Homophily of Network Ties and Bonding and Bridging Social Capital in Computer-Mediated Distributed Teams

Y. Connie Yuan
Geri Gay
Departments of Communication and Information Science
Cornell University

This research studied homophily of network ties in distributed teams in both task-related instrumental networks and non-task related expressive networks. Homophily of network ties was examined in terms of demographic and social characteristics, including gender, race, geographic location, and group assignment. Social network data were collected from 32 students enrolled in a distance learning class from two universities. MQAP regression analysis showed that homophily in gender and in race had no significant impact on the development of either instrumental or expressive ties. In instrumental networks, both homophily in group assignment and in location had significant impact on the development of network ties. In expressive networks, homophily in location had significant impact on the development of network ties, but the impact of homophily in group membership was only marginally significant. Further analysis of bonding ties with people of the same group and bridging ties with people from different groups showed that bonding social capital can exert significant influence on performance.

doi:10.1111/j.1083-6101.2006.00308.x

Introduction

Developing network relationships is one of the key factors influencing the effective functioning of small groups, particularly when groups are engaged in knowledge-intensive work. Unlike equipment or other major assets of an organization, knowledge assets can be tacit in that certain types of knowledge, e.g., task-specific experiences, cannot easily be articulated (Polanyi, 1967). Because tacit knowledge resides mainly in individuals’ minds, and sometimes within the connections among individuals (Argote & Ophir, 2002), it cannot easily be articulated or stored in documents. As a result, sharing tacit knowledge becomes very difficult.

To deal with this challenge in organizational knowledge management, scholars have suggested that management should devote resources to the building of network
ties as a key strategy in managing knowledge itself, because firsthand experience through close contacts is the only avenue for people to acquire tacit knowledge (Nonaka & Takeuchi, 1995). For distributed teams engaged in knowledge-intensive work, developing network ties becomes even more crucial because members of distributed teams have only limited opportunities to learn from observing others. Although communication and information technologies can provide a powerful platform to facilitate the process, it is network relationships that serve as the actual bonds that help people to overcome geographic constraints (Barab, Kling, & Gray, 2004a; Yuan, Gay, & Hembrooke, 2006). Given the vital role that network ties play in distributed knowledge-work teams, both students and practitioners of knowledge management need to develop a good understanding of what factors drive the formation of network ties.

Previous research has explored properties of social network relationships within and across distributed teams that belong to only one learning community. This research employed the theory of homophily to explain the development of network ties in task- and non-task-related communication networks. Homophily theory predicts that people are more likely to interact with individuals similar to themselves in respect to a variety of qualities and characteristics (McPherson, Smith-Lovin, & Cook, 2001; Monge & Contractor, 2003). In particular, extensive research has been done to investigate how homophily in age, gender, race, education, occupation, and values, for instance, influences the formation of network ties in communities, voluntary organizations, private businesses, etc. (see McPherson, et al., 2001 for a comprehensive review). Overall, the theory has received widespread support in diverse contexts.

The current study contributes to existing homophily research in several important ways. First, past research on network homophily in organizations mainly focused on demographic characteristics, such as gender or race (Ibarra, 1992; Mollica, Gray, & Trevino, 2003), while the current research explores homophily as it affects both demographic (race and gender) and social/contextual factors, such as group membership and location. Second, following the call by Mcpherson, et al. (2001), the current research explores homophily in multiple relationships. For that reason, we have adopted Ibarra’s (Ibarra, 1995; Ibarra & Andrews, 1993) conceptual framework for network tie classifications and have made a distinction between two types of network ties. Instrumental network ties are those developed for information, advice, and resource exchanges that are needed to accomplish tasks. Expressive ties are affective ties carrying either positive or negative emotions that are not necessarily task-related. The present research explores homophily of network ties in both types of relationships.

Third, past organizational research supporting the homophily hypothesis mainly studied only co-located teams. Our research is among the few studies that have investigated how demographic and social homophilias influence the formation of instrumental and expressive network ties in distributed teams. Finally, most existing homophily research focuses on examining the level of homogeneity or heterogeneity
of ties in a focal person’s social networks. Building on past research in social network analysis, the current research further explores how instrumental and expressive ties with similar versus dissimilar others bond or bridge relationships among people of same versus different characteristics, and consequently, influence performance.

The article is organized as follows. The following section reviews and outlines prior research. Based on past studies, we then propose a series of hypotheses about how demographic and social homophilies influence the formation of instrumental and expressive ties and how such ties, with both similar and dissimilar others, influence performance. The second section reports the results of testing each of these hypotheses using data collected from 32 students enrolled in a distance-learning, advanced engineering design class that required students to finish an integrated, intensive knowledge-creation group project in one academic year. The article concludes with discussion of the implications of the results, as well as limitations and directions for future research.

The Theory of Homophily

Monge and Contractor (2003) summarized two main lines of reasoning that support the theory of homophily, including Byrne’s (1971) similarity-attraction hypothesis and Turner’s (1987) theory of self-categorization. The similarity-attraction hypothesis predicts that people are more likely to interact with those with whom they share similar traits. The theory of self-categorization proposes that people tend to self-categorize themselves and others in terms of race, gender, age, education, etc., and that they use these categories to further differentiate between similar and dissimilar others. That is, Person A will perceive Person B to be more similar to him- or herself when Person B belongs to the same social categories as him- or herself than when Person B belongs to a different social category. In addition, because interpersonal similarity increases predictability of behavior and reduces communication apprehension (Ibarra, 1992), communications among similar others are more likely to occur.

Combining these postulates, the rationale for the theory of homophily is straightforward: “Similarity breeds connections” (McPherson, et al., 2001, p. 415) and “birds of a feather flock together” (p. 417). Empirical research has found strong support for the homophily hypothesis, particularly in terms of age (e.g., Feld, 1982), gender (e.g., Ibarra, 1992; Leenders, 1996), race/ethnicity (e.g., Mollica, et al., 2003), education (e.g., Marsden, 1987), status (e.g., McPherson & Smith-Lovin, 1987), etc. People who are homogeneous in age, ethnicity, educational level, and status are much more likely to interact with each other than with people who are heterogeneous in these respects. Because people are more comfortable interacting with similar others, we hypothesized that team members are more likely to exchange task-related information with people of the same gender and/or race.

Although most existing studies of network homophily investigated the issue in co-located situations, we predict that the findings from these studies are generalizable...
to distributed settings for the following two reasons. First, despite possible loss of certain nonverbal cues in communication, many information and communication technologies such as videoconferencing and webcams have the capability to emulate face-to-face communication and provide full access to a wide variety of information sources, including such observable features as people’s gender and race. Second, while media richness theory (Daft & Lengel, 1986) and social presence theory (Short, Williams, & Christie, 1976) emphasize how features of communication technologies impose intrinsic constraints on the amount of information transmittable via a particular communication technology, a more social approach to computer-mediated communication (CMC) emphasizes people’s creativity in using technologies adaptively. Both the social influence model (Fulk, 1993) and the hyper-personal model (Walther, 1996) argue that social norms, subjective perceptions, and individual agency can influence how a particular communication technology is used. Specifically, both models maintain that through creative usage, people have the capability to overcome the constraints intrinsic to technologies and thus make an otherwise lean medium capable of conveying rich social information. For instance, email has been considered a lean medium that is unable to convey rich social cues. Yet with the constant invention and widespread usage of different emoticons, people have learned to make their emotions explicit in text messages. Therefore, in the context of the current research, we believe that in distributed teams people are more likely to communicate with those of the same race and gender than with those differing in those traits, even though their social interactions are mediated by technologies and the possibility of concomitant loss of social cues.

In this research, we did not study homophily in education and status because distributed team members who participated in this study all had the same level of education, and the teams’ structures were flat, with no differentiation in hierarchical status. Therefore, we excluded these two variables in our analysis and proposed and tested only the following hypotheses:

**Hypothesis 1:** Members of distributed teams are more likely to form task-related instrumental ties with those who are of the same (a) gender and (b) race.

Network ties can provide not only instrumental resource exchange but also expressive emotional support (Haythornthwaite & Wellman, 1998). Past research found greater racial and gender homophily in expressive ties than in instrumental ones (Ibarra, 1992; McPherson, et al., 2001). Despite the usual skewed distribution of race and gender in the workplace, employees’ emotional support networks were found to consist of mainly same-gender/same-race confidants. Although it is difficult to find similar others when surrounded by white males, female minority employees still seek out same-gender/same-race others for emotional support because these people are more likely to be perceived as peers, and therefore as more capable of demonstrating empathy. Based on these findings and reasoning, we proposed that

**Hypothesis 2:** Members of distributed teams are more likely to form non-task-related expressive ties with those who are of the same (a) gender and (b) race.
While traditional homophily research mainly focuses on demographic characteristics (Monge & Contractor, 2003), recent research reveals that location and organizational membership (McPherson, et al., 2001) are important contextual factors influencing the emergence of homophilous ties. In distributed virtual teams, location and group membership become particularly important because it is the distance between people that results in a distributed team and it is group membership that defines the boundary of work relationships.

Group membership breeds network connections because group memberships create foci of activities around which people organize their social relations (Feld, 1981; McPherson, et al., 2001). The impact of group membership on building network ties can be strong enough either to override or lessen the differences that age, race, religion, and other individual characteristics can create (Marsden, 1990; McPherson, et al., 2001). Haythornthwaite (2001) found in her study on distance learning that the social network structure of a class converged on group assignment. That is, communications for collaborative work, exchanging advice, and socializing were much more frequent among people of the same groups than those of different groups. Based on these arguments and findings, hypotheses 3(a) and 3(b) are postulated:

**Hypothesis 3(a):** Members of distributed teams are more likely to form task-related *instrumental ties* with people of the same group.

**Hypothesis 3(b):** Members of distributed teams are more likely to form non-task-related *expressive ties* with people of the same group.

In addition to group membership, proximity in location can also exert a significant impact on the formation of network ties because common location provides common social context, opportunities, and motivations for interaction (Barab, MaKinster, & Scheckler, 2004; Feld, 1981). Monge and his colleagues (1985) found that employees often felt psychologically obligated to engage in face-to-face communications with those in close physical proximity. Distance, on the other hand, erodes shared context, familiarity, and friendship among group members (Armstrong & Cole, 2002; Hinds & Bailey, 2003). Past social network research has found that, as physical distance increases, the probability of face-to-face communication decreases significantly (Zahn, 1991). Allen (1977) claimed that, if one were farther than 30 meters apart from someone, one might as well be several miles apart. In distributed teams when team members are no longer co-located in the same physical place, proximity in location, then, functions to divide the teams into smaller cliques. Deprived of a common work context, it is more difficult to become familiar and friendly with team members, and the lack of causal, informal encounters makes people more likely to feel isolated (Connaughton & Daly, 2004). Although recent development in information and communication technologies has greatly eased communication across long distances, geographical separation remains a major hurdle of communication. Bradner and Mark (2002) found that in computer-mediated communication situations, people were less cooperative, less likely to be persuaded,
and more likely to deceive others when they perceived their partners were located in a different city. The effects of distance on cooperation, persuasion, and deception did not change with the use of different communication media, i.e., text-only instant messaging versus video-conferencing. Based on these arguments, it is hypothesized that

**Hypothesis 4(a):** Members of distributed teams are more likely to form task-related *instrumental ties* with people at the same location.

**Hypothesis 4(b):** People are more likely to form non-task-related *expressive ties* with people at the same location.

Finally, homophily is additive (McPherson, et al., 2001). Similarities in diverse demographic and social traits can accumulate to create greater homophily so that people having more attributes in common have a higher probability of forming network relationships. Based on these arguments, these two hypotheses follow:

**Hypothesis 5(a):** People are more likely to form *instrumental ties* when they share homophily in more traits.

**Hypothesis 5(b):** People are more likely to form *expressive ties* when they share homophily in more traits.

**Homophily and Bonding and Bridging Social Capital**

While the main argument of homophily theory is that people sharing similar traits are more likely to interact with each other, there is a reverse effect in that the likelihood of interacting with dissimilar others is reduced. In other words, homophily not only unifies, it also divides a network. Nonetheless, for people working in knowledge groups, the capability of reaching out to dissimilar others is as important as reaching out to similar others (Mollica, et al., 2003). This is particularly true for tasks that involve the creation of new knowledge (e.g., the engineering design task for this research), where diverse sources of information provide for a more fruitful intellectual exchange of new ideas (Argote, 1999; Argote & Ophir, 2002; Lewis, 2000).

Prior social network research has made a distinction between two types of network resources: *bonding* and *bridging* social capital. According to Gittell and Vidal (1998), who gave the original definitions of the terms, bonding social capital is “the type that brings closer together people who already know each other” (p. 15), and bridging social capital is “the type that brings together people or groups who previously did not know each other” (p. 15). Despite the increasing use and popularity of the two terms in network research, such conceptual classifications and definitions of the two terms have, however, not been widely adopted. The most widely accepted classification of the two terms focuses on group memberships. In this sense, bonding social capital refers to resources that people can obtain from within-group ties, while bridging social capital refers to resources that people can gain from their ties with people from the outside (Adler & Kwon, 2002; Kavanaugh, Reese, Carroll, & Rosson, 2005; Leonard, 2004; Putnam, 2000). Borrowing this terminology, as well as findings from related research, we further investigated how within-group and across-group
ties created by homophily in group membership (or the lack thereof) influence performance.

Past research has found that both bonding ties with people of the same group and bridging ties with people from outside the group are important sources for task-related resource exchange and non-task-related social support, albeit for very different reasons. Bonding social capital is important for knowledge creation because it reinforces group cohesion. New knowledge is more likely to be created only when people are comfortable expressing different ideas (Argote, 1999). However, successful brainstorming, a vital process for knowledge creation, does not happen automatically. If a group does not have the right dynamic that supports free exchange of ideas, expressing different opinions becomes difficult and can cause task and relational conflicts (Lewis, 2000). To avoid such conflicts, “group think,” or the convergence of thoughts among all team members, can happen when no one individual has sufficient fortitude to advance different proposals to the group (Arrow, McGrath, & Berdahl, 2000). Real breakthroughs in knowledge creation are therefore hard to achieve when everyone chooses to think the same way, especially in a public forum.

To create a comfortable environment for knowledge creation, it is important to build strong bonding ties within a group because bonding ties foster cohesion and trust among group members (Krackhardt, 1992). When individuals do not feel threatened or fear losing face, they are more comfortable exchanging ideas and, in turn, are more productive. One key indicator of successful knowledge creation is product performance (Argote, 1999), particularly in relation to such knowledge-intensive work as engineering design, in that better design creates better performance. Based on these arguments and findings, we formulated the following hypothesis:

**Hypothesis 6(a):** Bonding social capital derived from ties with people from the same group has a positive impact on performance.

In addition to bonding social capital, bridging social capital, which results from ties with people from outside the group, is also vital for knowledge creation. Kauffman (1993) proposes in his NKC model that the fitness of one organism is influenced not only by the density of links within the system, but also by the density of links without. Building on Kauffman’s research, organizational scholars have also come to the realization that groups should be perceived as complex adaptive systems that are reactive to changes in the external environment (Arrow, et al., 2000). That is, they are not self-sufficient entities that can be isolated from their operating environment (Baum, 1999; McKelvey, 1999; Yuan & McKelvey, 2004) because the environment can provide vital resources for groups’ survival and growth.

Burt (2001) has made a similar proposition in his effort to integrate Coleman’s closure argument and his structural hole theory. He argues that while it is important for groups to achieve high cohesion via closed, dense internal networks, it is also crucial for groups to exploit structural holes outside the group in order to gain access to unique, diverse information. Burt (1997) defines a structural hole as “separation
between nonredundant contacts” (p. 18), by which he means the absence of direct connections between any two network nodes. People filling such network holes can broker the relationships between these otherwise disconnected nodes and thereby obtain special information and control benefits. Since most groups are capable of achieving group cohesion, building and maintaining ties with people outside a group’s boundary lines becomes the key factor promising a group a competitive edge (Burt, 2001, p. 52). These boundary-spanning ties can serve as bridges connecting the focal group to more external resources. Past research has shown that boundary-spanning ties can provide an organization with useful information about legitimate practices of other organizations, help the organization better align its own strategies with critical directions in the industry, and thereby attain better performance (Geletkanycz & Hambrick, 1997).

In the context of the current research, we maintain that building strong connections with the environment is important for knowledge creation for two reasons. First, having ties outside the group can expand the focal group’s reach of ideas and keep the group connected with the environment. One key precondition for successful knowledge creation is having access to diverse sources of opinions (Argote, 1999; Argote & Ophir, 2002). While building a cohesive group is important, such cohesiveness can help fully exploit only internal resources. A group also needs to absorb external resources in order to expand its scope of knowledge and expertise. Previous network research has found that connections with others outside the group are important sources for new information. While weak ties are sufficient to locate well-codified knowledge, strong ties are necessary in order to learn tacit knowledge (Hansen, 1999).

Second, a group’s success in knowledge creation is evaluated by environmental standards. The environment provides not only rich sources of information, but also the calibrations with which to measure the performance of one particular system/group (Kauffman, 1993) in comparison to others. For instance, Yahoo! was considered one of the most popular search engines around the end of the last century, when the World Wide Web had just been created and when using human capabilities to evaluate the quality of web pages was still practicable. In recent years, however, Yahoo! has lost some of its market share to Google because (a) the World Wide Web has grown exponentially, beyond the ability of Yahoo! to service it, and (b) Google has offered a more powerful alternative to deal with the new challenges of the operating environment. In other words, good performance is relative. It is determined by the demands of the environment as well as by the quality of the product offered by the other players in the same competitive field. A top performer can easily lose his/her status when the operating environment changes or when competitors can offer a better solution. As a result, group members need to maintain strong connections with people outside the group in order to help the group remain competitive.

In our research, we thought it was important for students to maintain ties with other groups that had been assigned to work on the same engineering design project.
Our reasoning was based on the fact that students’ performance was evaluated by both the stand-alone quality of their respective engineering designs and by the comparative standings of their work in relation to other groups’ products. A group cannot earn a top grade if other groups produce more creative designs. Therefore, for these groups, building bridging ties was as important as building bonding ties. Based on these arguments, our final hypothesis is as follows:

**Hypothesis 6(b):** Bridging social capital derived from ties with people from different groups has a positive impact on performance.

**Methods**

**Design and Sample**

The data were collected from 32 students enrolled in a senior level engineering class. There were six female (19%) students and 26 (81%) male students in the class. Among these students, 23 (72%) were white and 9 (28%) were nonwhite. Students came from two universities (exactly 50% from each one) located in two different cities in the United States that are approximately 70 miles away from each other. Students attended lectures simultaneously co-taught by professors from both universities via a coded audio/video system. Students from the same campus had opportunities to meet face-to-face as frequently as they desired.

Five project teams were formed to design thermal protection systems for a next-generation space shuttle. Each team had between six to seven students, half from University A and half from University B. A special distributed learning support system was deployed to facilitate communications among students from both universities. The system supported email, instant messaging, group discussion, document sharing, and some other functionalities. Students were given high-end laptop computers with wireless connections for the duration of the academic year to ensure their access to the computer-supported collaboration system.

All surveys were introduced and collected through an independent “evaluation” group, which was comprised of a number of social scientists interested in studying group behavior in CMC situations. Members of this group were not involved in the preparation or delivery of course content. Students were informed that their participation was voluntary, that all responses would be strictly confidential and that the professor teaching the class would only be provided with the final data without students’ names.

**Measures**

**Homophily**

In the survey, we collected information on students’ gender, race, group assignment, and location (as indicated by the different universities they were attending). These attribute data were first imported into UCINET. Using the “Attribute – Exact Match” function in the program, these attribute data were transformed into square matrices measuring homophily in each of the four attributes. For instance, the
number “1” was assigned to a cell $C_{ij}$ of the gender homophily matrix, if node $i$ and $j$ were of the same gender. To test hypothesis 5, the four individual homophily matrices—all binary in value—were summed into one aggregate measure of overall homophily.

**Instrumental and Expressive Relationships**

Students were given a complete roster of all students enrolled in the distance learning class. Instrumental network relationships were measured by asking students to report the frequency of their task-related communications with other students in the class during a typical week, while expressive network relationships were measured by the frequency of non-task related communications. To deal with possible upward or downward biases in students’ self-reported data, the two matrices were symmetrized by taking the average of the responses from any pair of actors involved in a relationship (Borgatti, Everett, & Freeman, 2002). For instance, if person A reported communicating with person B four times a week, and person B reported communicating with person A eight times a week, the average frequency of communication between A and B would be six. The final symmetric matrix, which contained all averages, was used for further data manipulation and analysis.

**Bonding and Bridging Social Capital**

Bonding and bridging social capital were measured by evaluating the overall level of *within- and across-group connectedness* that any one group member had established with individuals of the same versus different groups in instrumental and expressive networks. Differing from some prior studies that focus on the binary existence vs. absence of network ties, we took into consideration both the existence and the value of these ties (frequency of communication) when calculating these connectedness measures. A person scored high on *within-group connectedness* when s/he communicated *frequently* with *many* people from the *same* group. A person scored high on *across-group connectedness* when s/he communicated *frequently* with *many* people from *different* groups. The technical procedures for obtaining these measures from UCINET 6.0 (Borgatti, et al., 2002) are presented in the Appendix.

*Within- and across-group* connectedness scores were calculated for both instrumental and expressive networks. The mean within-group connectedness was 17.97 (SD = 5.83) for instrumental relations and 8.13 (SD = 7.06) for expressive relations. The existing literature on bonding and bridging social capital does not make a distinction between expressive and instrumental relations. Therefore, we decided to combine instrumental and expressive relationships together when measuring bonding and bridging social capital. QAP correlation analysis showed that instrumental and expressive networks correlated at .41 ($p < .05$). This medium level of overlap between the two types of networks provided empirical support for data summation. Bonding social capital was then calculated by adding up the within-group connectedness in both the instrumental and expressive relations. The mean level of bonding social capital was 26.09 (SD = 10.97). The mean level of
across-group connectedness was 12.59 (SD = 10.22) for instrumental networks and 22.88 (SD = 22.74) for expressive relations. Bridging social capital was calculated by taking the sum of the across-group connectedness scores from both types of networks. The mean level of bridging social capital was 35.47 (SD = 28.14).

Performance
Performance was measured using each student’s final semester grade. Students’ final grades were jointly determined by professors teaching the class at both universities. The final semester grade was a weighted sum of a student’s individual performance on assignments, quizzes, and the group project. Individual group project grades were assigned by weighting the group project grade with an average peer assessment of each individual team member’s contribution to the group project in terms of time and team support. The mean final score for the semester was 86.71 (SD = 6.32). Descriptive statistics and zero-order correlations among these variables are summarized in Table 1.

Results
Hypotheses 1 through 4 investigated how race, gender, group membership, and location influenced the formation of instrumental and expressive network ties. Instrumental and expressive networks were regressed respectively on the four homophily matrices measuring similarities between any pair of people in the network along the four dimensions using the MQAP regression analysis supported by UCINET. The four homophily matrices were entered into the regression analysis simultaneously in order to obtain the pure effect of each matrix while controlling for the impact of the others. In addition, because the sample contained all senior students, to control for the effect of their past history of social interactions we have also included a matrix depicting their previous expressive network relations as a control variable. These expressive network relation data were collected at the beginning of the semester prior to the start of the group project. The results are presented in Table 2.

Table 1 Descriptive statistics and zero-order correlations

<table>
<thead>
<tr>
<th>Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Final grade</td>
<td></td>
<td>.37</td>
<td>.31</td>
<td>-.04</td>
<td>-.18</td>
<td>.39</td>
<td>-.16</td>
</tr>
<tr>
<td>2. Within-group (Instrumental)</td>
<td></td>
<td>-.44</td>
<td>.04</td>
<td>-.05</td>
<td>.82</td>
<td>-.03</td>
<td></td>
</tr>
<tr>
<td>3. Within-group (Expressive)</td>
<td></td>
<td></td>
<td>.02</td>
<td>-.02</td>
<td>.88</td>
<td>-.01</td>
<td></td>
</tr>
<tr>
<td>4. Across-group (Instrumental)</td>
<td></td>
<td></td>
<td></td>
<td>.37</td>
<td>.03</td>
<td>.66</td>
<td></td>
</tr>
<tr>
<td>5. Across-group (Expressive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.04</td>
<td>.94</td>
<td></td>
</tr>
<tr>
<td>6. Bonding social capital</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td>7. Bridging social capital</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>86.71</td>
<td>17.97</td>
<td>8.13</td>
<td>12.59</td>
<td>22.88</td>
<td>26.09</td>
<td>35.47</td>
</tr>
<tr>
<td>Std. deviation</td>
<td>6.32</td>
<td>5.83</td>
<td>7.06</td>
<td>10.22</td>
<td>22.74</td>
<td>10.97</td>
<td>28.14</td>
</tr>
</tbody>
</table>
Analysis showed that prior expressive network ties had significant influence on the pattern of relationships in both the new instrumental relations and new expressive relations. Its standardized regression coefficient was .40 \((p < .05)\) for instrumental relations and .70 \((p < .05)\) for expressive relations. Alone, it explains 16% of variance in instrumental relations and 47% of variance in new expressive relations.

After controlling for the impact of prior social interactions among subjects, neither gender nor race had significant impact on the formation of network ties for either the instrumental or expressive relationships. The standardized regression coefficient for the impact of gender on instrumental relations was .01 \((p > .05)\). For expressive relations, the standardized regression coefficient was -.02 \((p > .05)\). The standardized regression coefficient for the impact of race on instrumental relations was -.01 \((p > .05)\). The standardized regression coefficient for the impact of race on expressive relations was -.04 \((p > .05)\). Therefore, hypotheses 1(a) and 1(b), and 2(a) and 2(b), were all rejected.

For socio-contextual characteristics, after controlling for the impact of prior expressive relations, the standardized regression coefficient for the impact of group assignment on instrumental relations was .50 \((p < .05)\). The results lent strong support to hypothesis 3(a) that group assignment would significantly influence the formation of instrumental ties. The standardized regression coefficient for the impact of group assignment on expressive relations was .07 \((p = .06)\), only marginally significant. Hypothesis 3(b), which predicted a significant impact of group assignment on the formation of expressive ties, was, therefore, not supported.

Hypothesis 4(a) predicted that instrumental relationships were more likely to happen among people of the same location. After controlling for the impact of prior expressive relations on the formation of new ties, the standardized regression coefficient for the impact of location on instrumental relations was .11 \((p < .05)\). Therefore, hypothesis 4(a) was supported. Next, hypothesis 4(b) predicted that expressive relationships were also more likely to happen among people in the same location. The standardized regression coefficient for the impact of location on expressive relations was .09 \((p < .05)\). Therefore, hypothesis 4(b) was also supported.

With all four homophily matrices included in the analysis, together with the control matrix, the \(R^2\) (the proportion of the variance explained by the predictor variables) was .42 for instrumental networks with an increase of .26. For expressive networks, the \(R^2\) for the final model was .49 with an increase of .02. Because current MQAP regression analysis does not allow testing of nested hierarchical models, changes in \(R^2\) could not be tested for statistical significance. In network analysis, a small \(R^2\) or changes in \(R^2\) are not uncommon; however, the comparatively low \(R^2\) change for expressive network relations showed that future investigation is warranted to determine what factors did play a role in influencing the formation of new social support relations above and beyond preexisting ones.

Hypothesis 5(a) and 5(b) predicted greater likelihood of forming instrumental and expressive ties, respectively, when people share a high number of traits in common. Regressing the instrumental network matrix on the overall homophily
Table 2 Summary of UCINET MQAP analysis for hypotheses 1(a) - 5(b)

<table>
<thead>
<tr>
<th>Hypothesis1(a) – 4(b)</th>
<th>Instrumental Network</th>
<th></th>
<th>Expressive Network</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standardized β</td>
<td>p</td>
<td>Standardized β</td>
<td>p</td>
</tr>
<tr>
<td>Control: Prior expressive ties</td>
<td>.40</td>
<td>&lt;.05</td>
<td>.37</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Gender</td>
<td>.01</td>
<td>&gt;.05</td>
<td>-</td>
<td>.02</td>
</tr>
<tr>
<td>Race</td>
<td>-.01</td>
<td>&gt;.05</td>
<td>-</td>
<td>-.04</td>
</tr>
<tr>
<td>Group assignment</td>
<td>.50</td>
<td>&lt;.05</td>
<td>-</td>
<td>.07</td>
</tr>
<tr>
<td>Location</td>
<td>.11</td>
<td>&lt;.05</td>
<td>-</td>
<td>.09</td>
</tr>
<tr>
<td>R²</td>
<td>.16</td>
<td>.42</td>
<td>.47</td>
<td>.49</td>
</tr>
<tr>
<td>R² change</td>
<td>.26</td>
<td>.02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hypothesis5(a) – 5(b)</td>
<td>Homophily</td>
<td>.31</td>
<td>&lt;.05</td>
<td>.13</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>.09</td>
<td>.02</td>
<td>-</td>
</tr>
</tbody>
</table>
matrix, which was derived by summing up the four individual homophily matrices together, the resulting standardized regression coefficient was .31 ($p < .05$). The significant coefficients showed strong support for hypothesis 5(a), although the $R^2$ was not substantial ($R^2 = .09$). Regressing the expressive network matrix on the overall homophily matrix, the resulting standardized regression coefficient was .13 ($p < .05$). Although the standardized regression coefficient was smaller in magnitude in comparison to instrumental relationships, it was significant, supporting hypothesis 5(b). In this case, the $R^2$ was .02.

Hypothesis 6(a) and 6(b) predicted that both bonding and bridging social capital would have a positive influence on performance, albeit for different reasons. To test these two hypotheses, students’ final semester grades were regressed on measures of bonding and bridging social capital simultaneously. The standardized regression coefficient measuring the impact of bonding social capital on performance was significant ($\beta = .39$, $t(30) = 2.35$, $p < .05$). Therefore, hypothesis 6(a) was supported. The standardized regression coefficient measuring the impact of bridging social capital on performance was, however, nonsignificant ($\beta = -.16$, $t(30) = 2.35$, $p > .05$). Therefore, hypothesis 6(b) was rejected. The overall $R^2$ for the regression analysis was .16, where $F(1, 30) = 5.53$, $p = .03$.

**Discussion**

**Contribution**

The theory of homophily predicts that the likelihood of social interaction increases among people who share similar traits. The theory has received wide support in diverse contexts, and homophily has been found in many studies to be one of the major forces driving the emergence and maintenance of network ties (McPherson, et al., 2001). The current study aimed to contribute to existing homophily research in several ways. First, the current research explored homophily in terms of both demographic (gender and race) and social (group assignment and location) characteristics. Studying homophily along multiple dimensions is important because people self-categorize themselves using multiple traits. At work, social characteristics may easily suppress demographic ones in defining who we are because common tasks, common organizational affiliations, etc. bring people of different genders and races together. Second, following McPherson, et al.’s call, the current research explored homophily in more than one type of relationship. While traditional social network studies tended to downplay the importance of network content (Wellman, 1988), recent research has placed greater emphasis on the differences in network properties across different types of relationships (e.g., Burt, 1997; Podolny & Baron, 1997). Following Ibarra’s classification framework, we made a distinction between task-related instrumental relations and non-task-related expressive ones and examined homophilous relations in both types of networks. Third, the current research is one of the few studies that has explored homophily of ties in distributed teams. Aided by fast development in information and communication technologies, distributed
teams have become a popular form of organizing in both the workplace and in schools (e.g., distance learning). However, information technologies only provide a platform for communication among distributed team members. It is social relations that support real collaboration (Barab, et al., 2004a; Yuan, et al., 2006). It is important, therefore, to explore what forces drive the formation of network ties in distributed teams.

Finally, building on past research on social capital, the current research explored how bonding ties with people of the same group and bridging ties with people from different groups influenced performance. Homophily unifies and fragments a network simultaneously. While past homophily research focused more on ties with similar others, bonding and bridging social capital research expands the scope of existing homophily research. It also provides a rich vocabulary and a conceptual framework with which to explore the benefits of having bridging ties with those who do not share common group membership in addition to bonding ties with those who are in the same group. The current research further explored how both types of ties facilitate knowledge creation and improve group performance, albeit for very different reasons.

To test our hypotheses on homophily and the impact of bonding and bridging ties on performance, we used a sample consisting of 32 advanced engineering design students assigned to work together on a complicated task. The purpose of this advanced class was to replicate what students would encounter in actual work settings and organizations (NASA). Therefore, we believe that the results of this study can be generalized beyond a college student population. However, the generalizability of our research is limited because the results were obtained from one single sample.

Counter to our predictions, the impact of racial and gender homophily was non-significant for either instrumental or expressive ties. However, homophily in social characteristics demonstrated a strong impact. Specifically, location was found to have a significant influence on the formation of both instrumental and expressive relations, and group assignment significantly influenced the formation of instrumental ties. These results support Bradner and Mark’s (2002) earlier findings that location had a significant influence on social interactions in computer-mediated communications, and that the provision of different communication media—including rich media such as videoconferencing—to connect people from different locations did not shorten the “social distance” between them.

In a review of existing research on homophily, McPherson, et al. (2001) observed that racial homophily has the strongest impact on the formation of ties, followed by age, religion, education, occupation, and gender. They found these results to be fairly robust despite skewed distribution of the samples in terms of both gender and race. Like many homophily studies done on real organizations, our sample had more males than females, and more whites than nonwhites. In other words, our sample was representative of the typical demographic distributions of many organizations, particularly those high-tech companies engaged in intensive knowledge work. The
similarity in the distribution of research samples between existing homophily studies and our study increased our confidence in concluding that in distributed teams, homophily in social characteristics was indeed more important than homophily in demographic ones in influencing the formation of network ties. Distance between people can erode the uniting power of racial and gender homophily in the formation of both task-related and non-task related ties. Moreover, distance can also make the development of expressive ties among team members more difficult. Also counter to our predictions, we failed to observe a strong alignment of expressive ties along group assignments.

In testing the impact of bonding and bridging social capital on performance, we combined instrumental ties and expressive ties into one composite measure of how connected people were with their team members in terms of (a) the number of ties and (b) the frequency of communication with each tie. We predicted that both bonding ties and bridging ties would significantly influence performance because bonding ties could increase group cohesion and because bridging ties could bring in new information to facilitate knowledge creation. The results showed that only bonding ties had significant influence on performance. This suggests that in distributed teams, building a cohesive team is of primary importance.

Homophily and the Development of a Learning Community

In recent years, research on both educational learning and organizational learning has experienced a change in orientation from focusing on individual knowledge acquisition to focusing on the development of a learning community which has been shown to facilitate collective, participatory learning (Barab, Kling, & Gray, 2004b; Lave, 1997; Lave & Wenger, 1991). Scholars believe that a learning community can produce more fruitful learning experiences when its members can learn from each other through social interaction. Developing such a community requires free formation of network ties among all members of the community regardless of individual differences in demographic, social, or geographic backgrounds. But homophily is a double-edged sword: It unifies people sharing the same characteristics, but it divides people with different characteristics. In the context of community development, homophily simultaneously builds and divides a community. Acknowledging the double-edged nature of homophily has significant strategic implications for community development. Traditionally, people tend to focus exclusively on the rosy side of homophily, neglecting its potentially divisive power. As a result, a large, integrated community is hard to actualize when people of similar traits form cliques among themselves and are not encouraged to build connections across cliques.

In the present study, we investigated homophily in terms of both demographic and social characteristics. We found that in distributed teams, socio-contextual homophily exerted stronger influence than did demographic homophily on the formation of network ties. In terms of developing a learning community, this result is both encouraging and unsettling. Mollica, et al. (2003) found that racial homophily formed fast and persisted for a long time in co-located teams. Although
developing ties with people of the same gender or race may provide better emotional support, building teams mainly of strongly-tied Asians or African-Americans, or of males or females, should be discouraged under both co-located and distributed conditions. Effective team functioning requires that members have the capability of reaching out to diverse sources of information, regardless of gender or race. While racial or gender homophily is found to be prevalent in co-located teams, it is encouraging to see that in distributed teams, team members could reach out to diverse resources, regardless of race or gender.

However, the results from our research were also unsettling because the learning community was fragmented by location and group boundaries. There is no doubt that rapid development in communication technologies has made it possible for distributed team members to overcome some geographic constraints. However, the results from our research clearly confirmed Barab, et al.’s (2004a) concerns that (a) the development of a virtual learning community requires more than providing technical affordance, and (b) the development of the next generation of communication technologies should recognize the factors of both usability and sociability. When designing the audio/video platform used to support distance learning and group collaboration in our research, we focused more on supporting within group collaboration than on supporting across-group, within community communication. We also focused more on supporting task-related functionalities than on non-task related functionalities. It would be interesting to explore whether making the platform a fun place to be would boost social interactions. To address these issues, we propose to design and implement an intervention program that has the potential to facilitate community-wide social interactions. A more detailed discussion of this is provided in the section on directions for future research.

Limitations
One major limitation of the current study was the lack of completely co-located teams as comparisons. Many existing homophily studies have used co-located teams. For instance, Mollica, et al. (2003) found in their study of newcomer groups that racial homophily formed quickly among African-American minority students and persisted throughout the whole academic year, despite the school administrators’ consistent efforts to encourage across-race interaction. However, in our study, we did not find any significant impact of race on the formation of network ties. The differences in findings between the two studies might have been caused by the differences in samples. Mollica, et al. (2003) studied co-located students, while we studied distributed ones. Had we included co-located teams as control groups in our research, we could have determined more conclusively whether or not the differences in the findings between the two studies were indeed caused by differences in samples.

The second limitation of the current research was the small sample size. Although the sample of 32 students was not small in terms of collecting complete network data, statistical tests connecting network measures with other measures of
nodal attributes could have been made stronger if a larger sample size had been available.

**Directions for Future Research**

In the present study, we mainly focused on the theory of homophily to explain the formation of network ties. While homophily has been found to be one of the major reasons explaining the emergence of ties (McPherson, et al., 2001; Mollica, et al., 2003), other factors, for instance, self-interested expertise seeking and resource exchanges, can motivate people to initiate connections. Monge and Contractor (2003) summarized five families of social theories predicting the creation, maintenance, and dissolution of network ties, including the theory of homophily, theories of self/collective interests, and theories of resource exchange. Future research should, therefore, include more finely tuned questions to elucidate how other factors drive the emergence of network ties.

Second, in the current study, we found large differences in connectedness between instrumental and expressive relations—the intermediate measurements in the process of creating the final measurement for bonding and bridging social capital—in terms of both group assignment and location. For instance, within-group connectedness averaged 17.97 in *instrumental ties*, but only 8.13 in *expressive ties*; and across-group connectedness averaged 20.03 in *instrumental ties*, but 28.88 in *expressive ties*. This leads us to raise the following questions. Does the level of overlap between task-related instrumental ties and non-task-related expressive ties matter for knowledge creation? Should people keep their jobs and emotions separate? Previous small group research on the issue of group cohesion seems to favor developing ties that serve multiple purposes (Arrow, et al., 2000), regardless of whether the teams are co-located or distributed. Recent knowledge management research also advocates the development of strong ties, which usually serve multiple purposes, to support the sharing of tacit knowledge (Hansen, 1999). The current study seems to indicate, however, that groups can function equally well when task and socialization ties are kept separate. Future research needs to further identify and define the contingency factors that explain under what kind of conditions overlaps in network ties of different contents—expressive versus instrumental ties—really matter.

Third, the research team designed and implemented a computer-supported collaboration system to facilitate communication among distributed team members. Data analysis showed that location exerted the strongest impact on the formation of both task-related instrumental ties and non-task related expressive ties, over and above the impact of gender, race, and group assignment. In a related analysis, Yuan, et al. (2006) found that the impact of location was predominant and persisted for nearly a year. It is clear that future research should explore more about what intervention tools can be implemented to help people overcome the constraints of location. One possible tool would be an expertise/common interest identification system that helps members locate potential contacts with whom they can socialize either to address their individual knowledge needs or to share their common interests. We
believe that such a system can facilitate social interactions in a distributed community because the system provides not only technical functionalities to support communication, but also interesting topics and reasons to communicate. Without such a system it may take people from different locations a long time to identify common interests and/or topics for small talk, and an even longer time to develop social connections and a feeling of bondedness with each other and with the community.

Fourth, following Ibarra’s classification framework, we made a distinction between two types of network relationships: task-related instrumental ties and non-task-related expressive ones. Future research should seek to make finer distinctions among different types of instrumental and expressive relations. For instance, instrumental ties can include ties for expertise/advice seeking, knowledge exchange, etc., and expressive ties can include those for social activities, friendship, emotional support, etc. It would be interesting to explore whether different instrumental and expressive ties demonstrate the same pattern of network properties in terms of size, average strength of ties, reachability, network density, and so forth.

**Conclusion**

While information and communication technologies are vital for the effective functioning of distributed teams, network ties are more important because they provide the actual bond that helps to overcome geographic constraints. Therefore, uncovering the underlying mechanisms that drive the formation of network ties warrants increased research attention. The current research contributed to existing homophily research by studying homophilous ties in multiple relations (in both instrumental and expressive ties), along multiple dimensions (including gender, race, location, and group membership), and in distributed teams (versus co-located ones as found in most existing homophily research). The results showed that in distributed teams, homophily in social characteristics was more important than either racial or gender homophily in driving the formation of network ties.

**Acknowledgments**

This research was supported by a NASA research grant (NASA/NYS 3533168) and a Seed Grant (2006) from the Institute of Social Sciences from Cornell University. The authors thank Barry Davidson and Tony Ingraffea for their work on the design and data collection in the project.

**References**


Appendix

To Obtain Measures of Connectedness with People of the Same Group:
Step 1: Using Attribute → Exact Match function in UCINET 6.0, a proximity matrix was created for the attribute variable of group membership and location. In the group proximity matrix, for instance, a number of 1 in row 3 and column 5 means that Actor 3 and Actor 5 belonged to the same group.
Step 2: Multiply the specific proximity matrix with respective instrumental and expressive network matrices. The diagonals of the resulting matrices, which represent within-group connectedness in instrumental and expressive relations, were extracted from the matrices and copied and pasted into SPSS to create a measure for bonding social capital.
To Obtain Measures of Connectedness with People of the Different Group:
Step 1: Reverse-code the co-group matrix created in the above-mentioned procedure so that a number of 1 in row 4 and 5 means that Actor 4 and 5 were not in the same group.
Step 2: Multiply the reverse-coded matrix with instrumental and expressive network matrices. The diagonals of the resulting matrices, which represent across-group connectedness in instrumental and expressive relations, were extracted from the matrices and copied and pasted into SPSS to create a measure for bridging social capital.

About the Authors
Y. Connie Yuan is an Assistant Professor in Communication and Information Science at Cornell University. Her research interests include knowledge management, social network analysis, and computer-supported collaboration in distributed teams.

Address: 308 Kennedy Hall, Department of Communication, Cornell University, Ithaca, NY 14853 USA

Geri Gay is the Director of the Human Computer Interaction group and a professor and chair in the Department of Communication at Cornell University. She is also a professor in Computer and Information Science at Cornell. Her research interests focus on cognitive and social issues for the design and use of interactive communication technologies. Her research examines navigation issues, knowledge management, mental models and metaphors, social networks, collaborative work and learning, and design.

Address: Human Computer Interaction Laboratory, Information Science Bldg., 301 College Ave., Ithaca NY 14850 USA