



## Task Division for Team Success in Crowdsourcing Contests: Resource Allocation and Alignment Effects

Indika Dissanayake, Jie Zhang & Bin Gu

To cite this article: Indika Dissanayake, Jie Zhang & Bin Gu (2015) Task Division for Team Success in Crowdsourcing Contests: Resource Allocation and Alignment Effects, Journal of Management Information Systems, 32:2, 8-39, DOI: [10.1080/07421222.2015.1068604](https://doi.org/10.1080/07421222.2015.1068604)

To link to this article: <http://dx.doi.org/10.1080/07421222.2015.1068604>



Published online: 28 Aug 2015.



Submit your article to this journal [↗](#)



Article views: 128



View related articles [↗](#)



View Crossmark data [↗](#)

---

# Task Division for Team Success in Crowdsourcing Contests: Resource Allocation and Alignment Effects

INDIKA DISSANAYAKE, JIE ZHANG, AND BIN GU

INDIKA DISSANAYAKE is a Visiting Assistant Professor of Information Systems at the D'Amore-McKim School of Business at Northeastern University. She received her Ph.D. in Information Systems from the University of Texas at Arlington. Her research interests include crowdsourcing, social media, and virtual communities. Her research work has been presented in conferences such as International Conference on Information Systems (ICIS), Americas Conference on Information Systems (AMCIS), Decision Science Institute (DSI), and the Institute for Operations Research and the Management Sciences (INFORMS).

JIE (JENNIFER) ZHANG is an associate professor of information systems at the College of Business Administration, University of Texas at Arlington. She received her Ph.D. in computer information systems from the William E. Simon Graduate School of Business at the University of Rochester. She employs analytical and empirical techniques to examine a number of issues in electronic retail channels, online reputation and feedback systems, software pricing and licensing models, online consumer search and shopping decisions, website designs, social media, and crowdsourcing. Her research appears in *Information Systems Research*, *Journal of Management Information Systems*, *Journal of Economics and Management Strategies*, *Communications of the ACM*, and elsewhere. She serves as associate editor for *Decision Support Systems* and *Electronic Commerce Research*.

BIN GU is an associate professor of information systems at the W.P. Carey School of Business at Arizona State University. His research interests are in crowdsourcing and crowdfunding, online social media, user-generated content, IT business value, and IT governance. His work has appeared in *Management Science*, *Information Systems Research*, *Journal of Management Information Systems*, *MIS Quarterly*, *Journal of Retailing*, and *Decision Support Systems*. His research received the 2014 and 2012 Emerald Management Reviews Citations of Excellence Awards, the 2014 Americas

---

The authors are grateful to seminar participants at the University of Texas at Arlington, University of Florida, University of Texas at Dallas, and conference participants at AMCIS 2014 (Savannah), ICIS 2014 (Auckland), and HICSS 2015 (Kauai) for excellent feedback. They also benefited greatly from discussions with Mahmut Yasar, Radha Mahapatra, Greg Frazier, and Wendy Casper. The authors especially appreciate valuable comments from the editors Rob Kauffman, Eric Clemons, Thomas Weber, and the reviewers of the Special Section. Bin Gu acknowledges the financial support of the National Natural Science Foundation of China [Grant 71328102].

Conference on Information Systems Best Research-in-Progress Award, and the 2008 *Information Systems Research* Best Published Paper Award. He serves as senior editor for *MIS Quarterly* and as associate editor for *Information Systems Research*.

**ABSTRACT:** Advances in information technology bring changes to the nature of work by facilitating companies to go beyond the wisdom of their workforce and tap into the “wisdom of the crowd” via online crowdsourcing contests. In these contests, active and motivated individuals collaborate in the form of self-organized teams that compete for rewards. Using a rich data set of 732 teams in 52 contests collected from the crowdsourcing platform, Kaggle.com, from its launch in April 2010 to July 2012, we studied how the allocation of members’ social and intellectual capital within a virtual team affects team performance in online crowdsourcing contests. Our econometric analysis uses a rank-ordered logistic regression model, and suggests that the effect of a member’s social and intellectual capital on team performance varies depending on his or her roles. Though a team leader’s social capital and a team expert’s intellectual capital significantly influence team performance, a team leader’s intellectual capital and a team expert’s social capital do not. Further, we found that the alignment of a member’s social and intellectual capital within a team has a significant influence on team performance. Moreover, the intensity of the competition moderates the impact. When a contest is highly competitive, the social and intellectual capital alignment negatively affects team performance, and when the competitive intensity is low, this alignment positively affects team performance. Our findings provide insights into improving performance in team-based competitions in crowdsourcing communities.

**KEY WORDS: AND PHRASES:** crowdsourcing, crowdsourcing contests, econometrics, intellectual capital, social capital, social network analysis, team competition, virtual teams.

---

Information technology (IT) has enabled socially connected “crowds” to become partners of choice to find creative solutions for a variety of problems. Individuals in these crowds are often interacting and collaborating with people of similar interest through online communities. Organizations leverage some of these online communities, such as crowdsourcing platforms, to access expertise of the general public around the globe to seek solutions [38].

*Crowdsourcing* means “to publicly invite a large group of people to take a task that was traditionally performed by an employee or contractor in the form of an open call” [31, p. 1]. Online platforms, such as Kaggle.com for predictive modeling projects, IdeaStorm.com for idea generation, Threadless.com for product design, InnoCentive.com for research and development (R&D), and TopCoder.com for software development projects, often facilitate crowdsourcing. Evidence of the value of using online crowdsourcing communities to solicit the “wisdom of the crowd” in business practices abounds. For instance, about 51,000 people from 86 countries took the Netflix \$1M challenge, and the winning solution improved the accuracy of their movie recommendation algorithm by 10 percent [62]. Also, more than 57,000 online gamers, most of whom did not have

a molecular biology background, contributed to the identification of the structure of a particular protein within three weeks, even though the same problem has puzzled researchers at the University of Washington for years [54]. According to a survey conducted by the Marketing Executive Networking Group in 2009, 75 percent of company executives think that crowdsourcing is highly effective with respect to new product and service development [60]. Ten out of the eleven top global brands use crowdsourcing communities to find solutions for their business problems [18].

Given the importance of crowdsourcing communities, there is a growing interest in studying how to improve the performance of crowdsourcing contests. The extant literature has focused extensively on contest design [10, 40, 61, 65] or the effects of individual behavior on contest outcomes with the goal of maximizing the payoff [1, 6, 32, 46, 64]. These studies have demonstrated the advantages of crowdsourcing, such as lower cost [32], lower risk [10], higher quality solutions [24], and multiple alternative solutions [61]. All these studies have been conducted either at the individual or the contest level.

With the popularity of virtual teams in crowdsourcing content, there is an increasing interest in understanding team performance as opposed to individuals. In online crowdsourcing, members usually form groups or teams of motivated individuals looking to achieve common goals. Compared with the general workplace settings, crowdsourcing contests tend to involve greater complexity, a higher degree of innovation, more intense competitive pressure, and stricter time constraints. Therefore, crowdsourcing encourages teamwork. Web and social technologies bring changes not only to the ways a project workforce is sourced but also to the ways that a workforce is organized and coordinated. Digital media also allow individuals located in any geographic area to form self-organized “virtual teams” and coordinate with team members to work collectively to win the contest. Such virtual team members can share information, exchange ideas, brainstorm and negotiate alternative solutions, and make decisions together. Anecdotal evidence indicates that teams fare better in crowdsourcing contests. For example, the leading performers in the Netflix contest mentioned earlier were all teams, not individuals, and the size of the winning team increased over time from three in 2007 [48] to seven in 2009 [47].

It is well-known that team performance is not merely an aggregation of individual performance [16, 36]. Teams can generate positive synergy and outperform individuals when solving more difficult problems [16, 63]. To the best of our knowledge, no prior studies have considered teams and their performance in crowdsourcing, and the role of participants’ social capital (as “advantage created by a person’s location in a structure of relationships” [12, p. 5]) is also underinvestigated in the crowdsourcing literature (Table 1, presented later in the paper). In addition, crowdsourcing differs from most online communities (e.g., open source communities) in that crowdsourcing relies on competition among teams to engender the best solution [55], whereas teams in online communities are collaborative in nature. Thus, the existing online community literature does not apply directly to crowdsourcing either.

To fill in these gaps, we inspect the social network structure within self-organized virtual teams that compete in online crowdsourcing contests involving rewards.

We collected data for 732 teams participating in 52 contests on the crowdsourcing platform Kaggle.com, from its launch in April 2010 to July 2012. Using a rank-ordered logistic regression model, we first empirically examined how the *social capital* (SC) and *intellectual capital* (IC) of the team leader and team expert affect team performance. Intellectual capital refers to task-related skills, whereas social capital refers to their social connections within the team. In fact, our econometric analysis shows that social capital is more important for team leaders, whereas intellectual capital is more important for team experts. Next we adopted the concept of SC-IC alignment (i.e., the coefficient of correlation of SC and IC) [36] to further explore how the allocation of intellectual capital and social capital among team members affects team performance in crowdsourcing contests. We found that the alignment between members' *social capital and intellectual capital* (*SI alignment*) negatively affects team performance. This effect, however, is positive in a less competitive environment. That is, in highly competitive contests, teams perform better if the members with higher intellectual capital are not at the center of the group's social network. This finding supports the argument for giving team members different roles based on skill sets [21]. It suggests that under intense competition teams should have members with high intellectual capital concentrate on innovative work and allow members with high social capital to take charge of ensuring effective communication within the team. This provides new evidence to inform our understanding of the ideal team structure for crowdsourcing competitions. The results are robust to alternative measures of social capital and intellectual capital.

This article provides unique contributions to theory and practice. First, while all the prior studies in crowdsourcing context have been conducted either at an individual or contest level, to our knowledge, this study is the first that considers teams and their performance in crowdsourcing. As teams increasingly become key participants in crowdsourcing contests, our results provide key insights to crowdsourcing participants. Our data sample collected from the popular crowdsourcing website, Kaggle.com, suggests that a self-organized virtual team is a common way of participating in crowdsourcing projects: 16 percent (12 percent) of the teams have members from different countries (continents). Our sample also showed that teams have 6 percent higher odds of winning than individual participants (1 percent). Thus, individuals may wish to participate in a team in order to maximize the likelihood of winning in crowdsourcing.

Second, this article adopts a social network positioning to further explore the composition of team members with different social capital, while prior literature in crowdsourcing has mainly investigated the implications of solvers' intellectual capital (Table 1). Third, it also contributes to the growing literature on online communities by investigating value creation of a special online community with competing teams, whereas the online communities explored in prior literature have been mostly collaborative in nature.

Last, this study presents new empirical evidence on the effect of allocation of team resource on team performance in crowdsourcing contests, an important but hitherto

neglected topic in crowdsourcing research. It heeds the call for more research on resource alignment within teams in a broader context [36]. Specifically, we extended the Kane–Borgatti model of social capital and intellectual capital alignment to crowdsourcing communities by considering the moderation impact of environmental competition. We found that the alignment has a negative effect in highly competitive contests and a positive one in less competitive contests. Our result extended earlier studies suggesting that alignment has a positive effect in noncompetitive environments. The findings provide new empirical evidence on the theory of division of labor in the online crowdsourcing environment with competitive virtual teams [21] and insights on team management. That is, in a competitive setting, with everything else equal, a team with negative alignment of intellectual capital and social capital among its members will perform better than a positively aligned one.

## Literature Review

---

### Crowdsourcing Contests

In general, three parties are involved in a typical crowdsourcing contest. *Seekers* are typically companies looking for solutions; *solvers* refer to the open crowd or the members of a crowdsourcing community who provide the solutions, and *platform providers* offer an interface that allows interaction between seekers and solvers. As a new business model, crowdsourcing has the potential to transform work for organizations. Brabham [11, p. 76] claimed that crowdsourcing is a “problem-solving model . . . that can have profound influence on the way we solve our world’s most pressing social and environmental problems.” This Web-based business model allows companies to work with individuals outside geographical and organizational boundaries. Each crowdsourcing community focuses on a certain type of problem and solves them mainly through contests. Seekers request solutions to platform providers who then set up contests for solvers or crowdsourcing community members. As an incentive, winners earn some form of reward from the seeker. Community members compete among themselves individually or as teams. In some instances, companies directly post their problems to open crowds without going through platform providers [66].

Table 1 summarizes the literature on crowdsourcing. Previous studies have investigated the impact of contest and contestant characteristics on the outcomes of crowdsourcing contests, including the quality of the output [10], the solver’s project completion rate [65], the solver’s probability of winning the contest [46, 64], and the number of solvers participating in a contest [65]. Empirical evidence indicates that factors such as solver skill [1], solver effort [46], contest reward structure and amount [1], the total number of solvers [1, 10], and the solvers’ skill distribution [10] influence performance and the chances of winning. Furthermore, contest characteristics, such as the reward, duration, and complexity of the project, influence both the number of solvers who participate and the completion rates [65].

Table 1. Empirical Literature on Crowdsourcing/Open Contests

Literature	Social network perspective	Key research findings: IV → DV relationships	Data source	Unit of measure
Yang et al. [65]	No	reward, duration (+) → # subscribers time cost, description length (-) → # subscribers reward, Duration (-, +) → completion rate time cost (-) → completion rate	TaskCN.com	Contest level
Archak [1]	No	reward, contestant rating } (+) → project score experience, #opponents } max. opponent rating # requirements (-) → project score submission time (U-shape) → prob. of winning # prior winnings (+) → prob. of winning # contestants (-) → average score experience, submissions (+,-) → score degree centrality (+) → prob. of winning (friend, rival and co-opetitor network) SI alignment (-/+ ) → team performance Moderated by competitive intensity	TopCoder.com	Individual level
Yang et al. [64]	No		TaskCN.com	Individual level
Boudreau et al. [10]	No		TopCoder.com	Contest level
Huang et al. [32]	No		Threadless.com	Individual level
Mo et al. [46]	Yes		Zhubajie.com	Individual level
Our paper	Yes		Kaggle.com	Team level

To our knowledge, the crowdsourcing literature does not include any studies on teams, even though the latter represents a common way to participate in crowdsourcing community contests. A rare exception is Mo et al. [46], who examined the effect of social interactions in a crowdsourcing contest on a solver's chance of winning; however, they still focused on individual participants. Our work complements the literature in that we examined teams and the effect of resource allocation within a virtual team on its performance.

## IT and Virtual Team Performance

There is growing evidence that recent advances in IT dramatically change the nature of work within organizations. It is no longer necessary for employees to work in the same physical location. Digitally mediated collaboration techniques facilitate pooling expertise within and/or beyond the organization by eliminating geographical barriers [22]. Rapid developments in IT make it easier, faster, and more efficient to distribute work across geographic boundaries [30]. As a result, virtual teams have become popular among organizations. Johnson et al. [34, p. 29] noted that “we have moved away from working with people who are in our visual proximity to working with people around the globe.” Other studies have shown that using virtual teamwork to replace some co-located work can bring organizations several benefits, including enhanced responsiveness and flexibility [50], reduction of relocation time and cost [7], and better results through attracting the best individuals [43].

Despite these benefits, virtual teams also face substantial challenges due to potential miscommunication, lack of trust, and attrition. Virtual teams may delay or misinterpret communication due to geographic separation, which may exacerbate conflict, decrease trust and cohesion [33], drive up costs, and reduce project quality [14]. To mitigate such negative factors, virtual teams often involve more knowledge sharing [19] and require a central coordinator or leader [49].

Virtual or co-located general organizational teams are noncompetitive or collaborative in nature. Ebrahim et al. [22, p. 2654] defined *virtual teams* as “groups of geographically, organizationally, and/or time dispersed workers brought together by information technologies to accomplish one or more organization tasks.” In addition to these general virtual team characteristics, crowdsourcing teams are self-organized and competitive in nature. Specifically, they use the Internet to form teams and distribute their work.

In terms of studies, some have investigated the behavior of voluntarily organized virtual teams in the context of open source development communities. The literature has focused on the impact of team characteristics, such as prior collaboration ties [27] and diversity [20], on the decision to join a team. In this study, we focus on the impact of intellectual and social capital on team performance in virtual teams participating in crowdsourcing contests. Unlike teams in open source projects, they compete against one another. To the best of our knowledge, the effect of allocation of team members' intellectual and social capital on team performance



has not yet been investigated in crowdsourcing communities. This study takes the first step in doing so.

## A Social Network Perspective of Teams

Scholars have applied social network theory to investigate team behavior in many different contexts. Chen and Lim [13] developed a behavioral economics model that assessed whether contestants are averse to being responsible for the team's loss. A team-based contest can yield better effort than an individual-based contest. Singh and Tan [56] investigated network formation in open source software (OSS) teams to characterize stable and efficient structures. Several inefficient stable structures may exist, and an efficient stable structure may not always exist. Magni et al. [41] studied the influence of team network structure influences on an individual's technology use and reported that internal closure has a U-shaped effect. Balkundi and Harrison [4] conducted a meta-study to investigate the causal effect of a network structure on team performance in a general setting. They found that teams with densely configured interpersonal ties have higher task performance, and teams with leaders who are central in the teams' intragroup networks and teams that are central in their intergroup network tend to perform better. Sarker et al. [53] argued that a member's centrality in trust and communication networks enhances his or her performance. Hahn et al. [27] applied the social network perspective to study why and how past collaborative ties among developers affect their choice of new projects in OSS.

Kane and Borgatti [36] made an important contribution to the literature by integrating structural and resource perspectives on team networks. They showed that within a health-care organization, a group always performs better when the more proficient members are highly centralized in the communication and workflow network. This article extends Kane and Borgatti's [36] research on alignment between skill and network positions to a competitive virtual team setting that demands both innovative work and extensive communication within the team. We find more extensive results and explain how the competitive intensity moderates the relationship between the SI alignment and team performance.

## Research Model

---

### Team's Intellectual Capital and Social Capital

Team members bring two types of capital to their teams: social capital and intellectual capital. *Intellectual capital* refers to their task-related skills and knowledge gained from experience, learning and education [37], whereas *social capital* refers to the "advantage created by a person's location in a structure of relationships" [12, p. 5]. Following the literature, we operationalized these relations as ties that a team

member has with other members within a team [58]. These two kinds of capital are known to influence the collective's actions and its effectiveness [37].

The literature on team performance identified a team's intellectual capital or skills as an essential ingredient to team effectiveness and performance. Specifically, in software development teams, members with high intellectual capital play critical roles in the success of information systems projects [26]. Studies conducted in the context of open contests have shown that solvers' intellectual capital is a significant factor that determines their performance [1, 61] and their probability of winning in crowdsourcing competitions [10, 46, 64]. Salas et al. [52] indicated that effective utilization of members' task-related expertise contributes to high outcomes in expert teams.

Besides intellectual capital, prior research [2, 59] has also emphasized team members' social skills, that is, their ability to work in teams effectively. In particular, researchers have addressed the impact of the structural position of members within a network or social capital [9] on team performance. Sarker et al. [53] indicated that a member's centrality in trust and communication networks enhances member performance. Cross and Cummings [17] found evidence that betweenness centrality in information and awareness networks influences an individual's performance in knowledge-intensive work. Baldwin et al. [3] and Reagans and Zuckerman [51] have shown that more ties increase team performance. Mehra et al. [45] indicated that density of social network ties within a team enhance team performance.

Our context deals with predictive analytics tasks that are very intellectual in nature. We based intellectual capital on a solver's analytics skills and experience in analytics projects and tools. Members with high intellectual capital provide the necessary knowledge to solve the problems to which a seeker wants the solution in a crowdsourcing contest. Next, we based social capital on past collaboration ties with members within the team, as these ties indicate the likelihood that the team members can work well with one another. The model with the overall social ties within the entire community is tested later. Members with high social capital provide the necessary coordination within the team to allocate work and facilitate communications and information sharing. Thus, we propose:

***Hypothesis 1a (The Team Intellectual Capital Hypothesis):*** *A team's intellectual capital positively influences team performance.*

***Hypothesis 1b (The Team Social Capital Hypothesis):*** *A team's social capital positively influences team performance.*

## Leader's and Expert's Intellectual and Social Capital

As previously discussed, studies have shown that a team's intellectual and social capital positively influences team performance. This raises the question of whether each is equally important for all the roles within a team. To further explore this question we considered two extreme roles: leader and expert.

Kanawattanachai and Yoo [35] indicated that awareness of the expertise location influences team performance. More broadly, social connections allow team members to better coordinate and match tasks with expertise within a team. This coordination and matching task falls primarily on the team leader who serves as the center of communication and coordination. The literature reports that social network ties of team leaders have a significant and positive impact on team performance in offline settings [4, 44]. We expect it to hold for online crowdsourcing teams as well. Moreover, Mehra et al. [45] showed that a team leader's centrality within a team's social network is positively related to team performance.

In this context, the team leader forms a team, serves as a point of contact to the platform provider, and leads communication and coordination within the team. Thus, the leader's social capital proves valuable in forming a strong team in the first place. Second, it is important to allocate tasks based on member expertise and effectively manage team communication and coordination. Furthermore, our study looked at very competitive and time-sensitive contests. Thus, members often have a limited capacity to do multiple tasks. Even though both social capital and intellectual capital are important for team performance, we argue that as the center of the communication, the leader's social capital has more of an impact on team performance than his/her intellectual capital.

On the other hand, experts or team members with high intellectual capital provide the necessary technical knowledge to solve the task at hand. It is important to have them allocate the limited time they have to solving technical aspects of the problem. Hence, we argue that the intellectual capital of the team expert has a higher impact on team performance than his or her social capital. Thus, we propose:

***Hypothesis 2a (The Leader Social Capital Hypothesis):*** *A team leader's social capital matters more to team performance than intellectual capital.*

***Hypothesis 2b (The Expert Intellectual Capital Hypothesis):*** *A team expert's intellectual capital matters more to team performance than social capital.*

## Social-Intellectual (SI) Alignment

The team intellectual capital and social capital may be more than simple average of the individuals' social and intellectual capital combined. Besides the team leader's social capital and the team expert's intellectual capital, another important factor is the alignment between the two types of capital within a team, which is measured by the correlation coefficient of the two kinds of capital, ranging from negative to positive alignment [36]. A positive alignment suggests that members with a high level of knowledge or intellectual capital also have more network connections or social capital in the virtual team, while a negative alignment indicates separation of labor, such that members with high intellectual capital do not necessarily have high social capital, and vice versa. Kane and Borgatti [36] noted that a positive alignment

may facilitate learning transfer and exert peer influence in standalone teams, thereby improving team performance.

We applied the theory of the division of labor from Adam Smith, who identified several advantages of division of labor including reducing switching costs and improving performance through familiarization and learning (Smith 1937). The same principle has been applied to teams, regardless of their size, for example, Häußler and Sauermann (2014), and Lin [39]. In our context, while both social capital and intellectual capital are important to team performance, a member often has a limited capability to perform well both socially and intellectually in a competitive environment. The experts or members with high task-related skills bring necessary technical knowledge required for developing efficient algorithms that help in effectively analyzing and classifying large amounts of data. However, group communication and coordination can take a significant amount of time and distract members with high intellectual capital from focusing on solving the problem. For the same reason, members with high social capital can play a better role as facilitators of team coordination and communication. They can use their social connections and knowledge about members to recruit better individuals, allocate tasks based on member expertise, effectively handle communication and coordination within the group as well as between group and platform provider, and consolidate work in a timely manner.

Therefore, we argue that in a competitive environment, like crowdsourcing contests, a division of labor that allows members with high intellectual capital to focus on activities directly related to the task and members with high social capital to handle coordination and communication among team members positively influences team performance. The effect is likely to be intensified in a more competitive environment. Thus, we propose:

***Hypothesis 3 (The Team SI Alignment Hypothesis):*** *Negative alignment between social capital and intellectual capital (SI alignment) positively influences team performance in crowdsourcing contests.*

***Hypothesis 4 (The Competitive Intensity Hypothesis):*** *Competitive intensity moderates the relationship between SI alignment and team performance, such that the influence of negative alignment becomes stronger as the competitive intensity increases.*

Figure 1 summarizes the research model. We have not included “The Leader Social Capital (H2a) Hypothesis” and “The Expert Intellectual Capital (H2b) Hypothesis” because these are individual-level measures.

## Data Collection and Variable Definitions

---

### Data Collection

We used web crawlers to obtain data from a specialized crowdsourcing community platform that focuses on data analytics projects: Kaggle.com (see Figure 2). This

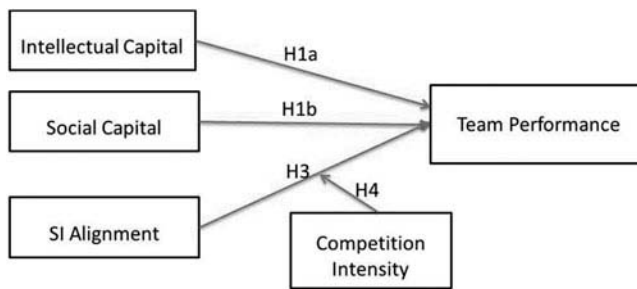


Figure 1. Research Model

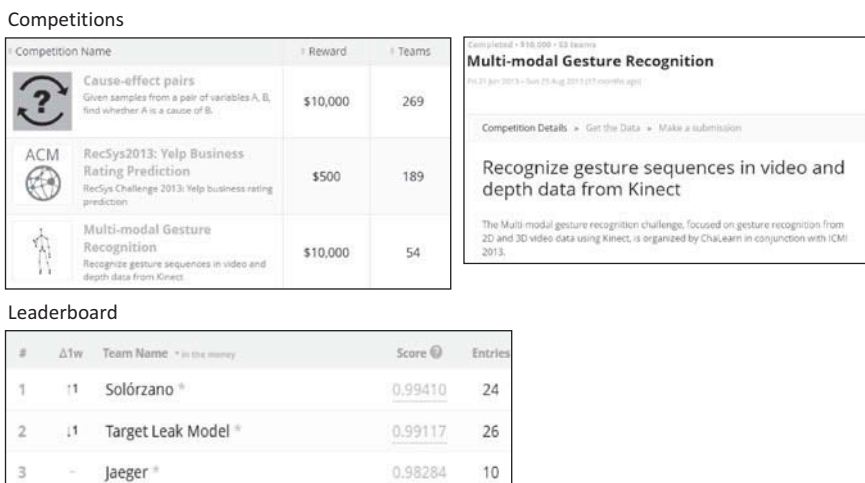


Figure 2. Screenshots of the Data Source: Kaggle.com

community consists of more than 100,000 data scientists from over 100 countries and 200 universities. They are experts in various quantitative fields, such as computer science, statistics, economics, math, and physics. Over the past few years, Kaggle has served many companies, including General Electric, Allstate, Merck, Ford, and Facebook, to improve sales forecasting, increase customer retention, reduce operating costs, accelerate product development, and gather information from social media.

Companies, government, and researchers provide data sets to Kaggle along with their problems and the amount of reward they are willing to pay the winners. In our data set, the reward amounts range from \$100 to \$100,000, with an average of \$16,050. Based on the requests, Kaggle sets up contests for the crowdsourcing community. Each participant or participating team can submit multiple solutions before the contest deadline. In our data set, contest durations range from 1 to 120 days, with an average of 74 days. Kaggle evaluates all submissions in real time using

Table 2. Variable Definitions

	Variable	Definition
Contest characteristics	Duration	Duration of the contest
	Award	Monetary reward to which the winning contestant(s) are entitled
Teams in each contest	# of participating teams	Total number of teams/individuals participating the contest
	Rank of each team	Real-time relative rank of each team in a given contest
	# of submissions	Number of submissions made to date by each team in a given contest
	Team Size	Number of participants in a team
Contestants profile	Team Leader	Leader of the team
	Team Expert	Member with the highest profile score
	Profile Score	Individual's profile score. This is a cumulative score based on the individual's relative performance in all the contests that he/she has competed.

a test data set and provides instant feedback, which includes information on the prediction accuracy of their model and their relative positions (e.g., rank) in the contest. The accuracy score and ranking represent unique and objective measures of the project quality, which is unavailable in most other crowdsourcing initiatives or team performance studies.

Teams that participate are self-selected teams. The team leader initiates the team's formation and is the only person who has authority to add new people to the team. Also, the team leader is the primary contact for Kaggle. The team leader usually has a higher number of prior connections with the team members.

For this study, we collected data on all of Kaggle's public contests since the launch of the community platform in April 2010 through July 2012. After eliminating contests without monetary rewards, teams with a single member, and an outlier with extremely high monetary reward, our final sample consists of 732 teams that participated in 52 contests. Our data set suggests that the chance of winning as a team is 4.8 times higher than as an individual after accounting for participation ratios. The participants' online profiles suggest that team members are geographically dispersed. Team members come from different countries or sometimes even different continents, which demonstrates that IT supports virtual team collaborations. The variables are described in [Table 2](#) and explained in the following subsections.

## Dependent and Independent Variables

*Team Rank.* We used team rank as a measure of team performance, the main dependent variable. The former refers to the relative ordinal position of a team in

a given contest. Because Kaggle provides real-time performance evaluation of solutions and ranks them based on their prediction accuracy, and a team may submit multiple solutions, we used their final ranking as the team’s performance measure.

Our primary independent variables are *intellectual capital* (IC), *social capital* (SC), and *social-intellectual* (SI) *alignment*. These team-level measures are based on individual-level intellectual capital and social capital.

*Individual’s Intellectual Capital.* We used individual skill scores from a person’s profile as a proxy for an individual’s intellectual capital. Kaggle uses a formula to calculate each individual’s skill score based on their performance in prior competitions. The maximum achievable score in a competition is derived from the total number of participants and the content level of difficulty. According to Kaggle, “the current formula for each competition splits the points among the team members, decays the points for lower finishes, adjusts for the number of teams that entered the competition, and linearly decays the points to 0 over a two-year period from the end of the competition.” Kaggle updates each individual’s skill scores after each competition. The individual skill scores in our data set ranged from 0 to 563,500, with a mean of 21,362.

*Individual’s Social Capital.* We operationalized social capital as ties that team members have to other members within the team. Following Hahn et al. [27], these ties are based on prior collaborations, that is, whether the members of a given team had worked in the same team in prior contests. We used two types of commonly used centrality measures based on ties to measure social capital: degree centrality and eigenvector centrality. The degree centrality ranges from 0 to 22 and eigenvector centrality ranges from 0 to 1.

*Degree Centrality.* In a network, *degree centrality* is defined as the number of ties incident upon a node [8, 23]. In this section, we consider only the connections within the internal network of a team in a given contest [27, 41]. In a later section, we extended the social capital to the connections within the overall Kaggle network. We defined an *individual’s degree centrality* as the number of prior ties he/she has had with other team members prior to the formation of the current team. If two members of a given team had collaborated in a virtual team in a previous contest, then they have a tie. A high degree centrality indicates a member has a large number of social connections (familiarity) with other team members based on past collaborations. Using subscripts  $i$  and  $j$  for team members,  $n$  for team size,  $t$  for contest, and  $tie_{ijt} = 1$  if members  $i$  and  $j$  have been in a team in any contest prior to  $t$  and 0 otherwise, the degree centrality of participant  $i$  in contest  $t$  is denoted as:

$$Degree\ Centrality_{it} = \sum_{j=-i}^n tie_{ijt}. \quad (1)$$

To investigate whether a member with more ties is the center of communication and coordination of the team’s network, we considered the leader role. Our data showed that the team leader has a significantly higher degree of centrality than the rest of the team members (see Table 3). Hence, team leaders are commonly

Table 3. Paired *T*-Test for Leader’s and Others’ Social Capital

	Mean	Std. error	Std. deviation	95% confidence interval	
Leader’s SC	.3078	.0413	1.1163	.2267	.3889
Others’ SC	.2941	.0401	1.0839	.2154	.3728
diff	.0137	.0066	.1783	.0008	.0266
mean(diff) = mean(Leader’s SC – Others’ SC)			$t = 2.0772$		
Ho: mean(diff) = 0			df = 730		
Ha: mean(diff) < 0		Ha: mean(diff) != 0		Ha: mean(diff) > 0	
Pr( $T < t$ ) = 0.9809		Pr( $ T  >  t $ ) = 0.0381		Pr( $T > t$ ) = 0.0191	

the ones who have collaborated previously with most of the other group members.

*Eigenvector Centrality.* We entered team internal social network data (adjacency matrices) based on past collaborative ties into UCInet, which was used to calculate an individual’s eigenvector centrality scores. Formally, *eigenvector centrality* is “the principal eigenvector of the adjacency matrix defining the network” [7, p. 62]. It assigns relative scores to all team members such that a member’s score is higher when he or she is connected to someone with higher degree centrality. In other words, the “node’s centrality is proportional to sum of centralities of its contacts” [36, p. 1067].

For a social network with an adjacency matrix  $A = (a_{v,t})$ , where  $a_{v,t} = 1$  if vertex  $v$  is linked to vertex  $t$ , and  $a_{v,t} = 0$  otherwise, the eigenvector centrality  $x$  is given by:

$$Ax = \lambda x. \quad (2)$$

where  $\lambda$  is the eigenvalue associated with eigenvector  $x$ .

The additional requirement that all the entries in the eigenvector be positive implies by the Perron–Frobenius theorem that only the greatest eigenvalue results in the desired centrality measure. The  $v$ th component of the related eigenvector gives the centrality score of the vertex  $v$  in the network.

Following the literature, we also accounted for average levels of intellectual capital in the group [4]. As we mentioned earlier, Kaggle maintains a skill score for each team member based on past performance. We calculated the *team intellectual capital* (IC) by taking the average of profile scores of all the members. Similarly, *team social capital* (SC) is calculated by taking the average of degree or eigenvector centralities of all the members in a team.

*Social-Intellectual (SI) Alignment.* Following Kane and Borgatti [36], we defined the *social-intellectual (SI) alignment* of a given team as the correlation coefficient between the members’ intellectual capital (skill) and their social capital (centrality) in the team social network. We used two types of centrality measures based on ties in our calculation of SI alignment: degree centrality and eigenvector centrality. This correlation measure has been used as an independent variable to measure performance in different contexts [36].



## Moderating and Control Variables

*Competitive Intensity.* We investigated the moderation impact of level competition by using the *Herfindahl Index* (HHI) to measure the level of competition (competitive intensity) in a contest. We adopted it from the marketing literature, where it has been commonly used to measure market competition. A higher HHI indicates a lower level of competitive intensity, and vice versa. In our context, HHI measures teams' intellectual capital in relation to the contest; we calculated it by taking the sum of the squares of weighted intellectual capital for all teams in the contest. Using subscript  $i$  for team and  $j$  for contest, the competitive intensity measure for contest  $j$  was calculated with the equation:

$$\text{Competition HHI}_j = \frac{\sum_{i=1}^n \text{TeamIC}_{ij}^2}{\left(\sum_{i=1}^n \text{TeamIC}_{ij}\right)^2}, \quad (3)$$

where  $n$  is the total number of teams, including single-member teams, participating in contest  $j$ .

To control for individual contest heterogeneity (e.g., contest rewards and duration), we included contest-specific fixed effects in the model. Based on the literature on crowdsourcing, we also controlled for team size and the number of submissions each team made (see [Table 1](#)).

*Submissions.* This is the number of solutions a team submitted in a given contest. Kaggle allows participants to submit multiple solutions throughout the duration of the contest. Prior studies [46] have shown that the number of submissions affects team performance. This figure ranges from 1 to 256, with a mean of 23.

*Team Size.* This is the number of members. Previous studies [15, 25] indicated that team size also affects team performance. It ranges from 2 to 40, with a mean of 2.9.

[Table 4](#) presents the descriptive statistics of the variables. To remove the scale effects, we took the natural log of all of the variables and added one. [Table 5](#) reports the correlation matrix of the log-transformed variables.

Table 4. Descriptive Statistics

Variable	Mean	Std. deviation	Min.	Max.
Team Rank	129.9590	166.7877	1	865
Team SC	0.2970	1.0866	0	22
Team IC	24,295.40	58,587.39	0	411,573
SI Alignment	0.0213	0.1560	-1	1
Submissions	22.9850	33.9122	1	256
Team Size	2.8538	2.2894	2	40

Table 5. Correlation Matrix

		1	2	3	4	5	6
1	Team Rank	1.0000					
2	Team SC	-0.1557	1.000				
3	Team IC	-0.1540	0.2617	1.0000			
4	SI Alignment	-0.0152	0.1723	0.0539	1.0000		
5	Submissions	-0.3529	0.0768	0.3960	0.0154	1.0000	
6	Team Size	-0.0255	0.1831	-0.0377	0.2154	0.1401	1.0000

## Results

To address the measurement scale issue, we used rank-ordered logistic regression [28, 42] to test our models and ran them on Stata 11.2. Since a low rank represents a high performance order, we used the *reverse preferential order* option. We also controlled for a contest’s specific characteristics (e.g., reward, duration) using contest-fixed effects. First, “The Intellectual Capital Hypothesis (H1a)” and “The Social Capital Hypothesis (H1b)” are validated. Second, “The Leader Social Capital Hypothesis (H2a)” and “The Expert Intellectual Capital Hypothesis (H2b)” are tested. Last, “The SI Alignment Hypothesis (H3)” and “The Competitive Intensity Hypothesis (H4)” are validated.

### The Impact of a Team’s Social Capital and Intellectual Capital on Team Performance

In the first-stage analysis, we investigated how a team’s social capital and intellectual capital influenced team performance. When addressing group capabilities, the common method is to use average or summed abilities of individual members [5]. Hence, we used team members’ average social and intellectual capital.

$$\begin{aligned} \logit(\text{TeamPerformance}_{ij}) = & \alpha_0 + \alpha_1 \text{TeamIC}_{ij} + \alpha_2 \text{TeamSC}_{ij} \\ & + \alpha_3 \text{Submissions}_{ij} + \alpha_4 \text{TeamSize}_{ij} + \delta_j + \varepsilon_{ij} \end{aligned} \quad (4)$$

where  $\alpha_k$  ( $k = 0 \dots 4$ ) represents the coefficients of the variables,  $\delta_j$  is the coefficient for the contest fixed effect, subscript  $i$  is for the team, and subscript  $j$  is for the contest.

Table 6 summarizes the results of the first-stage analysis. Models 1 and 2 test the direct impact of the team’s intellectual capital and social capital on team performance. Supporting Hypotheses 1a and 1b, both intellectual capital ( $\alpha_1 = 0.22$ ,  $p < 0.01$ ) and social capital ( $\alpha_2 = 0.27$ ,  $p < 0.05$ ) show positive and significant influence on team performance in Models 1 and 2, respectively. Our results suggest that both are important to good team performance, and thereby support “The Intellectual Capital Hypothesis (H1a)” and “The Social Capital Hypothesis (H1b).”

Table 6. Rank-ordered Logistic Regression Results (Team IC and SC)

Reversed Team Rank	Model 1	Model 2
Team IC	0.2251***	
Team SC		0.2709**
Submissions	0.8470***	0.9471***
Team Size	0.2924*	-0.0017
Log likelihood	-1,314.02	-1,351.11
Sample size	732	732

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .

Note: All independent and control variables are in natural log form.

### The Impact of a Team Leader's SC and IC, and an Expert's SC and IC on Team Performance

The team leader is responsible for communication and coordination activities within the team. Hence, we investigated how the leader's social capital and intellectual capital affect team performance in the context of crowdsourcing contests. As the center of the communication, the leader must have good social skills and become familiar with team members in order to communicate effectively and allocate tasks and resources based on the skills and capabilities of the team members. We derived a complete model (5) based on model (4):

$$\text{logit}(\text{TeamPerformance}_{ij}) = \alpha_0 + \alpha_1 \text{LeaderIC}_{ij} + \alpha_2 \text{LeaderSC}_{ij} + \alpha_3 \text{Submissions}_{ij} + \alpha_4 \text{TeamSize}_{ij} + \delta_j + \varepsilon_{ij}, \quad (5)$$

where  $\alpha_k$  ( $k = 0 \dots 4$ ) represents the coefficients of the variables,  $\delta_j$  is the coefficient for the contest fixed effect, subscript  $i$  is for the team, and subscript  $j$  is for the contest.

The results in Table 7, Model 1 show that the leader's social capital ( $\alpha_2 = 0.29$ ,  $p < 0.05$ ) has a positive and significant influence on team performance, while the leader's intellectual capital does not. Hence, "The Leader Social Capital Hypothesis (H2a)" is supported. One possible explanation is that the leader has a significantly higher number of ties when compared to other members in the team. Thus, the social capital brought by the leader helps in the formation of a better team through recruiting individuals with the necessary skills, the delegation of tasks based on member expertise, and the effective management of coordination.

Similarly, we investigated the effect of the expert's social capital and intellectual capital on team performance. As shown in Table 7, Model 2, an expert's intellectual capital has a positive and significant influence on team performance ( $\alpha_1 = 0.19$ ,  $p < 0.01$ ), while the expert's social capital proves insignificant in terms of performance. Hence, "The Expert Intellectual Capital Hypothesis (H2b)" is supported. Thus, the results suggest that although both kinds of capital are important

Table 7. Rank-ordered Logistic Regression Results (Leader and Expert)

Reversed Team Rank	Model 1	Model 2	Model 3a	Model 3b
	Leader	Expert		
Leader IC	0.0072		-0.0041	0.0360
Leader SC	0.2929**		1.0114**	0.8237*
Expert IC		0.1910***	0.1886***	0.2947***
Expert SC		-0.1000	-1.0601**	-0.8901*
Submissions	0.9458***	0.8592***	0.8580***	0.8321***
Team Size	-0.0096	0.1491	0.1388	0.1960
Log likelihood	-1,347.88	-1,318.71	-1,316.31	-961.52
Sample size	731	731	731	589

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .

Note: All independent and control variables are in natural log form. Model 3b did not consider those teams whose leaders are experts.

on average, they vary in different roles. As the results indicate, social capital is more important for leaders and intellectual capital is more important for experts. However, our current empirical data limit our capability to deepen our understanding of how the actual collaboration happens in the team.

In Table 7, Model 3a, we included both the leader's and the expert's social capital and intellectual capital, and assessed the impact on team performance. The findings demonstrate that the expert's intellectual capital has a significant positive impact on team performance ( $\alpha_1 = 0.19$ ,  $p < 0.01$ ), and the leader's social capital has a significant impact on team performance ( $\alpha_2 = 1.01$ ,  $p < 0.05$ ). Model 3b retested the model after removing experts who are also team leaders and produced consistent results.

### The Impact of SI Alignment on Team Performance

$$\begin{aligned} \text{logit}(\text{TeamPerformance}_{ij}) = & \alpha_0 + \alpha_1 SI_{ij} + \alpha_2 \text{TeamIC}_{ij} + \alpha_3 \text{TeamSC}_{ij} + \\ & \alpha_4 \text{Submissions}_{ij} + \alpha_5 \text{TeamSize}_{ij} + \delta_j + \varepsilon_{ij} \end{aligned} \quad (6)$$

where  $\alpha_k$  ( $k = 0 \dots 5$ ) represents the coefficients of the variables,  $\delta_j$  is the coefficient for the fixed effects, subscript  $i$  is for the team, and subscript  $j$  is for the contest.

Table 8 summarizes the results of the rank-ordered logistic regression of Equation 6. The effect of SI alignment based on degree centrality is given in Model 1, whereas the effects based on eigenvector centrality is given in Model 2. The effect of SI alignment is significant and negatively associated with team rank in both Model 1 ( $\alpha_1 = -0.61$ ,  $p < 0.05$ ) and Model 2 ( $\alpha_1 = -0.63$ ,  $p < 0.05$ ). Thus, "The SI Alignment Hypothesis (H3)" is supported. Our results suggest that teams perform poorly when high intellectual capital and high social capital were concentrated among a few members.

Table 8. Rank-ordered Logistic Regression Results (Team SI Alignment)

Reversed Team Rank	Degree Centrality	Eigenvector Centrality
	Model 1	Model 2
SI Alignment	-0.6123**	-0.6334**
Team SC	-0.1048	-0.2148
Team IC	0.2363***	0.2336***
Submissions	0.8495***	0.8508***
Team Size	0.4038**	0.2979
Log likelihood	-1,313.42	-1,313.58
Sample size	732	732

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .  
*Note:* Independent and control variables are in natural log form.

## The Moderation Impact of Competitive Intensity on Team Performance

We further investigated how competitive intensity affected the relationship between SI alignment and team performance by considering the moderating effect of the former on the latter.

Testing the model using different values of competition HHI, we noticed that when competition HHI increases, the coefficient of SI alignment changes from negative to positive. We sorted the data based on the HHI values and divided them into two equal number groups. We considered the group with lower competition HHI (i.e., higher competitiveness) as the base group. Then we introduced a dummy variable to represent a higher competition HHI (i.e., lower competitiveness) group.

$$\begin{aligned} \text{logit}(\text{TeamPerformance}_{ij}) = & \alpha_0 + \alpha_1 \text{SI}_{ij} + \alpha_2 \text{TeamIC}_{ij} + \alpha_3 \text{TeamSC}_{ij} \\ & + \alpha_4 \text{SI}_{ij} * \text{CompetitionHHI}_j + \alpha_5 \text{CompetitionHHI}_j + \\ & \alpha_6 \text{Submissions}_{ij} + \alpha_7 \text{TeamSize}_{ij} + \delta_j + \varepsilon_{ij}, \end{aligned} \quad (7)$$

where  $\alpha_k$  ( $k = 0 \dots 7$ ) represents the coefficients of the variables,  $\delta_j$  is the coefficient of the fixed effects, subscript  $i$  is for the team, and subscript  $j$  is for the contest.

The results are summarized in Table 9. The coefficient of SI alignment is negative and significant ( $\alpha_1 = -1.15$ ,  $p < 0.01$ ), which means negative alignment enhances performance in highly competitive environments. Moreover, the interaction effect of SI alignment and competition HHI is positive and significant as well ( $\alpha_4 = 1.18$ ,  $p < 0.05$ ). Thus, “The Competitive Intensity Hypothesis (H4)” is supported. The results suggest that the negative impact of SI alignment on performance will become weaker or may even become positive as the competitive intensity decreases. Hence, the impact of SI alignment on performance is different for competitive and

Table 9. Rank-ordered Logistic Regression Results (Competitive Intensity)

Reversed Team Rank	Competitive Intensity
SI Alignment	-1.1488***
SI * Competition HHI	1.1851**
Team IC	0.2282***
Team SC	-0.0893
Submissions	0.8523***
Team Size	0.3722**
Log likelihood	-1,311.56
Sample size	732

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .

Note: Independent and control variables are in natural log form.

noncompetitive environments. Since we used a rank-ordered logistic regression model *grouped by contests* option, we cannot estimate the coefficients for the contest-specific variables. Thus, we do not report the direct impact of competitive intensity in Table 9.

## Robustness Tests

In addition to the above main results, we also conducted various supplementary tests, including endogeneity tests of the social capital and intellectual capital with performance, and alternative measures of social capital and intellectual capital. All these checks consistently demonstrated that the results are robust.

## Endogeneity Tests

It is possible that the member's uncaptured individual characteristics play a role in their intellectual capital, thereby making the latter an endogenous variable due to omitted variables. Further, because we derived social capital based on past collaborations, and past performance can affect current and future collaborations, social capital is an endogenous variable as well. Hence, we used the Hausman test to determine whether both are, in fact, endogenous. Tables 10 and 11 report the endogeneity test results for team intellectual capital and team social capital, respectively.

To test whether team intellectual capital is endogenous, we used the adjacent opponent team's intellectual capital as an instrumental variable. The adjacent opponent team's intellectual capital has a significant and direct impact on the potentially endogenous variable for team intellectual capital ( $\alpha = 0.37$ ,  $p < 0.01$ ,  $t = 12.2 > 3.3$ ). Hence, the instrumental variable satisfies the relevant condition. Then we ran the first-stage regression and estimated the residuals of Model 1 in

Table 10. Hausman Test for Endogeneity of Team IC

DV	Model 1	Model 2	Model 3
	OLS	OLS	2SLS
	Team skill	Team Rank	Team Rank
Team IC		-5.4928**	-5.4928**
Opponent IC	0.3688***		
Submissions	0.6878***	-53.1103***	-53.1103***
Team Size	-1.0883 ***	17.6934	17.6934
<i>vhat</i>		2.722	

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .

Notes: All independent and control variables are in natural log form. Opponent IC is the instrumental variable for team IC. The residuals obtained from Model 1 is *vhat*; *vhat* is insignificant in Model 2. In the absence of evidence to the contrary we assume that team IC is exogenous. We also control for contest-specific effects.

Table 11. Hausman Test for Endogeneity of SC

DV	Model 1	Model 2	Model 3
	OLS	OLS	2SLS
	Ties	Team Rank	Team Rank
Team SC		-124.2429**	-124.2429**
Thanks Received	0.0674***		
Submissions	-0.0123	-59.6173***	-59.6173***
Team Size	0.2551***	35.0403*	35.0403*
<i>vhat</i>		93.7572	

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .

Notes: All independent and control variables are in natural log form. Number of thanks received is the instrumental variable for ties. The residual values obtained from Model 1 are values of *vhat*; *vhat* is insignificant in Model 2 though. In the absence of evidence to the contrary, we assume that the variable ties are exogenous. We also control for contest-specific effects.

Table 10 (*vhat*), included Model 2, and found that the coefficient of *vhat* is insignificant ( $p > 0.10$ ). Therefore, in the absence of evidence to the contrary, we assume that team intellectual capital is exogenous. Further, in Model 3 in Table 10, we report the coefficients of two-stage least squares regression (2SLS) using the opponent team's intellectual capital as the instrumental variable for team intellectual capital. In Model 3, 2SLS produces the same coefficient estimates as the OLS coefficient estimates in Model 2, thereby confirming that we properly performed regression testing for endogeneity. Since we have one instrumental variable for team intellectual capital, the model is exactly identified. Thus, we

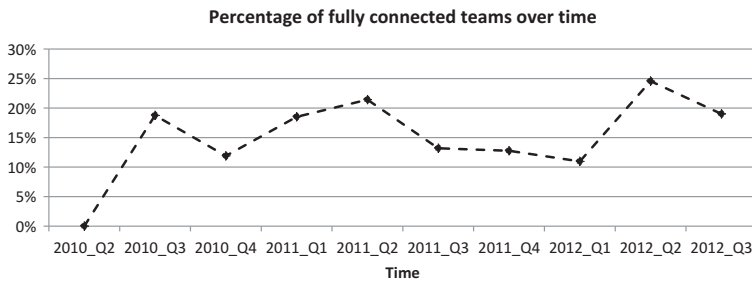


Figure 3. Plot of Team Composition over Time

cannot conduct an overidentification test to validate the exogeneity assumption of the instrumental variable.

In Kaggle, each contest has a forum where solvers share information and knowledge. Other solvers can appreciate the post by sending “thanks” to the contributor. In this context, “thanks” is similar to “likes” in other online social media platforms. It is an indication of social connections. Consequently, we used the total number of thanks received as an instrumental variable for social capital to conduct the Hausman test. Thanks received has a significant direct impact on potential endogenous variable ties ( $\alpha = 0.07$ ,  $p < 0.01$ ,  $t = 5.6 > 3.3$ ) and satisfies the relevant condition. Following the same procedure as above, we took the residuals of Model 1 in Table 11 (*vhat*), included Model 2, and found that the coefficient of *vhat* is insignificant ( $p > 0.1$ ). This means we did not find evidence indicating a violation of an OLS assumption by assuming social capital is exogenous. Model 3 also reports the 2SLS results with thanks received as an instrumental variable for team social capital. As we have only one instrumental variable for team social capital, the model is exactly identified once again. Thus, once again, we cannot conduct an overidentification test to check the efficiency of the instrumental variable.

We also plotted the fully connected teams over time to check whether the majority of teams connected to the same people or team composition remains sticky across contests. As shown in Figure 3, this does not seem to be an issue based on our current set of data.

#### Alternative Measures of Individual Intellectual Capital

As Kaggle’s formula for intellectual capital is based on a contestant’s prior performance, the measure may underestimate the intellectual capital of new contestants. Thus, we developed an alternative measure to allow for a robustness check. Based on a regression model with relatively mature solvers, and assuming that their profile scores truthfully reflect their intellectual capital, we first estimated individual intellectual capital as an expression of their profile, for example, experiences or number of projects completed, average ranking in past projects, tools, and the average of an adjacent opponent’s intellectual capital, with:



Table 12. Rank-ordered Logistic Regression Results (Alternative Measure of IC)

Reversed Team Rank	Degree Centrality
	Model 1
SI Alignment	-1.5670***
SI * Competition HHI	2.1613***
Team SC	-0.0821
Team IC	0.2057***
Submissions	0.9031***
Team Size	0.1522
Log likelihood	-1,296.64
Sample size	732

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .  
*Note:* All independent and control variables are in natural log form.

$$IC_i = \alpha_0 + \alpha_1 \#Contests_i + \alpha_2 AvgRank_i + \alpha_3 \#Tools_i + \alpha_4 AvgOpponentScore_i. \quad (10)$$

Then we retested our main model. The results are shown in Table 12. As expected, SI alignment has a negative and significant impact ( $\alpha = -1.57$ ,  $p < 0.01$ ) on team rank, whereas the interaction effect between SI alignment and competitive intensity has a positive and significant impact ( $\alpha = 2.16$ ,  $p < 0.01$ ) on team rank. Table 13 summarizes the results for teams with more than two members. The coefficient of SI alignment is negative and significant ( $\alpha = -1.63$ ,  $p < 0.01$ ) and the coefficient of the interaction effect of SI alignment and competition HHI is positive and significant ( $\alpha$

Table 13. Rank-ordered Logistic Regression Results (Alternative Measure of IC, Team Size&gt;2)

Reversed Team Rank	Degree Centrality
	Model 1
SI Alignment	-1.6283***
SI * Competition HHI	1.8889***
Team SC	0.0266
Team IC	0.2268***
Submissions	0.8606***
Team Size	0.4761
Log likelihood	-306.75
Sample size	258

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .  
*Note:* All independent and control variables are in natural log form.

= 1.89,  $p < 0.01$ ), which indicates that the results are robust when the team size is greater than 2.

### Alternative Measures of Individual Social Capital

As an alternative measure of individual social capital, we used weighted ties. We weighted them based on the number of times team members collaborated in previous contests. We reestimated the degree centrality and eigenvector centrality using these weighted ties. Table 14 summarizes the results of the rank-ordered logistics regression. Models 1 and 2 show the effect of SI alignment based on degree centrality and eigenvector centrality, respectively. The effect of SI alignment is significant and negatively associated with team rank in both Model 1 ( $\alpha = -0.90$ ,  $p < 0.01$ ) and Model 2 ( $\alpha = -0.92$ ,  $p < 0.01$ ), whereas the interaction effect of SI alignment and competition HHI is positively associated with team rank in both Model 1 ( $\alpha = 0.91$ ,  $p < 0.10$ ) and Model 2 ( $\alpha = 0.93$ ,  $p < 0.10$ ). The findings emphasize consistency when using an alternative measure of individual social capital.

Moreover, we investigated how the overall social capital (i.e., the collaboration ties of team members within the entire Kaggle platform) affects team performance. As shown in Table 15, results suggest that overall social capital positively influences team performance. However, the SI alignment based on the overall social capital of team members is not significantly related to team performance. This shows that internal social capital contributes more to communication and interactions within a team than overall social capital. Further, the interaction effect of competitive intensity and SI alignment based on overall social capital is not significant, either.

In addition to the aforementioned robustness tests, we tested the model with yearly and quarterly dummy variables separately to control for time effects and did not find evidence of the latter.

Table 14. Rank-ordered Logistic Regression Results (Weighted Ties)

Reversed Team Rank	Degree Centrality	Eigenvector Centrality
	Model 1	Model 2
SI Alignment	-0.8969***	-0.9162***
SI * Competition HHI	0.9140*	0.9347*
Team SC	0.0215	0.0353
Team IC	0.2231***	0.2181***
Submissions	0.8514***	0.8546***
Team Size	0.3546**	0.3256**
Log likelihood	-1,312.27	-1,311.93
Sample size	732	732

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .

Note: All independent and control variables are in natural log form.

Table 15. Rank-ordered Logistic Regression Results (Overall SC)

Reversed Team Rank	Degree Centrality
	Model 1
SI Alignment	-0.0852
SI * Competition HHI	-0.2901
Team SC	0.3523*
Team IC	0.2306***
Submissions	0.8377***
Team Size	-0.1264
Log likelihood	-1,313.17
Sample size	732

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$ .  
*Note:* All independent and control variables are in natural log form.

## Conclusion

### Key Findings

Given the rapid development of crowdsourcing technologies and practices, more data and evidence have become available to support closer scrutiny of design and managerial issues in crowdsourcing. This study takes an important step forward by examining how the stock of team members' social capital and intellectual capital affects team performance in crowdsourcing contests. We have some interesting and unique findings.

We started by examining how teams benefit from the intellectual capital and the social capital of members, and showed that teams benefit if their members have higher task-related skills or more connection ties with the other members.

In addition, we studied how the relationships involving social capital and intellectual capital with team performance differ with the roles of the members. The team leader and team expert are the extreme roles in the team with the greatest social capital and intellectual capital, respectively. On the one hand, we found that the team leader's social capital has more impact on team performance than the leader's intellectual capital, suggesting that a team performs better when coordination and communication roles are centralized to the member with the best connections with the other team members. On the other hand, our results indicate that experts' intellectual capital has a higher impact on team performance than does their social capital. In fact, their social capital has a weak, and even negative, impact on performance. This supports our argument that teams generally perform better when the experts are not centralized in the network in crowdsourcing contests.

To get a full picture of the research question, we extended the study to consider the roles in the entire team, that is, the alignment between intellectual capital and social

capital. We found that, after controlling for average levels of intellectual capital and social capital in the team, SI alignment has a negative effect on team performance. This implies that when everything else is the same, a team will perform better if the members who are more intellectually capable are not centralized in the network structure. Furthermore, our results also show that competition in these open contest environments drives the negative impact of SI alignment on performance. When competitive intensity is high, teams in which members with high intellectual capital focus on the main task fare better.

Our results are robust to various endogeneity tests, alternative measures, and alternative models. We used the closest-ranked neighbor's intellectual capital and the "thanks received" as instrument variables of team intellectual capital and social capital, respectively, and conducted endogeneity tests. The Hausman test failed to reject the exogeneity assumption of the variables.

We also addressed another concern about a team's choosing a difficult contest and losing the opportunity to achieve better performance in an easier contest, leading to an endogenous influence on the skill scores that are used to calculate the intellectual capital. First, the formula for calculating intellectual capital (skill-score) used by Kaggle.com has accounted for the total number of participating teams in the contest besides team ranking and experiences. Basically, a team performing well in a more complex and competitive contest will earn more skill points than if achieving a similar rank in a less competitive and complex contest. Thus that should have partially addressed this concern.

Second, we further and more formally investigated the impact of selected project complexity on our hypotheses by retesting the model in two ways: (1) with teams that have only selected more complicated (high award) contests, and (2) with teams stratified into groups of high and low past reward levels (a proxy for the complexity of the contest) with matched IC. All our main hypotheses were still supported, demonstrating that our conclusion is robust to a team's past contest selection.

To contrast the internal social capital concept that we used with the overall social capital beyond the team, we also tested the main model with overall social capital as a replacement. The results suggest that overall social capital was positively related to team performance because it increases the team's likelihood of choosing better members through broader global social ties. However, the alignment of overall social capital and intellectual capital did not significantly affect team performance, implying that overall social capital does not play a direct role in team coordination, as internal social capital does.

## Theoretical Implications

The literature in management has a long stream of studies on teamwork. However, due to limitations in data collection and the measurement of performance, there is a lack of research on the allocation of intellectual and social capital among members,

within a self-organized virtual team on team performance in the competitive setting. This research fills the gap by making the following theoretical contributions.

First of all, our study leveraged unique data from a leading predictive analytics crowdsourcing platform, Kaggle, that provides objective performance evaluation on every participating team in a contest. This rich data set allowed us to control for contest heterogeneity, and to use within-contest performance variations to assess the relationship between team social network structure and performance.

Second, as far as we know, this is the first study that focuses on team structures in crowdsourcing. These teams are of a special type that is self-organized and coordinated by individual members, who may come from geographically dispersed locations, without centralized controls. Thus, social skills in sharing and communicating information, as well as coordinating the entire team, become indispensable capital for such teams, in addition to task-related problem solving skills. These types of teams differ from the virtual teams inside a firm, and have not been extensively studied in the literature.

Third, we contribute to the growing literature on online communities by investigating team performance in a specific type of online community: the crowdsourcing community. Unlike the online communities in the prior literature that were collaborative in nature, crowdsourcing communities are competitive among teams. We showed that the relationship of SI alignment with performance varies with competitive intensity.

Last, we complement the research on online social networks by extending the concept of SI alignment to a competitive environment. Our results indicated that SI alignment significantly and negatively affects team performance in crowdsourcing contests. The highly competitive nature of crowdsourcing contests, in which hundreds of teams compete for the win, plays a role in the aforementioned effect. The competitive external environment made the division of the central network positions from high intellectual members critical in our setting. The results also suggested that in environments with low competitive intensity, the effect of SI alignment could be low or even opposite. Hence, our findings are consistent with the findings in the literature related to noncompetitive environments.

## Managerial Implications

Understanding our results concerning how the alignment of intellectual capital and social capital affects team performance, and how competition moderates this relationship, promises benefits for practitioners. The main insights offered by our study relate to dividing tasks and allocating roles within the team for better performance. This principle can be generalized beyond online crowdsourcing to any team management and community design settings. Given the average level of skills of the members, our results suggest that the ideal would be for each member to have his or her own strength in different dimensions of the skill sets that happen to be required by a certain task. In such a case, a team's resources should be allocated

to match the tasks with the skill strengths of team members in mind. As a result, every team member should be assigned a task that he or she is most proficient at, so everyone can contribute to the team project or goal of the community.

Our findings offer guidelines to participants in crowdsourcing contests as well as to contest organizers to improve the performance of virtual teams. As the results suggest, negative alignment of intellectual capital and social capital is beneficial in a crowdsourcing contest setting. Hence, digital platform providers can encourage this by offering advice to participants, making interface design changes, and directly managing team formation, rather than leaving it to the contestants to form their own teams.

Since we have found that competition affects performance in crowdsourcing communities, a similar approach can be adapted to organizational teams where applicable.

## Limitations and Future Research

This research has a number of limitations that can lead to future research. First, our study is based on data from a data-mining crowdsourcing community website. It will be useful to generalize our study to other types of crowdsourcing communities, as well as to general team competition scenarios. Second, our data include only information that is publicly available on the website, which limits our ability to capture the actual interactions among members within a team. In our leader and expert models, we saw that the impacts of a leader's intellectual capital and an expert's social capital were insignificant, though we expected them to be significant to a lesser degree. Thus, we encourage further investigation on how collaboration actually happens within a team. Additional methods, such as follow-up surveys, may result in richer data in future studies. Third, it will be useful to measure the complexity of the crowdsourcing tasks and the level of difficulty in coordinating members within a team. These elements need to be considered in the model because the aforementioned factors may moderate the need for the division of labor. Finally, it will be interesting to investigate how team performance changes over time.

## REFERENCES

1. Archak, N. Money, glory and cheap talk: Analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder.com. *Proceedings of the 19th International Conference on World Wide Web*. New York: Association for Computer Machinery, 2010, pp. 21–30.
2. Asproni, G. Motivation, teamwork, and agile development. *Agile Times*, 4, 1 (2004), 8–15.
3. Baldwin, T.T.; Bedell, M.D.; and Johnson, J.L. The social fabric of a team-based MBA program: Network effects on student satisfaction and performance. *Academy of Management Journal*, 40, 6 (1997), 1369–1397.
4. Balkundi, P., and Harrison, D.A. Ties, leaders, and time in teams: Strong inference about network structure's effects on team viability and performance. *Academy of Management Journal*, 49, 1 (2006), 49–68.

5. Bass, B. Individual capability, team performance, and team productivity. In M. D. Dunnette and E. A. Fleishman (eds.) *Human Performance and Productivity*. Vols. 1, 2, and 3. New York: Psychology Press, 2014, pp. 179–222.
6. Bayus, B.L. Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community. *Management Science*, 59, 1 (2013), 226–244.
7. Bergiel, B.J.; Bergiel, E.B.; and Balsmeier, P.W. Nature of virtual teams: A summary of their advantages and disadvantages. *Management Research News*, 31, 2 (2008), 99–110.
8. Borgatti, S.P. Centrality and network flow. *Social Networks*, 27, 1 (2005), 55–71.
9. Borgatti, S.P.; Jones, C.; and Everett, M. G. Network measures of social capital. *Connections*, 21, 2 (1998), 27–36.
10. Boudreau, K.J.; Lacetera, N.; and Lakhani, K.R. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science*, 57, 5 (2011), 843–863.
11. Brabham, D.C. Crowdsourcing as a model for problem solving. *Convergence: The International Journal of Research into New Media Technologies*, 14, 1 (2008), 75–90.
12. Burt, R.S. *Brokerage and Closure: An Introduction to Social Capital*. Oxford: Oxford University Press, 2005.
13. Chen, H., and Lim, N. Should managers use team-based contests? *Management Science*, 59, 12 (2013), 2823–2836.
14. Chidambaram, L., and Tung, L.L. Is out of sight, out of mind? An empirical study of social loafing in technology-supported groups. *Information Systems Research*, 16, 2 (2005), 149–168.
15. Cohen, S.G., and Bailey, D.E. What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, 23, 3 (1997), 239–290.
16. Cooper, D.J., and Kagel, J. H. Are two heads better than one? Team versus individual play in signaling games. *American Economic Review*, 95, 3 (2005), 477–509.
17. Cross, R., and Cummings, J.N. Tie and network correlates of individual performance in knowledge-intensive work. *Academy of Management Journal*, 47, 6 (2004), 928–937.
18. Crowdsourcing.org. 10 of the 11 best global brands use creative crowdsourcing. <http://www.crowdsourcing.org/editorial/10-of-the-11-best-global-brands-use-creative-crowdsourcing-/16935> (accessed on March 20, 2015).
19. Cummings, J.N. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science*, 50, 3 (2004), 352–364.
20. Daniel, S.; Agarwal, R.; and Stewart, K. J. The effects of diversity in global, distributed collectives: A study of open source project success. *Information Systems Research*, 24, 2 (2013), 312–333.
21. Davenport, T.H.; Thomas, R.J., and Cantrell, S. The mysterious art and science of knowledge-worker performance. *MIT Sloan Management Review*, 44,1 (Fall 2002), 22–30.
22. Ebrahim, N.A.; Ahmed, S.; and Taha, Z. Virtual teams: A literature review. *Australian Journal of Basic and Applied Sciences*, 3, 3 (2009), 2653–2669.
23. Freeman, L.C. Centrality in social networks conceptual clarification. *Social Networks*, 1, 3 (1979), 215–239.
24. Girotra, K.; Terwiesch, C.; and Ulrich, K.T. Idea generation and the quality of the best idea. *Management Science*, 56, 4 (2010), 591–605.
25. Guimera, R.; Uzzi, B.; Spiro, J.; and Amaral, L.A.N. Team assembly mechanisms determine collaboration network structure and team performance. *Science*, 308, 5722 (2005), 697–702.
26. Guinan, P.J.; Coopridge, J.G.; and Faraj, S. Enabling software development team performance during requirements definition: A behavioral versus technical approach. *Information Systems Research*, 9, 2 (1998), 101–125.
27. Hahn, J.; Moon, J.Y.; and Zhang, C. Emergence of new project teams from open source software developer networks: Impact of prior collaboration ties. *Information Systems Research*, 19, 3 (2008), 369–391.
28. Hair, J.F., Jr.; Black, W.C.; Babin, B.J.; and Anderson, R.E. *Multivariate Data Analysis*. 7th ed. Upper Saddle River, NJ: Pearson, 2010.

29. Häussler, C., and Sauer mann, H. The Anatomy of Teams: Division of Labor and Allocation of Credit in Collaborative Knowledge Production. (2014). <http://ssrn.com/abstract=2434327>
30. Hertel, G.; Geister, S.; and Konradt, U. Managing virtual teams: A review of current empirical research. *Human Resource Management Review*, 15, 1 (2005), 69–95.
31. Howe, J. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. New York: Crown Business, 2008.
32. Huang, Y.; Singh, P.; and Srinivasan, K. Crowdsourcing “Blockbuster” ideas: A dynamic structural model of ideation. *Proceedings of the 32nd International Conference on Information Systems*. Atlanta, GA: Association for Information Systems, 2011, pp. 19–22.
33. Jarvenpaa, S.L., and Leidner, D.E. Communication and trust in global virtual teams. *Organizational Science*, 10, 6 (1999), 791–815.
34. Johnson, P.; Heimann, V.; and O’Neill, K. The “wonderland” of virtual teams. *Journal of Workplace Learning*, 13, 1 (2001), 24–30.
35. Kanawattanachai, P., and Yoo, Y. The impact of knowledge coordination on virtual team performance over time. *MIS Quarterly*, 31, 4 (2007), 783–808.
36. Kane, G.C., and Borgatti, S.P. Centrality-IS proficiency alignment and workgroup performance. *MIS Quarterly*, 35, 4 (2011), 1063–1078.
37. Kor, Y.Y., and Sundaramurthy, C. Experience-based human capital and social capital of outside directors. *Journal of Management*, 35, 4 (2009), 981–1006.
38. Lakhani, K.R.; Fayard, A.-L.; Levina, N.; and Pokrywa, S.H. OpenIDEO. Case 612-066. Boston: Harvard Business School Press, February 2012.
39. Lin, Y.J. Division of labor in teams. *Journal of Economics and Management Strategy*, 6, 1 (1997), 403–423.
40. Liu, D.; Geng, X.; and Whinston, A.B. Optimal design of consumer contests. *Journal of Marketing*, 71, 4 (2007), 140–155.
41. Magni, M.; Angst, C.M.; and Agarwal, R. Everybody needs somebody: The influence of team network structure on information technology use. *Journal of Management Information Systems*, 29, 3 (2012), 9–42.
42. Marden, J.I. *Analyzing and Modeling Rank Data*. London: Chapman and Hall, 1995.
43. Martins, L.L.; Gilson, L.L.; and Maynard, M.T. Virtual teams: What do we know and where do we go from here? *Journal of Management*, 30, 6 (2004), 805–835.
44. Mehra, A. Distributed leadership in teams: The network of leadership perceptions and team performance. *Leadership Quarterly*, 17, 3 (2006), 232–245.
45. Mehra, A.; Dixon, A.L.; Brass, D.J.; and Robertson, B. The social network ties of group leaders: Implications for group performance and leader reputation. *Organization Science*, 17, 1 (2006), 64–79.
46. Mo, J.; Zheng, Z.; and Geng, X. Winning crowdsourcing contests: A micro-structural analysis of multi-relational networks. Paper presented at the 5th China Summer Workshop on Information Management. Harbin, China, June 2011.
47. Netflix Prize: Forum. Grand Prize awarded to team BellKor’s Pragmatic Chaos. <http://www.netflixprize.com/community/viewtopic.php?id=1537> (accessed on March 20, 2015).
48. Netflix Prize: Forum. Netflix Progress Prize 2007 awarded to team KorBell. <http://www.netflixprize.com/community/viewtopic.php?id=799> (accessed on March 20, 2015).
49. Pillis, E.D., and Furumo. Counting the cost of virtual teams. *Communications of the ACM*, 50, 12 (2007), 93–95.
50. Powell, A.; Piccoli, G.; and Ives, B. Virtual teams: A review of current literature and directions for future research. *SIGMIS Database*, 35, 1 (2004), 6–36.
51. Reagans, R., and Zuckerman, E.W. Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization Science*, 12, 4 (2001), 502–517.
52. Salas, E.; Rosen, M.A.; Burke, C.S.; Goodwin, G.F.; and Fiore, S.M. The making of a dream team: When expert teams do best. In K.A. Ericsson, N. Charness, P.J. Feltovich, and R. R. Hoffman (eds.), *The Cambridge Handbook of Expertise and Expert Performance*. New York: Cambridge University Press, 2006, pp. 439–453.



53. Sarker, S.; Ahuja, M.; Sarker, S.; and Kirkeby, S. The role of communication and trust in global virtual teams: A social network perspective. *Journal of Management Information Systems*, 28, 1 (2011), 273–310.
54. Savage, N. Gaining wisdom from crowds. *Communications of the ACM*, 55, 3 (2012), 13–15.
55. Seltzer, E., and Mahmoudi, D. Citizen participation, open innovation, and crowdsourcing: Challenges and opportunities for planning. *Journal of Planning Literature*, 28, 1 (2012), 3–18.
56. Singh, P.V., and Tan, Y. Developer heterogeneity and formation of communication networks in open source software projects. *Journal of Management Information Systems*, 27, 3 (2011), 179–210.
57. Smith, A. *The Wealth of Nations*, Cannan, E. (ed.), New York: Modern Library, 1937.
58. Stevenson, W.B., and Radin, R.F. Social capital and social influence on the board of directors. *Journal of Management Studies*, 46, 1 (2009), 16–44.
59. Sudhakar, G.P.; Farooq, A.; and Patnaik, S. Soft factors affecting the performance of software development teams. *Team Performance Management*, 17, 3/4 (2011), 187–205.
60. Sullivan, E. A group effort: More companies are turning to the wisdom of the crowd to find ways to innovate. *Marketing News*, 44, 2 (2010), 22–28.
61. Terwiesch, C., and Xu, Y. Innovation contests, open innovation, and multiagent problem solving. *Management Science*, 54, 9 (2008), 1529–1543.
62. Vance, A. Fight club for geeks. *Business Week*, January 9, 2012, pp. 37–38.
63. Woolley, A.W.; Chabris, C.F.; Pentland, A.; Hashmi, N.; and Malone, T.W. Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330, 6004 (2010), 686–688.
64. Yang, Y.; Chen, P.Y.; and Banker, R. Impact of past performance and strategic bidding on winner determination of open innovation contest. *Proceedings of the 21st Workshop on Information Systems and Economics*. Atlanta, GA: Association for Information Systems, 2010, pp. 11–12.
65. Yang, Y.; Chen, P.Y.; and Pavlou, P. Open innovation: Strategic design of online contests. *Proceedings of the 20th Workshop on Information Systems and Economics*. Atlanta, GA: Association for Information Systems, 2009, pp. 14–15.
66. Zheng, H.; Li, D.; and Hou, W. Task Design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15, 4 (2011), 57–88.