikipedia, a vast majority of changes made to Wikipedia is done by registered users. Within the community of registered users of Wikipedia, anyone who makes changes to articles may be considered an editor. Editors make up the largest proportion of Wikipedia users, recently approaching 7 million members. Administrators are promoted from the group of editors after going a stringent week long democratic process in which users discuss and vote on their fit for the administrator position. From within the administrator position, a very select group of users are then promoted to Bureaucrats, and have a very public role as the policy developers, mediators, and community leaders. The focus of this paper will be on the promotion of editors to administrators.

Editors

Editors in Wikipedia may add content, make changes, delete content, and change an article back to a previous version if unwanted changes are made by another editor (commonly referred to as vandalism). Editors often

maintain "watch lists" of articles they have an interest in developing so that if any changes are made to those articles by other editors they can review the changes and decide if they are of sufficient quality and have a neutral point of view as required by encyclopedia standards. An editor's role in Wikipedia is characterized by developing the content, and an editor's merit is often judged by bringing an article up to peer-reviewed quality standards in Wikipedia such as that of a "Good Article", or the gold standard in Wikipedia a "Featured Article" which will at some point be featured on the front page of Wikipedia.

Administrators

To help coordinate work in Wikipedia in the face of a large and largely unsupervised workforce, Wikipedia has developed an elaborate system of policies that lay out "Wikipedian" values such as civility and neutral point of view and work procedures such as those for deleting pages, resolving disputes and electing administrators (Butler et al., 2008). It also empowered a class of administrators with special tools to enforce some of the policies. For example, these administrators can ban other editors from working in the encyclopedia, protect pages from changes, and often serve as formal and informal mediators in resolving disputes. Administrators in Wikipedia differ from regular editors in two main ways. First, they have additional tools for page deletion and user maintenance, and regular editors request assistance from administrators at a public notice board for such issues as vandalism, page protection, inappropriate usernames, conflict between users, and page deletion. Second, administrators have elevated social status, as trusted editors who understand policy and its application. Administrators in practice are well regarded as acting in the community's best interest and having enough experience to use their tools wisely. Approximately 2700 editors have been nominated for administrator status since 2001 with an overall promotion success rate of 53%.

Developing a Measurement Model of Managerial Performance

The first step in our approach was to identify the qualities and behaviors that are important for administrators within Wikipedia. To become an administrator, an editor must undergo a week of scrutiny known as the Request for Adminship (RfA), during which the community builds consensus about the candidate's experience and suitability for the position. The process consists of three parts: an introductory nomination statement, the nominee's answers to questions about past and future behavior, and statements of support, opposition, or neutrality by community members. These statements of support/oppose/neutral are public, can be made by any member of the community, and typically contain a rationale behind their decision.

A list of 15 desirable behavioral categories were identified by combining (1) a qualitative analysis of the voter's discussion of 50 candidates and (2) a community guide for desirable behavioral categories in administrators[1] (see Table 1 for complete list). From this list of desirable behavioral categories, a list of 40 potential measures was developed through informal discussions with administrators and community members, a qualitative analysis of the voter discussions, and previous research examining measurement of work in Wikipedia (Kittur et al., 2007; Kriplean et al., 2008).

After developing a list of behavioral categories (e.g. knowledge of policy, helping with chores, strong edit history) and a list of potential measures for those categories (e.g. editing policy pages, reverting vandalism, number of article edits) we constructed a survey which was administered to current administrators, voters of previous Request for Adminship's, or previous candidate for promotion. We put a call out on the administrator's Noticeboard (a central communication hub for current administrators), the Village Pump (a general discussion board within Wikipedia) and the Request for Comment page (a general request page for contributors to review a policy or discussion). 91 contributors participated in the survey.

The objective of the first part of the survey was to determine which behavioral measures would map onto the construct of the behavioral categories. We operationalized this by using an online card sort task having participants be given a list of 15 categories and 40 measures. Their task was to drag each of the measures onto a category which it best measured.

The objective of the second part of the survey was to determine which behavioral categories were "central, important, or critical" to being an editor or administrator. Participants filled out a 5 point scale as to how relevant

---

[1] http://en.wikipedia.org/wiki/Wikipedia:Guide_to_requests_for_adminship

each category was to being a good editor, and another 5 point scale as to how relevant each category was to being a good administrator.

Behavioral Category	Measure
Interaction with others	Posting to other user's talk pages
	Posting a reply to talk page comments on your own talk page
	Posting to an administrator's talk page
	Having an administrator post on your own talk page
	Using civil and polite language in discussions
	Using rude or impolite language in discussions (Negatively Coded)
	Receiving barnstars for social support
	Receiving barnstars for good disposition
	Mentoring others
	Welcoming new users
Article quality	Editing articles that become "featured"
	Editing articles that become "good"
	Edit persistence (how long your contributions remain before they are changed)
Knowledge of policy	Citing policy in discussions
	Editing policy pages
	Editing policy talk pages
	Using Wikipedia Jargon
Community involvement	Receiving barnstars for community involvement
	Being a member of a Wikiproject
	Leading a Wikiproject
Fighting vandalism	Posting vandalism alerts
	Posting warning messages to others
	Receiving barnstars for border patrol work
Conflict management	Participating in dispute resolution (e.g. Arbitration Committee)
	Informal mediation experience
Varied experience	Distribution of edits across topics
	Editing in many different namespaces
Edit summary usage	Percentage of edits with summary
	Edit summary length
Trustworthiness	Having blocks on your account (Negatively Coded)
	Having your edits reverted (Negatively Coded)
Helping with chores	Receiving barnstars for metacontent work
	Receiving barnstars for administrative work
Strong edit history	Edits to articles
	Edits to article talk pages
	Having your articles appear on watch lists
Decision making through consensus	Article talk length
	Posting to administrator's Noticeboard
	Participating in Requests for Comment
	Participating in deletion discussions
Time in Wikipedia	Months since first edit
Commitment to Wikipedia	Enabled Email Address
Neutral Point of View	Ratio of subject to objective language used

Table 1. Survey Results – Matching Measurements to Behavioral Categories

Table 2. Survey Results – Matching Categories to User Types

Category			Mean	Mean of Difference*	*p*
Editor Behaviors	Article quality	Administrator	3.05	-1.217	<.001
		Editor	4.18		
	Strong edit history	Administrator	3.76	-.241	.113
		Editor	4.00		
	Neutral point of view	Administrator	4.29	-.089	.399
		Editor	4.38		
Administrator Behaviors	Edit summary usage	Administrator	3.72	.449	<.001
		Editor	3.27		
	Decision making through consensus	Administrator	4.54	.671	<.001
		Editor	3.87		
	Interaction with others	Administrator	4.52	.722	<.001
		Editor	3.80		
	Trustworthiness	Administrator	4.81	.785	<.001
		Editor	4.03		
	Conflict management	Administrator	4.30	.899	<.001
		Editor	3.41		
	Commitment to Wikipedia	Administrator	4.03	.937	<.001
		Editor	3.09		
	Community involvement	Administrator	3.63	.973	<.001
		Editor	2.66		
	Time in Wikipedia	Administrator	3.85	1.025	<.001
		Editor	2.82		
	Knowledge of policy	Administrator	4.63	1.051	<.001
		Editor	3.58		
	Varied experience	Administrator	4.04	1.192	<.001
		Editor	2.85		
	Fighting vandalism	Administrator	3.74	1.231	<.001
		Editor	2.51		
	Helping with chores	Administrator	4.21	1.551	<.001
		Editor	2.65		

Scale: 1=Not Relevant, 2=Slightly Relevant, 3=Moderately Relevant, 4=Relevant, 5=Very Relevant

*Calculated as Mean(Relevant for Administrators – Relevant for Editors) within subjects. Negative values indicate more relevance for Editors.

Results

Using data from the 91 participants in the card sort we used a Ward's cluster analysis technique combining the categories and the measurement items and clustered them based on how often they were placed together by participants. The overall measurement model clustered well for 12 of the 15 categories with an inter-rater reliability Kappa of .81 for those categories. The remaining three categories have only one measurement per category. Results of the cluster analysis are shown in Table 1.

To determine which categories are of most important for measuring administrator performance relative to editor performance we did a within-subjects t-test examining the difference in ratings of relevance of each category. For example, if a participant rated "knowledge of policy" as Very Relevant (a value of 5) for administrators but Not Relevant (a value of 1) for editors, the difference would be a value of 4. We then test whether the average rating for that category has a difference score that is statistically significantly different from zero. If they are found to have a positive value, they are most relevant for administrators, a negative value would indicate they are more relevant for editors. The results from participant's ratings of categories are shown in Table 2.

The results of this study provide several key insights on management behaviors in open collaboration communities. The first is that the behaviors leaders engage in are significantly different than those of typical contributors. The two behaviors that are most different are helping with chores (metacontent and administrative work) and fighting vandalism (posting alerts, warning messages, and border patrol work). This suggests that a manager's value comes from doing smaller day to day administrative tasks rather than being a centralized coordinator and monitor of production work. By contrast, building consensus (decision making through consensus) and conflict management, while highly valued within the community, are not the primary role of the management, rather, those tasks rest on the shoulders of both editors and administrators. Similarly, it is interesting that as an administrator the quality of your writing is of much less value to the organization than for an editor. Once an administrator is promoted their value lies not in the quality of the production content they create, but in the administrative tasks they perform.

Research Model

Drawing from the discussion of promotion in traditional organizations, we extend the framework into the open collaboration setting. We know from the offline setting that promotion functions as both an incentive for employees to work harder (Lazear and Rosen, 1981) and as a selection mechanism to fill a needed role (O'Flaherty and Siow, 1992). Additionally, research examining promotion and governance in the FLOSS context has found a combination of social and technical performance factors that influence the promotion of leadership (Fleming and Waguespack, 2007; O'Mahony and Ferraro, 2007). It follows that promotion within Wikipedia will be more likely when the candidate has a strong record of both production experience (editor behaviors) and management experience (administrator behaviors) prior to coming up for promotion.

H1: Management experience prior to promotion will have a direct positive relationship with whether a not a candidate is successfully promoted.

H2: Production experience prior to promotion will have a direct positive relationship with whether a not a candidate is successfully promoted.

The Wikipedia community has clear expectations that once a candidate is promoted they are responsible for a different set of tasks than an editor may be. We see from our initial survey results that the community has expectations that behaviors such as writing quality articles are less valued, and activities such as attending to metacontent and administrative work are now highly valued. We posit that once candidates are promoted they will shift their focus to perform on activities relating to management (administrator behaviors), and shift their focus away from production activities (editor behaviors).

H3a: Candidates who are promoted will increase their management performance.

H3b: Candidates who are promoted will decrease the production performance.

Methodology

We sampled a set of 1502 candidates who went through the Request for Adminship process from January 2006 to October 2007. The behavioral measures (listed in Table 1) are continuing to be mined from the archives of each candidate's history. Data and analysis should be completed for presentation in December of 2010. We will utilize a structural equation model (SEM) to measure the latent constructs within the research model using the survey results in Table 1 as a guide for an exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Initial results examining the first measures collected look promising, with an overall model RMSEA of .045 and CFI of .86.

Conclusion

Governance and leadership in open collaboration communities plays a crucial role in the coordinating and making key decisions that have a long term impact. As these communities continue to grow, the need for policy, centralized decision making, arbitration of conflict, and administrative tasks will only increase. This study makes three significant contributions to a growing stream of research on governance and leadership in open collaboration. First, it extends and validates previous models of leadership promotion tested in the FLOSS context into the related but

different context of Wikipedia (Fleming and Waguespack, 2007; O'Mahony and Ferraro, 2007). Second, it develops a specific measurement model for exploring pre-promotion and post-promotion performance of both management and production related tasks. This measurement model can be easily adapted for future researchers to explore organizational theories of management in open collaboration communities. Lastly, this paper proposes a testable research model examining the impact of pre-promotion performance dimensions on post-promotion adaptation to a new role.

Figure 1. Research Model
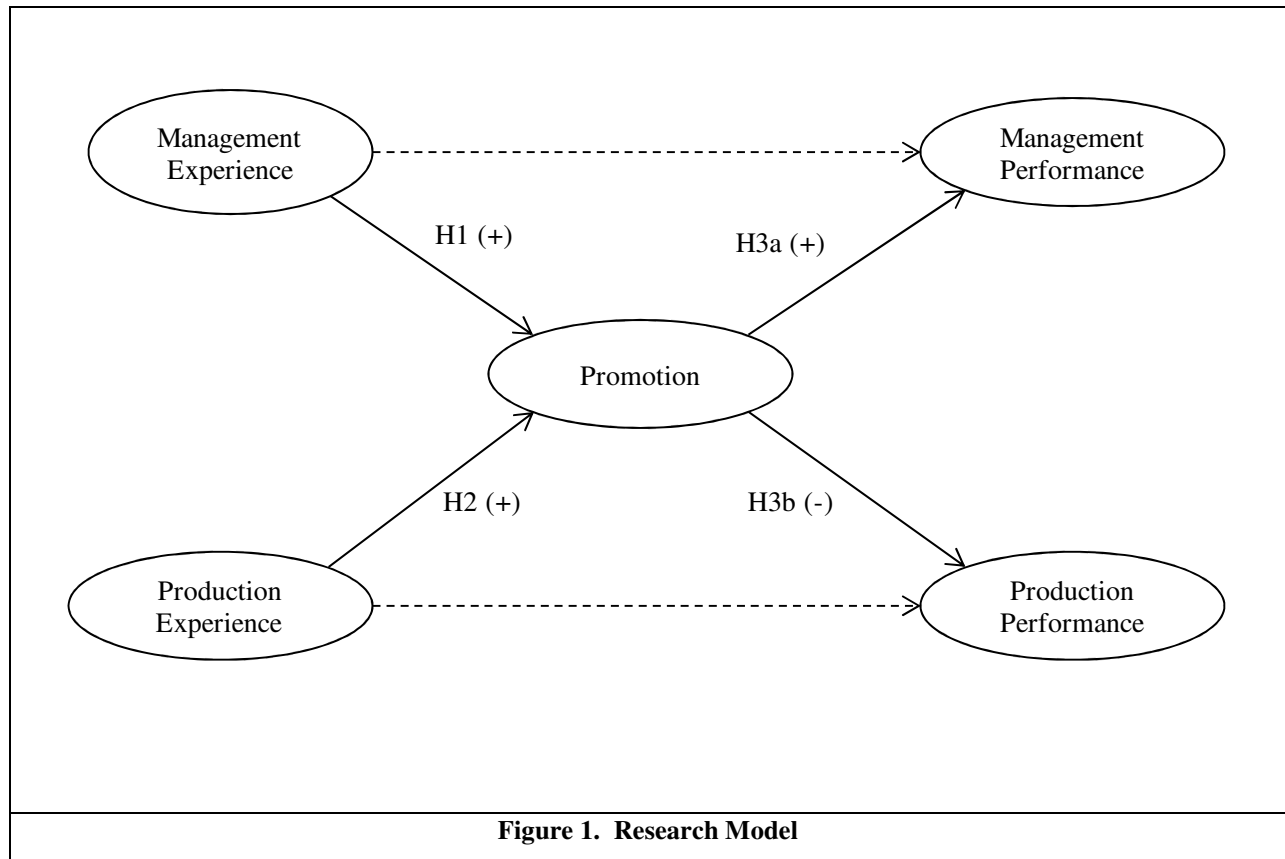
References

Ashforth, B. E., & Saks, A. M. (1995). Work-role Transitions: A Longitudinal Examination of the Nicholson Model. *Journal of Occupational and Organizational Psychology*, *68*, 157-175.

Bryant, S. L., Forte, A., & Bruckman, A. (2005). Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia. Sanibel Island, Florida.

Butler, B., Joyce, L., & Pike, J. (2008). Don't Look Now, But We've Created a Bureaucracy. Florence: ACM Press.

Crowston, K., & Howison, J. (2005). The social structure of Free and Open Source software development. *First Monday*, *10*(2).

Crowston, K., Howison, J., & Annabi, H. (2006). *Information Systems Success in Free and Open Source Software Development : Theory and Measures*. Information Systems.

Drucker, P. F. (1954). *The Practice of Management*. New York: Harper and Row.

Fayol, H. (1916). *Administration industrielle et générale; prévoyance, organisation, commandement, coordination, controle*. Paris: Dunods.

Fleming, L., & Waguespack, D. M. (2007). Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities. *Organization Science*, *18*(2), 165-180.

Forte, A., & Bruckman, A. (2008). *Scaling Consensus: Increasing Decentralization in Wikipedia Governance*.

Gibbons, R., & Waldman, M. (2006). Enriching a Theory of Wage and Promotion Dynamics inside Firms. *Journal of Labor Economics*, *24*(1), 59-107.

Giles, J. (2005). *Internet encyclopedias go head to head*. *Nature* (Vol. 438, pp. 900-901).

Grewel, R., Lilien, G. L., & Mallapragada, G. (2006). Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems. *Management Science*, *52*(7), 1043-1056..

Hertel, G., Niedner, S., & Herrmann, S. (2003). Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel. *Research Policy*, *32*, 1159–1177.

Kittur, A., Chi, E., Pendleton, B. A., Suh, B., & Mytkowicz, T. (2007). Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie. San Jose: ACM Press.

Kittur, A., Suh, B., Pendleton, B. A., & Chi, E. H. (2007). He Says, She Says: Conflict and Coordination in Wikipedia (pp. 453-462). San Jose, CA.

Kriplean, T., Beschastnikh, I., & McDonald, D. W. (2008). Articulations of WikiWork: Uncovering Valued Work in Wikipedia through Barnstars. San Diego, CA: ACM.

Lakhani, K. R., Wolf, R. G., Feller, J., Fitzgerald, B., & Hissam, S. (2005). Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects (pp. 3-22). Cambridge, MA: MIT Press.

Lazear, E., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy*, *89*(5), 841. University of Chicago Press.

Milgrom, P., & Roberts, J. (1992). *Economics, Organization and Management*. Englewood Cliffs, N.J: Prentice Hall.

Mintzberg, H. (1971). Managerial Work: Analysis from Observation. *Management Science*, *18*(2), B97-B110.

Mintzberg, H. (1990). The Manager's Job: Folklore and Fact. *Harvard Business Review*, *68*(2), 163-176.

Nicholson, N. (1984). A Theory of Work Role Transitions. *Administrative Science Quarterly*, *29*(2), 172-191.

O'Flaherty, B., & Siow, A. (1992). On the job screening, up or out rules, and firm growth. *Canadian Journal Of Economics*, *25*(2), 346-368.

O'Mahony, S., & Ferraro, F. (2007). The Emergence of Governance in an Open Source Community. *Academy of Management Journal*, *50*(5), 1079-1106.

Roberts, J., Hann, I., & Slaughter, S. (2006). Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Project. *Management Science*, *52*(7), 984-999.

Shah, S. K. (2006). Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. *Management Science*, *52*(7), 1000-1014.

Stewart, D. (2005). Social status in an open-source community. *American Sociological Review*, *70*(5), 823–842. American Sociological Association.

Suh, B., Convertino, G., Chi, E. H., & Pirolli, P. (2009). The Singularity is Not Near: Slowing Growth of Wikipedia. In *WikiSym* (Vol. 1). Orlando, FL.

von Krogh, G., Spaeth, S., & Lakhani, K. R. (2003). Community, Joining, and Specialization in Open Source Software Innovation: A Case Study. *Research Policy*, *32*(7), 1217-1241.