Software Tool Collaboration on nanoHub: Team Composition Implications for Performance

First Author Name (Blank if Blind Review)
Affiliation (Blank if Blind Review)
Address (Blank if Blind Review)
e-mail address (Blank if Blind Review)
Optional phone number (Blank if Blind Review)

Second Author Name (Blank if Blind Review)
Affiliation (Blank if Blind Review)
Address (Blank if Blind Review)
e-mail address (Blank if Blind Review)
Optional phone number (Blank if Blind Review)

ABSTRACT

The portion of scientific papers produced by teams has increased tremendously in recent years, and papers produced by teams are cited more frequently and received more highly. This expansion in the scope and quality of teamwork has been partly attributed to improvements in communication technology and cyber-infrastructure. This study examines collaboration within the nanoHub.org workspace, a cyberinfrastructure-enabled virtual organization hosted by Purdue University. The study tests theoretical mechanisms established in research of offline team collaboration. Results indicate that nanoHub teams operate under similar principles, but with some interesting exceptions. The study concludes by exploring ways in which nanoHub features might be added or improved to further foster collaboration both within and outside the workspace.

Author Keywords

Collaboration, online community, social computing

ACM Classification Keywords

H.5.3 Collaborative Computing.

INTRODUCTION

The role of teams in the production of knowledge within the sciences has been a subject of academic interest for quite some time now. The recent emergence of several factors has prompted researchers to devote substantial efforts to the investigation of scientific collaboration. Acedo et al [1] reported that the percentage of scholarly articles produced by teams, as well as the average size of the teams and the quality of their production, have all been steadily increasing. Knowledge accrual and the rise of complexity within many disciplines call for specialization of expertise and division of labor, making group work a necessity [4]. The growth in the incidence and productivity of team-centered science has made investigations of the mechanisms of effective teamwork and collaboration increasingly relevant.

A primary explanation for the growth in collaboration is the increase in potential facilitative benefits offered by computing and information technology [7]. Information technology gives collaborators the opportunity to share information more quickly and easily than ever before. Yet few studies examine the specific ways in which this general facilitation manifests itself in improved productivity. Identifying and operationalizing with precision the factors which enhance or hinder team performance (and the relationships between them) is a challenge that scholars have yet to tackle.

This study explores the factors that contribute to team effectiveness in the production of analytical and simulation tools on nanoHub.org, an online collaborative workspace developed to facilitate research and learning in the area of nanotechnology. The study proceeds in two parts. First, existing findings from the literature on team formation and performance are reviewed and summarized. The relationships outlined by previous research are then explored in the context of the nanoHub platform. This study demonstrates that, by and large, nanoHub collaborations follow the patterns observed in offline teamwork.

After presenting these findings, the ways in which the unique features offered by the nanoHub site can be used to examine more closely the mechanisms that assist and hinder collaboration are discussed. These insights are then used to develop further recommendations for enhancements to the nanoHub features and interface.

THEORIES AND HYPOTHESES

Literature Review – Teams

Research on the factors that influence team success has produced varying results [4, 8, 10]. Studies do not find that there is a generically "optimal" team size nor that any single factor serves as a consistent predictor of team success. Despite apparent inconsistencies, scholars have forged a basic consensus on the general factors that drive team performance.

Research in teams and collaboration addresses the question in terms of the marginal costs and benefits of adding additional team members. This approach is particularly relevant to collaborative teams (such as in nanoHub) or the production of scholarship, where team members generally have control over who they work with and projects have a limited time frame. In these settings, new team members

The research is supported by < Blank for Blind Review >

are believed to bring three benefits – expertise, information processing ability, and diversity of perspective as well as one cost -- an additional need for coordination [4].

Larger teams can benefit from a wider pool of expertise. Acedo et al [1] suggest that with the increasing specialization in the sciences, "it may be necessary to combine the skills of two or more investigators in order to maintain a given level of quality of published work." Larger teams are also better equipped to deal with difficult tasks that might require a large investment of resources. Koch and Sneider [9] find evidence that larger software development teams are able to marshal greater effort from teammates.

These two benefits of adding teammates can be considered functions of team members' human and social capital -- the stock of existing skills, knowledge and resources the teammates possess prior to their collaboration [8]. Teammates can also bring benefits in what Katz and Lazer refer to as process gains. Process gains are the synergies that result from productive teamwork. They can be considered as the complement to process losses, the inefficiencies that stem from team dysfunction.

One process gain that has been identified in the team collaboration literature is the benefit of diversity of teammate experiences and points of view. Though diversity can be a function of teammate attributes normally associated with human and social capital, it represents a process variable because it can only be measured for particular instantiations of teams. An individual does not carry "diversity" with them; diversity is a property of a team and the manner in which it works together. More diverse teams can be more creative and can stave off the stifling conformity of groupthink [6], increasing team performance.

The process benefits of diversity also suggest a set of process costs. These costs, generally termed the costs of coordination, apply to all teams, but theory argues that coordination costs increase as team size and team diversity increase [4, 12]. For example, as the number of individuals in a team increases, the number of pairs or dyads within the team increases exponentially. It follows that the probability that two team members disagree grows exponentially with team size. Looking at individual member productivity, Nan et al [11] suggest that it may be higher in smaller groups, as those collectives manage to avoid excessive communication costs. Enhanced performance in small teams is expected in particular for projects with low level of modularity.

Despite the importance of coordination costs in the theory of effective collaboration, few studies have operationalized this variable in an examination of real teams. Researchers have, however, demonstrated that there is an effect of previous collaboration on team assembly. When forming new teams, people are more likely to prefer partners who are already familiar from earlier work on joint projects. Prior knowledge reduces uncertainty; successful team projects increase the social capital of their creators, which may in turn enhance the popularity of (and the community support for) the new collaborative endeavor [5].

This tendency towards repeated ties between team members is presumed to be an indication of the importance of lowering coordination costs. Yet this method for reducing coordination costs brings a trade-off. Guimera et al [4] find that previous collaboration is negatively associated with the quality of a team's output. They suggest that teams with large numbers of repeat ties are unlikely to produce many innovative ideas, because shared experiences of members may homogenize their perspectives. When individuals work together, they tend to develop shared understandings and interpretations. These shared interpretations facilitate coordination, i.e. they reduce coordination costs, but they can also reduce creativity by eliminating criticism, doubt, and the exploration of new ideas, encouraging groupthink. Even if social cohesiveness does initially enhance group performance [18], with time it may have a negative effect, especially in the context of intellectual or creative tasks [16].

In their work on team composition in interdisciplinary projects, Cummings and Kiesler [3] expand on the implications of previous collaboration, introducing the concept of collaborative tie strength – the intensity of working relationships among team members. In the context of scientific work, collaborative tie strength is related to the existence of previous and ongoing engagement in knowledge production (exchange of ideas, shared research projects, co-authored articles, etc.). Cumming and Kiesler take into consideration the effects of homophily and proximity in addition to previous teamwork experiences as mechanisms that increase the strength of links between collaborators. According to their findings, larger geographic distance as well as disciplinary differences are negatively associated with collaborative tie strength.

Considering the ambiguous contribution of previous collaboration and the difficulty researchers have had in directly measuring coordination costs, it is not surprising that the findings of scholars investigating the relationship between diversity and team performance are inconsistent. In general, types of diversity which are more likely to cause coordination problems (e.g. geographic heterogeneity) are expected to negatively impact team performance. On the other hand, diversity that is likely to be associated with relevant, critical, alternative perspectives generally contributes positively to team performance [15]. Most measures of diversity will, however, include both factors.

Testing Established Findings

The preceding review suggests that research findings in the field of team collaboration can be summarized by a few important principles that involve the following factors: team size, team geographic diversity, team institutional diversity, and previous collaborative experience between team members These factors are tested using the nanoHub data to assess the extent to which these relationships hold in the online environment provided by virtual organizations

Team Size. As described above, previous literature suggests that larger teams outperform smaller teams.

H1: Larger team size is associated with greater team performance.

The general finding that larger teams create better products does not assure that the hypothetical addition of a team member will improve team performance. As teammates are added to a team, coordination costs rise. The addition of coordination costs can have one of two effects on team performance. First, additional coordination costs can cause teams to slow their production or abandon it altogether. That is, teams that are too large to coordinate their activities may not be able to produce any tools at all, or may disband before they have produced a tool. Second, additional coordination costs can hinder team performance, leading to the production of less successful tools.

As team size increases, teams possess improved capability but also increased cost of coordination. These factors are, however, unlikely to increase at the same rate. While teams are likely to gain immediate advantages from adding a few members with unique expertise, this marginal value should diminish as more and more team members are added. At the same time, the addition of additional team members is likely to add ever increasing costs of coordination. For example, if costs of coordination are proportional to the number of possible communication links between team members, these costs increase exponentially as teams expand in size. This argument suggests that increased team size may increase team performance to a local maximum, after which coordination costs dominate with the addition of new team members:

H2: Increases in team size improve team performance to up to an optimal team size, above which further increases in team size reduce team performance.

Team Diversity, Geographic. Geographic diversity adds to a team's coordination burden. As Cummings and Kiesler point out, geographic distance between team members negatively impacts their collaborative tie strength as they have fewer opportunities for shared experiences. Even in the context of virtual teams where physical distance is less of an issue [17, 13], researchers have pointed out the high coordination cost of international partnerships. Problems associated with cultural divides and managing cooperation across time zones can harm team performance [14].

H3a: Higher geographic diversity is associated with lower team performance.

Team Diversity, Institutional. On the other hand, institutional diversity can minimize the degree to which teams will fall into a rut of groupthink. Collaboration between individuals that reside at different universities should bring the benefit of new ideas and perspectives due to exposure to different colleagues and different norms and routines. Thus, if geographic barriers can be removed or controlled for, it is possible that team performance will increase with university diversity. nanoHub's unique features, in particular workspace sharing, permit individuals at different locations to collaborate in software development as though they were sharing a desk. Thus, we propose the following as a research question:

H3b: Higher university diversity is associated with higher team performance.

Prior collaborative relationships. A larger number of previous collaborative relationships should both reduce the costs of coordination and reduce team creativity. While the overall impact on performance is thus uncertain, it follows that teams with a large portion of members that have worked together before are likely to produce tools faster and with fewer errors. The tools they produce, however, will be less creative or innovative.

H4a: A greater portion of previous collaboration will be associated with a greater useability of tools.

H4b: A greater portion of previous collaboration will be associated with a lower originality of tools.

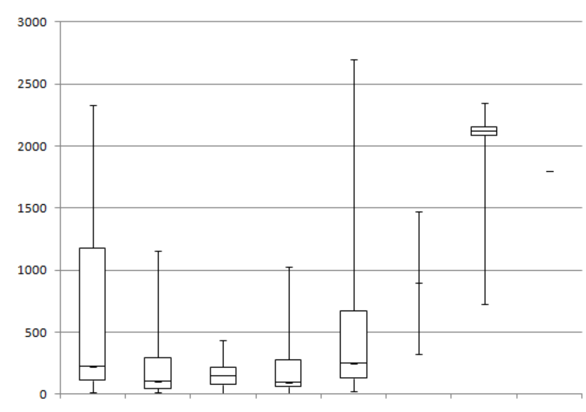

Figure 1. Boxplot for the numbers of users by team size.

MODEL AND RESULTS

nanoHub is a cyberinfrastructure-enabled virtual organization hosted by Purdue University that enables scientists and engineers to develop software tools and performance simulations for nanotechnology research. It is a global resource for nanoscience and technology, created by the NSF-funded Network for Computational Nanotechnology. The tools are hosted and running on computer clusters that are equivalent to 100 personal computers provided through the nanoHub. Without downloading complex software and building their own simulation environment, researchers around the world can run the simulations and do their experiment online with an ordinary Web browser on their personal computers even on cell phones. Therefore, nanoHub allows researchers to spend more time doing research and less time and money developing simulations.

nanoHub facilitates software tool based collaboration in three ways. First, by providing a platform for software simulation, nanoHub reduces the cost of using the tools and creates a community of tool developers and users. Second, through a shared workspace, tool users can transfer their running applications to others without interruption. Therefore a collection of users can perform the simulation virtually together to share results, or for troubleshooting. Finally, developers can build virtual teams on nanoHub to simplify the software development process.

This paper studies the software development teams in nanoHub and examines the impacts of team characteristics on their outcomes. As shown in Figure 1, the total numbers of users for the solo author tools has a bigger variance than that of team outcomes. This indicates that team collaboration may have different behavior compared to single person software development. Therefore we select 114 published software tools developed by teams (i.e. at least two persons) as the samples. Figure 2 shows the membership structure of all teams. Blue squares indicate tools and red circles indicate persons; a link between a person and a tool indicates that the person is a member of the team develops the tool.

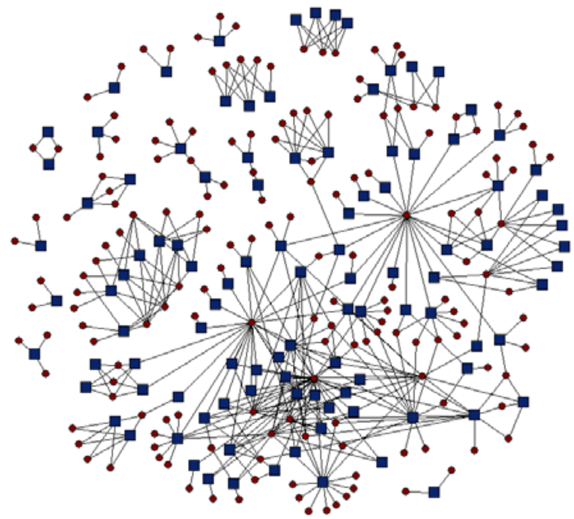

Figure 2. nanoHub team-user affiliation network.

Measures

To evaluate the team performance we use three perspectives to measure the outcomes associated with the tools they develop: the numbers of users, citations, and positive ratings.

- *Number of users*: Because all software tools are running on the computer clusters, nanoHub has accurate logs about who runs which application for how long. The total usage was computed as the number of unique users who used each tool. The number of users measures the general use-ability of a tool among all users in nanoHub including researchers, practitioners, and students.
- *Number of citations*: Academic citation is a standard measure to evaluate the impact on scientific research and the level of originality. nanoHub collects information of citations for each tool from published databases. This list is augmented by users who can upload the citations they find. All information is publicly available for all tool developers and nanoHub users.
- *Number of positive ratings*: Like many community portals, nanoHub provides a rating function: each user can voluntarily rate online tools using a 5-level scale. In nanoHub users rarely gave negative ratings, i.e. the ratings below 3 points, and some tools did not get any rating. Therefore ratings above or equal to 3 points can be taken as a measure of users' positive feedback on the tools.

There were five measures of team structures. While there were changes over time in team composition, these measures were computed on the most recent list of team members.

- *Team size:* The size of tool development teams in nanoHub ranged from 2 to 8. 38 teams were of size 2, 29 of size 3, 21 of size 4, 19 of size 5, 2 of size 6, 4 of size 7, and 1of size 8. Dummy variables, Size_2, Size_3, Size_4, Size_5, and Size_6+, were used in the analysis to control the effect of team size.
- *Previous collaboration:* Some individuals developed multiple tools together in nanoHub. The previous collaboration experience in the team was computed as the ratio of e total number of pairwise collaborations among members based on prior tool development projects on nanoHub to the number of possible pairwise collaborations among all members in the team.
- *University diversity:* University diversity was computed as a Blau's diversity index [reference] based on the total number of unique universities of the team members and

their distribution. For example, if all members are at the same university, the university diversity is 0; if all members in a 4-person team are from different university, the university diversity is 0.75.

- *Country diversity*: Blau's index was also used to estimate the country diversity of team members. Most teams reside in the United States; only ten teams have members from more than one country.
- *Relation to Purdue University*: Because nanoHub is developed and hosted at Purdue University, the researchers within the university have better support and access to the facility. Hence the percentage of members affiliated with Purdue University was used as a control variable.

Each tool is designed for a specific type of application. The following five metrics were used to estimate the nature and complexity levels of the tools.

- *Open source*: Most tools in nanoHub are proprietary. Only six tools provide source code. A dummy variable was used to identify the tool.
- *Number of versions*: A team can keep working on the tool and provide new updates to nanoHub users. The number of versions is used to estimate the update frequency.
- *Number of tags*: Each nanoHub user can tag the tools with a keyword. The number of unique tags a tool received measures the semantic complexity of the tool.
- *Average session time*: The average number of hours used in user simulation sessions estimates the usage complexity.
- *Average CPU time*: The average number of CPU hours used in simulation sessions estimates the computational complexity.

Variables	Min	Mean	Max	Std. deviation
# Users	6	354.23	2699	522.61
# Citation	0	2.87	92	11.74
# Positive ratings	0	0.98	16	2.17
# Days online	84	691.15	1319	363.07
Size	2	3.42	8	1.40
Pre-collaboration	0	1.03	11	1.91
University diversity	0	0.17	0.75	0.24
Country diversity	0	0.03	0.625	0.12
Relation to Purdue	0	0.46	1	0.45
Open source	0	0.05	1	0.22
# Visions	1	3.52	51	5.13
# Tags	2	13	55	7.89
Avg session time	0.05	2.04	86.14	8.07
Avg CPU time	0	0.13	2.97	0.38

Table 1. Descriptive statistics of variables.

Descriptive statistics of the variables are reported in Table 1.

Results

Because all three outcome measures are count variables accumulated in the total time period since the tools were published on nanoHub, the number of days online is used as a control measure of exposure and Poisson regression is used to test the hypotheses.

Model 1 tested the impact of team size and diversity on the team performance measured by the numbers of users, citations, and positive ratings. Tool attributes including Open source, Number of versions, Number of tags, Average session time, and Average CPU time are added to control the different characteristics of the team outcomes. As reported in Table 2, team size has significant positive impacts on all three team performance measures. Tools produced by larger teams have more users, citations, and positive ratings. Hypothesis 1 is supported.

Previous collaboration has a significant negative impact on the number of users and no significant impact on the numbers of citations and positive ratings. Therefore, Hypotheses 4a and 4b are not supported.

University diversity has a significant positive impact on the number of users and Country diversity has a negative impact. Both Hypotheses 3a and 3b are supported when the team performance is measured based on the usage. However, for citations and positive rates, none of them are supported.

Among the control variables, Relation to Purdue and the number of tags have positive impacts on the numbers of users and citations.

Variables	# Users	# Citation	# Positive ratings
Size	0.11*** (.004)	0.57*** (.08)	0.19** (.09)
Pre-collaboration	-0.06***(.004)	-0.12 (.11)	-0.15 (.10)
University diversity	1.10*** (.03)	0.40 (.54)	-0.67 (.61)
Country diversity	-0.26*** (.04)	1.82***(.59)	0.02 (.85)
Relation to Purdue	0.75*** (.02)	0.83** (.35)	0.32 (.28)
Open source	0.03 (.02)	3.07***(.20)	0.13 (.52)
# Visions	0.03***(.001)	-0.0001 (.02)	0.07*** (.01)
# Tags	0.03***(.001)	0.03***(.01)	0.02** (.01)
Avg session time	0.01*** (.001)	0.02 (.02)	-0.001 (0.03)
Avg CPU time	-0.002**(.001)	0.04***(.01)	0.009 (.02)
Constant	-2.27	-10.13	-7.99
# Days online	(exposure)		
Log likelihood	-7799.77	-150.91	-133.03
Pseudo R2	0.53	0.76	0.28

* indicates p<0.10; ** indicates p<0.05; *** indicates p<0.01

Table 2. Results of Model 1.

Model 2 studied the non-linear impact of team size. All variables in Model 1 are included except Team size, which is replaced with four dummy variables indicating the categories of team sizes. Table 3 shows that all results are consistent with model 1. Moreover, the coefficients of the

Variables	# Users	# Citation	# Positive ratings
Pre-collaboration	-0.02***(.004)	-0.13 (.11)	-0.15 (.11)
University diversity	0.88***(.03)	0.96* (.51)	-1.05 (.65)
Country diversity	0.06 (.04)	2.33***(.61)	0.40 (.87)
Relation to Purdue	0.73*** (.02)	1.07***(.37)	0.39 (.29)
Open source	-0.09***(.02)	2.56***(.18)	-0.01 (.52)
# Visions	0.04***(.001)	0.009 (.03)	0.07***(.01)
# Tags	0.02***(.001)	0.04***(.009)	0.01 (.01)
Avg session time	0.01***(.001)	0.02 (.02)	-0.004 (0.03)
Avg CPU time	0.001 (.001)	0.02* (.01)	0.03 (.02)
Size = 3	-0.53***(.02)	0.03 (.43)	-0.62* (.35)
Size = 4	-0.47***(.02)	0.27 (.43)	-0.19 (.37)
Size = 5	0.01 (.02)	1.56***(.32)	-0.18 (.36)
Size >5	0.47***(.02)	2.15***(.34)	1.02**(.43)
Constant	-1.62	-8.90	-7.03
Days online	(exposure)		
Log likelihood	-6811.51	-156.96	-128.27
Pseudo R2	0.59	0.75	0.31

* indicates p<0.10; ** indicates p<0.05; *** indicates p<0.01

Table 3. Results of Model 2 (big group impact).

team size dummy variables clearly show the non-linear effect of team size. Although bigger teams have better performance in general, teams with size 3 and size 4 perform worse that the teams with two members (the base class). However the effect is opposite to the prediction in Hypothesis 2.

DISCUSSION

Discussion of results

As previous evidence shows, collaborations have an increasing influence on the generation and distribution of scientific knowledge. In keeping with these findings, our analysis suggests a number of relevant impact measures illustrating how collaboration is influencing both developers and users. These mechanisms of team assembly have proven to play an important role in the outcome of collaborative projects.

Preliminary analysis from Model 1 supports previous research, demonstrating higher outcome impact as team size increases confirming hypothesis 1 - Team Size. However, detailed analysis of team size in Model 2 illustrates a non-linear relationship of these attributes. This trend reflects that teams of size two outperform individuals, though slightly larger groups of three to four members show a decline in impact over the pair-team and groups of size five or more show an increase over the impact of a pair-team. These findings support hypothesis 2 by illustrating the trend to performance increase and decrease thresholds.

This phenomenon may be explained through the balance of collaboration costs and benefits. Teams with two members have an advantage in terms of minimizing collaboration costs as communication remains focused and local to two sources, while teams with over five members thrive on collaboration benefits such as diversity of experience or knowledge, and simply more opportunity for labor and production as well as quality control. However, teams that consist of three to four members suffer from fewer benefits that appear to be offset by higher collaboration costs. The collaboration costs might include scheduling, communicating using technology that is more generalizable or understandable, and adapting language to fit diverse linguistics; in addition, the collaboration benefits exist, but in less prominence than with larger groups. The ability to achieve the goal and share the product benefit from big teams, but not enough to maximize on the cost in coordinating these efforts.

The diversity impacts pose interesting variations across teams as well. The results indicate that teams where members are from multiple countries are more likely to develop tools that are more widely cited although that are not likely to have a larger number of users. As these findings partially support hypothesis H3a, it becomes important to focus on the meaning of outcome. Measuring outcomes through citations supports this hypothesis, while the use of user access contradicts this hypothesis. Similar support exists in team diversity as measured by university association for hypothesis H3b. Teams which have members from diverse universities are more likely to develop tools that are accessed by more users. This results from users at each of the represented universities being more likely to use a tool co-developed by one of their university colleagues.

The number of versions associated with the development of a tool also impacts the number of users that access that tool. This might be a result of users sharing interests and needs with developers by providing detailed feedback as well as offering wishlists for desired features. As more users access and review tools, each version of these resources becomes increasingly accurate and reliable. Improvements through each iteration bring better reviews and new users with new observations and goals for the tool, again cycling the growth process.

Previous collaboration between pairs of members among the team also tends to influence the impact of the resulting tool. Teams with previous collaboration show lower impact on both user rates as well as citation rates than those teams with fewer instances of previous collaboration. These findings contradict both posits of hypothesis H4 - previous collaboration increases impact. This negative impact may be the result of team member time constraints. As multiple team members work with other team members to produce one tool, they often are simultaneously producing another. Attempting to disperse their efforts with the same collaborators on different projects may lead to infringing meetings, or project discussion overlap and interfere with both productivity and quality. Similar contributors may also lead to monotonous tool generation. That is, if the same team is creating multiple tools they may be losing or lacking diversity brought from the addition of new members, ideas, or topics.

Another important area that needs to be investigated involves the distribution of the collaboratively produced content. Institutional and disciplinary diversity of team members may lead to work that is more widely understood by multiple communities. Within the group network, however, this heterogeneity may create barriers to communication and understanding. Bringing together multiple worlds will occasionally result in misaligned definitions, expectations, and intentions. Finding the balance that allows exploitation of the benefits offered by team diversity while at the same time minimizes the negative consequences of the increased coordination cost is a complex and challenging task.

The results of this study indicate that research should, therefore, devote more attention to the implications and possibilities offered by virtual team formation. As scientific collaboration becomes an increasingly prevalent practice, it is important to understand not only how teams work together, but also how teams come together to produce and disseminate knowledge. In order to maximize the impact and scope of newly generated resources, scholars need to be aware of the specifics of group formation and their effect on team performance.

Future Directions

Compared to traditional ways of scientific collaboration, software production in nanoHub has several important characteristics. It not only connects geographically distributed scientists in the form of an online community, but also provides them with a shared work space. Additionally, as developers have the opportunity to observe the usage of and feedback on their tools in real time, teams can flexibly respond and incorporate feedback. Teams no longer produce a single, best version of a tool. Rather, they seek to produce an interesting, usable version that others can try out and rate, providing suggestions for further development. In general, the findings suggest that these features of nanoHub might have the potential to create a conducive virtual environment for scientific collaboration. The findings also point out the ways in which the design of similar collaborative systems could be further improved. As team size demonstrates a non-linear relationship with team performance, it could be helpful to implement a recommendation system that suggests an optimal team size according to the task type, team member diversity, and other contextual factors. Similarly, such a system could also make suggestions for potential collaborators, based on the objective of the project, respective expertise and background of the collaborators and existing members. Collectively, these recommendation algorithms can contribute to building an intelligent workspace to help achieve optimal team work in the scientific community.

Several future research directions are particularly promising. First, as the findings suggest a non-linear relationship between team size and collaboration performance, an important future research direction is to closely examine the possible mechanisms contributing to such a relationship in online shared work spaces. For example, this study hypothesized that coordination cost rises as the number of collaborators increases. Future studies can quantify coordination costs by examining the communication logs and work flow of collaborative tasks.

Second, collaborative relations captured in this dataset constitute only one single layer of the complex social fabrics connecting these nano scientists. The models reported here did not examine the effect of previous collaborations outside of nanoHub. As several studies have found [4, 5], performance of teamwork is often associated with the breadth and strength of previous interactions. Future work should incorporate previous collaborations such as joint publications into the models.

Third, the performance of team collaboration may also be subject to other contextual factors, such as task complexity and modularity. Studies have found that teamwork is preferred when tasks could be decomposed into modular activities that could be carried out in parallel and require minimal input from other processes [15]. Future work must examine the characteristics of producing these scientific tools on nanoHub and how they affect the relationship between team size and performance.

CONCLUSION

The results of this study have shown complicated and significant relationships among team assembly and product impact in regards to tools produced in the nanoHub community. Though diversity and collaboration bring great value to teams with expanded perceptions, processing power, publicity, and levels of expertise, the cost of collaboration may offset these benefits in many specific circumstances. The technology available in the nanoHub community allows for a unique relationship between the tool's users and developers, as well as an infrastructure to analyze this relationship's dynamics through iterations of applications, development teams, and user needs. With a greater understanding of these team assembly strategies and tool usage practices, online communities may better model and implement applications to support productive user behaviors.

REFERENCES

1. Acedo, F., Barroso, C., Rocha, C., & Galán, J. (2006). Co-authorship in management and organizational studies: an empirical and network analysis. Journal of Management Studies, 43(5), 957-983.
2. Blau, P. Inequality and heterogeneity. Free Press New York, 1977.

3. Cummings, J., & Kiesler, S. (2007). Who works with whom? Collaborative tie strength in distributed interdisciplinary projects. Paper presented at the Third International e-Social Science Conference.
4. Guimera, R., Uzzi, B., Spiro, J., & Amaral, L. (2005). Team assembly mechanisms determine collaboration network structure and team performance. Science, 308(5722), 697-702.
5. Hahn, J., Moon, J. Y., & Zhang, C. (2008). Emergence of new project teams from open source software developer networks: Impact of prior collaboration ties. Information Systems Research, 19(3), 369-391.
6. Janis, I. (1972). Victims of groupthink: Houghton Mifflin Boston.
7. Jones, B., Wuchty, S., & Uzzi, B. (2008). Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. Science, 322(5905), 1259.
8. Katz, N., & Lazer, D. (2003). Building effective intra-organizational networks: The role of teams. Retrieved from http://ksg.harvard.edu
9. Koch, S., & Schneider, G. (2002). Effort, co-operation and co-ordination in an open source software project: GNOME. Information Systems Journal, 12(1), 27-42.
10. Kretschmer, H. (1985). Cooperation structure, group size and productivity in research groups. Scientometrics, 7(1), 39-53.
11. Nan, N., Kumar, S., & Wang, L. (2007). Collaboration Network Structure in Open Source Software Projects. Paper presented at the Sixth AIS SIGSAND Symposium on Research in System Analysis and Design Tulsa, USA.
12. Pendharkar, P., & Rodger, J. (2007). An empirical study of the impact of team size on software development effort. Information Technology and Management, 8(4), 253-262.
13. Pieper, J., Grace, J., & Dill, S. (2009). Team analytics: understanding teams in the global workplace. Paper presented at the 27th international conference on Human factors in computing systems, Boston, MA, USA.
14. Powell, A., Piccoli, G., & Ives, B. (2004). Virtual teams: a review of current literature and directions for future research. The DATA BASE for Advances in Information Systems, 35(1), 7.
15. Stewart, G. (2006). A meta-analytic review of relationships between team design features and team performance. Journal of Management, 32(1), 29.
16. Stokols, D., Misra, S., Moser, R., Hall, K., & Taylor, B. (2008). The Ecology of Team Science. American Journal of Preventive Medicine, 35(2), 96
17. Wuchty, S., Jones, B., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036.
18. Yang, H., & Tang, J. (2004). Team structure and team performance in IS development: a social network perspective. Information & Management, 41(3), 335-349.