MINING FACE-TO-FACE INTERACTION NETWORKS USING SOCIOMETRIC BADGES: PREDICTING PRODUCTIVITY IN AN IT CONFIGURATION TASK

Lynn Wu
MIT Sloan School of Management
lynnwu@mit.edu

Ben Waber
MIT Media Laboratory
bwaber@media.mit.edu

Sinan Aral
NYU Stern School of Business & MIT Sloan School of Management
sinan@stern.nyu.edu

Erik Brynjolfsson
MIT Sloan School of Management
erikb@mit.edu

Alex "Sandy" Pentland
MIT Media Laboratory
sandy@media.mit.edu

DRAFT – May 7, 2008 – Comments Welcome

* We are grateful to seminar participants at IS Winter Conference, University of Utah for valuable comments and to Cisco Systems, Hewlett Packard, the MIT Media Laboratory and the MIT Center for Digital Business for generous funding.

MINING FACE-TO-FACE INTERACTION NETWORKS USING SOCIOMETRIC BADGES: PREDICTING PRODUCTIVITY IN AN IT CONFIGURATION TASK

Abstract

Social network theories (e.g. Granovetter 1973, Burt 1992) and information richness theory (Daft & Lengel 1987) have both been used independently to understand knowledge transfer in information intensive work settings. Social network theories explain how network structures covary with the diffusion and distribution of information, but largely ignore characteristics of the communication channels (or media) through which information and knowledge are transferred. Information richness theory on the other hand focuses explicitly on the communication channel requirements for different types of knowledge transfer but ignores the population level topology through which information is transferred in a network. This paper aims to bridge these two sets of theories to understand what types of social structures are most conducive to transferring knowledge and improving work performance in face-to-face communication networks. Using a novel set of data collection tools, techniques and methodologies, we were able to record precise data on the face-to-face interaction networks, tonal conversational variation and physical proximity of a group of IT configuration specialists over a one month period while they conducted their work. Linking these data to detailed performance and productivity metrics, we find four main results. First, the face-to-face communication networks of productive workers display very different topological structures compared to those discovered for email networks in previous research. In face-to-face networks, network cohesion is positively correlated with higher worker productivity, while the opposite is true in email communication. Second, network cohesion in face-to-face networks is associated with even higher work performance when executing complex tasks. This result suggests that network cohesion may complement information-rich communication media for transferring the complex or tacit knowledge needed to complete complex tasks. Third, the most effective network structures for "latent" social networks (those that characterize the network of available communication partners) differ from "in-task" social networks (those that characterize the network of communication partners that are actualized during the execution of a particular task). Finally, the effect of cohesion is much stronger in face-to-face networks than in physical proximity networks, demonstrating that information flows in actual conversations (rather than mere physical proximity) are driving our results. Our work bridges two influential bodies of research in order to contrast face-to-face network structure with network structure in electronic communication. We also contribute a novel set of tools and techniques for discovering and recording precise face-to-face interaction data in real world work settings.

Keywords: Social Networks, Face-to-Face Communications, Information Worker, Productivity

1. Introduction

Information workers now account for more than 70% of the U.S. labor force in the U.S. economy (Apte & Nath 2004). As the information content of work has increased, the distribution and use of information has become central to the performance of workers and organizations. Advances in information technology have given workers access to a tremendous amount of information to solve complex problems, and new modes of electronic communication have facilitated changes in the way information workers seek and obtain information. How information workers obtain, transfer and utilize information, through both social and technological means, is central to how these workers make decisions, perform tasks, and deliver information-based products and services.

Two broad theoretical perspectives inform our understanding of information transfer in information-intensive work settings: social network theory (e.g. Granovetter 1973, Burt 1992) and information richness theory (e.g. Daft & Lengel 1986). While social network theories explain how network structures covary with the diffusion and distribution of information, they largely ignore characteristics of the communication channels through which information and knowledge are transferred. Information richness theory on the other hand focuses explicitly on the communication channel requirements for different types of knowledge transfer but ignores the population-level topology through which information is transferred within a network. We bridge these two theories in order to understand what types of communication channels are most effective in transferring knowledge in different kinds of social networks. In particular, we examine whether optimal network structures in face-to-face interaction networks are the same as or different from those structures found to be optimal in electronic communication networks such as email. We hypothesize that the different modes of communication reflect differences in the type of information being transferred—for example tacit versus explicit knowledge—and that these differences are reflected in the structure of networks that most effectively transfer different types of information and knowledge.

Social network theories have been instrumental for understanding how information workers leverage their network potential to achieve better work performance. Although researchers have been creative in obtaining social network data from multiple sources, methodologies such as surveys, questionnaires, and self-reports have been predominant. Unfortunately, data collection based on self-reporting, which requires that subjects recall their social networks from memory, are often biased toward the inclusion of strong links (Marsden 1990). It has been shown that most network data from surveys and questionnaires tend to produce better quality for close and strong ties than for distant and weak ties. In addition, although subjects are able to describe their local network in general terms, they fail to describe details such as discussion topics or the timing of interactions (Marsden 1990). While numerous methods try to address these issues, the bias inherent in self-reports remains problematic. To bring social network analysis to the next level, it has become imperative to collect more reliable data collected in real time.

Ubiquitous email access gives researchers the opportunity to solicit real-time email communication data. Since email archives record detailed communication logs, such as who has emailed whom, the exact time of the interaction, and the content of the exchange, using email archives to construct social networks allows researchers to eliminate bias introduced through survey methods. Aral, Brynjolfsson, and Van Alstyne (2006, 2007) and Aral & Van Alstyne (2007) report on one of the first empirical studies to use email data to understand how social networks impact information worker productivity. They find that structurally diverse networks with an abundance of structural holes are most effective for gathering diverse information, which in turn improves information worker productivity.

Although email has become an important communication tool over the last fifteen years, face-to-face conversations remain an important and in many cases predominant mode of communication. In fact, information workers may use face-to-face conversations to transfer and process fundamentally different types of knowledge than those transferred via email. Consequently, the network structures that are most effective for improving work performance in face-to-face networks may be different from those in email social networks. In addition to studying electronic communication networks such as email, it is therefore essential to explore and contrast the types of structures in face-to-face networks that are most effective for accessing and transferring information and improving worker productivity. Unfortunately, until now, recording precise and reliable data on face-to-face interaction networks has been difficult. To fill this gap, we employ a new data collection method that utilizes Sociometric badges developed at the MIT Media Laboratory to record real-time patterns of face-to-face interactions between employees in real-world work settings over time (Waber et al., 2007). These data enable us to analyze both proximity and conversation networks amongst a group of IT workers who execute configuration tasks during the sale and delivery of commercial IT hardware. Collecting actual face-to-face interaction data, we are able to introduce information richness theory into social network analysis to contrast the communication modes used to actualize social networks in information intensive work settings. Using these data, combined with project and accounting data on the relative performance of the workers, we evaluate which face-to-face network structures best predict higher performance, and whether these structures differ from those found to predict productivity in the context of electronic communication networks such as email and phone communication.

Our analyses uncover four key results. First, we demonstrate that face-to-face networks are indeed different from electronic networks in terms of their relationship to worker productivity. In face-to-face networks, cohesion is positively correlated to higher worker productivity, while the opposite is true in email communication. Second, network cohesion is also associated with even higher performance when executing complex tasks. This suggests that cohesive networks using information-rich media may be particularly effective in transmitting complex knowledge needed to solve complex problems. Third, we show that the most effective network structures for "latent" social networks (those that characterize the network of available communication partners) differ from "in-task" social networks (those that characterize the network of communication partners that are actualized during the execution of a particular task). Lastly, the effect of cohesion is much stronger in face-to-face networks than in physical proximity networks, demonstrating that information flows in actual conversations (rather than mere physical proximity) are driving our results. Although our results do not firmly identify the direction of causality, our panel data estimates eliminate bias from any unobserved time-invariant factors that may confound our results. Furthermore, on-site visits and interviews support our empirical evidence, as employees corroborated their use face-to-face conversations to communicate complex and embedded knowledge. Our results demonstrate the importance of face-to-face social networks in predicting worker productivity even as email communication becomes ubiquitous. Such evidence is important for managers who face increasingly global and geographically dispersed work environments, as electronic communication networks alone may not be enough to transfer complex knowledge needed for complex tasks.

2. Theory

Prior research has shown that factors such as technology use, geographic dispersion and organizational structure significantly affect information transfer effectiveness (Argote 1999, Contu & Willmott 2003, Hansen 1999). Other research emphasizes the importance of using different transfer mechanisms for different types of knowledge (Slaughter & Kirsch, 2006). In this study, we link two broad theoretical perspectives, social network theory (e.g. Granovetter 1973, Burt 1992) and information richness theory (e.g. Daft & Lengel 1986), to understand what types of networked social structures are most conducive to transferring knowledge and improving work performance across different communication media. Specifically, we contrast new evidence on face-to-face networks with prior results on email networks in information intensive work. By elevating face-to-face network data collection to comparable standards of accuracy and precision in electronic communication data, we open new avenues for true comparisons across heavily theorized media choices.

2.1 Information Richness

The first broad perspective on information transfer is information richness theory, which provides a foundation for understanding how media choice affects information and knowledge transfer. Daft and Lengel (1986) combined two dimensions of information processing—uncertainty and equivocality—into a single framework to understand the implications of information richness on media choice. In their framework, media is assigned to a richness scale, where media that require a long time to transfer knowledge and through which it is difficult to resolve divergent perspectives have low richness, while media that enable reductions in equivocality and allow managers to process complex information are considered richer. Face-to-face communication is a rich medium because it provides multiple social cues through both natural language and body language, and can therefore greatly reduce equivocality (Daft & Lengel 1986). Following face-to face communication, the hierarchy in decreasing ability to transfer social cues and feedback is telephone, electronic mail, letter, note, memo, special report, and finally, flier and bulletin. According to this framework, when the information to be transferred is equivocal or uncertain, rich media such as face-to-face meetings are most effective for clarifying ambiguous events and developing common ground for mutual understanding. On the other hand, when information requirements are unequivocal and relatively simple, less rich media such as memos are sufficient (Daft & Lengel 1986).

Much work in the Information Systems literature has built on information richness theory to elucidate other important factors that influence media choice (e.g. Sproul & Kiesler 1986, Markus 1994, Walther 1995). However, social network theories have rarely been examined together with information richness to explain media choices and their effectiveness in the context of the topological network structures of communicative interaction. It is therefore important to bridge these theories to examine what types of media support what types of social network structure.

2.2 Network Structure and Knowledge Exchange

Social network analysis provides a rich class of theories to explain how information workers obtain knowledge and information. Two properties in particular have been instrumental in understanding the impact of social networks on work performance: structural diversity and social cohesion (Burt 1992, Coleman 1988). Some social scholars believe that a structurally diverse network is the most effective structure for improving performance in information intensive work. Burt (1992) shows that individuals with structurally diverse networks spanning multiple structural holes are more successful in terms of wages and promotion (Burt 1997). He attributes these performance differences to actors' ability to access and gather diverse pools of information from diverse social groups. Aral, Brynjolfsson and Van Alstyne (2006) demonstrate that structural diversity is associated with higher levels of economic productivity for information workers. Other studies also demonstrate an association between network diversity and performance and infer that diverse contacts provide access to novel information (e.g. Ancona & Caldwell 1992, Sparrowe et al. 2001, Reagans & Zuckerman 2001, Cummings & Cross 2003). By analyzing email communication networks, message content and employee performance, Aral & Van Alstyne (2007) demonstrate that networks with structural holes deliver diverse and novel information and that access to novel information explains a significant portion of the variance in productivity – more so for instance than traditional human capital.

Although structurally diverse networks are beneficial for exposing actors to different kinds of information, they are less effective at transferring complex knowledge. Knowing where the knowledge resides is very different from understanding the knowledge itself (Hansen 1999). When information is simple, explicit or declarative, using a structurally diverse network with many weak ties to transfer knowledge may be sufficient, as the information is easily transferred between actors. However, when information is complex or tacit a cohesive network with strong ties may be more effective for transferring knowledge (Hansen 1999). Cohesion can induce cooperative behavior which facilitates knowledge transfer, especially for complex tasks. If the source decides not to cooperate, his behavior may tarnish his reputation and consequently restrict his ability to interact with others in the cohesive group in the future. Stronger relationships among parties in a cohesive network also foster norm, trusts and reciprocity which induce the source to commit more time and energy to transfer information (Granovetter 1992). Furthermore, a dense web of third-party relationships in a cohesive network reinforces learning since it allows the same information to be presented using multiple perspectives, creating better understanding.

In summary, social network theories expect diverse networks to improve productivity when the ability to access diverse information is important. In contrast, cohesive networks are most effective for work performance when transferring complex tacit knowledge is important.

2.3 Combining Information Richness Theory with Social Network Theory

The central concept linking social network theories and information richness theory is knowledge complexity. Social network theories emphasize the impact of structural properties on an actor's ability to obtain and transfer knowledge, while information richness theory focuses on identifying the appropriate media with which to most effectively transfer different types of information and knowledge. Two dimensions of knowledge are typically used to characterize complexity: codifiability and interdependence. Codifiability refers to the degree to which knowledge can be fully documented in writing at the time of knowledge transfer (Brynjolfsson 1994, Hansen 1999). It has been shown that tacit knowledge is often associated with low levels of codifiability, making it difficult to articulate and transfer (Polanyi 1966, Nelson & Winter 1982). The second dimension, information interdependence, measures whether knowledge is part of a larger system of interrelated concepts (Teece 1986, Winter 1987). When knowledge is embedded in a system, information transfer can be particularly challenging, as it requires transmitting knowledge of the larger system in addition to the specific knowledge.

Information richness theory expects the use of rich media to transmit uncertain and equivocal knowledge, and less rich media to transfer simple, codifiable knowledge. Although face-to-face communication is typically the most costly communication channel (in time, effort and energy), it is preferred when transferring complex knowledge because it can facilitate the resolution of confusion and misconception. Network theories predict that structurally diverse networks with weak ties are beneficial for accessing novel and diverse information. However, knowledge accessed through diverse networks, typically characterized by an overabundance of weak ties (Granovetter 1973, Burt 1992, Aral & Van Alstyne 2007), is often limited to simpler, codifiable knowledge in part because weak ties have a limited ability to transfer tacit, embedded information (Hansen 1999). As complex knowledge is more difficult to transfer, a cohesive network may be even more important for complex knowledge transfers.

The Effect of Network Cohesion in Face-to-face Networks

As face-to-face communication offers the richest medium for transferring complex knowledge, cohesive networks are in the best position to fully utilize the power of face-to-face communication. Specifically, cohesive face-to-face networks are more effective in transferring complex knowledge for three reasons. First, cohesive networks are more likely to develop trust among actors. Second, the absorptive capacity in a cohesive network facilitates knowledge transfer. Third, the redundancy inherent in cohesive networks allows actors to receive information through multiple perspectives, easing knowledge transfer.

In a cohesive network, actors are more likely to trust each other. Since transferring knowledge requires the cooperation of the source, it is important to convince the source that the transfer would not negatively affect them. Without trust, the source may simply refuse to pass on the knowledge to the recipient. However, when there is a strong tie between them or a dense web of third party ties around the relationship, the source may be more willing to initiate the transfer. Consequently, knowledge transfer between the source and the recipient in a cohesive network is more likely as cohesion creates cooperative motivation and removes competitive impediments to information transfer by increasing trust between parties (Granovetter 1992, Reagans & McEvily 2003). Trusting the recipient, the source has greater incentive to be of assistance and is typically more available to help (Granovetter 1982). Face-to-face communication offers a rich medium where actors can quickly establish familiarity and rapport through both natural language and body language. These rich social cues allow actors to easily understand each other and to develop the trust needed to transfer complex knowledge between them.

However, even when the source is willing, the process of knowledge transfer may not be straightforward due to the inherent complexity of the information (Hansen 1999). Absorptive capacity is important for recognizing the value of new information and the ability to assimilate and apply the information to solve complex problems (Cohen & Levinthal, 1990). Facilitating communication between individuals especially across social boundaries is important. Furthermore, effective communication for transferring complex knowledge requires shared language and social norms (Cohen & Levinthal 1990). A cohesive network can increase the absorptive capacity in a network as repeated communication allows actors to develop relationship-specific communication heuristics that ease knowledge transfer (Hansen 1999). With more frequent communication, actors are less inhibited from seeking information and asking for clarification in a cohesive network, and accordingly, they are more likely to understand how to correctly use the information sooner. Face-to-face communication offers the maximal knowledge transfer in each information exchange by using communication channels with the richest social cues, which improves absorptive capacity in a short amount of time.

Although cohesive networks have been criticized for generating redundant information, redundancy could be a powerful instrument for effectively transferring tacit knowledge. Redundancy is not simply duplication of existing knowledge but, in fact often helps to create a common cognitive ground that can help individuals sense what others are struggling to articulate (Nonaka 1990, 1994, Grant 1996). Consequently, cohesive networks can facilitate tacit knowledge transfer, by allowing the same information to be repeated multiple times from different perspectives. Face-to-face communication further aids actors in developing the common cognitive background needed for clear articulation of complex knowledge and therefore improved work performance.

We expect face-to-face networks to require different network structures to transfer fundamentally different types of knowledge when compared to email networks. Structurally diverse networks that use less rich media such as email are beneficial for obtaining diverse sources of information and consequently improving worker productivity (Aral & Van Alstyne 2006, Aral & Van Alstyne 2007). Based on information richness theory and social network theories, cohesion (rather than diversity) in face-to-face networks should improve work performance as face-to-face communication is typically used to transfer more complex, embedded knowledge, and because network cohesion aids complex knowledge transfers. We therefore hypothesize that network cohesion is positively associated with work performance in a face-to-face networks.

> *Hypothesis 1a: Cohesion in face-to-face networks, measured by network constraint, is correlated with stronger information worker performance.*

Although face-to-face communication in a cohesive network may have significant impact on work performance in a variety of tasks, it may be especially beneficial for more complex tasks. Basic tasks that require relatively simple information can often be solved without the necessity of face-to-face communication. When information can be written into succinct rules that workers can easily follow and when tasks are relatively simple, cohesion in face-to-face networks may provide less additional benefit. However, when workers face complex tasks that presumably require access to tacit and embedded information, manuals prove to be less useful and workers must to turn to face-to-face conversations with colleagues to access the desired information. Because face-to-face communication is the richest medium that can most effectively transfer complex knowledge between actors, and because cohesive networks are most effective for transferring complex knowledge, a cohesive face-to-face network may be especially helpful in transferring tacit or embedded knowledge used for the execution of complex tasks. Thus, we hypothesize:

> *Hypothesis 1b: Cohesion in face-to-face networks, measured by network constraint, is more helpful for completing complex tasks than simple tasks.*

The Effect of Indirect Contact in Face-to-face Networks

Although having a cohesive face-to-face network is beneficial for transferring complex knowledge, the ability to access a broad range of information in the network, be it complex or simple, is still important. In addition to the topology of the surrounding network, an actor's direct and indirect contacts can affect her ability to search for relevant information (Hansen 2002). While direct contacts enable actors to immediately initiate information transfers, indirect contacts quicken the search process because actors learn of information and opportunities in the network through word of mouth. Individuals can therefore leverage their web of third-party ties to obtain the desired information. However, indirect contacts can also distort information content. When information gets passed through long path lengths (Freeman 1979), the chance of distortion is particularly high, as people tend to misunderstand or misinterpret information (Collins and Guetzkow 1964, Huber an Daft 1987, Gilovich 1991, Hansen 2002). Imprecise or inaccurate information can have a negative performance impact on the focal actor. Acting on vague information obtained indirectly, the focal actor may need to use her ties to connect to the original source of the information, only to find it was not what she sought. Eliminating misleading information is costly, as verifying each incorrect lead wastes valuable time and effort. When an actor has relatively short path lengths to other experts, not only is she exposed to less information distortion, she can also access knowledge experts more quickly. Two network properties measure the relative path length between actors: betweenness centrality and network reach.

Betweenness centrality measures how often an actor is positioned on the shortest path between other pairs of actors in the network (Freeman 1979). When an information worker is positioned in the network where she can access other actors quickly, she is more likely to be in the most effective position (Freeman 1979, Brass & Burkhardt 1992, Burt 2992, Hansen 2001) and is more likely to access more novel information more quickly. While betweenness centrality measures the positional advantage in a network for accessing information, *network reach* measures the degree to which an actor in a network can reach everyone else in the network. An actor with broad network reach is less affected by information distortion, since the path lengths to other actors in the network are relatively short. Consequently, network reach can facilitate information transfer and improve work performance. We, therefore, hypothesize that high betweenness centrality and broad network reach are beneficial for accessing critical information and consequently improving an actor's work performance.

> *Hypothesis 2a. High betweenness centrality and broad network reach in face-to-face networks are correlated with stronger work performance.*

Transfers of complex information may experience even greater distortion than transfers of simple knowledge, as complex information is inherently more difficult to understand and is more likely to be misinterpreted. The ability to access many indirect contacts efficiently can be especially helpful in obtaining complex information needed to complete more complex tasks. Occupying network positions with high betweenness centrality enables workers to quickly access valuable information. This can be particularly important in completing more difficult - tasks which tend to require more information to complete. Broad network reach can reduce information distortion and promote knowledge transfer by influencing an actor's ability to effectively access complex ideas in the network. Workers with broad network reach are exposed to more views and perspectives, allowing them to understand information from different angles and to frame information in ways others can understand. Thus, we hypothesize that broad network reach and high betweenness centrality are particularly beneficial for complex tasks that require both more information and information that is inherently more complex and therefore more difficult to transfer and absorb.

> *Hypothesis 2b: High betweenness centrality and broad network reach in face-to-face networks can be especially beneficial in completing complex tasks.*

The Effect of Direct Contacts in Face-to-face Networks

Direct contacts are actors one can communicate with directly. Measured using network size, direct contacts are often associated with easing knowledge transfer by removing information distortion caused by middlemen. They can therefore be effective in transferring complex knowledge. While a cohesive face-to-face network provides a good infrastructure for transferring complex information, relying on direct links may further ease the transfer, since actors may already have developed relation-specific heuristics to communicate with each other (Hansen 1999, Uzzi 1997). However, while direct links ease information exchange, maintaining direct links is costly, especially in face-to-face networks which require time-consuming conversations and co-presence (Burt 1992, Uzzi 1997, Hansen 1999).

The tradeoff between direct and indirect contacts depends largely on the nature of the information being transferred. If the task requires complex information, direct links can give extra power in transmitting knowledge in cohesive face-to-face networks. However, if the information is simple and can be articulated in writing, maintaining direct contacts may be too expensive to justify the maintenance cost. In this case, face-to-face conversations do not aid transfer, but take time away from task completion activities. Thus, we hypothesize that network size (maintaining many contacts) has a negative average effect on work performance, as transferring simple knowledge through direct contacts is expensive. However, a large network may be justified when workers are executing complex tasks that require more information and multiple sources of corroboration.

Hypothesis 3a: On average, network size has negative effect on work performance.

Hypothesis 3b: Network size has positive effect on work performance when solving complex tasks.

3. Background and Data

We studied an IT configuration facility with 37 employees whose primary job is to guide, solicit and capture clients' IT configuration requirements, and to produce IT products according to those specifications. Interviews indicate that the data configuration process is information-intensive, requiring employees to quickly analyze the feasibility of specifications and build the system. Our interviews also indicate that talking to others is particularly helpful in improving a worker's overall understanding of the whole system. This can be viewed as workers using face-to-face communication to transfer embedded knowledge. Each configuration task is a single-person task and is randomly assigned given a workload constraint, much like a series of queued tasks (Aral, Brynjolfsson & Van Alstyne 2006).

To measure worker performance, we collected data on 911 configuration tasks during the experimental period of 25 working days (more than one month's activities at the facility). For each task, we gathered data on the task duration, difficulty level, the number of follow-ups, and information about the employee who performed the task. Although some of the tasks took less than a day to finish, tasks that took more than one day deserve special consideration as we cannot assume the worker is working on the task 24 hours a day. To better approximate the completion time of tasks that span multiple days, we assumed an 8-hour work day. Our interviews with staff indicate that employees typically follow this work schedule and rarely stay late or work on weekends to catch up. Although task completion time is only one dimension of work performance, it is an important outcome in the computing industry (Eisenhardt & Tabrizi 1995), and in this organization employees are formally evaluated on this metric.

<table>
<tr><td rowspan="6">
Figure 1: The Wearable Sociometric Badge</td><th>Capabilities of Wearable Sociometric Badge</th></tr>
<tr><td>Recognizing common daily human activities (such as sitting, standing, walking, and running) in real time using a 3-axis accelerometer (Olguin Olguin & Pentland, 2006).</td></tr>
<tr><td>Extracting speech features in real time to capture nonlinguistic social signals such as interest and excitement, the amount of influence each person has on another in a social interaction, and unconscious back-and-forth interjections, while ignoring the words themselves in order to assuage privacy concerns (Pentland, 2005).</td></tr>
<tr><td>Performing indoor user localization by measuring received signal strength and using triangulation algorithms that can achieve position estimation errors as low as 1.5 meters, which also allows for detection of people in close physical proximity (Sugano, Kawazoe, Ohta, & Murata, 2006; Gwon, Jain, & Kawahara, 2004).</td></tr>
<tr><td>Communicating with Bluetooth enabled cell phones, PDAs, and other devices to study user behavior and detect people in close proximity (Eagle & Pentland, 2006).</td></tr>
<tr><td>Capturing face-to-face interaction time using an IR sensor that can detect when two people wearing badges are facing each other within a 30°-cone and one meter distance. Choudhury (Choudhury, 2004) showed that it was possible to detect face-to-face conversations of more than one minute using an earlier version of the Sociometric badge with 87% accuracy.</td></tr>
</table>

To collect face-to-face and physical proximity interactions, we utilized the wearable Sociometric badge, a sensing device that collects behavioral data from many individuals over time (Waber et al. 2007). Our data collection method deserves special note. Instead of using surveys that have traditionally been used to construct social networks, we recorded every face-to-face interaction between workers using the Sociometric badge and continuously logged physical proximity to others, as well as many other behavioral features. The "wearable badge" form factor is particularly useful in organizational contexts. First, most organizations already require individuals to wear identification badges with embedded RFID. It is not hard to extend the sensing functionality of these badges further with accelerometers, IR transceivers, and microphones. Second, wearable badges are less obtrusive than sensors that have to be in physical contact with the user or require a long setup period to function. The success of IT products that employ this form factor for wearable sensors, such as the nTag (http://www.ntag.com/) and Vocera systems (http://www.vocera.com/) implies that this technology is acceptable to users in a wide variety of contexts. The capabilities of the wearable Sociometric badge and a picture of the wearable badge are shown in Figure 1. Waber et. al. (2007) provides detailed analysis of how the badge is used to detect worker interactions.

3.1 Network Variable Construction

Network size is simply the number of direct contacts one has. In face-to-face networks, a direct link between two actors exists when they engage in at least one conversation during the experimental period. Physical proximity networks, on the other hand, are a broader measure of direct links where network size counts an interaction between actors when they either engaged in a conversation or when they were physically (within ten meters) of each other. The *volume of interactions* measures the total interactions an actor has with anyone else in the network. This differs from network size as it counts all communication incidents regardless of with whom the actor has interacted. For example, an actor who communicates 100 times to a single person in the network would have the same volume of interactions as someone who communicates with 100 different people once. The network size of the former case is one, but in the latter case it is 100. While both variables measure the number of direct interactions between actors, network size may have a stronger effect than the volume of interactions in accessing and transferring complex information. Since high volume of interaction may also only involve a small group of actors, frequent interaction with the same person may be redundant and may not add value for knowledge transfer.

Table 1: Network Characteristics and Description

Network characteristics	Description
Network size	The total number of contacts with whom an actor exchanges at least one message
Volume of interactions	The total number of face-to-face interactions an actor experiences
Betweenness centrality	The probability of an actor that falls on the shortest path between any two other actors
Cohesion (constraint)	Degree to which an actor's contacts are connected to each other
Reach	The number of other people an actor can reach in two links or less

Betweenness centrality $B(n_i)$ measures the probability that an individual i will fall on the shortest path between any two other individuals in a network (Freeman 1979), where $g_{jk}(n)$ is the number of shortest geodesic paths from i to j that pass through a node n, while g_{jk} is the number of shortest geodesic paths from i to j:

$$B(n_i) = \sum_{j<k} g_{jk}(n_i) \Big/ g_{jk} ;$$

As shown in the hypothetical network in Table 3, actor 7 is located in a relatively more central position than actor 12. As actor 7 is closer to three different groups of actors, her betweenness centrality is much higher than actor 12.

Network reach measures the degree to which any member of a network can reach everyone else in the network. We measure 2-step reach which calculates the number of actors that an individual can reach in the network in 2 steps. We choose 2-step reach because our network is small enough that all actors are able to connect to everyone else in the network in three steps or less. Actor 7, located in the center of the network in Table 3, can reach eight other employees in two steps and therefore has a higher network reach than actor 12 who can only reach five others.

Network constraint C_i measures the degree to which an individual's contacts are connected to each other. P_{ij} is the proportion of *i*'s network time and energy invested in communicating with j. Network constraint can be used as proxy for measuring network cohesion (Burt 1992), and network diversity is simply computed as 1-C. In the hypothetical network in Table 3, C_{12} is much higher than C_7, because friends of actor 12 are more likely to be friends with each other than friends of actor 7. We construct network characteristics for both face-to-face and physical proximity. The network topologies are shown in Figures 2 and 3, and the summary statistics are shown in Tables 2 and 3.

$$C_i = \sum_j \left(p_{ij} + \sum_q p_{iq} p_{qj} \right)^2, \quad q \neq i, j.$$

3.3 Control Variable Construction

In addition to network structure, we posit two broad factors that may influence the task completion rate besides network variables: characteristics of tasks and individual workers. *Characteristics of tasks*: as harder tasks take longer to finish, task difficulty is strongly correlated with time to completion. We include two controls for task difficulty: *task complexity* and *the number of follow-ups.* Managers determine the *task complexity* and assign one of three levels to each task—basic, complex, or advanced—when the task is created. Basic tasks are the most common and rudimentary job, while difficult tasks take considerably more time on average. *The number of follow-ups* represents complexity measured during task execution. When a task is difficult and the requirements are unclear, employees may need further clarification. Accordingly, a task with many follow-ups tends to take longer to finish.

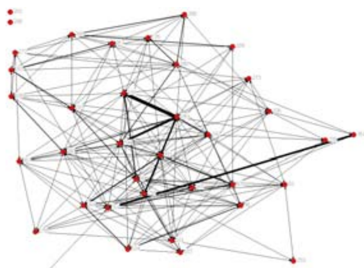


Figure 2: Face-to-face Conversation Network

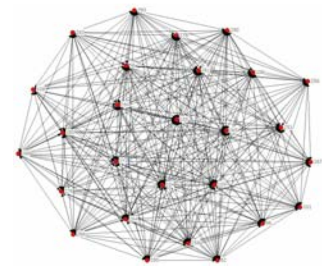
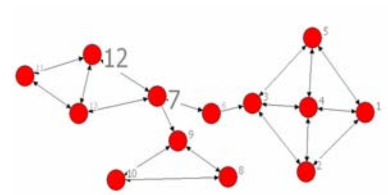


Figure 3: Weighted Face-to-face and Physical Proximity Network

Table 2: Summary Statistics F2F Network Variables

Latent Network Variables					
Variable	Obs	Mean	Std. Dev.	Min	Max
Interactions	931	526.62	421.52	156	2701
Network size	931	11.44	3.47	1	20
Betweenness	931	1.49	1.38	0	8.95
Constraint	931	0.53	0.19	0	1.33
2-step reach	931	86.73	7.52	0	94.44
In-Task Network Variables					
Interactions	937	26.71612	39.80165	0	287
Network Size	937	3.315902	3.10302	0	15
Betweenness	937	1.510955	2.450147	0	17.422
Constraint	937	0.523816	0.388709	0	1.9
2-step reach	937	33.23671	26.96877	0	94.29
3-day Network Variables					
Interactions	132	30.9697	48.84059	0	287
Network Size	132	3.666667	3.513695	0	13
Betweenness	132	1.648174	2.453405	0	10.729
Constraint	132	0.695031	0.268236	0.25	1.9
2-step reach	132	33.93879	28.11835	0	88.57

Table 3: Network Measures for a Hypothetical Network



Direct Contacts	Size(7)= 4 Size(12)= 3
Indirect Contacts	Btw(7)= 33 Btw(12)=6
	Reach(7)=67% Reach(12)=41%
Constraint	Constr(7)=0.47 Constr(12)=0.84

Characteristics of individual. We included controls for human capital using functional titles that classify employees into 3 categories: manager, pricing strategist and configuration specialist. While managers may be knowledgeable about the entire system, they are less likely to be intimately familiar with day-to-day configuration routines Although all three types of worker perform configuration tasks in our sample, the configuration specialist is most prepared to execute the configuration and we expect the complete tasks more quickly and accurately. Although we lack complete demographic data of workers, we infer some worker characteristics from the badge data. By measuring the tonal variance of workers, we can infer how animated a person is at the time (Pentland 2006). The animation of a worker's voice may give us indications about his general enthusiasm or motivation (Basu 2002, Pentland 2006). Summary statistics and correlations are listed in Tables 4 and 5.

Table 4: Summary Statistics for Worker and Task Characteristics

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Task Completion time (minutes)	1201	515.9159	968.8949	1	12281
Functional Title	1157	1.57822	0.502736	1	3
Task Complexity	1201	1.437968	0.761374	1	3
Number of Follow-ups	1217	4.612983	3.270522	0	21
Voice Animation	931	6703509	6056288	85	2.89E+07

Table 5: Pair Wise Correlations Between Independent Variables for the F2F Network

	Function	Complexity	Follow up	Animation	Volume	Direct links	Btw	Constraint	Reach
Function	1.00								
Complexity	-0.52	1.00							
Follow up	-0.52	0.42	1.00						
Animation	-0.14	0.04	0.03	1.00					
Interactions	-0.17	0.12	0.13	0.25	1.00				
Size	-0.31	0.16	0.26	0.43	0.62	1.00			
Betweenness	-0.27	0.16	0.17	0.43	0.71	0.87	1.00		
Constraint	-0.28	0.13	0.17	-0.17	-0.28	-0.44	-0.38	1.00	
Reach	-0.07	0.04	0.16	0.38	0.15	0.69	0.54	-0.46	1.00

4. Empirical Methods & Results

Combining task performance data and network data, we empirically test whether face-to-face and proximity networks are correlated with productivity and performance. Time to task completion measures how fast a person can finish a given task and based on our interviews, speed is a good measure of work performance in this setting. The accuracy or quality of configurations is also an important measure but only 20 of the 1217 tasks in our sample contain errors and 90% of those errors were caused by server configuration issues that are largely outside the control of individual workers. Since the majority of the tasks are completed correctly, completion time is a good metric for work performance. Although multitasking can increase total task throughput, it could confound the use of duration as the only performance measure as it increases the average time spent per task (Aral, Brynjolfsson & Van Alstyne 2007). However, in this setting, multitasking is not possible since tasks are assigned to workers one at a time. Consequently, task duration can be used as an overall measure of work performance. In order to complete a task faster, it is essential that workers have both basic configuration information, as well as information about the overall system. The latter is especially important for completing complex tasks. While most of the basic tasks are routine jobs for which referencing a manual is sufficient, complex tasks require significantly more information and more complex information. Our interviews indicate that communicating with peers is particularly beneficial for understanding the overall system, highlighting the importance of face-to-face communication for transferring tacit and embedded knowledge during complex tasks.

Since our dependent variable is the number of minutes it takes to complete a task, the model specification follows a duration model. We use a hazard rate model of the likelihood of a project completing at time *t*, conditional on it not having been completed earlier. The Cox proportional hazards model is used to examine the effect of network characteristics on project completion rate:

$$HazardRate\ (R) = f\,(size, betweennes\ s, cohesion\ , reach\ , taskComple\ xity\ , jobTitle\ , followups\ , tones)$$

$$R(t) = r(t)^b\ e^{\beta X}$$

where $R(t)$ represents the project completion rate, t is project time in the risk set, and $r(t)^b$ is the baseline completion rate when all the independent variables are set to zero. In this duration model, the effects of independent variables are specified in the exponential power, where β is a vector of estimated coefficients on a vector of independent variables X. β has a straight forward interpretation, where | β-1| represents the percentage increase (or decrease) in project completion rate associated with a one unit incrase in the independent variable depending on whether β-1 is positive (or negative). We tested this specification using both face-to-face and proximity-based network characteristics (Figures 2 and 3). The thickness of the lines in the graphs indicates the number of interactions between workers. As shown in the figures, there are more interactions between workers in the physical proximity-based network than in the face-to-face network because when people are engaged in conversation they are by definition close to each other, whereas two people who are not talking could still be in close proximity. Therefore, the face-to-face network is a subset of the proximity network.

We test the effects of four face-to-face network attributes on the speed of task completion: size, volume, reach, and cohesion. First, we use a single cross-sectional network over the entire experimental period to compute network variables. Constructing a network over the entire period allows us to assess the 'latent social network' that a worker can potentially leverage when completing a task. In addition, for every task in our sample, we construct an 'in-task social network' that includes only the interactions that took place while the worker was performing that particular task. For example, if a task takes 3 days to complete, the in-task network for that task is computed using the interactions of every worker for those 3 days. Task-specific social networks help us explore whether in-task networks differ from workers' latent networks of available contacts and how these differences affect performance.

Model 1 in Table 6 shows the effect of the cross-sectional or latent network on worker productivity. Unsurprisingly, complex tasks and tasks requiring more follow-ups display longer completion times on average. Interestingly, tonal variation, a proxy for employees' level of enthusiasm and motivation, has no effect on work performance. As predicted, network cohesion is positively correlated with work performance. Instead of reducing speed and productivity, as in email networks (Aral & Van Alstyne 2007), a one-standard-deviation increase in network constraint in face-to-face networks is associated with a 9.5% increase in the speed of task completion, demonstrating that cohesive ties in a face-to-face network are *more* conducive to productivity than diverse ties. We suspect that the information transmitted in face-to-face networks is inherently different from that which is transferred in email networks. It appears that the advantages of using face-to-face communication to transmit complex knowledge are enhanced in cohesive networks. Similarly, indirect contacts are positively correlated with the task completion rate.

Although the coefficient on betweenness centrality is not statistically significant, a one percent increase in network reach is associated with a 4% increase in the speed of task completion, demonstrating the power of indirect contacts in obtaining information. Lastly, direct contacts seem to have either no effect or a negative effect on the task completion rate. The total number of face-to-face interactions has a minimal impact on the time to finish a task. However, an additional network contact is associated with an 8% decrease in the average speed of task completion, demonstrating the potential cost in time, effort, and energy of maintaining face-to-face relationships.

Although the result using the cross-sectional latent network shows promising evidence supporting our hypotheses, the results may be driven by unobserved variation in individual characteristics, such as employees' inherent ability or ambition. Without comprehensive demographic data, it may be premature to attribute the performance differences to social networks alone. We therefore constructed a panel data set of in-task face-to-face networks for each task performed during the experiment. We employ the Cox proportional hazards model using fixed effects and random effects specifications to eliminate variance explained by any time-invariant characteristics of individual employees that could affect performance. The results are shown in Table 6, Models 2 and 3 respectively. The coefficients from the random effects model are roughly the same as with the latent cross-sectional network. Although the coefficients in the fixed effects model diminish in size, the signs of those coefficients and statistical significance retain powerful evidence supporting our hypotheses. We discuss these findings in greater detail below. We first address one modeling concern which led us to estimate a third family of specifications beyond the latent and in-task network estimations.

One possible concern in using task-specific networks to infer the effect of social networks on task completion time is that tasks of longer duration may generate larger networks and a greater volume of interactions by construction. As a result, task-level networks may be less meaningful without controls for the time a network is allowed to grow. To eliminate this potential bias in analyzing task-specific networks, we set the number of days that can be used to build a social network. With this constraint, observations are no longer task-centric but relate to a person over a fixed period of time. Instead of task completion time, the dependent variable is the average completion time for all tasks started within a fixed time period. Instead of task-specific network variables, we take into account interactions only within the same period, creating a more traditional balanced longitudinal, panel data set of the 37 workers measured using the same time periods for each worker. The summary statistics for all three types of face-to-face network variables are shown in Table 2.

$$AvgDuration_{it} = \alpha + \beta_1 AvgTaskComplexity_{it} + \beta_2 Network_{it} + \beta_3 Individual\_Characteristics_i + \varepsilon_{it}$$

We chose spells of different durations (1, 3 and 5 day panels) and the length of the periods do not affect our results. Table 6 shows the results of a linear regression on the balanced panel with both random and fixed effects where the periods are fixed at 3 day intervals. The results from task-specific networks and 3-day networks (Table 6, Models 2-5) are qualitatively similar for both random and fixed effects models, demonstrating that the network size bias in task-specific networks is minimal. Network constraint continues to have the strongest effect on worker performance even after eliminating time-invariant factors such as individuals' inherent ability. As shown in Model 3 of Table 6, the fixed effects model demonstrates that a one standard deviation increase in network constraint is associated with a 6% increase in the task completion rate. The evidence from both longitudinal in-task networks and latent cross-sectional networks demonstrates that cohesion in face-to-face networks is correlated with higher productivity supporting Hypothesis 1a. We speculate that this result holds because the information transmitted in face-to-face networks is inherently different from what is transferred in email networks. The advantage of using face-to-face communication to transmit complex knowledge is enhanced through cohesion, which allows workers to assimilate knowledge more effectively and in a timely manner. This is in direct contrast to email social networks, in which structural diversity has been shown to be more effective (Aral, Brynjolfsson & Van Alstyne 2006, Aral & Van Alstyne 2007).

Overall, the empirical results also support the argument that indirect contacts have a positive effect on work performance (Hypothesis 2a). Interestingly, the two measures of indirect contacts (betweenness centrality and network reach) have different effects on worker productivity depending on which network model is used. The cross-sectional network approximates the latent network of potential contacts that workers can leverage while completing a task, while the in-task network only contains interactions workers actualized while the task was being performed. Using the latent network, a 1% increase in network reach is associated with increasing the speed of task completion by 4%, whereas the coefficient on betweenness centrality is positive but statistically insignificant. Network reach measures the ability to connect to other employees and to eliminate information distortion caused by intermediaries. High network reach in the latent network improves the task completion rate because latent networks represent the constellation of colleagues an employee can potentially contact for information. A broader network reach in potential contacts is particularly effective in reducing information distortion because an employee can choose the shortest path to task-relevant experts among those potential contacts while performing a particular task. Betweenness centrality in the latent network is a proxy for path lengths to potential contacts. Path lengths in potential contact networks are less relevant to performance than path lengths in in-task networks because distortion occurs only when long path lengths are actually chosen to collect information, not when path lengths of potential contacts are long. We therefore expect betweenness in the latent network to be a noisy proxy for path lengths.

As expected, for the in-task network in which information seeking paths are actualized for each task, a one standard deviation increase in betweenness is correlated with 13 to 16% faster task completion, but the coefficient on network reach produces no effect. The positive coefficient on betweenness centrality for the in-task network implies that the ability to quickly access information needed for the task at hand is important for task completion. On the other hand, network reach shows no significant effect in in-task networks. We suspect this is because reach is only relevant for how far my potential contacts reach into the organization, but if employees can find the information relevant to a given task in a nearby local network neighborhood, then actualizing an in-task network of little reach is unlikely to hamper performance.

Table 6: The Effect of F2F Networks on Work Performance

Network Type	Cross-sectional Latent Network	Panel: In-Task Networks		Panel: 3-Day Network Panels		Physical Proximity Network
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent Variable	Completion Rate	Completion Rate	Completion Rate	Average Duration	Average Duration	Completion Rate
	Hazard	Hazard -RE	Hazard-FE	Linear RE	Linear FE	Hazard
Task	0.793***	0.755***	0.747***	51.98	19.42	0.774***
Complexity	(0.043)	(0.041)	(0.043)	(89.500)	(102.000)	(0.042)
Follow-ups	0.854***	0.857***	0.844***	109.8***	114.6***	0.863***
	(0.013)	(0.012)	(0.013)	(20.400)	(22.100)	(0.012)
Configuration	2.272***	2.504***	2.672***	-567.0***	--	2.047***
Specialist	(0.230)	(0.230)	(0.420)	(198.000)	--	(0.190)
Tonal	1.000	1.000	1.000	6.29E-07	--	1.000
Variation	(0.000)	(0.000)	(0.000)	(0.000)	--	(0.000)
InteractionsVolume	1.000**	0.997**	0.996***	-2.425*	-1.756	1.000
	(0.000)	(0.001)	(0.001)	(1.400)	(1.580)	(0.000)
Network	0.917***	0.951**	0.961*	139.3***	136.4***	1.002
Size	(0.026)	(0.024)	(0.024)	(45.500)	(49.300)	(0.017)
Network Cohesion	1.095**	1.116***	1.063*	-87.18*	-93.68*	0.855**
	(0.050)	(0.043)	(0.043)	(57.900)	(63.200)	(0.047)
BetweennessCentrality	1.092	1.163***	1.128***	-175.3***	-185.7**	0.959
	(0.089)	(0.047)	(0.047)	(65.500)	(69.800)	(0.049)
2 Step	1.039***	1.001	1.001	-11.08**	-11.92**	--
Reach	(0.010)	(0.003)	(0.003)	(4.450)	(4.810)	--[1]
Constant				184.9	-211.8	
				(230.000)	(238.000)	
Observations	911	911	911	93	93	911
Standard errors in parentheses *** p<0.001, ** p<0.05, * p<0.1						

The number of direct contacts has either a limited or a negative effect on the task completion rate for both cross-sectional and longitudinal models. Although the total number of face-to-face interactions has a minimal impact on the rate of task completion, an additional network contact in the in-task fixed effects model is associated with a 4% decrease in the task completion rate. Since most of the tasks in our sample are simple tasks, the negative impact of network size on work performance suggests a potential cost in time, effort, and energy to maintain face-to-face relationships. Disruptions during task execution can be especially distracting, as the cognitive cost of switching tasks can impede the rate of task completion (Aral, Brynjolfsson & Van Alstyne 2006). Our results show this effect is more pronounced for basic tasks that are relatively simple and do not require extensive face-to-face counseling. Since workers have less need to seek knowledge from other members of the group when performing basic tasks, the

[1] We excluded network reach in physical proximity models since there is little variation in reach levels among actors. Since physical proximity-based networks have many more interactions between actors than in face-to-face networks (Figure 2, 3), their topology also tend to be denser, where 90% of the actors achieved 100% reach level while the remaining 10% achieved 94.75% level. Therefore, to avoid multicollinearity problems, we eliminated network reach in our model.

delay in task completion is likely due to unproductive or interruption-driven communication from other colleagues. Since interruptions are costly to workers, we may also expect the total number of interactions to negatively affect work performance. In support of this argument, the volume of interactions for the in-task network has a slight negative effect on task completion, suggesting interactions with anyone during task execution has a cognitive disruption cost. The total number of interactions has no impact on work completion in the latent network however, implying that interactions outside of the task have little disruptive impact on the average task duration, demonstrating the importance of separating in-task and latent networks and interaction characteristics.
Lastly, when we compare face-to-face networks with physical proximity networks, we see (in Model 6) that most of the coefficients in the physical proximity network are insignificant, demonstrating that face-to-face conversations matter more than physical proximity alone. Conley and Udry (2005) find similar results when studying the effect of social networks effect on the use of fertilizer in Ghanaian pineapple farms.

4.1 The Effect of Network Structure on Completing Complex Tasks

As cohesive networks enable more effective transfers of complex knowledge (Reagans & McEvily 2003, Hansen 1999) we expect cohesion to be more important when employees are engaged in complex tasks. Given the cost of face-to-face interactions in time, effort, energy and interruption, we also expect additional interactions during task execution (e.g. in-task networks) to reduce the speed of project completion, but to help increase the speed of project completion on complex tasks that require more information, advice and tacit guidance from colleagues. For complex tasks we expect the benefits of interaction to outweigh the costs, whereas for simple tasks we expect there to be less benefit to interaction, while still creating costs. We also expect network size to follow the same pattern – costly for simple tasks but beneficial for complex tasks that require more support from colleagues. To test these expectations, we add interaction terms between task difficulty levels, the volume of interactions and network variables. The results in Table 7 lend broad support to our expectations with one interesting deviation.

The interaction of task complexity and network reach is statistically significant and has a positive effect in both the latent network and the 3-day network, lending support to Hypothesis 2b. For the latent model, a 1% increase in network reach is associated with a 3% greater increase in the rate of completing a complex task (when compared to tasks on average) while for the 3-day model with fixed effects, a one percent increase in network reach is associated with an 18 minute decrease in the average completion time. Network reach, the ability to access colleagues in the network in 2 steps or less, can be beneficial in a latent network of available contacts for contacting the right expert to access and assimilate the knowledge needed to complete the task. Consequently, network reach is especially helpful for completing difficult tasks that require more information, advice and tacit guidance from colleagues. In addition, network reach eliminates intermediaries and can reduce information distortion, which is particularly important when transferring complex information. We expect network reach to have less significance in analyses of in-task networks, as in-task networks measure networks of colleagues that have already been chosen as contacts during the execution of a particular task. The ability to reach a large portion of the network is less relevant once the most appropriate colleagues have already been chosen. In fact, we find the interaction terms for task complexity and network reach to be insignificant in analyses of in-task networks.

We also find that more interactions are costly on average during task execution with an additional interaction correlated with reducing the project completion rate by about 1%. However, more interactions are beneficial when tasks are complex and require more guidance. The interaction terms in Models 2 & 3 show that for complex tasks, one additional interaction increases the project completion rate by between .4 - .5%. We suspect that for complex tasks the benefits of interaction outweigh the costs, whereas for simple tasks there are fewer benefits to interaction, while still creating costs.

We expected network size to follow a similar pattern to interaction volume, with the maintenance cost of face-to-face interactions with more people reducing the project completion rate on average but increasing the speed of project completion for complex tasks. We found strong evidence of the cost of network size on the completion rate in Table 6, but in Table 7 we see this is true even for complex tasks. The interaction terms in Models 4 & 5 show that larger networks increase the average duration of complex tasks as well. This contrasts the interactions volume result, demonstrating that *more* interactions with *fewer* people are the most beneficial for increasing the speed of work. We suspect that employees who seek information from a greater number of colleagues not only experience a cost to those interactions, but are also not finding the information they are looking for and thus are seeking advice from additional colleagues. Our interviews corroborate this finding as employees report having to contact more people when they can't find the guidance they are seeking. More interactions with the right colleagues are helpful on complex tasks, but seeking advice from many colleagues is not only costly but also signals an inability to find the information necessary to complete the task quickly. It could also be that more interactions with fewer colleagues generate a higher degree of mutual understanding and conversational rapport that facilitates more efficient transfers of complex knowledge.

Table 7: The Effect of F2F Networks on Completing Complex Task--Complexity Using Job Difficulty

Network type	Cross-sectional Latent Network	Panel: In-Task Networks		Panel: 3-Day Network Panels		Physical Proximity Network
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dependent Variable	Completion Rate	Completion Rate	Completion Rate	Average Duration	Average Duration	Completion Rate
	Hazard	Hazard -RE	Hazard-FE	Linear RE	Linear FE	Hazard
Task	0.0695**	0.760***	0.718***	206.8	130.4	0.718
Complexity	(0.079)	(0.071)	(0.068)	(216.000)	(242.000)	(0.380)
Follow-ups	0.855***	0.854***	0.840***	99.89***	98.57***	0.861***
	(0.013)	(0.012)	(0.013)	(21.300)	(24.200)	(0.012)
Configuration	2.204***	2.559***	2.799***	-582.2***	--	2.117***
Specialist	(0.240)	(0.240)	(0.440)	(210.000)	--	(0.210)
Tonal	1.000	1.000	1.000	4.13E-06	--	1.000
Variation	(0.000)	(0.000)	(0.000)	(0.000)	--	(0.000)
Interaction	1.000	0.989***	0.990***	-2.25	0.689	1.000
Volume	(0.000)	(0.003)	(0.003)	(4.650)	(5.350)	(0.000)
Network	0.972	1.021	1.020	-105.2	-155.3	0.997
Size	(0.066)	(0.059)	(0.057)	(133.000)	(149.000)	(0.042)
Network	1.006	1.069	1.014	231	250.2	0.855
Cohesion	(0.100)	(0.089)	(0.087)	(207.000)	(238.000)	(0.100)
Betweenness	1.225	1.106	1.027	199.9	251.2	0.857
centrality	(0.220)	(0.110)	(0.099)	(214.000)	(248.000)	(0.100)
2-Step	0.991	1.001	0.999	16.67	16.65	1.000
Reach	(0.023)	(0.006)	(0.006)	(12.900)	(14.600)	(0.000)
ComplexityX	1.000*	1.005***	1.004**	-0.15	-1.416	--
Interact. Vol.	(0.000)	(0.002)	(0.002)	(2.330)	(2.650)	--
ComplexityX	0.958	0.956	0.963	154.4**	176.9**	0.999
Network Size	(0.038)	(0.033)	(0.033)	(75.500)	(83.900)	(0.021)
ComplexityX	1.080	1.031	1.037	-200.4*	-216.1*	1.070
Cohesion	(0.071)	(0.052)	(0.054)	(125.000)	(143.000)	(0.066)
ComplexityX	0.966	1.034	1.067	-247.9**	-282.2*	0.982
Betweenness	(0.092)	(0.058)	(0.062)	(125.000)	(145.000)	(0.081)
ComplexityX	1.032**	1.000	1.002	-17.68**	-17.58**	--
Reach	(0.016)	(0.004)	(0.004)	(7.620)	(8.530)	--
Constant				3.345	-267.9	
				(409)	(466)	
Observations	911	911	911	93	93	911
Standard errors in parentheses, *** p<0.001, ** p<0.05, * p<0.1						

As expected, the interaction with betweenness centrality is insignificant in the latent model, since short path lengths in actualized in-task networks are those that we expect to be beneficial. We have strong evidence, displayed in Table 6, that betweenness in in-task networks is correlated with faster project completion. When the interaction term with task complexity is introduced, we observe several interesting results. First, the average positive effect of betweenness on the task completion rate decreases to zero and the interaction terms with task complexity are positive in Models 2-5 demonstrating statistically significant reductions in the average duration of complex tasks in Models 4 & 5. These results provide consistent (albeit weak) evidence that higher betweenness is more beneficial for complex tasks.

The coefficient of the interaction term between network cohesion and task complexity is positive but statistically insignificant. However, if we use the number of follow-ups as the measurement for task complexity, this interaction becomes significant[2], lending partial support for hypothesis 1b. Lastly, the physical proximity network displays no significant effect on task completion, suggesting that face-to-face conversations are more important than physical proximity when completing complex tasks.

5. Discussion and Conclusion

[2] The table that use the number of follow-ups to interact with network variables is not displayed here due to space constraints. The results are qualitatively similar to what is shown in Table 7. However, the interaction terms with network cohesion is statistically significant, demonstrating that an additional interaction is positively associated with a 4% increase in task completion speed.

We use new tools and methodologies to collect precise real time data on face-to-face interactions in an IT configuration facility. By matching data obtained through the use of wearable Sociometric badges with detailed performance data from the firm's accounting records we are able to test the effects of face-to-face interaction networks on individual information worker performance. Although detailed data on electronic interactions (e.g. email, phone logs, instant messaging) has become readily available in recent years, our ability to record network data for face-to-face interactions has lagged behind. The tools and methods presented in this paper give researchers important new opportunities for collecting fine grained data about the flow of information and knowledge in face-to-face interaction networks in real organizations, opening new avenues for research into social networks, knowledge management and IT use in organizations and elevating data collection on face-to-face networks to the standards of accuracy and precision displayed in electronic communication data.

We also make important theoretical contributions to Information Systems research. Until now social network theories (e.g. Granovetter 1973, Burt 1992) and information richness theory (Daft & Lengel 1987) have been used independently to understand knowledge transfer in information intensive work. Social network theories explain how network structures covary with the diffusion and distribution of information, but largely ignore characteristics of communication channels. Information richness theory focuses explicitly on communication channel requirements for different types of knowledge transfer but ignores the population level topology through which information is transferred in a network. We bridge these two sets of theory to understand what types of social structures are most conducive to transferring knowledge and improving performance in face-to-face communication networks.

Our research uncovers four main results. First, optimal face-to-face communication networks display very different topological structures compared to email networks. In both cross-sectional and longitudinal models of face-to-face networks, network cohesion is associated with higher productivity, while the opposite is true in email communication. We suspect that information transmitted in face-to-face networks is more tacit, complex and embedded than information transferred through electronic channels, and that the advantages of using face-to-face communication to transmit complex knowledge are enhanced by cohesion which increases norms of trust, effective communication heuristics and absorptive capacity through the provision of multiple perspectives on a problem. Second, we find that cohesion in face-to-face networks is especially effective when solving complex problems, suggesting that cohesion complements information-rich communication media for the effective transmission of complex tacit knowledge when conducting complex tasks. Third, we show that the most effective network structures for "latent" social networks (those that characterize the network of available communication partners) differ from "in-task" social networks (those that characterize the network of communication partners that are actualized during the execution of a particular task). We find betweenness centrality is important for in-task networks, as occupying a central position in the network accelerates the speed of obtaining information required to complete the task. On the other hand, network reach is more important in the latent network than in the in-task network, as the latent model represents the network of potential contacts a worker could leverage when completing a task. We find direct contacts have a negative impact on task-completion for in-task networks, as the cognitive cost of interruptions is high during task execution, but that more interactions with fewer people speed project completion for complex tasks, which require more information and guidance from colleagues. Finally, the effect of cohesion is stronger in face-to-face networks than in physical proximity networks, demonstrating that information flows in actual conversations (rather than mere physical proximity) are driving our results.

There are two main limitations of our work. First, we have no access to email, phone or IM traffic. Without data on other communication channels, it is difficult to make direct comparisons of the magnitudes and directions of our results across different communication media. However, our on-site interviews indicate that the information transferred in face-face conversations in our firm may be fundamentally different from that which is transferred in electronic media. Second, although our longitudinal models allow us to control for variance explained by any time-invariant characteristics of employees, our results may still be biased by unobserved time-varying characteristics such as media choice at different points during a task or simultaneity. Although we do not make causal interpretations of our parameter estimates, our fixed effects analyses which control for omitted variables that could explain our results, combined with interview evidence, suggest that a causal interpretation is plausible.

Caveats aside, our results represent some of the first evidence measuring the effects of a face-to-face communication networks on information worker performance. Using innovative technology to record face-to-face interactions, we link information richness theory and social network theories to show that in contrast to email networks, cohesive networks in a rich communication medium such as face-to-face interaction are associated with higher employee performance. The unique characteristics of face-to-face networks highlight the need to distinguish them from other types of communication networks, particularly when analyzing their effects on productivity and performance.

References

Ancona, D.G. & Caldwell, D.F. 1992. "Demography & Design: Predictors of new Product Team Performance." Organization Science, 3(3): 321-341.

Apte, U., Nath, H. 2004. "Size, structure and growth of the U.S. economy." Center for Management in the Information Economy, Business and Information Technologies Project (BIT) Working Paper.

Aral, S., Brynjolfsson, E., & Van Alstyne, M. 2006. "Information, Technology and Information Worker_Productivity: Task Level Evidence." Proceedings of the 27th Annual International Conference on Information Systems, Milwaukee, Wisconsin.

Aral, S., Brynjolfsson, E., & Van Alstyne, M. (2007) "Productivity Effects of Information Diffusion in Networks." Proceedings of the 28th Annual International Conference on Information Systems, Montreal, CA.

Aral, S., & Van Alstyne, M. 2007. "Networks, Information & Social Capital" International Conference on Network Science 2007

Argote, L. 1999. Organizational Learning: Creating, Retaining and Transferring Knowledge. Kluwer Academic Publishers, Boston, MA

Boorman, Scott A. 1975 "A combinatorial optimization model for transmission of job information through contact networks." Bell Journal of Economics, 6: 21 6-249.

Brass, D. J., M. E. Burkhardt. 1992. Centrality and power in organizations. N. Nohria and R. G. Eccles, eds. Networks and Organizations: Structure, Form and Action. HBS Press, Boston, MA, 191–215.

Burt, R. 1987. "Social Contagion & Innovation: Cohesion versus Structural Equivalence." *American Journal of Sociology*, 92: 1287-1335.

Burt, R. 1992. "Structural Holes: The Social Structure of Competition." Harvard University Press, Cambridge, MA.

Burt, R. 1997. "The Contingent Value of Social Capital", Administrative Science Quarterly, Vol 42. No. 2

Burt, R. 2000. "The network structure of social capital" In B. Staw, & Sutton, R. (Ed.), Research in organizational behavior (Vol. 22). New York, NY, JAI Press.

Burt, R. 2004. "Structural Holes & Good Ideas" American Journal of Sociology, (110): 349-99.

Coleman, J.S. 1988. "Social Capital in the Creation of Human Capital." American Journal of Sociology, (94): S95-S120.

Cohen, W., Levinthal, D., 1990, "Absorptive Capacity: A New Perspective on Learning and Innovation", Administrative Science Quarterly 35(1) pp 128-152.

Conley and Udry (2005), "Learning about a New Technology: Pineapple in Ghana, Yale University Working Paper

Contu, A., H. Willmott. 2003. Re-embedding situatedness: The importance of power relations in learning theory. Organizational Science 14(3) 283–296.

Cummings, J., & Cross, R. 2003. "Structural properties of work groups and their consequences for performance." Social Networks, 25(3):197-210.

Daft, R. L. and Lengel, R.H. "Organizational Information Requirements, Media Richness and Structural Design," Management Science, 32, 5, 1986: 554-571.

Eisenhardt, K., Tabrizi, B., 1995, "Accelerating adaptive processes: Product Innovation In the global computer industry." Administrative Science Quarterly, 40: 84-1 10.

Freeman, L. 1979. Centrality in social networks: Conceptual clarification. Social Networks 1(3) 215–234. Garguilo, M., and A. Rus 2002 "Access and mobilization: Social capital and top management response to market shocks." Working paper, INSEAD.

Gilovich, T. 1991. How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life. Free Press, New York.

Granovetter, M., 1973 "The strength of weak ties." American Journal of Sociology, 6: 1360-1380.

Granovetter, M., 1978. “Threshold models of collective behavior.” American Journal of Sociology (83:6):1420-1443.

Granovetter, M., 1982 "The strength of weak ties: A network theory revisited." In P. V. Marsden and N. Lin(eds.), Social Structure and Network Analysis: 105-1 30.

Granovetter, M. 1985. “Economic Action & Social Structure: The Problem of Embeddedness.” American Journal of Sociology (91):1420-1443.

Granovetter, M. 1992. “Problems of Explanation in Economic Sociology.” In N. Nohria & R.G. Eccles (eds.), Networks & Organizations: 25-56. Harvard Business School Press, Boston.

Grant, R. 1996. Prospering in dynamically-competitive environments:Organizational capability as knowledge integration.

Organ.Sci. 7(4) 375–387. Hinds, P.J. and Mortensen, M., 2005 “Understanding Conflict in Geographically Distributed Teams: The Moderating Effects of Shared Identity, Shared Context, and Spontaneous Communication,” Organization Science, 16, 3: 290-307.

Hansen, M. 1999. "The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits." Administrative Science Quarterly (44:1): 82-111.

Hansen, M. 2002. "Knowledge networks: Explaining effective knowledge sharing in multiunit companies." Organization Science (13:3): 232-248.

Markus, M.L., 1994, “Electronic Mail as the Medium of Managerial Choice,” Organization Science, 5, 4: 502-527.
Marsden, P. 1990. “Network Data & Measurement.” *Annual Review of Sociology* (16): 435-463.

McCain, B., O’Reilly, C., Pfeffer, J. 1983. “The effects of departmental demography on turnover: The case of a university”. Acadamy of Management J. 26 626–641.

Nelson, R., and Winter, G., 1982 “An Evolutionary Theory of Economic Change.” Cambridge, MA: Belknap Press.

Nonaka, I. (1990). “Redundant. overlapping organization: A Japanese approach to managing the innovation process”, California Management Review, Spring, pp. 27-38.

Nonaka, I. “A Dynamic Theory of Organizational Knowledge Creation,” Organization Science (5:1), 1994, pp. 14-37.

O’Reilly, C., 1980 “Individuals and Information Overload in Organizations: Is More Necessarily Better?” The Academy of Management Journal, Vol. 23, No. 4. (Dec., 1980), pp. 684-696.

O’Reilly, C., Caldwell, D., Barnett, W. 1989. Work group demography, social integration, and turnover. Administrative Science Quarterly. 34 21–37.

Podolny, J., 1994 "Market uncertainty and the social character of economic exchange." Administrative Science Quarterly, 39: 458- 483.

Podolny, J., Baron, J. 1997. "Resources and relationships: Social networks and mobility in the work-place." American Sociological Review (62:5): 673-693.

Podolny, J. 2001. “Networks as the Pipes and Prisms of the Market.” American Journal of Sociology, (107:1): 33-60.

Polanyi, M. 1966. “The Tacit Dimension.” New York: Anchor Day Books.

Reagans, R. & McEvily, B. 2003. “Network Structure & Knowledge Transfer: The Effects of Cohesion & Range.” Administrative Science Quarterly, (48): 240-67.

Reagans, R. & Zuckerman, E. 2001. "Networks, diversity, and productivity: The social capital of corporate R&D teams." Organization Science (12:4): 502-517

Slaughter, S. and Kirsch, L. 2006. “The Effectiveness of Knowledge Transfer Portfolios in Software Process Improvement: A Field Study.” Information Systems Research, Vol. 17, No. 3, September 2006, pp. 301–320

Sparrowe, R., Liden, R., Wayne, S., & Kraimer, M., 2001. “Social networks and the performance of individuals and groups.” Academy of Management Journal, 44(2): 316-325.

Sproull, L. and Kiesler, S. 1986, “Reducing Social Context Cues: Electronic Mail in Organizational Communication,” Management Science, 32, 11, 1492-1512.

Sproull, L. and Kiesler, 1991a, "Computers, Networks, and Work," Scientific American, September, 84-91.

Sproull, L. and Kiesler, 1991b “Connections: New Ways of Working in the Networked Organization, Cambridge, MA: M.I.T. Press.

Trevino, L.K., R. L. Daft and R. H. Lengel.1991, "Understanding Managers' Media Choices: A Symbolic Interactionist Perspective," in J. Fulk and C. Steinfield (Eds.), Organizations and Communication Technology, Newbury Park, CA: Sage, 71-94.

Teece, David J. 1986 "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy." Research Policy, 15: 285-305.

Waber, B.N., Olguín Olguín, D., Kim, T., Mohan, A., Ara, K., and Pentland, A. 2007. “Organizational Engineering using Sociometric Badges” International Conference on Network Science, New York, NY.

Walther, J. 1995, “Relational Aspects of Computer-Mediated Communication: Experimental observations over time”, Organization Science, Vol6, No.2

Winter, S., 1987 "Knowledge and competence as strategic assets." In DavidJ. Teece (ed.), The Competitive Challenge: 159-184. Cambridge, MA: Ballinger.

Uzzi, B. 1996. “The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect.” American Sociological Review, (61):674-98.

Uzzi, B. 1997. “Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness.” Administrative Science Quarterly, 42: 35-67.

Zenger, T., Lawrence, B. 1989. “Organizational demography: The differential effects of age and tenure distributions on technical communication.” Academy of Management J. 32 353–376.