Peer-recognition and Performance in Online Crowdsourcing Communities
(Extended Abstract)

Indika Dissanayake
University of Texas Arlington
indika.dissanayake@mavs.uta.edu

Jie Zhang
University of Texas Arlington
jie.zhang@uta.edu

Feirong Yuan
University of Texas Arlington
fyuan@uta.edu

Jingguo Wang
University of Texas Arlington
jwang@uta.edu

Abstract
Recent advances in information technology bring significant changes to the nature of social interactions and information exchange. Physical face-to-face communications are slowly replaced by online virtual communities. Motivated by this phenomenon, this research investigates how voluntary community involvement and self-disclosure behavior impact a member’s peer-recognition and task performance within a virtual crowdsourcing competition community. We collected secondary data from the discussion forums of a specialized crowdsourcing platform that focuses on data analytics projects. Our results reveal that a member’s community involvement improves both the peer-recognition of the member in the community and his/her performance ranking. Our findings have strategic implications to participants of virtual crowdsourcing communities and other professional online communities.

1. Introduction
Recent advances in information technology have facilitated the rapid growth of online discussion communities, which facilitate sharing knowledge and information among the community members. Without formal governance, individuals voluntarily participate in and contribute to the discussion communities. The communities also depend on the members’ pro-social behaviors and involvement for growth and success. In particular, an emerging form of virtual professional communities, such as Internet-based communities of practice, have become more popular and attract a large group of individuals with a shared practice but are geographically distributed to exchange knowledge [14]. Individual participants of these communities are likely to learn from their peers, develop professional skills, build social connections, and/or gain peer-recognition [15, 7]. Many organizations also encourage their employees to participate in these professional communities.

Prior research on online communities has primarily focused on the participation motives [1, 14, 18], and few studies examined the impact of community involvement activities [9, 7]. There remain issues to be investigated. First, studies have proposed various arguments to explain users’ participation motives based on their self-interest, altruism and psychological bonds [3]. However, scholars rarely validate whether participants truly realized these benefits. There lacks enough empirical evidence on the impacts of community involvement on both intellectual capital and social capital. This may be because of the difficulty in assessing those metrics and collecting empirical data. Second, sharing knowledge among members in professional discussion communities, though are likely to benefit the entire community including the contributors, can potentially result in knowledge spillover and weaken his/her competitiveness and performance ranking. Very scarce attempt has been made to empirically verify this question. Third, in online discussion communities, members can control the release of their personal information such as geographic location, education background, and so forth. The amount of information disclosed will presumably influence the peers’ trust and recognition and therefore affect members’ achieving their social capital or reputation. However, to the best of our knowledge the role of self-disclosure has not been investigated in crowdsourcing communities.

Using data of individuals participating in innovative contests and affiliated discussion forums from one of the largest data analytic crowdsourcing platforms, Kaggle.com, we empirically validates the realization of the perceived benefits of individuals’ involvement in professional discussion communities by
showing how solver’s community involvement influences their performance and peer-recognition.

2. Literature review

2.1. Pro-social behavior

This study focuses on pro-social behavior or “voluntary community involvement” of people participating in online virtual communities, mainly demonstrated through sharing information and knowledge, and asking and answering questions by posting messages in discussion forums. Pro-social behavior is defined as “voluntary intentional behavior that results in benefits for another” [11]. These behaviors are led by pro-social motives such as altruism (desire to benefit others), egoism (desire to benefit one-self) or combination of both [11].

Extant researches have been extensively investigating the motivations behind pro-social behaviors in online communities. Perceived benefits such as internal satisfaction enhance reputation, obligation to reciprocate, and help to advance community will lead individuals to contribute in knowledge sharing communities [13].

The research on pro-social behaviors suggests that giving to, sharing knowledge with, and helping others may indeed influence the success of the contributor. Podsakoff et al. [8]'s showed that organization citizenship behaviors had positive effects on both organizational-level outcomes and individual-level outcomes. In the context of open source software development, scholars shown that people who help others by answering questions in virtual software support communities enjoys benefits such as learning benefits, reputational benefits, and other benefits related to advancing the group (as cited in [11]). Evidence abound that pro-social behaviors in any domain, not just open source software development, may influence individuals' skills, performance, and career developments. Yet, to this date, we still know little about how community involvement will influence an individual’s own performance in online communities, especially when other members of the community are competitors.

2.2. Self-disclosure behavior

In this paper we focus on voluntarily self-disclosure behaviors of individuals participate in online virtual communities. Voluntarily self-disclosure is defined as “the revealing of personally relevance experience, thoughts, and feelings to others” [12]. Many online communities offer profiles where users can disclose their personal information such as photos, location, experience, and so forth.

There are numerous costs and benefits associated with these voluntary self-disclosure behaviors in cyber space. Evidence shows that self-disclosure is an important factor when forming relationships. It increases interaction, promotes interpersonal bonds [17], improves friendships [12], influences likings and closeness (as cited in [10]), and contributes to the communication behavior and relational closeness [16, 12].

Most of the previous studies investigate the antecedent of self-disclosure behavior/intention. To best of our knowledge, very few studies have been investigated the consequences of the self-disclosure. Our study hopes to bridge this gap in the literature.

3. Research model

Kaggle is dealing with prediction modeling competitions. Tasks are very intellectual and demands innovative work. Pro-social literature showed that when driven by pro-social motives, people were able to come up with creative ideas [4]. Due to competitive nature crowdsourcing contests, the community involvement comes with a cost. By considering nature of the virtual crowdsourcing context and applying the pro-social literature to this context, we anticipate an overall positive relationship between community involvement and individual performance. Further, Pro-social research has identified reputation gain as one possible reason for people to engage in prosocial behavior [2,5,6]. Hence, we hypothesize that the solvers’ community involvement positively related to their performance and peer-recognition.

In our setting disclosure of identity related information is voluntarily. Individuals may decide to fully or partially display information related to their name, education, skills and so forth or to be completely anonymous. When receive information from a person with identity that will increase the trustworthiness of the information. Hence, we hypothesize that the solvers' degree of self-disclosure is positively related to their peer-recognition.

4. Data

We collected data from a specialized crowdsourcing platform that focuses on data analytics projects Kaggle.com. Companies, government, and researchers provide datasets to Kaggle. Based on their requirements, Kaggle sets up contests. Each participant can submit multiple solutions with in the contest duration. Kaggle evaluates all submissions in real time
using a test dataset. Each contest has a live score card where solvers receive information on the prediction accuracy of their model and their relative positions (e.g. ranks) in the contest.

Each solver can maintain online profile in kaggle website. Profile shows their cumulative performance score. In these profiles, solvers can decide to be completely anonymous, partially or fully disclose their true personal information such as name, education, location, favorite software tools, experience, and so forth. In addition, each contest has a forum where solvers can participate in discussion threads in various topics. When an individual make a post in a forum, if other solvers find it helpful, they recognize the contribution by sending thanks to the contributors. Summary statistics of number of posts made by an individual, number of thanks received and so forth are also reflected in their profiles.

For this study, we collected data on 10312 number of solver profiles that reflects competitions they have participated since the joining of Kaggle platform through July 2012.

4.1. Variables (dependent, independent, control)

Our main dependent variables are performance and peer-recognition. We used contest’s profile score as the measure of their performance and total number thanks received as a proxy for peer-recognition.

Our main independent variables are community involvement and self-disclosure. We used total number of posts made by an individual on forums as proxy for community involvement. We considered disclosure of location, education, and experience to measure the degree of self-disclosure. We controlled for the total number of contests participated by each individual.

5. Results

We tested our research model using OLS and simultaneous equation modeling techniques. Our results indicate that community involvement positively influences both performance and peer-recognition of the contributor. Contrary to our expectation we did not find any direct effect of self-disclosure on peer-recognition.

Also, we tested the model using alternative measure of community involvement. Our main results are consistent for these alternative models, alternative measures of community involvement, and different random samples of data confirming the robustness of our results.

6. Discussion

This study helps understanding how the community involvement improves the online status of the contributor in terms of contest performance and peer-recognition in the context of crowdsourcing competitions. Our results indicate that community involvement positively influences both performance and peer-recognition of the contributor.

Our findings have number of theoretical and practical implications. First, our rich data set allow us to measure the performance of the individual’s contributing to the community. Second, our study contributes to the literature on pro-social behavior, and self-disclosure behavior by applying these to crowdsourcing competitions settings. Pro-social behavior in our setting involves high cost, because individuals compete against each other. Third, our findings extend growing research on virtual community contributions by exploring important outcome of community contribution to the contributor. While prior research has addressed the motivation behind the community contributions, little research has addressed the actual benefits from community contribution to the contributor.

Crowdsourcing platform providers can use these insights when designing crowdsourcing contests. Improving the participants’ performance help to improve the quality of wining solution and attract more seekers to their platforms. Managers interested in developing and sustaining intellectual exchange through virtual communities can use these insights when developing strategies and mechanisms to encourage and facilitate knowledge contribution of members. In organizational settings, project managers can use these insights to enhance performance of individuals in their work teams. Participants of crowdsourcing competitions can use these findings to improve their performance and gain social capital through enhancing recognition in the community.

7. References


