Analyzing Online Teacher Networks: 
Cyber-Networks Require Cyber-Research Tools

Mark S. Schlager,¹ Umer Farooq,¹ Judith Fusco,² Patricia Schank,² & Nathan Dwyer³
¹Microsoft  ²SRI International  ³Atlassian

ABSTRACT
We argue that new frameworks, tools, and techniques are needed to understand and maximize the benefits of emerging online social networks of K-12 educators (and students). We base our argument on conceptual and methodological obstacles inherent in existing research approaches that severely limit theory building and empirical exploration of learning in online social networks. We present preliminary data exploring bridges—network members who belong to two or more groups—to illuminate both the power and limitations of current tools and techniques for studying online social networks. Using extant data from Tapped In, a large online network of K-12 education professionals, we show that bridges participate more than nonbridges in both synchronous and asynchronous communication modes. These preliminary findings raise a number of fundamental questions that, we argue, are beyond the capability of most education researchers and evaluators to address rigorously and cost-effectively. We then propose a research agenda designed to create and validate a new generation of research tools and techniques that enable more comprehensive and useful analysis of the massive amounts of heterogeneous data generated automatically by online teacher networks. Our long-range goals are to help researchers ask more incisive and convergent research questions and help school leaders and teachers support, learn, and collaborate with one another more effectively in cyber-enabled professional communities. We conclude with a discussion of the deeper questions these data raise and the challenges we must overcome to answer them.

INTRODUCTION
We are in a period of transition to a world in which all human networks will be mediated through cyber-enabled (e.g., web, Internet2, and mobile) technologies (Tapscott & Williams, 2006).¹ Information and communication technologies that power complex social systems are rapidly scaling up and becoming integral to daily life around the globe. The ubiquity of online social networking in popular culture and the business sector heralds the promise of online social networks for education (Resnick, 2002; Leana & Pil, 2006; Penuel & Riel, 2007; National School Boards Association, 2007; National Science Foundation, 2005). The popularity of cyber-enabled social networking among youth and teachers of the net generation is undeniable. The Pew Internet and American Life project (Lenhart, Madden, Macgill, & Smith, 2007) reports that

¹ For purposes of this article, we use the terms cyber-enabled and online interchangeably. Online is used more frequently in the traditional literature, while cyber-enabled is gaining popularity. Neither term implies that interaction need be exclusively technology mediated or that face-to-face interaction is not desirable.
more than half of all of online American youth ages 12–17 use online social networking sites. A more focused study by the National School Boards Association (2007) reports that 96% of students with online access have used social networking technologies, and more than 50% talk online specifically about schoolwork. Regarding teachers, the NSBA reports that in districts where structured online professional communities exist, participation by teachers and administrators is quite high. These reports and others (e.g., Atkins et al., 2003; Computing Research Association, 2005; Fulton, Yoon, & Lee, 2005; Jenkins, Clinton, Purushotma, Robinson, & Weigal, 2006) envision a future in which cyber-enabled social networks become a central context for student and teacher learning and a catalyst for instructional improvement.

To harness the power of this societal transformation to serve teaching and learning, we need to understand the phenomenon and unlock the value it holds. The NSBA study tells an optimistic but cautionary tale, reminding the research community that education administrators and policymakers will permit access to social networking only when the strong educational value and purpose of such networks can be demonstrated. Research must help education communities convert the current enthusiasm for online social networking into reliable evidence of how, when, and why online social networks do, and do not, advance learning, and we must develop scalable and replicable models that maximize the value and benefits of emerging social networking models and technologies.

This paper presents a long-term research agenda aimed at meeting this challenge through the development of new analytical frameworks and more integrative and automated methods and tools that can rapidly mine and reliably analyze the massive amounts of data generated by cyber-enabled social networks. We begin our exploration with a basic question in social network analysis (SNA): What constitutes a meaningful relation or tie between individuals? Our goal is to answer this question using a combination of traditional SNA methods and new, more scalable methods that use automatically recorded interaction data.

Applying traditional social network analysis methods can be problematic in large-scale cyber-enabled social networks, which typically do not have a well-defined structure and are therefore difficult to put boundaries around. Moreover, SNA methods are limited in their ability to identify the precursors of and other enabling factors for social capital or trace how social capital is fostered and leveraged (Leana & Pil, 2006). Cyber-enabled social networks offer the ability to capture and analyze a more complete and objective record of peoples’ actions and interactions automatically over time. However, digital interactions are not simple to mine or interpret. In addition, interaction data are missing a key ingredient of SNA: judgments about the strength of social ties.

**RESEARCH CONTEXT: PROFESSIONAL NETWORKS IN TEACHING**

The context of our research is analysis of professional networks in teaching (Lieberman, 2000), such as professional learning communities (PLCs; Dufour, 2004; Stoll, Bolam, McMahon, Wallace, & Thomas, 2006), teaching communities of practice (CoPs; McLaughlin &
In press, Journal of Teacher Education.

Talbert, 2001), and networks with other labels. Our agenda is based on three research-supported propositions:

1. Professional networks among educators are a growing trend, and research is beginning to show the important roles they play in improving education.
2. Cyber-enabled networks hold great promise for supporting teachers’ development of new knowledge and practices, but their activities and value are not well understood.
3. A major obstacle to realizing the promise of cyber-enabled teaching networks reliably and at scale is a lack of appropriate frameworks, tools, and techniques for studying them.

Social Network Research in K-12 Teaching

Education practitioners and researchers have been creating and studying teacher communities for more than a decade (see, for example, DuFour, 2004; McLaughlin & Talbert, 2001; Grossman, Wineburg, & Woolworth, 2001; Stein, Silver, & Smith, 1998; Stoll et al., 2006). Kruse and her colleagues (1995) described the key components of professional communities and suggested that PLCs become “a major rallying cry among reformers, rather than a secondary whisper.” Historically, research on teacher networks has been dominated by richly descriptive but intrinsically localized ethnographic accounts of school- or district-based networks. Variations on the concept have also been central to professional development and school reform interventions (Loucks-Horsley, Hewson, Love, & Stiles, 1998; Putnam & Borko, 2000; Stein et al., 1998; Stein & Coburn, 2005). This work has shown that teacher networks, in different forms, are effective alternative and supplemental interventions to traditional workshops and institutes for learning content and pedagogy.

A parallel body of research has focused on online teacher communities (Schlager, Fusco, & Schank, 1998, 2002; Farooq, Schank, Harris, Fusco, & Schlager, 2008; Renninger & Shumar, 2002; Barab, Kling, & Gray, 2004), inspired by both the descriptive literature on face-to-face teacher communities and the quantitative social network research literature outside K-12 education (e.g., Koku & Wellman, 2004; Wellman & Gulia, 1999; Hinds & Kiesler, 2002; Sproull & Kiesler, 1991; Carroll & Rosson, 2003; Hine, 2000; Lampe, Ellison, & Steinfield, 2007; Preece & Maloney-Krichmar, 2003). These studies have shown that, online, teachers can, under certain conditions, interact more frequently, build more diverse networks, and gain more equitable access to human and information resources not available locally. In addition, the quality of dialogue online has been shown to be equivalent to and in some cases better than face to face. These studies have been limited, however, by their focus on small groupings within larger communities (selected because they are part of an intervention or display some “interesting” attribute), which may not represent the larger community.

Such accounts serve to show the field what is attainable under particular conditions, interventions, or contexts, but we are still unable to rigorously measure their value, much less predict, guide, or replicate results reliably or at scale. Conspicuously missing from the literature are the many more failures to document meaningful outcomes, achieve sustainability, and scale

---

2 We deliberately chose the more generic term professional network for our purposes to avoid philosophical arguments about what constitutes a PLC, CoP, and the like. We intend this work to apply equally well to any flavor of teacher network and to student networks as well.
beyond an initial cohort. Fullan (2006) concludes his assessment of the PLC literature with the recommendation that education researchers “should take them more seriously.”

One reason to take professional networks seriously is that a new line of research is beginning to show that “networks matter for school change” (Penuel & Riel, 2007). Recently, quantitative SNA methods have made their way into teacher learning literature in the form of social capital research in face-to-face contexts. SNA (Wasserman & Faust, 1994; Wellman & Berkowitz, 1998) is a mathematical approach to analyzing the interactions and relations of entities (in our case, people) in a network and creating a visual representation, called a sociogram, mapping the structure and strength of relationships. When the goal is to be able to make changes to improve information flow among members of a network, the sociograms are extremely valuable as a tool to show interactions and relationships explicitly (Cross, Parker, & Borgatti, 2002).

Using SNA, researchers can analyze a network at multiple points in time to understand how expertise, information, and resources flow between the individuals in the network. A growing body of work (Frank, Zhao, & Borman, 2004; Frank & Zhao, 2005; Leana & Pil, 2006; Parsad, Frechtling, May, & Shapiro, 2007; Penuel & Riel, 2007; Penuel, Sussex, Korbak, & Hoadley, 2006) shows that the study of educators’ social ties, the resources they obtain from them (social capital), and the communication tools that facilitate the flow of expertise and resources can help us better understand how to support and promote school change. This body of work is breaking new ground in correlating the existence of social capital with outcomes that matter to practitioners and policymakers (e.g., higher student test scores, more helping behaviors, and successful implementation of reform practices).

The Promise of Cyber-Enabled Professional Networks in Education

Research on face-to-face social networks (cited above) has clear and strong implications for demonstrating the purpose and value of cyber-enabled networks in teaching. For example, research suggests that it is not necessarily important for each individual to be connected to every other person in an organization; most important is to have people connect to the right experts for the information they need. Because teachers have a limited amount of time, fostering the correct ties (often between novices and experts) is important. Limited resources are best spent giving those with the expertise and the propensity to exchange information more time to connect with others. Teachers who are connected weakly to others outside their tight group of colleagues (alternatively called bridges, boundary-spanners, or brokers) are beneficial to the network because through the weak ties, they are connected to and spread different ideas and resources. A related finding is that teachers benefit both when they talk with teachers similar to themselves (e.g., teachers of the same grade or discipline) and teachers with whom they have less in common.

All these findings are, in principle, more readily, reliably, and cost-efficiently achieved at scale through networks that are cyber enabled. If the findings above can be replicated in cyber-enabled networks, teachers everywhere will benefit. New teachers, isolated teachers, and those in underserved schools, where expertise and resources are least available, will benefit the most. If we can rigorously define the structures and trace the outcomes of face-to-face teacher networks, can we not do the same for cyber-enabled networks? The answer, it turns out, is that it is harder than one might think (Garton, Haythornthwaite, & Wellman, 1997).
The Need for New Social Network Analysis Tools and Techniques

Research on cyber-enabled social networks in K-12 education is nascent, being informed by work in progress in related fields. For example, Haythornthwaite (2001) offers a theory of tie strength, taking into account multiple media (e.g., e-mail, online systems, chat, and instant messaging), the uses they afford, and the way they are used in relation to tie strength. Her study of online interactions suggests that cyber-enabled interactions foster the development of weak ties. Weak ties are important for bringing in new ideas and connecting people so information can travel through a network. Others have created sociograms using data collected automatically by different technological systems (Wellman, 2001; Tyler, Wilkinson, & Huberman 2003; Park, 2003; Park & Thewall, 2003; Herring et al., 2005). For example, a recent study of 7 million cell phone users shows just how important weak ties are for maintaining the integrity of a network and helping information flow (Onella et al., 2007).

When used in conjunction with qualitative or ethnographic accounts, SNA techniques help show where information is and is not flowing and suggest where interventions might improve information flow. For example, supporting teachers in professional development requires being able to know who needs support, who has resources that are valuable, and who has the inclination to support others. This laborious approach can be problematic in large-scale cyber-enabled social networks. Cyber-enabled networks typically do not have a well-defined structure and are therefore difficult to put boundaries around; they can be an offshoot of a face-to-face network or an amalgam of many other networks; the reliable contact with members needed to apply SNA methods is difficult to achieve. Moreover, the methods are limited in their ability to identify the precursors and other enabling factors for social capital or trace how social capital is fostered and leveraged (Leana & Pil, 2006).

In traditional social network analysis, researchers map the structure of the network—who is connected to whom (ties) and how strong the relationship is (weights). This is commonly done by asking every member of a well-defined network (e.g., all teachers in a school or district) a set of questions that require them to make judgments about their relationships with others (e.g., whom they trust, go to for help, are closest to) in an interview or survey. This allows the people surveyed to choose their responses and weight their view of the importance of the relationships. Surveying people takes time and requires a response from every member of the group being studied. To measure changes over time, this laborious process must be repeated.

Cyber-enabled social networks offer the ability to capture and analyze a more complete and objective record of peoples’ actions and interactions automatically over time. Automatically collected data alleviates the problem of response rate (Garton et al., 1997) and could potentially speed up the creation of sociograms and lower the cost when tools are available to extract and process the data. But digital interactions are not simple to mine or interpret. Interaction is distributed across space, time, and media, and the data obtained through instrumentation come in a variety of formats. Even basic interpretation of these data can require tracing thousands of individual paths of activity and points of intersection across multiple tools. Finally, interaction data are missing the second key ingredient of SNA: judgments about the strength of social ties. In an online situation, it cannot be assumed that a discussion between two teachers, even on a topic of great importance, constitutes a meaningful relationship that they would report on a questionnaire designed for SNA. Yet the actual interactions between people may provide more objective and reliable measures of different kinds of social ties than retrospective self-report. We know that relationships develop online, but we do not know how to convert automatically
collected data into the essence of relationships in a network. Whether we can understand what constitutes a meaningful relationship or tie between individuals from automatically recorded interaction data and create reliable sociograms from such data are important unanswered research questions (Carley, 2003). To realize the promise of cyber-enabled social networks in education, we must build a new generation of methods and tools that bridge multiple research traditions and types of data.

**MOTIVATING A NEW RESEARCH AGENDA**

The growing body of SNA research on face-to-face teacher communities is beginning to show compelling evidence of how teachers’ social ties, the resources they obtain from them (social capital), and the communication tools that facilitate the flow of expertise and resources can help support school reform interventions, promote peer mentoring, and even improve student test scores. Cyber-enabled educator networks have the potential to achieve similar (some advocates claim more powerful) results at greater scale by overcoming the constraints and limitations of local networks, thereby enabling teachers everywhere, particularly where expertise and resources are least available, to gain access to the benefits of professional networks. Yet little hard evidence exists that cyber-enabled teacher networks have the same properties or can achieve the same outcomes.

**Goldmines of Digital Social Network Data**

Our approach builds on our prior efforts to understand the processes and outcomes of online communities through a unique asset that, in principle, should enable us to detect highly robust and reliable patterns of data. Ten years ago, we began the Tapped In (www.tappedin.org) online teacher community research test bed program with funding from the National Science Foundation. Our premise was that researchers cannot understand how mature networks of educators function online, and the value they provide, without a mature, well-functioning online network of teachers to study. (Then, as today, most online teacher communities never reached the point of maturity.) Ten years later, Tapped In is a large (20,000 members annually), complex (500+ active groups annually), self-organizing (run by community volunteers), and actively evolving global network of education professionals. Nearly 40% of members are K-12 teachers, 9% are graduate students, 6% are university faculty, 5% are preservice teachers, and the remaining 40% represent 27 other education-based occupations (librarian, principal, etc). Membership is free for all individuals and small groups.

One of the technological features the Tapped In community platform offers to all members is the ability to create and support groups. Groups can be private (invitation only), moderated (join by request or invitation), or public and may be open to all Tapped In members or just members affiliated with Tenant organizations. More than 800 groups currently are in Tapped In, ranging in size from two to hundreds of members. All Tapped In groups have text chat, a threaded discussion board, a repository area to share web bookmarks and files, and other tools to support

---

3 By *hard*, we mean evidence that is robust, reliable, and replicable.

4 Tenants are organizations that pay an annual fee to set up an entire building on the Tapped In virtual campus.
group work. Some groups use chat almost exclusively as their means of communication, while others rely on the threaded discussion board to keep in contact. Of course, there are groups that use both synchronous and asynchronous communication.

Within groups, people take on different roles in support of the group’s purpose. The first role is the owner of the group. This person creates the group and invites people to join. The owner can either perform all the moderation duties or enlist others to help. Moderators are usually given more “powers” within a group. For example, moderators can approve people who wish to join a group, change the look and feel of the group room, delete items, and set up messages that are part of the group room. Members of a group have the powers the owner and moderators assign to them. This flexibility on permissions allows group rooms to work in diverse situations, for example, in cases where colleagues are collaborating (democratic situation) and in class situations where the professor is the leader (hierarchical situation). As the network has evolved, it has generated several gigabytes of social interaction data offering an unprecedented lens into self-directive, naturally evolving teacher networks that typical online teacher professional development projects—with fixed boundaries, goals, and time spans—cannot offer.

**Finding Needles in Social Network Haystacks**

Over the years, our research team has conducted traditional ethnographic studies of small groups (Derry, Gance, Gance, & Schlager, 2000; Gray & Tatar, 2004; Schlager et al., 2002) to illustrate what is possible for teachers to accomplish online. For example, in some of our early work (well before instant messaging became ubiquitous), Derry and colleagues demonstrated that it is possible for a research group to engage in productive discourse through online text-based chat, and Schlager and colleagues demonstrated that it is possible for teachers to have meaningful professional development conversations in text-based chat. Gray and Tatar demonstrated that it is possible for one member of a large online community to move from the periphery to the center and documented the changes that occurred in that person’s professional life.

Taking a self-critical view of our research, our findings (and those of several other researchers who have studied groups in Tapped In and other online communities) have told us little about the network as a whole or the majority of groups and individuals in it. We do not know how representative the phenomena or patterns we uncovered are in the broader network of groups in our corpus or how to foster (or discourage) them. Nor do these studies lead to robust design principles from which developers and network leaders can make reliable decisions. We have also conducted quantitative analyses of much larger portions of the Tapped In population, which reveal tantalizing patterns of interconnections that people can achieve online but offer no insights into the causal relationships, outcomes, and benefits of those connections for learning or instructional practice. In summary, the small qualitative studies have found a few interesting needles in our large haystack, but they cannot tell us how many needles of that type are in the stack, how many other types of needles we might find, or anything about the haystack itself. The

---

5 The Tapped In research team did not “engineer” the social meta-network that has evolved or the groups within it, and we have not been involved in the day-to-day activities since Version 2 went online in 2003.
larger quantitative analyses tell us where in the haystack we might profitably look for needles, but under closer inspection they may just as likely be straight pins, safety pins, or nails.

Clearly, the solution would be to use both quantitative and qualitative tools, and here we arrive at the crux of the problem—the time and cost of conducting cycles of quantitative and qualitative analyses of large, evolving cyber-enabled social networks using outdated tools and techniques designed for taking snapshots of small, well-boundary, static groups of educators. Few research or evaluation teams have the time, resources, and range of expertise needed to analyze the gigabytes of data generated by today’s social networks, even if they had access to the data.6

**Leveraging Automated Analysis Tools**

Building on our prior work and advances in analytical tools made by colleagues (e.g., Suthers, Dwyer, Medina, & Vatrapu, 2007a, 2007b; Suthers, Dwyer, Vatrapu, & Medina, 2007; Suthers, Medina, Vatrapu, & Dwyer, 2007), we envision a suite of analytical methods and software tools that resolve the tradeoff between large- and small-scale analysis. Imagine a researcher first applying social network analysis tools to a set of log files from an online network to understand the extent to which network members who belong to multiple subgroups are having an effect on knowledge diffusion across groups in the network. The software automatically identifies when users post and read messages, access shared resources, and contribute to chat sessions. The software also identifies networks of relationships between actors and the events they attend. The resulting data are sufficient for the researcher to gather aggregate statistics on message-posting behavior and identify people who bridge across groups.

The researcher then applies a semantic analysis tool to a subset of the original corpus to test empirically whether the same ideas appear in the discourse of different groups in which the bridges are members. Finally, the researcher uses a data-mining tool on those particular dialogues found to have ideas in common to identify specific, temporally ordered patterns of behavior in the cross-group postings by bridges. The interaction patterns of individuals in a group and groups are tied to specific collaboration tools in a system, and the researcher is able to activate a visualization of each person’s activity across multiple groups over time as he or she propagated an idea.

Once the researcher has developed a model of how the social and technical elements of the system support this behavior across several instances (the processes the individual and group use), he or she can query the data set for additional examples of the behavior as well as counterexamples where the community infrastructure limits or interferes with idea propagation. Given these examples and counterexamples, the researcher can create and test hypotheses about the role of the network’s sociotechnical design in support of bridging behavior.

The proposed new analytic tools will enable researchers to examine both the products of the group (what is typically analyzed currently) and the processes that contributed to the development of the products. In focusing on both product and process, researchers will know whether a group is successful by certain metrics and also why a group is or is not successful. Researchers will be able to determine whether certain outcomes are associated with particular

---

6 Most online learning and community platforms do not capture and store the kinds of data needed to trace interaction patterns.
processes and make suggestions to group leaders to help increase their chances for success in their online endeavors.

**AN EXPLORATORY STUDY**

As a first step toward realizing this vision, and one that reinforced the need for better tools, we analyzed 5 months of Tapped In data on more than 600 human *bridges* (also called brokers or boundary spanners)—people who belong to two or more groups within the network. The importance of bridges has been widely acknowledged in literature on social networks and community computing. Bridges are important to creating and maintaining extended social networks; people who are weak ties between different community groups facilitate the pooling of social capital (Granovetter, 1973). Bridges are also important for action; when people use communication media, such as the Internet, they become able to educate other community members and organize, as needed, for collective action (Kavanaugh, Reese, Carroll, & Rosson, 2005).

The original conception that people who belong to multiple groups act as bridging ties can be attributed to Simmel’s (1971) and Granovetter’s work on the strength of weak social ties. Granovetter (1973) hypothesized that weak ties—relationships characterized by infrequent interaction—are wide ranging and therefore more likely than strong ties to serve as bridges across social boundaries. Bridges are important for enhancing social capital. Social capital produced through civic participation in various groups is critical to the functional well-being of the society (Putnam, 2000). Bonding social capital is accomplished through strong ties within groups, whereas bridging social capital is accomplished through weak ties between groups. The latter form of social capital facilitates the integration of multiple and often diverse groups into larger social units (e.g., communities). Putnam (2000) argues that bridging social capital is important for increasing information flow between groups and sustaining healthy communities over time.

For the purposes of our argument in this paper, we highlight two empirical studies of bridges in physical communities that ultimately seek to understand how bridges participate in online communities and their demographic characteristics. The first study by Burt (1995) examines the role of bridges in brokering ideas between groups and making successful decisions in organizations. The second study by Kavanaugh, Reese, Carroll, and Rosson (2005) examines the role of bridges in organizing for collective action in physical communities. Burt analyzed managers at Raytheon, a large American company and military contractor, referring to them as brokers who bridge a social gap or structural hole between two groups. Burt interviewed several hundred managers and asked them to write down ideas for improving the company’s logistical and management operations. He had two Raytheon executives rate the ideas. Burt found that the highest ranked ideas came from managers who had contacts outside their immediate work group. Those managers, or brokers, had contacts that spanned structural holes in the organization. His study suggests that maintaining contacts in multiple groups is highly strategic and enables an individual to be more effective at sharing knowledge and information more aligned with the

---

7 We are using the terms bridges and brokers interchangeably because their exact differences are not important for the purpose of the literature review. We encourage readers to follow up on the references for details.
goals of the organization. Kavanaugh et al. (2005) examined differences in participation between bridges and nonbridges in Blacksburg, Virginia (USA). They found that bridges have higher levels of community involvement, civic interest, and collective efficacy than nonbridges. Further, leader bridges (people who bridge community groups through leadership roles) have higher levels of community participation and civic interest than member bridges (people who bridge community groups but are only members, not leaders).

The starting point for the current study investigating the characteristics and implications of bridges in online communities was Kavanaugh and colleagues’ study. Our work replicates their conceptual and analytical approach in trying to understand the differences in participation levels between bridges and nonbridges in online communities. The current study differs from Kavanaugh’s work in that it uses actual participation data from the online system rather than self-reports of participation. The gap in understanding online and face-to-face communities motivates our overarching research question: What are the characteristics and implications of people who bridge groups in online communities?

Our hypothesis is that bridges are as important in online communities as they are in traditional face-to-face communities; however, we have no empirical evidence to support that claim. Our work seeks to test the hypothesis and to document how bridges participate in online communities, whether their participation is different from that of nonbridges (members of a single group), what the demographic characteristics are of bridges, and related questions. This initial study explored the following research questions:

1. How does online participation vary between bridges and nonbridges in synchronous and asynchronous communication mediums?
2. How does online participation vary between different order bridges in synchronous and asynchronous communication mediums?

**Research Methods**

In the context of the research questions, two main categories of people were in the online community, bridges and nonbridges. Bridges were Tapped In members who were part of two or more groups in Tapped In. Nonbridges were Tapped In members who were part of only one group in Tapped In. We further divided bridges into three subcategories: (1) two-order bridges (2OB), Tapped In members who were part of exactly two groups in Tapped In; (2) three-order bridges (3OB), Tapped In members who were part of exactly three groups in Tapped In; and (3) four-plus-order bridges (4+OB), Tapped In members who were part of four or more groups in Tapped In. We wanted to consider online participation separately for synchronous (chat) and asynchronous (discussions) communication to capture any differences in the two modes of communication.

In the first investigation, the primary dependent variable was overall online participation. The two main categories for participation were (1) content participation and (2) After School Online (ASO) colloquium attendance. Content participation refers to the cumulative amount of content that members contributed in Tapped In. For the synchronous communication medium (chat), content participation was the total number of chat messages that Tapped In members uttered. For the asynchronous communication (discussions), content participation was the total number of discussion posts that Tapped In members had written. In both chat and discussion, the dependent variable was cumulative online participation, not participation associated with the groups.
Tapped In members can utter chat messages and author discussion posts outside the groups in which they have membership. That is, it is possible (but unlikely) that members have a high level of participation for chat and discussions but have never contributed any content within groups in which they are members.

The second category for online participation was ASO attendance, defined as the total number of ASO sessions Tapped In members attended. ASO is a mechanism through which Tapped In members involved in different groups can connect easily. ASO sessions are 1-hour public chat sessions led by a member of the community on a topic of his or her choosing. Forty to 50 sessions are held in a typical month. ASO attendance was included because it is different from content participation within groups. ASO sessions are a way for Tapped In members to share knowledge with other members and have a voice in the larger network. ASO attendance data show that members logged in to an ASO session; the data do not imply that a member actively participated (e.g., uttered a chat message) in the session. Although ASO session attendance is independent of how Tapped In members are affiliated with groups, we conjectured that bridges might be more active in ASO attendance than nonbridges because they may be more inclined to share with other members of the larger network.

Data Sampling Strategy

Five months of Tapped In data, January through May 2007, were analyzed. Groups that interacted primarily in synchronous chat-based sessions were separated from groups that primarily interacted in asynchronous discussions. The sampling strategy identified 184 chat-based groups and 200 discussion-based groups in which bridges participated. Groups set up by teachers for their K-12 students and groups with minimal activity were eliminated from the data set. For discussion-based groups, an active group was defined as one with at least as many discussion posts as members. In other words, a discussion-based group with on average at least one discussion post per member was considered active. The definition of active chat-based groups was created in a similar manner. Interaction in chat-based groups is quantitatively and qualitatively different from that in discussion-based groups (synchronous versus asynchronous interactions), so it is not possible to equate one chat utterance per member to one discussion post per member as the criterion. Accordingly, the ratio of all chat utterances to discussion board posts in Tapped In was used: the total number of discussion posts across all groups in Tapped In (39,072) and the total number of chat utterances across all groups in Tapped In (507,540). The ratio of chat utterances to discussion posts was approximately 13. That is, there are 13 chat utterances for every discussion post in Tapped In. (Intuitively, the ratio made sense to Tapped In staff and members, because more chat messages must be uttered to be equivalent to a discussion post. Chat messages are short and discussion posts tend to be longer.) Thus, an active chat-based group was defined as one with an average of at least 13 chat utterances per member.

Data Analysis

The data were explored using standard statistical analyses of two-sample t tests and ANOVA (analysis of variance). To determine whether there was a difference in online participation of bridges and nonbridges (Research Question 1), two-sample t tests were used. To understand whether there was a difference in online participation between different order bridges (Research
Question 2), one-way ANOVA was used. When the dependent variables were categorical, as in the case of exploring demographic characteristics of bridges, chi-square tests were used.

**Results**

The sampling strategy identified 123 active chat-based groups and 134 active discussion-based groups in the 5-month period. Of those groups, 33 were active in both synchronous and asynchronous modes. For both sets of groups, those using synchronous or asynchronous communication, members who were bridges and nonbridges were identified. *Chat-based bridges* refer to the bridges identified in the synchronous communication medium, that is, Tapped In members who were part of two or more groups among the 123 chat-based groups. *Discussion-based bridges* refer to the bridges identified in the asynchronous communication medium, that is, Tapped In members who were part of two or more groups among the 134 discussion-based groups.

For the synchronous communication medium, there were 330 chat-based bridges and 1,967 chat-based nonbridges. For the asynchronous communication medium, there were 329 discussion-based bridges and 1,977 discussion-based nonbridges (Table 1). There were 45 bridges who were both chat based and discussion based. Among the bridges (both chat-based and discussion-based), we also identified two-order (2OB), three-order (3OB), and four-plus-order (4+OB) bridges (Table 1).

<table>
<thead>
<tr>
<th>Synchronous communication medium</th>
<th>Chat-based bridges</th>
<th>n = 330</th>
<th>2OB N = 217</th>
<th>3OB N = 55</th>
<th>4+OB N = 58</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chat-based nonbridges</td>
<td>n = 1,967</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asynchronous communication medium</th>
<th>Discussion-based bridges</th>
<th>n = 329</th>
<th>2OB N = 211</th>
<th>3OB N = 91</th>
<th>4+OB N = 27</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discussion-based nonbridges</td>
<td>n = 1,977</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Breakdown of bridges.

(The breakdown for the 58 chat-based four-plus-order bridges was the following: 31 four-order bridges, seven five-order bridges, seven six-order bridges, one seven-order bridge, three eight-order bridges, three 11-order bridges, two 14-order bridges, one 15-order bridge, one 19-order bridge, one 33-order bridge, and one 54 order bridge. The breakdown for the 27 discussion-based four-plus-order bridges was the following: 17 four-order bridges, two five-order bridges, two six-order bridges, three seven-order bridges, one 10-order bridge, and two 12-order bridges.)

**Online Participation of Bridges and Nonbridges**

The first research question explored how overall online participation varied between bridges and nonbridges in synchronous and asynchronous communication mediums.
Chat-based bridges. A two-sample *t* test between chat-based bridges and nonbridges with content participation (number of chat utterances) and ASO attendance (number of ASO sessions) as dependent variables was performed. Before this, a Levene’s test for equality of variances was performed to determine whether the variances of the population samples were equal; the equal variances condition was violated. Thus, the two-sample *t* test for unequal variances was used. Table 2 shows the comparisons between the means. (Hereafter, an asterisk [*] in the table implies a statistically significant difference.) A statistically significant difference (*t*(340) = −5.038, *p* < .001) was found between chat-based bridges and nonbridges for content participation in the synchronous communication medium. For the dependent variable of ASO participation, the equal variances condition was also violated. A two-sample *t* test for unequal variances revealed *t*(329) = −3.707, *p* < .001, showing more attendance in ASO sessions by members who bridged groups than for members who did not.

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bridges</td>
</tr>
<tr>
<td>Content participation</td>
<td>185.35</td>
</tr>
<tr>
<td>(number of chat utterances)*</td>
<td></td>
</tr>
<tr>
<td>ASO attendance</td>
<td>19.52</td>
</tr>
<tr>
<td>(number of ASO sessions)*</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Comparison of participation between chat-based bridges and nonbridges.

Discussion-based bridges. We ran the same tests as above for discussion-based bridges and nonbridges, with content participation (number of discussion posts) and ASO participation (number of ASO sessions) as dependent variables. Again, for both dependent variables, the equal variances condition was violated. The two-sample *t* test for unequal variances was *t*(355) = −6.188, *p* < .001, showing discussion-based bridges had a significantly higher level of content participation than nonbridges for the asynchronous communication medium (Table 3). For ASO attendance, the two-sample *t* test for unequal variances was *t*(335) = −2.023, *p* < .02, showing discussion-based bridges attended more ASO sessions than nonbridges.

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bridges</td>
</tr>
<tr>
<td>Content participation</td>
<td>16.09</td>
</tr>
<tr>
<td>(number of discussion posts)*</td>
<td></td>
</tr>
<tr>
<td>ASO attendance</td>
<td>9.72</td>
</tr>
<tr>
<td>(number of ASO sessions)*</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Comparison of participation between discussion-based bridges and nonbridges.

Online Participation of Different Order Bridges

The second research question explored how online participation varied between different order bridges in synchronous and asynchronous communication mediums.
Chat-based bridges. A one-way ANOVA explored differences between two-order, three-order, and four-plus-order chat-based bridges with content participation (number of chat utterances) and ASO attendance (number of ASO sessions) as dependent variables. For content participation, the equal variances condition was violated, so we used the Welch statistic. A statistically significant difference ($F(2,84) = 4.5, p < .05$) was found for content participation and post hoc comparisons using the Games-Howell test (an alternative to the Tukey’s test when variances are unequal) revealed that four-plus-order chat-based bridges participated significantly more than two-order chat-based bridges. Table 4 shows the comparisons between the means. (‘Diff’ indicates statistically significant post hoc comparisons.)

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2OB</td>
</tr>
<tr>
<td>Content participation (number of chat utterances)*</td>
<td>116.53 (Diff 4+OB)</td>
</tr>
<tr>
<td>ASO attendance (number of ASO sessions)*</td>
<td>3.95 (Diff 4+OB)</td>
</tr>
</tbody>
</table>

Table 4. Comparison of participation between different order chat-based bridges.

For ASO attendance, the equal variances condition was also violated, so again the Welch statistic was used. There was a statistically significant difference ($F(2,75) = 5.687, p < .005$) in terms of ASO attendance (Table 4). Post hoc comparisons using the Games-Howell test revealed that four-plus-order chat-based bridges participated significantly more than both two-order and three-order chat-based bridges.

Discussion-based bridges. The same tests were run for discussion-based bridges, with content participation (number of discussion posts) and ASO participation (number of ASO sessions) as dependent variables, and again the equal variances condition was violated for both dependent variables. For content participation, a significant effect was found ($F(2,59) = 8.754 ,p < .001$), and a post hoc comparison using the Games-Howell test revealed that four-plus-order discussion-based bridges participated significantly more than both two-order and three-order discussion-based bridges. For ASO attendance, there were no significant differences ($F(2,62) = 2.381, p > .1$) between two-order, three-order, and four-plus-order discussion-based bridges (Table 5).

<table>
<thead>
<tr>
<th></th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2OB</td>
</tr>
<tr>
<td>Content participation (number of discussion posts)*</td>
<td>11.19 (Diff 4+OB)</td>
</tr>
<tr>
<td>ASO attendance (number of ASO sessions) (ns)</td>
<td>6.80</td>
</tr>
</tbody>
</table>

Table 5. Comparison of participation between different order discussion-based bridges.
Discussion

The results show that bridges have greater content and ASO participation than nonbridges in both synchronous and asynchronous communication mediums. Based on actual participation measures of members, the results validate the claim made by Kavanaugh et al. (2005) for online communities. In addition, the results demonstrate that the claim holds true for both synchronous and asynchronous communication mediums. The results also reflect a general upward trend for participation by higher order bridges.

That bridges participate overall more than nonbridges as measured by number of contributions is not a surprise. Belonging to more than one group gives bridges more opportunity to contribute to more threads of discourse. This finding is nonetheless important for understanding and managing social network activity. One might intuit that the more groups a member belongs to, the greater the member’s overall level of participation. However, an alternative hypothesis could just as well be supported. That alternative hypothesis is that network members have a limited amount of time and inclination for online participation, and the expectation would be they simply redistribute a fixed level of participation, contributing less and less to each group (or contributing to some groups but not other groups) as they join additional groups. This does not appear to be the case. Overall participation does rise, suggesting that people who join more groups do not simply redistribute their effort, but we have not yet studied the within-group participation of these bridges. A next analysis could be to determine how bridging affects contribution in any one group and whether we should be encouraging members to join multiple groups more actively. Do bridges communicate more in group A as a function of belonging to group B? Do they communicate more of one type of information, give more help, or answer more questions? Are more extreme bridges (four-plus) somehow different from other bridges?

The greater level of attendance at ASO sessions by the members who bridged groups is congruent to the finding from Kavanaugh et al. (2005) with group leaders who were bridges having higher community involvement and civic interest in face-to-face communities. ASO sessions are public, community-wide events not group events. The finding that higher order discussion-based bridges did not have higher participation in the synchronous ASO sessions is most likely an effect due to preference of asynchronous communication online. So, although being a discussion-based bridge predicts more participation in ASO chat sessions, the disposition to attend ASO sessions does not grow with the number of groups a person belongs to if that person prefers asynchronous communication. To take advantage of the goodwill of bridges, it might make sense for Tapped In to conduct asynchronous ASO discussion events to accommodate bridges who prefer that medium.

We do not know why people who bridge multiple groups also attend more public events, but the implications for teacher professional development could be significant. Some next questions include whether the spread of innovation can be traced through the posts bridges make: Is there evidence of bridges sharing what they learned in ASO sessions with members in their group? Do bridges provide advice on a new topic or teach others a new method as they attend ASO sessions? What, if any, negative implications are there of people who span the boundaries and identity of groups? Finding these patterns and contrasting these competing “intuitions” was possible only with a large and stable data set. Yet these are only superficial questions, and they still required significant researcher time to extract the necessary data from the large data set.
New tools are needed to ask the kinds of deeper, more interesting research questions these findings raise and more thoroughly understand online social networks and communities.

The above study examined only 5 months of more than 10 years of data. It only used data from groups defined as “active.” The study tells nothing about bridges who belong to inactive groups. Looking for patterns in all the groups will take more time and require the new tools. Moreover, with statistical methods alone, the processes of the members and groups are completely invisible. The simple yet still time-consuming steps to get the quantitative participation data cannot give all the answers we need. Currently, it would be extremely expensive and nearly impossible to examine the postings and chat sessions from approximately 300 groups. It would only be through the examination of content that we could determine how knowledge is passed between groups.

Although the study provides tantalizing hints about the critical roles that bridges play in knowledge generation and diffusion within and across hundreds of groups, many other questions are generated, and we do not have the tools needed to trace the actual flow of knowledge, to understand the developmental path to becoming a bridge, or to understand if there are patterns of interactions that are occurring inside of groups that foster the development of bridges. Finding interesting results on large numbers of groups leads back to conducting content analysis on small samples, because that is all that is possible given resources, and then once again not being able to draw generalizable conclusions. For example, we would like to be able to determine the developmental path to becoming a bridge, not just for some small sample, but all of Tapped In’s bridges. Do bridges typically become a member of a community, participate in sessions, and via legitimate peripheral participation (Lave & Wenger, 1991) become more central and take on leadership roles, or do they enter as a leader and immediately begin fostering groups and leading sessions? Do the paths members take to leadership vary by occupation? Are there patterns of interactions that regularly occur in groups with bridges in them? The above analysis is also unable to show whether and how bridges have a qualitative impact on the community, for example, by generating and sharing ideas and resources. How does the information flow? Does ASO attendance benefit the groups or do bridges bring in information to ASO that then gets diffused through the community?

CONCLUSION: CREATING CYBER-ANALYSIS TOOLS FOR UNDERSTANDING CYBER-NETWORKED LEARNING

Research must help science, technology, engineering, and math (STEM) education communities convert the current enthusiasm for cyber-enabled social networking into reliable evidence of how, when, and why cyber-enabled social networks do, and do not, advance learning, and we must develop scalable and replicable models that maximize the value and benefits of emerging social networking models and technologies. Education researchers and evaluators must overcome conceptual and methodological obstacles that limit exploration of the frontiers of learning in cyber-enabled social networks.

As described above, school and district administrators and outside experts can use social network analysis to understand the diffusion of reform knowledge and practices, identify problems, and deploy resources more effectively. Applying traditional social network analysis methods can be problematic in large-scale cyber-enabled social networks, which typically do not have a well-defined structure. Moreover, the methods are limited in their ability to identify the
precursors and other enabling factors for social capital or trace how social capital is fostered and leveraged (Leana & Pil, 2006).

We have argued that more productive and powerful tools and methods are needed to enable leaders of reform interventions to understand the educator networks and thereby create more effective and cost-efficient learning communities. Many districts, institutions of higher education, and other organizations create online teacher mentoring and professional development programs or online adjuncts to face-to-face programs and find they have few metrics to interpret or assess what is happening beyond basic countable requirements and laborious analysis of the discourse teachers produce. Results are often disappointing from both the programmatic and research perspectives. Traditional evaluations frequently find that teachers do not participate as expected or discuss topics expected by program leaders. Researchers have no way to predict or do anything to correct these eventual outcomes in time to make a difference. Most evaluators and instructors analyze online participation using frequency data (e.g., number of posts over time), sometimes supplemented by functional (e.g., posts whose purpose is to guide, give feedback, pose a question, reflect; Bonk & Kim, 1998; Schlager et al., 2002) and topical (e.g., posts about classroom management, assessment, STEM concepts; Jaffe, Moir, Swanson, & Wheeler, 2006) categorizations of discourse data. Quantifying participation tells us little about the quality of reflection, collaboration, or dialogue, and analyzing discussion postings at the end of a project is too late to do anything to improve the dialogue. There is no way to rigorously understand why a group was not successful beyond intuitions and anecdotal experiences with small groups.

Social networks are defined by their structure (the evolving relationships among members and subgroups) and the activities of those members using tools and other artifacts. We argue that structure and activity must be comprehended simultaneously in order to understand the evolution of a network and the trajectories of its members. A network’s structure and how it changes are the products of the members’ activity, and the activity is interpretable only in the context of the dynamically changing social structure. In this way, structure and activity are mutually constitutive. Structure is the product of activity, and that new structure constrains the activities that can occur. With each new activity in a social network comes an opportunity for changes to the structure of the network. With each change in structure, only certain activities can occur.

The nature of sociotechnical networks requires an expanded conception of where social capital resides (Resnick, 2002). In a technology-mediated context, social capital is not solely built on direct communication with persons known to each other. Sociotechnical networks enable knowledge and expertise to be shared indirectly as well, for example, when a message is read by someone other than the person to whom it was originally addressed or a digital artifact is contributed for use by whomever comes across it. Therefore, it is necessary to build the analysis of online interaction on a conception of “tie” that includes but is broader than ties conceived of as interpersonal relationships or communications directed from one person to another.

The fundamental relationship of interest is that one person has produced something of value that is then accessed by another person, as enabled by the sociotechnical network. Suthers and his colleagues have called this relationship “uptake”: the act of someone taking up the product of another (Suthers, 2006; Suthers et al., 2007b; Suthers et al., 2007). Uptake is an abstraction; in order to recognize potential uptake events in log files we need to identify specific relationships between the acts of participants, as mediated by the digital manifestations and traces of these acts. Suthers’ work seeks to identify contingencies between acts that suggest that one act had something to do with another and therefore that the potential for production of value that resides in the sociotechnical network has been realized. Suthers and his colleagues have explored
different types of contingency relationships (Suthers, 2006; Suthers et al., 2007a, 2007b; Suthers et al., 2007), including media dependencies, temporal and spatial proximity, representational similarity, and semantic overlap.

We need the ability to trace and understand technology-mediated interaction on a larger scale than has previously been attempted. We must address the practical issues of integrating heterogeneous data from across multiple interactional media and making temporally and spatially distributed interaction data available for analysis while scaling up to larger data sets. We need a unifying conceptual framework and data structure for representing large sets of heterogeneous data and making the data available for analysis through consistent and well-defined methods.

Cyber-enabled social networks offer the ability to capture and analyze a more complete and objective record of peoples’ actions and interactions automatically over time and could potentially speed up and lower the cost of analysis when tools are available that can extract and process the data. However, digital interactions are not simple to mine or interpret. Interaction is distributed across space, time, and media, and the data obtained through instrumentation come in a variety of formats. In addition, interaction data are missing judgments concerning the strength of social ties, a key ingredient of social network analysis, but they offer instead nuanced and detailed information on the actual contacts through which these ties form. We need to learn to interpret such data.

Whether we can understand what constitutes a meaningful relation or tie between individuals from automatically recorded interaction data is an important unanswered research questions (Carley, 2003). We must develop new analytical frameworks and more integrative and automated methods and tools that can rapidly mine and reliably analyze the massive amounts of data generated automatically through heterogeneous tools used to support cyber-enabled social networks (e.g., chat, cell phones, blogs, wikis, social networking sites, threaded discussions, mailing lists, and shared artifact creation).

We advocate a research agenda to develop and validate analytical tools and techniques for measuring the structures, activities, and impact of cyber-enabled social networks in teaching and learning using automatically generated digital data from large sets of heterogeneous data. We must construct software tools to trace online activities across groups and time in a form that is accessible for different classes of social network research (e.g., sequential analysis of activity, structural analysis of social relationships, semantic analysis of discourse). And we must validate the quality of analysis produced using new tools against traditional social network analysis methods. The tools we envision will help those working online to trace the processes as well as products (participation data and the content) of network participation using automatically collected data to determine what “invisible” work is occurring within and between online groups, among novices and experts. From this, we can begin to reliably identify patterns that lead to failure or success, both for individuals and entire networks.

ACKNOWLEDGMENTS

We thank Dan Suthers for his substantial contributions to the ideas presented, Christine Halverson for her conceptual insights, and Jasmine Lopez and Ruchi Bhanot for technical assistance with our analyses.

In press, Journal of Teacher Education.
REFERENCES


