

INSEAD

The Business School
for the World®

Faculty & Research Working Paper

The World is Not Small for Everyone:
Inequity in Searching for Knowledge
in Organizations

Jasjit SINGH
Morten HANSEN
Joel PODOLNY
2010/11/ST/EFE
(Revised version of 2009/49/ST/EFE)

**The World is Not Small for Everyone:
Inequity in Searching for Knowledge in Organizations**

By

Jasjit Singh*

Morten T. Hansen**

Joel M. Podolny***
**

February 2010

Revised version of INSEAD Working Paper 2009/49/ST/EFE

* Assistant Professor of Strategy at INSEAD, 1 Ayer Rajah Avenue, Singapore 138676 Ph: +65 6799 5341 E-Mail: jasjit.singh@insead.edu

** Professor of Entrepreneurship at INSEAD, Boulevard de Constance, Fontainebleau 77305 France Ph: +33 (0) 1 60 72 40 20 E-Mail: morten.hansen@insead.edu

*** Professor at Apple University, 1 Infinite Loop, MS 301-4AU, Cupertino, CA 95014, USA
Email: podolny@apple.com

A working paper in the INSEAD Working Paper Series is intended as a means whereby a faculty researcher's thoughts and findings may be communicated to interested readers. The paper should be considered preliminary in nature and may require revision.

Printed at INSEAD, Fontainebleau, France. Kindly do not reproduce or circulate without permission.

ABSTRACT

We explore why some employees may be at a disadvantage in searching for information in large complex organizations. The “small world” argument in social network theory emphasizes that people are on an average only a few connections away from the information they seek. However, we argue that such a network structure may benefit some people more than others. Specifically, some employees may have longer search paths in locating knowledge in an organization—their world may be large. We theorize that this disadvantage is the result of more than just an inferior network position. Instead, two mechanisms—periphery status and homophily—jointly operate to aggravate the inefficiency of search for knowledge. Employees who belong to the periphery of an organization by virtue of being of minority gender, lower tenure or poor connectedness to experts have limited awareness of who knows what in an organization and a lower ability to seek help from others best suited to guide the search. When they start a search chain, they are likely to engage in homophilous search by contacting colleagues like themselves, thus contacting others who also belong to the periphery. To search effectively, employees on the periphery need to engage in heterophilous search behaviors by crossing social boundaries. We find support for these arguments in a network field experiment consisting of 381 unfolding search chains in a large multinational professional services firm. The framework helps explain employees’ unequal pathways to the knowledge they seek, a poorly understood yet important type of organizational inequity in an information economy.

Keywords: Social networks; Small world; Network search; Homophily; Knowledge sharing; Inequity

INTRODUCTION

The small world argument in social network research holds that individuals are connected to one another through a seemingly small number of intermediary contacts. In the early research on the small world phenomenon, Travers and Milgram (1969) reported that an average of only five intermediaries (six steps) seemed to connect strangers in the U.S. (Guare 1990). While some subsequent research focused on average path lengths between individuals (Korte and Milgram 1970, Killworth and Bernard 1978, 1979 Lin et al. 1978), more recent research has analyzed the properties of the network structure underlying a small world network and has revealed an overarching network structure that is characterized by high local clustering and short global separation (Watts and Strogatz 1998, Kogut and Walker 2001, Baum et al. 2003, Davis et al. 2003, Uzzi and Spiro 2005, Fleming et al. 2007). Actors are on average connected through few intermediaries because of this underlying network structure.

While the focus on the overarching network structure with short average path lengths has yielded important insights about the properties of the networks as a whole, the mere existence of short paths does not guarantee that individual actors searching through these networks would be able leverage these paths. In fact, Kleinberg (2000) shows that, under fairly general assumptions, often no search algorithm exists that can help actors realize the shortest paths by relying on local information only. The search problem is further complicated by heterogeneity among actors, as different actors may differ in their cognitions regarding the overall network. For example, Krackhardt (1990) demonstrates that high-power employees tend to have more accurate cognition of the overall advice network in an organization. Yet empirical research has paid limited attention to individual-level variables that may explain resulting differences in search outcomes among individuals within a small world network.

Thus, while we know a fair amount about the average path length in small world networks (e.g, Dodds et al. 2003) and the structural properties that give rise to them (Watts 1999), we know much less about reasons for individual variations within them. As a consequence, current small-world network theories cannot explain well why some individuals and may be at a disadvantage in searching for information. This is especially the case for search among employees *within* an organization, as most

research on small world focuses on organizations (e.g, Kogut and Walker 2001, Schilling and Phelps 2007) or individuals outside of an organizational context (e.g., Dodds et al. 2003, Uzzi and Spiro 2005, Fleming et al. 2007; but see Adamic and Adar 2005 and Killworth et al. 2006 for exceptions). Understanding how employees in certain social categories may be systematically worse off in searching for information in an organization is important, because such disadvantages may negatively affect their work performance, with negative consequences for career progressions and pay (Uzzi et al. 2007).

To address the paucity of network research on individual variation in intra-organizational search behaviors, we explore the following question: why do some employees have longer search chains to locate an expert in their organization than do others? To answer this question, we develop a framework of unfolding search chains in organizations wherein an original searcher sequentially contacts intermediaries in order to locate an expert on a specific work-related topic. This chain can be characterized by its direction (i.e., who the searcher goes to for help in locating experts) and its length (i.e., the number of intermediaries needed to reach an expert). We draw on and expand two organization theory concepts to develop the framework. First we use the concept of core-periphery status in theories of power and stratification in organization research (e.g., Kanter 1977, Brass 1984): employees who belong to the periphery of an organization by virtue of attributes such as network centrality, gender, race or tenure may find themselves cut off from critical information flows and thus encounter difficulties in conducting efficient searches. Second, using the homophily concept (Lazarsfeld and Merton 1954, Ibarra 1992, 1997, McPherson et al. 2001), we seek to explain the direction of the search chain and why the length of the chain may be longer for employees on the periphery of an organization's social structure.

In this approach, we combine two lines of research in the social network tradition—the small world methodology and traditional intra-organization network research (e.g., Burt 1992, Hansen 1999, Tsai 2001, Borgatti and Cross 2003, Reagans and McEvily 2003, Burt 2004). In the latter body of research, studies tend to conceptualize and measure employees' positions in an organizational network, through concepts such as centrality and structural holes, and link those properties to outcomes, such as having good ideas, being promoted, and finishing tasks quickly. Yet this line of research does not unpack

“the black box” of individual behaviors that lie between network structure properties, on one hand, and outcomes, on the other. In contrast, we analyze the effects of employees’ positions in the organization’s network structure on individual behaviors—the actual steps employees follow in the search chain (who they contact and how long their search chain will be). This approach thus advances intra-organization network research by linking the “macro” level of network structure to the “micro” level of individual search behaviors.

To simplify our analysis, we limit our discussion to search for individual knowledge experts within an organization. We focus on the use of network contacts and intermediaries to locate experts and exclude considerations of the use of information technology and knowledge databases. We situate the analysis in one particular setting, that of searching for experts in large, multi-office and multinational professional services firms, such as law, consulting or investment banking firms (cf. Podolny 1993, Hansen and Haas 2001, Phillips 2001, Rogan 2008). In these firms, professionals often need to draw on the expertise of colleagues in order to develop a solution for a client, but they may not know *ex ante* who the experts are on a particular topic and in which offices they reside.

Our empirical setting is a large multi-national management consulting company with 3,150 professionals located in 50 offices in 34 different countries at the time of our study. We conducted a field experiment and began by determining the identity of the firm’s experts on four consulting topics. We then randomly initiated 381 individual search chains, starting with 381 *different* individuals randomly drawn from the firm’s employee base, and asked survey respondents to name an expert on a topic or someone who could help them identify an expert. Using the “snowball method,” we followed up with all new named contacts, who then submitted new contact names, and so on, until a chain reached an expert or stopped due to lack of response. To analyze the effect of individual positions in the network, we collected data on the firm’s complete network structure, allowing us to study the effects of network structure on individual search behaviors.

SEARCH CHAINS AND INDIVIDUAL-LEVEL VARIATION

Following the early research on the small world phenomenon (Milgram 1967, Travers and Milgram 1969), subsequent research in this area has pursued three related lines of inquiry with different levels of analysis.

In the first line, researchers have sought to verify whether the number of intermediaries required is indeed as small as that found by Milgram and colleagues and have conducted studies to compute average path lengths in various populations, including an urban area (Lin. et al. 1978), two cities (Korte and Milgram 1970), inventors in the U.S. (Fleming et al. 2007), and board of directors in German companies (Kogut and Walker 2001). Notwithstanding methodological issues (Kleinfeld 2002), a partial answer seems to be that the number of steps is indeed surprisingly small. For example, a large global population study using the Internet demonstrated a median path length of five to seven steps even after taking into account incomplete chains (Dodds et al. 2003).

The finding that paths to a target are surprisingly short on average raises the issue of why this is the case. In a second line of research, Watts and Strogatz (1998), following an early lead by Pool and Kochen (1978) and Kochen (1989), shifted the level of analysis to the properties of the overall network that give rise to short average path lengths. They proposed that small world networks exhibit two features: individuals are grouped into local clusters with high density, and these clusters are in turn linked globally through a few ties to others outside the cluster. Thus, an individual may only have relationships to others in the cluster but is nevertheless linked globally because of occasional global ties held by others in the cluster. This explains how individuals can have short path lengths in networks with overall low density.

Studying average path lengths and overall network properties do not, however, address much the issue of why there may be large variations among actors in a small world network. The Watts and Strogatz model may explain why the *average* path length is relatively short but it does not explain why some individuals or a certain category of individuals, such as women, may require more steps to complete

a chain than do other categories. A related but different line of research is required to understand the small world problem at the individual level of analysis.

In a third line of research, scholars have examined how individuals may find it difficult to traverse subgroup boundaries to reach a target: Caucasian individuals starting a chain experienced more difficulties reaching an African American target than a Caucasian (Korte and Milgram 1970); crossing racial boundaries was less likely to be attempted and less likely to be effective (Lin. et al. 1978); and low-income individuals failed to get messages through to targets in higher income groups (Kleinfeld 2002). Crossing social boundaries appear to make the world larger, and some subgroups find it more difficult to complete chains than do others. These studies, however, have not shed much light on why this may be the case and have not deployed predictive models that may explain why individuals belonging to a sub-group engage in search strategies that turn out to be less effective.

Search strategies

Individual chain lengths may vary because of different individual-level search strategies. One approach is to use clues about the eventual target and select intermediaries who are in close physical proximity to the target, who are in a similar profession as the target, or who are closest in the hierarchy to the target in an organization (Barnard et al. 1982, Adamic and Adar 2005). Relying on medium to weak ties and professional ties has also been shown to increase the completion rate of chains (Dodds et al. 2003). While a good beginning on which to build, these approaches to individual search strategies have several limitations and need to be extended, in three ways.

First, rationality-based models assume that individuals may indeed pursue a certain search strategy, such as selecting an intermediary who is higher up in the organizational hierarchy than themselves (Adamic and Adar 2005). But this approach presumes that individuals will cross organizational and social status boundaries, something they may be reluctant to do because of social barriers. In some sense these models are “under-socialized” by analyzing search approaches that do not take into account the social, organizational and demographic context within which search takes place.

Second, these small world studies are premised on search for a pre-determined person in a specified location (i.e., a target), such as Travers and Milgram's stockbroker who lived in the Boston area. Thus, much of the analysis has centered on search strategies based on clues about the target (Travers and Milgram 1969, Killworth and Bernard 1978, Barnard et al. 1982, Dodds et al. 2003, Adamic and Adar 2005). In an organization context, however, actors often do not know the exact identity of the actor who may possess the information they need (cf. Stuart and Podolny 1996, Katila and Ahuja 2002, Denrell et al. 2004). Search in an organization context is therefore often best understood as search without knowing *ex ante* the end destination.

Third, actual search behaviors are embedded within a social network structure that to some extent govern the direction and efficacy of search, but to empirically analyze this requires data on both actual search behaviors (i.e., an unfolding search chain) *and* the overall network structure that exists prior to a given search attempt. Existing studies have data on either the overall network structure (e.g., Watts and Strogatz 1998, Adamic and Adar 2005, Fleming et al. 2007) or the actual search (e.g., Travers and Milgram 1969, Dodds et al. 2003), but few have analyzed the effects of network structure on actual individual searches, including that within large, complex organizations. As one exception, Killworth et al. (2006) studied choice of contacts within a social network of 105 telephone survey interviewers in a small organization, but this study again analyzed reaching pre-determined targets and did not include unfolding search chains across subunit boundaries in an organization.

Thus, given our interest in understanding individual variations in the direction and efficacy of search for knowledge in large, complex organizations, we develop a model that extends extant research in several significant directions: we allow for the possibility that search is inherently "social" in that an employee may choose intermediaries based on organizational, social, and demographic considerations; searchers for knowledge may not know *ex ante* who the experts are or where they are located; and we analyze the effects of a firm's network structure on actual unfolding search chains.

SEARCH CHAINS WITHIN LARGE, COMPLEX ORGANIZATIONS

We apply the logic of search chains to the context of searching for knowledge in large and distributed organizations. Specifically, we investigate how searchers try to reach subject matter experts in a large management consulting company, which is representative of multi-office professional services firms in general. As discussed in detail later, an individual employee's ability to search for information through interpersonal networks is particularly important in such firms (this may not be true in other kinds of settings where an adoption of IT-based knowledge management system might suffice to connect employees with expertise). For example, consider a company employee looking for an expert on a particular topic related to solving a client problem, such as setting transfer prices between subsidiaries in the client's firm. She would like to locate a colleague in her company who is an expert on transfer pricing, but she does not know *ex ante* who that may be. Her first decision in the search process is to decide whether she wants to guess who may be an expert and contact that person directly. If she is correct, she would have located an expert in only one step, as illustrated in chain A in Figure 1. Alternatively, she could decide to contact a colleague and ask him to point her in the direction of an expert. In this scenario, the quality of the advice given by the intermediary influences the length of the search chain. The intermediary may point to the expert right away (chain B in Figure 1) or to someone else who is not an expert but who in turn points to someone else (e.g., as in chain C), and so on, until the chain ends with the identification of an expert or terminates prematurely.

Understanding why some searchers are able to find an expert in only a few steps requires an analysis of both the factors explaining why searchers may be able to identify an expert in one step (chain A in Figure 1) and, failing that, the process by which searchers are able to use as few intermediaries as possible (chain B vs. C in Figure 1). In the first scenario, search is to a large extent cognitive or "asocial" (Gavetti and Levinthal 2000), in the sense the searcher is examining the information she already knows in trying to identify in her mind who the experts may be on a certain topic. In the second scenario, search is not only cognitive but also social in that it involves enlisting the help of individuals acting as intermediaries to point the searcher toward the expert. We examine the two scenarios separately.

Identifying an expert in the first step of a chain: the role of peripheral status

Why would some searchers be able to identify an expert in one step while others would not? As research on transactive memory has shown, employees develop an awareness of “who knows what” in an organization (Wegner 1987; Wegner et al. 1991, Moreland et al. 1996, Austin 2003, Borgatti and Cross 2003, Schulz 2003, Brandon and Hollingshead 2004). Applied to our context, the existence of transactive memories suggest that employees may know who the experts are on a certain topic in the organization, but this information is likely to be unevenly distributed among employees. In particular, employees on the periphery of an organization’s social structure may be at a relative disadvantage in locating experts because they may have poorer information about the identity of experts than do members of the core. As a long line of organization research has shown, some groups of organizational members tend to be at a disadvantage in terms of access to timely information circulating within organizations (e.g., Kanter 1977, Mintz and Schwartz 1981 and 1985, Brass 1985, Ibarra 1992, 1995).

Organizations often exhibit a core-periphery pattern, with a core group of people—a dominant coalition, an elite, or a majority—exerting the most influence or decision making authority (Brass 1984 and 1985, McPherson et al. 2001). While position in the formal hierarchy to some extent determines members of this elite group, the degree to which a member belongs to the core or exists at the periphery depends on other factors as well. Extant research has typically focused on an employee’s position in social networks, gender, ethnicity, and length of tenure as categories along which organizational periphery is defined (e.g., Kanter 1977, Brass 1985, Ibarra 1992). In this paper we concentrate on network centrality, tenure, and gender to define the degree to which an employee belongs to an organization’s periphery. We define members as having a periphery status to the extent that they have a low network

expert-related centrality, have short tenure, and whose gender is underrepresented. The periphery status of employees is to some extent empirical and depends on the distribution of employees within a particular organization.¹ For example, only 19.5% of professionals in our empirical context were women, implying that women are on the periphery in this context, but this may not be the case in other organizations.

Expert-related centrality. One way of characterizing the degree to which an employee belongs to the periphery in an organization is to assess their position in the organization's social network (Brass 1984). Centrally placed individuals belong to the core while non-central actors belong to the periphery (McPherson et al. 2001). Therefore, the first variable we consider is a searcher's centrality with respect to location of a group of experts in the organization network. To understand how a searcher is linked through the network to a set of relevant experts, we distinguish between being central to *everybody* in the organization and to a confined group of experts. For a particular search, it is more important for a searcher to be central vis-à-vis experts than all employees: a person can have a central position vis-à-vis the overall set of employees in an organization, yet have a low centrality vis-à-vis a particular group of experts. To implement this distinction, we draw upon the well-established concept of closeness centrality in networks (Freeman 1978) and define "expert-related centrality" as the length of the path through the network that has the lowest number of intermediaries between a searcher and a set of topic experts.² The variable *network distance to expert* is the converse of expert-related centrality: the higher the number of intermediaries between a searcher and a group of topic experts in the network, the higher the network distance to expert.

Information about identities of experts is likely to travel through these pathways, like pipes through which information travels, including information about "who knows what" (Podolny 2001).

¹ While it is beyond our scope here to explore this empirically, one can imagine the underlying mechanisms driving the effects of being on the periphery to be shaped by constructs such as power, which is known to be correlated with individuals' cognition of intra-firm networks (Krackhardt, 1990) as well as social behavior (Keltner et. al., 2003).

² This is the same as the smallest of the geodesics between a focal employee and any of the experts on a topic. The geodesic is the shortest path through the network, e.g., a direct tie is a geodesic of one, a path through one intermediary is a geodesic of two, and so on. To probe the distinction between centrality vis-à-vis everybody or a group of experts, we also implemented a network distance measure based on geodesics to all employees in the organization, but this measure was not significant in our models and had no predictive power.

If the focal employee and an expert have worked with each other before on a project and thus have a direct tie (a path of one step), it is likely that the employee will know that the expert is in fact an expert on the topic in question. If the focal employee has no direct relation to the expert but they have both worked previously on different projects with a third person (i.e., a path of two steps), information about who is an expert needs to flow via this third person. For example, in our qualitative interviews, a consultant in the San Francisco office, who was connected to an Information Technology expert in Chicago through one intermediary, told us how information flowed to him via this link: “I have spoken to him [intermediary] on several occasions, and he told me about some of the projects in IT that he had been working on with Mike [expert], so when I needed that specific expertise I had some idea who to call.” In this case, the consultant in San Francisco knew that Mike was an expert because he had picked up that information while working with the intermediary on something else. Thus he knew the identity of the target via the intermediary, information he could act on when he needed to find an expert on IT.

In other situations, however, the original searcher may not know the identity of the target. Let’s suppose the consultant in San Francisco had not picked up the name “Mike.” When he needed an IT expert, the consultant could have contacted the intermediary, who would then point to Mike. The short network distance between Mike and the San Francisco-based consultant (via one intermediary) makes such a search easier.

However, the diffusion of information between two individuals in any network can be expected to fall as the network distance between the two increases (Singh 2005). Specifically, in our setting, information about who is an expert on what subjects is likely to become distorted, biased, or entirely lost as more intermediaries are needed in order to pass it on. Employees may misunderstand each other when exchanging information, and intermediaries may neglect to mention relevant pieces of information, forget details, filter or deliberately withhold some information (Miller 1972, O’Reilly 1978, Huber 1982, Huber and Daft 1987). The consequence is that a focal employee’s knowledge about who is an expert on different subjects becomes imprecise, incomplete and perhaps altogether incorrect to the extent that many intermediaries have been involved in funneling information through the network. When it comes time to

conduct a search, a focal employee who has a long network distance to relevant experts is therefore less likely to be able to pinpoint the identity of an expert on a particular subject than one with short network distance. As one interviewee, who had a long path of two intermediaries to an expert, told us when we mentioned the expert's name after our experiment had been concluded: "I have never heard that name [expert], I have no idea about this person."

Tenure. The second periphery variable we consider is a searcher's tenure. Employees' company tenure (i.e., number of years employed in the focal company) is likely to affect their ability to identify an expert in one step. Long-tenured employees are likely to have a more central network position than short-tenured ones, as they have had more time to build network connections in the organization (cf. Chatman and O'Reilly 1994, Harrison and Carroll 1998). Beyond this network effect, however, tenure may also be associated with a larger repertoire of knowledge about the body of expertise embedded in the organization. As employees accumulate experience in an organization, they also accumulate knowledge about who knows what. In professional services firms, for example, experience is related to project work on a range of topics. As employees work on more projects as time passes, they accumulate more project experience and with that more knowledge about project topics. Assuming that they do not work on the same project topics every time, this experience should translate into more knowledge about project topics, including knowledge about who knows the most about these topics. Because of this experience effect associated with tenure, short-tenured employees are less likely to identify an expert in the first step of a search than are long-tenured individuals.

Gender. A searcher's gender is also likely to affect the chances of identifying an expert in the first step of a search chain. For our analysis, we consider the situation in which men are the majority of the employees and the upper echelon of managers. In these contexts, women are likely to experience a worse information flow than men, including information about the identities of experts in an organization. This may be partly due to a worse position than men in the work-flow network (Brass 1985, Ibarra 1992, Podolny and Baron 1997). Beyond this task-related network effect, however, women may also be excluded from social circles in an organization, including social activities at work and after-job activities

that strengthen interpersonal ties and increase the rate of communication (Kanter 1977). Moreover, to the extent that men occupy the key positions in the organization and constitute a dominant coalition that also includes knowledge experts, women may find it difficult to develop interactions with these organizational elite and thus be cut off from the information flow emanating from it (Brass 1984 and 1985).

Women may also sort themselves into, or be sorted into, different types of work (Ibarra 1997). In our context, women may prefer to stay “local” in their geographically local office for family reasons, increasing the chances that they work on the same types of projects with the same colleagues over again. As one female consultant in a U.S. West Coast office related: “The partner really wanted me to go to new York and head the project over there, but I had a one-year old a home and didn’t want to, so I said no...I got myself on to a project working for a local client that I had worked with before.” Working locally with a small group of clients and colleagues in turn creates a narrow accumulated knowledge base, leading to less knowledge than men’s about the wider distribution of expertise in the organization. For these reasons, women may have less cognitive awareness than men about who knows what in the organization, making it more difficult to pinpoint a knowledge expert in the first step of chain.

In short, our arguments can be summarized in a hypothesis:³

***Hypothesis 1.** Searchers who have low expert-related centrality, have short tenure, or are in the gender minority are less likely to pinpoint an expert in the first step of a search chain.*

Impediment to effective search through intermediaries: the role of homophily

Many times employees may not know who the expert on a topic is. They can then ask someone to help them identify an expert. Employees may undertake “social” search by contacting intermediaries and asking them to point them in the direction of experts (Gavetti and Levinthal 2000). In this situation, the challenge is not to know the identity of experts but to identify an intermediary who could be in a position to help point to an expert—a “connector” or a “bridge” (Gladwell, 2002). As specified in the prior

³Our arguments suggest a cognitive rationale for why some employees may know the identity of experts while others do not. But even if employees know the identity of experts, some may be reluctant to contact them, perhaps because they feel intimidated approaching someone who is likely a senior partner. As discussed later, whether they feel comfortable approaching experts is, however, a separate issue from whether they know their identities.

section, employees who have a high expert-related centrality, long tenure, and are in the gender majority should be in the best position to act as an intermediary, because they have a high degree of awareness about who knows what in the organization. Thus, a rational search strategy would be to approach an intermediary who belongs to the organizational core as defined by these characteristics.

As research on stratification in organizations has shown, however, employees may instead prefer to socialize and connect with people who are like themselves along demographic characteristics such as race and gender (Lazarsfeld and Merton 1954, Ibarra 1992, McPherson et al. 2001, Ruef et al. 2003). Applying this logic to our setting, searchers who choose to contact an intermediary in search are likely to contact someone who are like themselves: Searchers on the organizational periphery are likely to contact an intermediary who also is also on the periphery, while searchers who are a member of the organizational core are likely to contact another member of the core.

One reason for this homophilous search behavior is familiarity: two colleagues who share one or more traits (e.g., gender, tenure, race, nationality) may prefer to interact with each other because of ease of communication based on common attitudes and world-views. Another reason is perceived safety or trust: asking someone for help in finding experts reveals a lack of one's knowledge and exposes oneself to the risk of an unfavorable judgment (Edmondson 1999). It is safer to contact an intermediary with whom one is more familiar based on similarities than someone who is different but could potentially be in a better position to act as an intermediary (Hansen and Lovas 2004, Casciaro and Lobo 2008). One junior consultant in Melbourne revealed in a follow-up interview: "I suppose the reason why I thought of Matt [a fairly junior person picked as the intermediary] is that Dave and John are all vice presidents...I had personally worked with Matt before so I probably just felt just a little more comfortable."

Thus, rather than approaching an intermediary belonging to the core of the organization, a female searcher will likely contact another woman, a junior consultant will likely reach out to a fairly junior

colleague, and a consultant with low network centrality will likely contact another with fairly low network centrality.⁴ Following these arguments leads us to our second formal hypothesis:

Hypothesis 2. *Searchers are more likely to select intermediaries with whom they share a characteristic (expert-related centrality, tenure, gender) than intermediaries with whom they do not.*

Homophilous search is naturally constrained, however, by the availability of similar others in an organization (McPherson et al. 2001). Women, for example, have fewer chances than men to develop relations with same-sex colleagues in organizations dominated by men (Ibarra 1992). McPherson et al. (2001) calls this a baseline (or induced) homophily that is determined by the opportunity set. Thus, it is quite possible that an employee on the organizational periphery will contact members of the core simply because there are few other peripheral members to select as intermediaries. However, to the extent that homophilous search operates (called “inbreeding” or choice homophily, McPherson et al. 2001), we should expect to see members of a sub-group select intermediaries who are also members of a sub-group in a greater proportion than what would be expected based on the baseline. For example, in our empirical context, 19.5 % of the consultants were women: inbreeding homophily suggests that female searchers will select female intermediaries in more than 19.5% of the cases (in fact, our empirical results revealed that female searchers chose female intermediaries in 38.5% of the cases).

Is homophily-based search disadvantageous for members on the organizational periphery? This hinges on whether intermediaries who are part of the periphery perform worse in their role as an intermediary than those who are part of the core. We argue that they will. The reasons for this are the same as those for Hypothesis 1. An intermediary on the organizational periphery is likely to suffer from the same exclusion issues as an original searcher who is part of the periphery. Intermediaries with low expert-related centrality are less likely than intermediaries with a high degree of centrality to identify an

⁴ While a searcher can easily identify the tenure and gender of potential intermediaries, detecting whether someone has high or low network centrality is more difficult. In our empirical context, consultants had some hunches as to who were in the “in-crowd” as defined by network relations. For example, well-networked junior consultants would go to lunches with partners or work with them in their offices; those were visible cues that created boundaries between central and peripheral actors in the network.

expert on a topic. An intermediary with short tenure is less likely to have accumulated enough experience to identify an expert than one with a long tenure. And a female intermediary is more likely to suffer from social exclusion and lack of integration into the dominant coalition than is a male one, leading to a lower chance of pinpointing an expert.

Furthermore, if an intermediary chooses not to try to pinpoint an expert but instead points to a second intermediary, the periphery effect is likely to manifest itself in this scenario too. Like the original searchers, intermediaries are also likely to tend toward homophily, pointing to a second intermediary that is like themselves along the centrality, tenure and gender dimensions. The unfolding search chain will therefore reproduce social structure, like ripples in a pond, leading to a differential effect for chains that begin on the periphery versus in the core.

We have argued that members on the organizational periphery are likely to have longer search chains because they have low cognitive awareness of who the experts are (and thus are likely to experience difficulties in naming an expert in the first step of the search) and base their searches on homophily (which is a poor search strategy for peripheral employees). If this is correct, one issue becomes, how can peripheral members improve their search for experts in an organization? A simple answer is to try to “cross over:” employees with short tenure can contact intermediaries with longer tenure; female searchers can contact male intermediaries, and employees with low centrality can contact centrally placed intermediaries. Such crossing over strategies are likely to be more beneficial to the extent that they occur earlier in the chain as they will redirect the subsequent steps. Crossing over should help mitigate the negative effects of homophily-based searches for members on the organizational periphery.

These arguments lead us to two related hypotheses predicting the number of search steps in a search chain required to reach an expert:

Hypothesis 3a. *Searchers (original and intermediaries) who have low expert-related centrality, short tenure, or are in a gender minority tend to have longer realized search paths.*

Hypothesis 3b. For searchers who belong to the organizational periphery, those who name members belonging to the core as the next search step tend to have shorter realized search paths than those who don't.

DATA AND METHODS

Our research site is a global management consulting firm with 50 offices in major cities worldwide, covering 34 countries around the world. At the time of our study, the firm had 2,800 consulting staff and 350 partners, for a total of 3,150 “line” consultants. After we had negotiated with a senior partner to conduct this study and received the go-ahead from the CEO, we spent six months designing and implementing a field-based small world study. We conducted several preliminary discussions with the consulting staff and spent several months in the company as a participant observer before developing and implementing surveys and extracting information from the firm’s databases. We also conducted nine follow-up interviews with consultants who participated in the surveys.

Setting

The 50 local offices were the primary organizing unit in the firm, and each office was staffed with a set of partners, senior managers, and other consulting staff. Each office had over a period of time accumulated consulting expertise that reflected the type of work that had been done in that office. What had emerged over the years was therefore a mosaic of knowledge dispersed across offices and consultants, with some individuals becoming experts on particular topics because of the issues they had dealt with during their client work. The topic experts were to a large extent not officially appointed experts with any formal responsibility or title, but were regular line consultants who emerged as informal and unofficial experts. Because of the dispersion and unofficial position of experts and the continued development of new knowledge due to ongoing projects, locating individuals with particular and up-to-date expertise was a major problem. To organize some of this vast knowledge pool, the firm had over the

past decade developed 11 “practice groups,” each of which was responsible for organizing the knowledge pertaining to either an industry (e.g., financial services) or a topic (e.g., information technology). The firm had also over the past five years implemented an electronic knowledge management (KM) system that stored prior sanitized client presentations and discussion documents on particular topics.

While the offices were the primary organizing unit, project-based teams were the primary work unit. Consultants joined these temporary teams, which normally lasted from two to six months, to work on a particular engagement. At the end, when the engagement was finished and the team was dissolved, they would leave to join another team for a new project. When a partner had sold a project to a client, the team was formed. Team composition was determined on a project-by-project basis: factors such as expertise, experience, availability, and geographic location all played a role in determining the composition of a given team, which typically included 4-7 consulting staff.

At the outset of a project, team members spent time getting up to speed on the particular topics covered by the project and often used the first couple of weeks to search the firm for experts on relevant topics. They typically downloaded relevant documents from the electronic KM system and also contacted the official practice group coordinators relevant to the project at hand. More than these formal sources of information, however, consultants relied heavily on informal, personal contacts with other consultants, frequently asking for advice on relevant topics, industries, and companies. As a consultant told us during preliminary interviews, “using KM is a first good start but you really get to the experts by asking around.”

Informal inter-personal relations were important in searching for knowledge and were to a large extent a by-product of joint work on past projects. While consultants developed informal relations with one another because they worked in the same office, came from the same university, or were part of the same incoming “class” by starting working at firm at the same time, working on the same projects was a main determinant of the formation of informal work-related relationships. Project work was typically very intense, with each consultant working upward of 80 hours per week on a project and often traveling with other team members to visit clients, becoming well acquainted with one another during the project. Thus, team membership was a major determinant of work-based interactions in the consulting practice.

Small-world study design

Using Travers and Milgram's (1969) original small-world study as a departure point, we designed a small-world study that had four components: (i) selection of four knowledge areas or topics; (ii) identification of topic experts for each of the four topic areas; (iii) selection of a random set of original searchers (i.e., consultants who started a search chain); and (iv) a chain-based survey methodology based on a "snowball" method.

(i) Selection of four knowledge topics. To limit the study, we first decided to narrow the scope of the possible consulting topics employees could search for. Because we wanted to have a relatively high number of original searchers per topic area, we chose to limit our analysis to four topic areas. The tradeoff in deciding the number of topics was that, while choosing only one topic area might not be representative, choosing too many would have reduced the number of survey respondents per topic. To select topic areas, we used the taxonomy of topics that had been developed by the firm's knowledge managers to categorize electronic documents on the KM system. This taxonomy was the firm's most comprehensive effort to categorize the topics covered by client projects. The list comprised five hierarchical levels of topics, ranging from the most general to very specific topics, such as from "marketing and sales" (the most general level), "marketing strategy," "advertising," "media planning," to "website measurement research" (the most specific level). We excluded the two most general levels, because they did not provide sufficient specificity for consultants to even initiate a search (e.g., it was too general to ask someone whom they would contact for knowledge on "marketing and sales"). Field interviews and preliminary checks of the data also revealed that the two most specific levels often did not contain any experts and had few associated electronic documents, making these less suitable for our purposes. Thus, as a practical matter, we focused on the middle, or third level, in the hierarchy (e.g., "advertising" in the example above). This level contained 227 topics.

Field interviews indicated that topics varied significantly in terms of the overall volume of knowledge and concentration of experts on a topic. While we remain agnostic as to how these two parameters affect search, we nevertheless want to incorporate these dimensions by selecting topic areas

that vary along these dimensions. To achieve this, we computed volume and concentration measures using data on electronic documents in the KM system. We measured volume of knowledge as the total number of electronic documents stored in the KM system on a particular topic. To compute concentration, we derived a Herfindahl index of the degree to which the electronic documents on a topic were authored by few individuals. Formally, this was calculated as $H = \sum_i p_i^2$, where p_i is the fraction of consultant i 's documents among all documents on a topic. This measure ranges from $1/n$ to 1, where n is the total number of authors on a topic. If one author had written all the documents on a topic, this measure is 1.

Figure 2, which depicts the volume of electronic documents and the degree of concentration of authors for the 227 topics, reveals that there is a fairly high correlation between the two measures ($r = -0.45$). As volume increased, knowledge in the form of electronic documents typically became increasingly dispersed among a greater number of individual authors in this company. Allocating the topics into four quadrants based on the mean-level of the two dimensions, we selected four topics: Enterprise Resource Planning Systems or ERP (i.e., consulting on how companies should develop information systems and integrate them with their strategy); Asset Productivity (i.e., review of and recommendations for improving returns on assets in a company); Transfer Pricing (i.e., issues around internal transfer prices in large multiunit companies); and Advertising (i.e., review of and recommendations for a company's advertising strategy and campaigns).

As summarized in Table 1 and graphically depicted in Figure 2, these four topics represent various combinations of volume and concentration. While ERP has a fairly high volume but is dispersed (i.e., upper-left quadrant in Figure 2), Asset Productivity has a somewhat low volume and is also fairly dispersed (i.e., lower-left quadrant). Transfer Pricing has a low volume but is fairly concentrated (i.e., lower-right quadrant). Advertising is the topic closest to the upper-right quadrant in Figure 2, by having a high volume and not being too dispersed. These four topics also meet an additional criterion of being easy to understand: pretests with several consultants showed that they immediately knew what the topics were about (although they did not necessarily know who the experts were).

(ii) Identifying topic experts prior to starting search chains. Because the firm's experts on the four topics did not occupy any formal expert roles and did not have any official title that acknowledged their expertise, they were not immediately visible. Searchers could therefore not just simply look them up in electronic "yellow pages" but had to ask for help from others in identifying the experts in case they did not themselves know the exact identity of experts. The essence of the search process was therefore to identify experts and not to move messages toward a pre-identified target, as is common in small-world studies. This setting thus satisfied our argument that a model of search chains in an organization should not be premised on searchers knowing the targets *ex ante*. But it posed a challenge for us as well, since we also did not have access to any existing source of information that would list the identities of experts on different topics. But we needed to know who the experts were so that we could determine whether a chain reached one. To identify the experts, we therefore relied on a systematic nomination process involving several iterations, the idea being that we would generate a list of who the experts were but we would not reveal this to the searchers.

Importantly, to ensure that the process of finding experts would be independent from the process of searching for them, we relied on different data sources and surveys. First, as a starting point, we identified an initial batch of "suspected experts" on the four topics by analyzing the project and KM databases for the previous five years. We identified consultants with the greatest project experience (by using data from the project database, which listed the number of times a consultant had worked on a project topic), the largest number of authored KM publications, and having the most "hits" in the KM database in terms of downloads of their publications (which help measure how influential their documents were on that topic). Specifically, an individual would have to meet one of the following two criteria to make our initial list of suspected experts: (a) be among the top 20% who had worked on the topic in terms of total number of projects on that topic; or (b) be among the top 20% individuals based on an individual's total number of authored documents *and* be among top 20% in terms of total number of downloads attributed to those documents. The first rule states that those consultants with the greatest relevant project experience "under their belt" have greater likelihood of being experts on a topic. The

second rule states that individuals who have authored most KM documents that are also widely read have greater likelihood of being experts on a topic. Using these rules, the initial list of “suspected” experts ranged from 14 to 29 per topic, for a total of 79 (see Table 2a).

Next, we sent each of these “suspected experts” a survey that asked them to rate their own expertise in this area and answer the following question (using the Advertising topic as an example): “Whom in [company name] would you identify as an expert on Advertising (name up to 5 persons)?” They could not nominate themselves. This step generated a list of consultants nominated as experts by the initial batch of suspected experts. In total, 78 individuals received at least one nomination in this round. We then sent out a second round of the same survey to all 50 nominees who had not already been surveyed, yielding 34 new nominees. Finally, we completed a third round of surveying, yielding only 9 new nominees across the four topics. The generation of new nominees had thus dwindled to between 1 and 3 per topic after the third survey, indicating that the total of 144 consultants who were surveyed converged in their assessment of who the experts were on these four topics.

Based on this information, we used the cut-off of four or more nominations to determine a real expert in the firm, leading to between 2 and 8 experts per topic, for a total of 26 experts (see the last column in Table 2a). Hereafter, we refer to these individuals as the experts in the firm. We conducted several follow-up checks to verify that using a cut-off of 4 nominations was indeed a right way of identifying experts. Many of the other nominees had received just 1 or 2 nominations, and could easily be dismissed as not being experts. The only doubt was about the handful that received 3 nominations, but on closer inspection none of them seemed worth considering either: they had been nominated mostly based on just the number of projects done related to the topics, with any demonstrated expertise beyond that being substantially lower than those with four or more nominations. Thus the picture was rather clear as to who should be judged as experts. Nevertheless, as yet another check, we sought reactions to our expert list from three “elders” in the company, long-time partners who sat at the center of the firm's knowledge flows by virtue of their roles (e.g., one was head of the company's “think tank”.) While their views could also be biased or incomplete, they could detect whether there was something completely wrong with our

selection. Each one reviewed the short-list and they only highlighted one person (in Advertising) who seemed suspect to them (upon further investigation we kept this person). As a final check, we repeated a similar verification process with leaders of each practice group pertaining to the four topics, and their reactions further validated that our list of experts was indeed correct.

One might still worry that we might have missed some real experts that were completely hidden in the organization. We think this highly unlikely because we started the nomination procedure with a very wide net, by surveying consultants who had worked and authored on these topics. Starting with this wide search makes it unlikely that a real expert on a topic had gone completely unnoticed in the firm.

Table 2b lists the demographics and office location of the experts. It is worth noting that the fraction of male experts is 92.3%, while the fraction of men among all employees is 80.5%. Similarly, the average tenure of experts is 9.7 years, while the average employee tenure in the company is only 3.8 years. Both of these suggest that members on the periphery have a lower probability of being experts than one would expect if experts were just randomly distributed among all employees. The experts were clearly part of the “core” of the organization.

(iii) Selecting and surveying original searchers. We randomly selected 96 individuals for each topic to start a search chain (hereafter called *original searchers*), yielding 384 original searchers or 381 usable ones as three individuals on the list had left the company. Thus, we were able to start 381 chains. Of the 381 submitted surveys, 241 (63%) responded. The original searchers were each sent an email survey from the office of the senior partner sponsoring the study (see Appendix 1 for a list of questions). We asked the original searcher to name *one* individual whom the searcher would contact as a topic expert if he or she were to do a client case on the topic (hereafter called a “contact”).⁵ If the searcher did not

⁵Our question did not ask the respondent to distinguish between who they would name as an expert (awareness) and who they would feel most comfortable contacting (familiarity and trust). This distinction was not especially important in our empirical context, as the role of familiarity was less important in approaching experts: there was a strong norm that it is acceptable to contact topic experts for their expertise and those experts were supposed to respond, making it less intimidating to approach them. In contrast, contacting someone to ask who the expert may be (the intermediary function) meant interrupting a colleague who is not an expert and asking for advice on finding an expert, something that could be intimidating for some consultants on the periphery, such as junior consultants. In this instance, going to colleagues similar to oneself was deemed safer.

know whom to contact as a topic expert, we also asked who he or she would contact as someone who could help identify an expert (hereafter called an “intermediary”). Intermediaries thus differ from contacts in that they are not considered topic experts but act as intermediaries who may point the original searchers to a topic expert. Of the 241 responses, 40 identified an expert right away and 139 provided a name of a contact that allowed us to follow the chain by sending e-mails to the named person. The remaining 62 respondents who returned the survey to us did not complete it by filling out a name and we could therefore not follow through with those chains. (We control for non-responses in our statistical analyses by running multinomial regression models where a non-response is one of the outcomes.)

(iv) Chain-based survey. Of the original searchers who responded to the survey, 40 identified one of our pre-identified experts, and the chain thus stopped, as the original searcher had already reached an expert. We conducted a follow-up survey for the other chains where a contact name was provided. In a second round of surveys, we submitted the same survey as in the first round to everyone who was named as an intermediary in a chain. This led to 114 individuals being surveyed in the second round. We repeated these steps in a third round (surveying an additional 31 people) and a fourth round (surveying an additional 3 people), at which point all chains had either reached an expert or could not make further progress due to non-responses or incomplete surveys (i.e., no name was given in the survey). In all, 529 surveys were conducted in four rounds. Table 3a summarizes the results for each survey round.⁶ In total, we received 356 responses, giving an overall survey response rate of 67.3%. There were no significant differences between respondents and non-respondents in terms of tenure (with respective means of 5.4 and 5.3 years, $t = -0.0741$), age (36.4 and 39.3 average years respectively, $t = 1.56$), proportion women (0.14 and 0.20 respectively, $t = 1.62$), and proportion partner (0.20 and 0.14, $t = 1.22$).

Sample construction

⁶ Since each survey pertained to search for an expert in a specific topic area, individuals named as contacts or intermediaries in multiple topic areas were asked to fill in one survey per topic area. Therefore, the 529 surveys involved 511 unique individuals (including original searchers, intermediaries and contacts).

We constructed a data set in which each surveyed individual for each chain was included as a separate observation, starting with the original searcher, then the first named contact or intermediary, and so on, up to and including the last person surveyed for that chain (including non-respondents as observations). This approach is akin to the methodology of “spell splitting” whereby a chain is broken into its constituent parts of all individuals belonging to that chain (Tuma and Hannan 1984). The resulting data set is summarized in Table 3b.

Because searchers from different chains (but on the same topic) sometimes named the same intermediaries, these intermediaries could appear in multiple chains: 23 individuals were named as intermediary by two searchers, 3 individuals were named thrice, 3 individuals were named four times, 1 individual was named five times, another individual was named six times, and a final one was named seven times. To construct complete chains, we therefore transformed the original data on 529 actual surveys (summarized in Table 3a) into a sample of 582 observations (summarized in Table 3b) used for further analysis.⁷ In subsequent regression analysis where observations from all chain steps are included, we assign different observations different weights to ensure that duplicate counting of a survey entered as multiple observations does not bias the results.⁸

Dependent and independent variables

In our initial analysis, we use a dependent variable, *chain length*, defined as the number of steps in the search chain starting from an originating individual to an expert. Based on this variable, we created another variable, *remaining chain length*, defined as the number of remaining steps in the search chain starting from *any* individual (originator or intermediary) to an expert. We also use another dependent variable, *reached expert*, which is an indicator for whether or not an original searcher was able to correctly name an expert in the first step of the chain.

⁷ One might reasonably ask why we did not exclude multiple observations for the same individual. The reason is that one of the central control variables is the individual’s position in the sequence of steps in a search chain: individuals who were named more than once occurred at different steps in the chains.

⁸ Specifically, this means weighting each observation so that the sum of the weights for a survey observation (i.e., a unique individual-topic combination) always sums to 1. For example, if a survey of an individual appears 4 times as different observations, the weight assigned to each observation for that survey would be 0.25.

We employ a binary variable, *woman*, to capture whether an original searcher or an intermediary was a woman (1) or a man (0). We use a variable, *tenure*, to denote the respondent's tenure, measured as the number of years that the individual has been employed in the company.

Creating our network measure, *network distance to expert* (i.e., the converse of expert-related centrality), was a more elaborate procedure. Importantly, we needed to construct a network measure that relied on network data measured *prior* to a particular search attempt; otherwise we would be confounding cause and effect relationships. We therefore relied on two entirely different data collections for the search chain measures and for the network distance (centrality) measure, for which we used historical network data that existed prior to the individual search attempts in our study.

Specifically, we used the affiliation network generated by the consultants' project work history to create this measure. As our preliminary field interviews revealed, task-mandated interactions for consultants arise most directly from project assignments; consultants worked in teams to serve clients, and accordingly a consultant's history of team membership is likely to be an important determinant of the relations that the consultant developed over time. As a general principle, a consultant (ego) who has worked with another consultant (alter) on a project is also indirectly connected to others with whom the alter (but not ego) has worked, and so on. Any two individuals in the work-flow network could therefore be connected through paths of various lengths in the work-flow network. By implication, a consultant also has established work-flow network paths to the experts. The longer the ego's path length to the experts *prior* to any given search effort, the more intermediaries lie between the experts and ego in the work-flow network. We define *network distance to expert* for individual i in topic area T as $D(i, T) = \min_{j \in E(T)} d(i, j)$, where $E(T)$ is the set of all experts in topic area T , and $d(i, j)$ is the length of the minimum path (geodesic) between individual i and individual j in the work flow network.

To construct the network and compute path lengths to experts, we extracted information from the company's time and billing databases, which comprised information on all projects in the past several years and the names of consultants who had worked on each project. This approach allowed us not only to

construct a complete network but also to use information about the network that existed prior to our search experiment. We used information from three years prior to our study, and created an affiliation measure in which an affiliation tie between two consultants was recorded if they both had spent at least 40 hours working on at least one common project.⁹ We used 40 hours as a cut-off to exclude so-called “sales leads,” which were non-project assignments on which consultants worked for a few hours; these assignments involved working on a small assignment (such as collecting data on a potentially new client) and did not involve in-depth teamwork from which relations could be formed. Our exclusion is consistent with the sharp distinction drawn in the firm between real project work and sales lead work.

We counted 95,578 symmetric ties among 4,533 consultants who had worked for the firm during this time, giving a density of 0.009 for this network. This network exhibits the small-world properties of high clustering and short average path lengths.¹⁰ It is also quite decentralized.¹¹ Based on this network, we constructed the variable, *network distance to expert*, as the shortest distance from a given individual to the closest expert on a topic (Wasserman and Faust 1994).¹² This variable can be seen as the converse of closeness centrality (high network distance means low closeness centrality).

Finally, we created three dependent variables to study the homophily effect. The indicator variable, *next is low network distance to expert*, takes on a value of 1 if the next person named in the chain has a below-median network distance to the experts, and 0 otherwise. The variables, *next is high*

⁹ Consulting projects tended to be relatively short, so the resulting network ties were not as long-lasting as in many other settings (e.g. collaboration between inventors or academics). In any case, consultants’ expertise developed significantly year by year, and the expertise from more than three years ago was probably also less relevant and possibly obsolete. On the other hand, these relationships were not short-lived either: we asked several consultants about how long they thought project relations “stuck” and found the consensus to be around 3 years.

¹⁰ Kogut and Walker (2001) compared the two measures of small-world properties (clustering and path length) for several networks. Using their method, our network appears to exhibit small-world properties. See Appendix II for detailed calculations for our network and a comparison with small worlds networks reported in previous studies.

¹¹ The distribution of degrees in this network is not very wide: mean=42.0; median=36.0; maximum=219; standard deviation=30.6; and skew=1.24. Formally, Freeman’s group degree centralization measure (Wasserman and Faust, 1994, p. 180) is 0.039, which is far closer to the lower theoretical extreme of 0 (for a network where all degrees are equal) than for the upper extreme of 1 (for a star graph).

¹² This measure thus captures the *minimum* network distance between the focal employee and one of the experts on the topic in question. We also specified an *average* distance measure between the focal employee and all experts on the topic, and our results remained the same for both measures.

tenure, and *next is male*, are indicator variables defined as 1 when the next person named in the chain is at above-median tenure or a man, respectively.

Control Variables

We use a control variable, *self-reported expertise*, to allow for the possibility that individuals who consider themselves to have high expertise on the topic might be more successful in reaching the experts (see Appendix 1 for the question asked to solicit this information). To account for the possibility that being in the same location as an expert may facilitate search, we also include a variable, *expert in same office*, defined as an indicator variable for whether or not an employee and an expert are in the same geographic office.¹³ As it could potentially be harder to look for experts in one or more of the four topic areas, we also included indicator variables for three of the topic areas - *advertisement*, *asset productivity* and *enterprise resource systems*, with the omitted category being *transfer pricing*. To account for the fact that search chains might progressively get closer to experts as they unfold, analysis using a sample from multiple steps also includes a control variable, *step*, which measures the sequential position that a respondent occupies in the chain (i.e., has a value of 1 if the respondent is the original searcher, 2 if the respondent is the first intermediary, and so on).

Models

In the analysis of *chain length* and *remaining chain length*, we have a choice between using completed chains only (where the final length is known) and also including non-completed chains. We chose the latter approach in order not to bias the results toward completed chains. Non-completion of chains might hold useful information, as surveys returned with the respondent being unable to suggest the next person may reflect a failure in completing the search.¹⁴ To include such observations, our empirical

¹³ In analyses not reported in the paper, we also tried a control based on the physical distance between the office of an employee and that of the nearest expert, but the findings did not change.

¹⁴ The same cannot be said of the chains that were not completed because the survey was not returned at all, as the respondent might have named the next person if he or she had returned the survey. So we do not include such chains in the sample. In checks not reported here, we ensured that our main findings are robust to including these chains as “failures” in the sample as well as to using a two-step procedure where the first step is a Heckman correction for failure to respond to the survey. All findings are also robust to only focusing on successful chains.

models for *chain length* and *remaining chain length* use an ordered logit regression model where such failures are coded as a chain length of “99” (i.e., longer than any successful chain: the exact magnitude does not matter as long as it is large enough, since an ordered logit only uses the ordinal ranking of the outcomes).

In our analysis of the dependent variable, *reached expert*, we again include not just observations with a next name given but also those where no such names exists, hence avoiding potential response biases. Specifically, we estimate a multinomial logit model with four possible outcomes: (i) the individual responds to the survey and named a contact or intermediary who was not an expert; (ii) the individual responds to the survey and named an expert; (iii) the individual returns it without including a name for the next contact; (iv) or the individual fails to return the survey at all. This approach controls for the potential selection bias of including completed search chains only (Kleinfeld 2002, Dodds et al. 2003). A similar multinomial logit model is employed when considering the homophily-related dependent variables *next is low network distance to expert*, *next is high tenure* and *next is male* in analyzing a sample that includes not just observations with a name given but also observations with no next person identified.¹⁵

RESULTS

Table 3b shows the distribution of chain lengths for completed and non-completed chains. Among the 107 completed chains, the average path length (i.e., number of steps) to reach an expert was 1.89, with a median length of 2.¹⁶ However, these numbers do not account for incomplete chains and can therefore be somewhat misleading. Specifically, since failure to respond to a survey can happen at each

¹⁵ In analysis not reported in the paper, we also employed just logit models based only on observations where a name was actually given. The main results remained very similar.

¹⁶ It may seem like a small number that only 107 (28%) of 381 started chains were completed, but this is consistent with prior research (Kleinfeld, 2002). This happens because a chain risks being terminated in any survey round. Even though we had high survey return rates in every round (63% to 100%), chains' likelihood of completion is a function of the products of survey completion rates.

step in a chain, longer chains are more susceptible to not ending up in the observed sample of completed chains since there are more opportunities (steps) at which these can get dropped due to non-response. Following an adjustment procedure similar to that employed by Dodds et al. (2003), we find a corrected estimate for chain length one would expect if all surveys were returned to be 2.37.¹⁷

Descriptive statistics for our main variables are shown in Table 4. Before we test our formal hypotheses, we show the results from a baseline analysis predicting the chain length for a sample of 241 chains that were initiated as a subset of the 381 original searchers returned our initial survey. Consistent with our baseline expectations, the results in Table 5 show that searchers who had long network distance to expert, short tenure, or were women had *longer* search chains. The result is robust to including these variables individually (columns 2, 3 and 4 respectively) as well as to including these three variables together in a model (column 5). Thus, the gender and tenure effects are not simply due to the possibility that the searcher had greater expert-related centrality in the work-flow network.

To test Hypothesis 1, we analyzed whether an original searcher would identify an expert in the first step. The multinomial logit approach models four competing outcomes: naming an expert, naming an intermediary, returning a survey without any name, and not returning the survey at all. The results in Table 6 reveal that long network distance to expert, short tenure, or being a woman leads to a lower likelihood of finding an expert in the first step of a search. These findings lend support to Hypothesis 1.

Turning to Hypothesis 2, we next test the tendency to engage in homophilous search. As shown in Table 7, the dependent variable used in columns 1 and 2 is a binary variable for whether the next person in a chain has a low (i.e., below median) network distance to expert. The results of this analysis reveal that a searcher's network distance to expert negatively and significantly affects whether network distance to expert for the next person in the chain would be low. In other words, searchers with low expert-related centrality are more likely to name a next person who also has a low centrality. Interestingly, the tenure

¹⁷ This calculation is sensitive to the assumption one makes regarding probability of reaching an expert at any given step for the chains that encountered the non-response issue. Specifically, our calculation assumes that these chains would have the same probability of reaching the target in a given step as the completed chains we actually observe. If, in fact, their probability of reaching an expert at any given step were systematically lower, the corrected value for average chain length would be even larger.

and woman variables are also significant in the results as displayed in column (2) in Table 7. The lower the tenure of the searcher, the *lower* the likelihood of naming someone with a low network distance to expert. Likewise, if the searcher is a woman, the next person she picks is less likely to have a low network distance to expert than if the searcher had been a man.

Columns (3) and (4) in Table 7 use tenure of the next person in the chain as the dependent variable: the estimates in column (4) support the homophily argument that a searcher's tenure positively predicts the likelihood of naming someone with high (i.e., above median) tenure as the next person.

Finally, columns (5) and (6) in Table 7 use gender of the next person in the chain as the dependent variable. The results support the homophily argument that a female respondent is significantly less likely than a male respondent to name a man as the next person. A cross-tabulation of descriptive statistics also revealed a tendency for gender homophily: Male searchers picked another male in 89% of the cases (with the remaining 11% picking a woman). In contrast, 72% of female searchers picked another woman as an intermediary (and 28% picked a man). That is well above the "baseline" in this organization where 19.5% of the consultants were women. Overall, these findings support Hypothesis 2.

We now turn to testing Hypothesis 3. The results reported in Table 8 predict *remaining chain length*, which is measured as the length of the chain from any position in the chain to the completion of the chain. This analysis extends the sample from Table 5, where only the 241 chains initiated by original searchers returning the survey were included, to 440 observations corresponding to all observations pertaining to these chains. As in Table 5, an ordered logit allowed us to include the non-completed chains in the analysis, while clustering of observations from different steps of the same chain takes their non-independence into account in computing standard errors. The results in column (1) of Table 8 reveal that long network distance, short tenure, and being a woman all lead to a longer remaining chain length from any step of a chain. These results lend support to Hypothesis 3a.

To test Hypothesis 3b, we examined whether peripheral members that "crossed over" had shorter chain length than those who did not. Columns (2) through (5) of Table 8 report findings on such crossing over for the three dimensions. Column (2) starts with an independent variable, *decrease in network*

distance to expert, that measures the network distance to expert score for the searcher *less* the network distance score for the next person she named. The idea is to test whether a searcher who picked someone with lower network distance than herself had shorter chain length. As the results reveal, this variable is indeed negative and significant, lending support to an argument that searchers with a high network distance who “cross over” to intermediaries with lower network distance have shorter search chains.

In column (3) of Table 8, we similarly analyze the effect for tenure by creating a variable, *increase in tenure*, that measures the next person’s tenure *less* the searcher’s tenure. The idea is test whether short-tenured consultants who “crossed over” to longer-tenured intermediaries have shorter search chains, a conjecture we again find to hold in line with Hypothesis 3b.

Finally, column (4) in Table 8 examines whether women who named men as intermediaries had shorter chains than women who did not. The negative and significant coefficient is again consistent with the argument that women who “cross over” to naming a man as an intermediary have shorter search chains than women who do not. Column (5) shows that crossing over continues to be valuable even when all three dimensions are considered simultaneously in a single regression model. Overall, all the findings therefore provide strong support for Hypothesis 3b: crossing over leads to shortening of chain lengths.

DISCUSSION

The main finding in this paper is that members of the periphery in an organization—those with low expert-related network centrality, short tenure, and women in our setting—experienced a form of information disadvantage by having longer search chains to reach experts than did members belonging to the core. Their world of search was relatively large. Two theoretical mechanisms explain why this was the case. First, members on the periphery had lower cognitive awareness about “who knows what” in an organization: they fared worse than members of the core in pinpointing an expert in the first step of a

search chain. Second, members on the periphery also selected intermediaries on the basis of homophily (i.e., intermediaries who were also on the periphery), worsening their search process. However, those members of the periphery that “crossed over” and selected intermediaries from the core shortened their search chain as compared with those who did not.

These findings cast new light on the important phenomenon of access to information in organizations. It has often been asserted that employees in large, complex organizations have uneven access to information and knowledge, but there is scant theory about, and empirical analysis of, employees’ search processes that may explain inequality of access. The theoretical framework presented in this paper provides insights into why some organizational members experience greater difficulties searching for information than do others. It also helps explain the direction of search—the different pathways that different employees take to find the information they need to accomplish tasks.

Before we discuss the contributions to extant research, it is worth noting several limitations of our empirical research. We analyzed one organization only, raising issues about the extent to which our findings can be generalized. Obviously, our framework does not apply to work contexts in which employees typically do not have to engage in search for information to accomplish their tasks. Even for the settings where such search is required, our model is most applicable when search needs to take place primarily through people. For example, it does not apply when relevant search can be conducted just through information systems, as when employees are able to get all the information they need from a knowledge management system. Finally, our approach does not generalize to settings where all employees know who the experts are *ex ante*. If everyone knew who to contact on specific topics, there is no need to start any search chains where the aim is to identify an expert.

Our study also has a few limitations due to a small sample size in some of the analyses. While we started a large number of chains (381), the sample size became lower as the chains progressed. While our response rate for each survey was high (ranging from 63.3% to 100%), the number of chains completed was only 107 because the response rate for completed chains is the *product* of the response rates for the survey rounds. The more survey rounds required to complete a chain, the lower the overall response rates

for the chains. This posed some problems when we analyzed sub-groups, especially women. Of the 69 chains started by women, for example, only 12 were completed. Subsequent studies could improve on our study of women in particular by stratifying the initial sample on gender. We did not do this because we were also interested in studying other effects, but such a design can be used to conduct more fine-grained analysis. For example, with a large sample of women, researchers could study the conditions that lead some women to cross over and choose men as intermediaries.

Despite our sample size preventing us from dividing the sample into smaller groups for further analysis, our main framework and hypotheses received strong support from the empirical analysis. While acknowledging the limitations of our settings, we believe that our results have important implications for research on small worlds, intra-organizational networks, and inequities in organization.

Implications for small world and organization network research

Our research expands existing small world studies in significant ways. While the small world literature in social network research has focused on network structures that give rise to short average path lengths, this line of research has paid much less attention to explaining why individuals within such networks vary in how well they search. Our paper reverts attention to the individual level of analysis but also incorporates network structure elements by theorizing about how network structure, in the form of expert-related network centrality, affects the length and direction of an individual's search chain. Our conceptualization of search thus situates individual search chains within a network structure. This allows us to "socialize" search in the sense that employees' search is shaped by their network position, the social category to which they belong, and demographic characteristics that explain the direction of a search chain. As our results show, employees do not base their search on rational considerations alone, as when a junior consultant would select the seemingly most effective choice of a senior partner as an intermediary, but rather on what is safe, appropriate, and familiar. This approach expands existing research on search chains considerably, because it shows how employees actually search in a context of social constraints. It calls for research to move beyond considering rational search strategies only and take into account the

rich social fabric surrounding search steps. For example, one fruitful avenue of subsequent research would be to further disentangle the mechanisms of cognitive awareness and relational considerations (such as familiarity and trust) in determining *who* gets contacted in a search chain. Our methods and setting did not allow us to ask the original searchers who they would *ideally* contact as experts and who they would feel most *comfortable* contacting as experts in the first step of the search chain. It's one thing to know who the experts are; it is another to feel safe to approach them. This line of inquiry would build on existing research that examines the role of affect in determining who contacts whom in organization (e.g., Casciaro and Lobo, 2008).

Our study also opens up for some new and interesting avenues in social network research in organizations. To our knowledge, this study is the first that combines more typical intra-organization network research with the small world method, by studying the effects of individuals' position in an organization's network structure (centrality) on the direction and length of actual search paths. While researchers have often collected intra-organizational network data and studied the effect of network properties on task outcomes (e.g., Burt, 2004; Hansen 1999), existing research has stopped short of studying the actual unfolding search processes. *How* employees actually search has remained a black box. Our approach, in contrast, opens up this black box by analyzing the unfolding steps in a search chain. In some sense, our framework provides the theoretical mechanisms for understanding why a given organizational network structure affects task outcomes based on the mechanism of search. Building on our approach, social network researchers who collect intra-organizational network data can add a field experiment of search chains and analyze additional mechanisms. For example, while we focused on only one network structure property (expert-related centrality), subsequent studies can study the effects of non-redundancy in ego-centric networks (Burt 1992) on unfolding search behaviors. It would be interesting to analyze, for example, whether intermediaries who occupy structural holes in the network perform better as intermediaries than those who do not in terms of reducing the length of search chains.

Implications for inequity in organizations

Our study also contributes to the body of research devoted to understanding inequity in organizations. Much research has been conducted on various forms of inequity, ranging from overt forms of discrimination to more subtle and implicit forms of inequity. The type of information disadvantage we studied can be considered in the subtle and implicit category. It is an innocuous form: members of the organizational periphery have longer search chains than do members of the core because they have less cognitive awareness about who knows what. This factor, combined with the fact that they contact others like themselves (who in turn face the same disadvantage), prolongs search. There is seemingly no actor who carries out any discriminatory behavior. Rather, the roles of network structure, gender, tenure, and homophily operate to lengthen search. In fact, it is unlikely that the employees we studied who had longer search chains because of their periphery status even recognized that this was the case. For example, women who need one more step than men in finding an expert are unlikely to know this as there is no explicit manifestation of the inequity, in contrast to highly visible outcomes such as promotions.

One possible criticism of this form of inequity is that it doesn't matter much. Having to contact one more person in a search chain to find an expert may not seem like much of a burden. However, the extent to which this poses problems depends on the context. In our study, this impediment likely mattered substantially. Because the consultants worked under severe time pressure and needed information at the beginning of a project task, locating an expert, quickly, likely explained the ability to assimilate high-quality information, leading to different work performances. While we did not empirically analyze these consequences, subsequent studies can link search performance to task outcomes.

To the extent that the kind of information inequities suggested here are widespread, our study has important managerial implications. It highlights a need to think broadly about "equal opportunity" to also encompass equitable access to information in organizations. Undoubtedly, wider information availability through IT-based solutions might at least partially overcome some of the information disparities.¹⁸ However, such benefits from formal IT-based systems vary across contexts. In settings such as the

¹⁸ For example, Ding et al. (2009) find that diffusion of Internet has disproportionately helped subgroups that were previously disadvantaged in academic science: women scientists, early-to-mid career scientists, and those employed by mid-to-lower-tier institutions.

professional service firm we examined, the ability to leverage inter-personal networks to look for hidden information is likely to remain critical even after implementation of formal IT-based knowledge management solutions. In fact, the company we study had put in place such a system and spent significant resources on it; yet, several years after its implementation, the primary mode of finding expertise was still to ask people.

Extrapolating from our setting, we speculate that reliance on interpersonal networks remains crucial when a firm's knowledge cannot be easily codified and stored in databases, when it changes quickly (making it difficult to keep track of who knows what), and when it is distributed across people who are not official experts. This calls for managers to recognize that formal IT systems are rarely substitutes for inter-personal networks. The implication is that managers need to help members on the periphery develop their networks. For example, one approach would be to help more peripheral members become reasonably good connectors between peripheral and core employees. Such connectors may not necessarily be perfect connectors (providing the shortest path), but they may provide "good enough" pathways to experts in an organization.

In conclusion, as many organizations are increasingly dependent on their knowledge to compete, how that knowledge is distributed and how employees are able to search for that knowledge become crucial to understand, yet extant theory offers few models for how employees actually conduct searches within large distributed organizations. Our model of network search, which not only explains the direction of search chains but also their efficacy, seeks to close that gap in organization theory.

REFERENCES

- Adamic, L., E. Adar. 2005. How to search a social network. *Social Networks*. **27** 187-203.
- Austin, J.R. 2003. Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology*. **88** 866-878.
- Barnard, H.R., H.R. Bernard, P.D. Killworth, C. McCarty. 1982. Index: an informant-defined experiment in social structure. *Social Forces*. **61** 99-133.

- Baum, J.A.C., A.V. Shipilov, T.J. Rowley. 2003. Where do small worlds come from?. *Industrial and Corporate Change*. **12** 697-725.
- Borgatti, S., R. Cross. 2003. A relational view of information sharing and learning in social networks. *Management Science*. **49** 432-445.
- Brandon, D.P., Hollingshead, A.B. 2004. Transactive memory systems in organizations: matching tasks, expertise, and people. *Organization Science*. **15** 633-644.
- Brass, D.J. 1984. Being in the right place: A structural analysis of individuals influence in an organization. *Administrative Science Quarterly*. **29** 518-539.
- Brass, D.J. 1985. Men's and women's networks: A study of interaction patterns and influence in an organization. *Academy of Management Journal*. **28** 327-343.
- Burt, R.S. 1992. *Structural Holes*. Cambridge, MA: Harvard University Press.
- Burt, R. 2004. Structural holes and good ideas. *American Journal of Sociology* **110**(2) 349-99.
- Chatman, J.A., C.A. O'Reilly. 1994. Working smarter and harder: a longitudinal study of managerial success. *Administrative Science Quarterly*. **39** 603-627.
- Davis, G.F., M. Yoo, W.E. Baker. 2003. The small world of the American corporate elite, 1982-2001. *Strategic Organization*. **1** 301-326.
- Denrell, J., N. Arvidsson, U. Zander. 2004. Managing Knowledge in the dark: An empirical study of the reliability of capability evaluations. *Management Science*. **50**(11) 1491-1503.
- Ding, W.W., S.G. Levin, P.E. Stephan, A.E. Winkler. 2009. The Impact of Information Technology on Scientists' Productivity, Quality and Collaboration Patterns. Working Paper.
- Dodds, P., R. Muhamad, D. Watts. 2003. An experimental study of search in global social networks. *Science*. **301** 827-829.
- Edmondson, A. 1999. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*. **44** (4) 350-383.
- Fleming, Lee, C. King, A. Juda. 2007. Small worlds and regional innovation. *Organization Science*. **18**(6) 938-954.
- Freeman, L. 1978. Centrality in social networks: Conceptual clarification. *Social Networks*. **1** 215-234.
- Gavetti, G., D. Levinthal. 2000. Looking forward and looking backward: cognitive and experiential search. *Administrative Science Quarterly*. **45** 113-137.
- Gladwell, M. 2002. *The Tipping Point: How Little Things Can Make a Big Difference*. Back Bay Books.
- Guare, J. 1990. *Six Degrees of Separation*. Random House.
- Hansen, M. T. 1999. The search transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*. **44** 82-111.
- Hansen, M.T., M.R. Haas. 2001. Competing for Attention in Knowledge Markets: Electronic Document Dissemination in a Management Consulting Company. *Administrative Science Quarterly* **46**: 1-28.
- Hansen, M. T., B. Lovas. 2004. How do multinational companies leverage technological competencies? Moving from single to interdependent explanations. *Strategic Management Journal* **25**(8) 801-822.
- Hargadon, A., Sutton, R. I. 1997. Technology brokering and innovation in a product development firm. *Administrative Science Quarterly* **42**(4) 716-749.

- Harrison, J. and G. Carroll. 1998. Organizational demography and culture: Insights from a formal model and simulation. *Administrative Science Quarterly* 43: 637-667.
- Huber, G. 1982. Organizational information systems: determinants of their performance and behavior. *Management Science* 28 135-155.
- Hubert, G., R. Daft. 1987. The information environment of organizations. In Jablin, F., L.L. Putnam, K.H. Roberts and L.W. Porter (eds), *Handbook of Organization Communication*. Sage, Beverly Hills, CA, 130-163.
- Ibarra, H. 1992. Homophily and differential returns: sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*. 37 422-447.
- Ibarra, H. 1995. Race, opportunity, and diversity of social circles in managerial networks. *Academy of Management Journal*. 38 673-703.
- Ibarra, H. 1997. Paving an alternative route: gender differences in managerial networks. *Social Psychology Quarterly*. 60 91-102.
- Kanter, R.M. 1977. *Men and Women of the Corporation*. New York: Basic Books.
- Katila, R., G. Ahuja. 2002. Something old, something new: a longitudinal study of search behavior and new product introduction. 45 1183-1194.
- Keltner, D., D.H. Gruenfeld and C. Anderson. 2003. Power, approach, and inhibition. *Psychological Review* 110 265-284.
- Killworth, P., H. Bernard. 1978. Reverse small world experiment. *Social Networks*. 1 159–192.
- Killworth, P., H. Bernard. 1979 A pseudomodel of the small world problem. *Social Forces* 58: 477-505.
- Killworth, P.D., C. McCarty, H. Bernard, M House. 2006. The accuracy of small world chains in social networks. *Social Networks*. 28 85-96.
- Kleinberg, J. 2000. The small-world phenomenon: An algorithmic perspective. *Proc. 32nd ACM Symposium on Theory of Computing*.
- Kleinfeld, J.S. 2002. The small world problem. *Society*, January/February.
- Kochen, M. (ed). 1989. *The Small World*. Norwood, New Jersey: Ablex Publishing Corp.
- Kogut, B., G. Walker. 2001. The small world of Germany and the durability of national networks. *American Sociological Review*. 66 317-335.
- Korte, C., S. Milgram. 1970. Acquaintance links between white and negro populations: application of the small world method. *Journal of personality and social psychology*. 15 101-108.
- Krackhardt, D. Assessing the Political Landscape: Structure, Cognition, and Power in Organizations. *Administrative Science Quarterly*. 35 (2) 342-369.
- Lazarsfeld, P.F., R.K. Merton. 1954. Friendship as a social process: a substantive and methodological analysis. In Berger, M. (ed), *Freedom and Control in Modern Society*. New York: Van Nostrand, 18-66.
- Lin, N., P. Dayton, P. Greenwald. 1978. *The urban communication network and social stratification: A small world experiment*. In B. D. Ruben (ed), *Communication Yearbook, Vol. 1*. New Brunswick: Transaction Books, 107-119.
- McPherson, M., L. Smith-Lovin, J.M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*. 27 415-44.
- Milgram, S. 1967. The small world problem. *Psychology Today*. 2 60-67.

- Miller, J. G. 1972. Living systems: the organization. *Behavioral Science* **17** 1-8.
- Mintz, B., M. Schwartz. 1981. Interlocking directorates and interest group formation. *American Sociological Review*. **46** 851-868.
- Mintz, B., M. Schwartz. 1985. *The Power Structure of American Business*. University of Chicago Press.
- Moreland, R., L. Argote, R. Krishnan. 1996. Socially shared cognition at work: transactive memory and group performance. In Nye, J. and A. Brower (eds), *What's Social About Social Cognition?* Thousand Oaks, California: Sage, 57-85.
- O'Reilly, C. 1978. The intentional distortion of information in organization communication: a laboratory and field approach. *Human Relations*. **31** 173-193.
- Phillips, D. 2001. The promotion paradox: Organizational mortality and employee promotion chances in Silicon Valley law firms, 1946–1996. *American Journal of Sociology*. **106** (4) 1058-98
- Podolny, J.M. 1993. A status-based model of market competition. *American Journal of Sociology*. **98** (4) 829.
- Podolny, J.M. 2001. Networks as the pipes and prisms of the market. *American Journal of Sociology*. **107** 33-60.
- Podolny, J.M., J. Baron. 1997. Resources and relationships: Social networks, mobility, and satisfaction in the workplace. *American Sociological Review* **62** 673-693.
- Pool, I., M. Kochen. 1978. Contacts and influence. *Social Networks*. **1** 5–51.
- Reagans, R., B. McEvily. 2003. Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*. **48** 240– 267.
- Rogan, M. 2008. Too close for comfort? The effect of competitive overlap and relationship duration on client retention post-acquisition. INSEAD Working Paper.
- Ruef, M., H.E. Aldrich, N.C. Carter. 2003. The structure of founding teams: homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review*. **68** 195-222.
- Schilling, M.A., C.C. Phelps. 2007. Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science*. **7** 1113-1126.
- Schulz, M. 2003. Pathways of relevance: exploring inflows of knowledge into subunits of multinational corporations. *Organization Science*. **149** 440-459.
- Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Science* **51** 756-770.
- Stuart, T., J. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal*. **17**.
- Travers, J., S. Milgram. 1969. An experimental study of the small world problem. *Sociometry*. **32** 425-443.
- Tsai, W. 2001. Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal*. **44** 996 – 1004.
- Tuma, N., M. Hannan. 1984. *Social Dynamics: Models and Methods*. Academic Press: New York.
- Uzzi, B., J. Spiro. 2005. Collaboration and creativity: The small world problem. *American Journal of Sociology*. **111** 447–504.

- Uzzi, B., L. AN. Amaral, F. Reed-Tsochas. 2007. Small-world networks and management science research: a review. *European Management Review*. **4** 77-91.
- Wasserman, S., K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge.
- Watts, D., S. Strogatz. 1998. Collective dynamics of small-world networks. *Nature*. **393** 440-442.
- Watts, D.J. 1999. Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*. **105** 493-528.
- Wegner, D.M. 1987. *Transactive memory: a contemporary analysis of the group mind*. In G. Mullen, G. Goethals (eds). *Theories of Group Behaviors*. Springer-Verlag, New York, 185-208.
- Wegner, D.M., R. Erber, P. Raymond. 1991. Transactive memory in close relationships. *Journal of Personality and Social Psychology*. **61** 923-929.
- Zander, U., B. Kogut. 1995. Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. *Organization Science*. **6** 76-92.

Figure 1. Illustration of search chains for knowledge experts in an organization

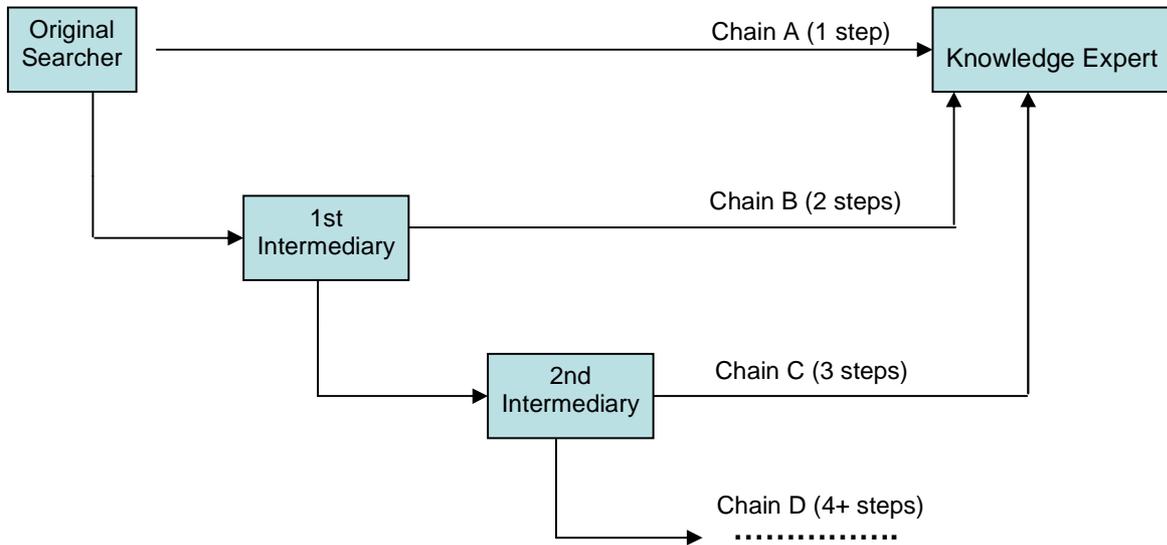


Figure 2. Plot of volume and concentration for 227 topic areas

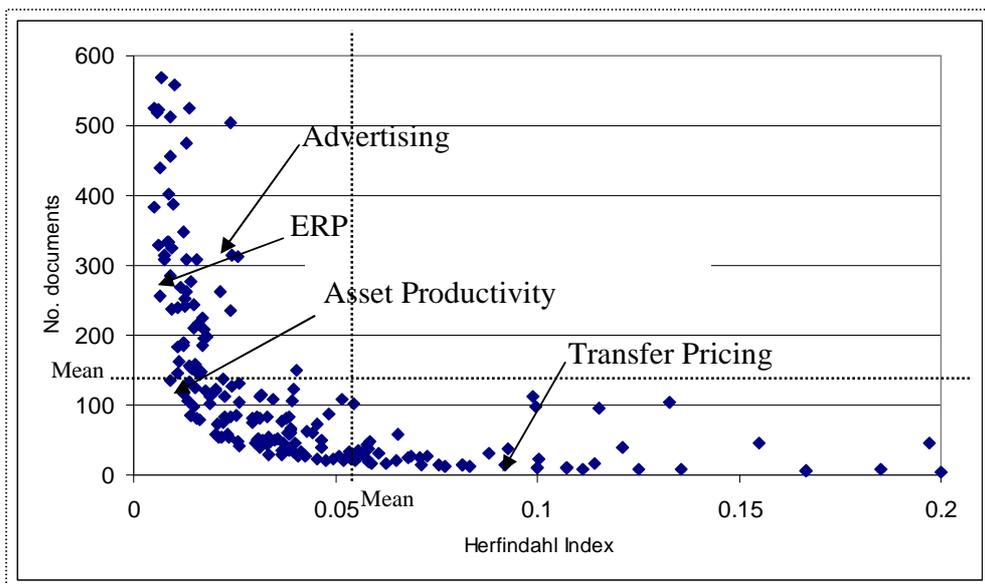


Table 1. Descriptive statistics for the selected four topics

Topic	No. documents uploaded in the database	Herfindahl index (based on authors and documents)	Quadrant in fig. 2 for which representative
Advertising	313	0.026	Upper-right (closest case)
Asset Productivity	148	0.017	Lower-left
Enterprise Resource Systems	269	0.011	Upper-left
Transfer Pricing	14	0.092	Lower-right
Mean for all 227 topics	153	0.033	
Std Dev for all 227 topics	219	0.090	

Table 2a. Results for nomination of experts

Topic	Number of initially surveyed "suspected experts"	1st survey: individuals receiving at least one nomination	2nd survey: new individuals receiving at least one nomination	3rd survey: new individuals receiving at least one nomination	Final results: individuals receiving at least three nominations
Total across all 4 topics	79	78	34	9	26
Advertising	29	24	12	3	8
Asset Productivity	18	26	7	2	7
Enterprise Resource Systems	14	23	8	3	9
Transfer Pricing	18	5	7	1	2
Cumulative number nominated		78	112	121	
Total surveyed in round		79	50	15	
Response rate for round		75%	90%	80%	

Table 2b. Descriptive statistics for the experts

Topic	No. of core experts	Average no. of nominations received	Average tenure in company (years)	% Male	No. of different offices	No. of different countries
Advertising	8	5.6	9.7	87.5%	6	6 – Japan, Austria, France, Hong Kong, Germany, US
Asset Productivity	7	7.4	10.7	85.7%	5	3 - Germany, France, US
Enterprise Resource Systems	9	5.9	9.3	100%	7	3 – US, Austria, Germany
Transfer Pricing	2	4.5	7.8	100%	2	2 – US, Australia
Total across all 4 topics	26	6.1	9.7	92.3%	14	7
Statistics for all employees in the company			3.8	80.5%	50	34

Statistics for the entire company were calculated based on the company's personnel database, which includes 3,150 consultants who worked at the company at the time of the study.

Table 3a. Details of the surveys conducted for examining realized search paths

	No. surveyed in 1st round	No. surveyed in 2nd round	No. surveyed in 3rd round	No. surveyed in 4th round	Total for all rounds
Total number of surveys	381	114	31	3	529
advertising	96	28	6	0	130
asset productivity	96	29	5	1	131
enterprise resource systems	94	27	7	0	128
transfer pricing	95	30	13	2	140
Total number of responses	241	89	23	3	356
responses naming an expert	40	32	7	1	80
responses naming an intermediary	139	45	10	0	194
responses with no name returned	62	12	6	2	82
Survey response rate	63.3%	78.1%	74.2%	100.0%	67.3%

Notes. Since each actual survey pertained to search for an expert in a specific topic area, individuals named as intermediaries multiple times within the same topic area needed to be asked to fill in the survey for that topic area only once. However, individuals named as intermediaries in multiple topic areas were asked to fill in one survey per topic area. In total, the 529 surveys involved 511 unique individuals (including original searchers or intermediaries).

Table 3b. Details of the search chain paths resulting from the surveys

	Step = 1	Step = 2	Step = 3	Step = 4	Step = 5	Total for all steps
Number of chains potentially reaching this step	381	137	49	13	2	582
Number of chains actually reaching this step	241	112	41	13	2	409
search ends successfully as expert named	40	46	14	7	0	107
search continues with intermediary named	139	49	13	2	0	203
search terminates unsuccessfully as no name returned	62	17	14	4	2	99

Notes. To construct a dataset where each step of a chain is an observation, we transformed the original data on 529 actual surveys from Table 3a into a sample of 582 observations. This approach is akin to the methodology of “spell splitting” whereby a chain is broken into its constituent parts of all individuals belonging to that chain (Tuma and Hannan 1984). In cases where the same individual was named in multiple chains, this implied creating multiple observations for the person’s corresponding survey. Because searchers from different chains (on the same topic) sometimes named the same intermediaries, these intermediaries could appear in multiple chains: 23 individuals were named as intermediary by two searchers, 3 individuals were named thrice, 3 individuals were named four times, 1 individual was named five times, another individual was named six times, and a final one was named seven times. Data coding errors for two chains led to 139 intermediaries named in step 1 actually leading to only 137 usable potential chains in step 2.

Table 4. Descriptive statistics for the original survey respondents (N=381)

	Mean	Std Dev	Min	Max	1	2	3	4	5	6	7	8
1 <i>network distance to expert</i>	2.09	0.60	1	4	1.00							
2 <i>tenure</i>	3.83	3.52	0.15	21.76	-0.15	1.00						
3 <i>woman</i>	0.18	0.39	0	1	0.00	-0.07	1.00					
4 <i>self-reported expertise</i>	1.77	1.70	0	7	-0.05	0.18	-0.10	1.00				
5 <i>expert in same office</i>	0.15	0.36	0	1	-0.35	-0.04	-0.06	0.12	1.00			
6 <i>advertisement</i>	0.25	0.43	0	1	-0.18	0.02	0.04	-0.06	-0.01	1.00		
7 <i>asset productivity</i>	0.25	0.43	0	1	-0.02	0.04	0.01	0.07	0.20	-0.34	1.00	
8 <i>enterprise resource systems</i>	0.25	0.43	0	1	-0.06	-0.03	0.00	-0.04	-0.04	-0.33	-0.33	1.00

Table 5. Determinants of chain length

	(1)	(2)	(3)	(4)	(5)
	Model: Ordered Logit				
	Dependent Variable: chain length				
	Ordered Logit	Ordered Logit	Ordered Logit	Ordered Logit	Ordered Logit
	chain length	chain length	chain length	chain length	chain length
<i>network distance to expert</i>		0.642*** (0.242)			0.537** (0.251)
<i>tenure</i>			-0.139*** (0.034)		-0.126*** (0.035)
<i>woman</i>				0.775** (0.371)	0.821** (0.382)
<i>self-reported expertise</i>	-0.339*** (0.103)	-0.332*** (0.104)	-0.303*** (0.105)	-0.324*** (0.104)	-0.281*** (0.107)
<i>expert in same office</i>	-0.002 (0.340)	0.295 (0.359)	-0.166 (0.346)	0.003 (0.343)	0.105 (0.370)
<i>advertisement</i>	-1.110*** (0.409)	-0.821* (0.421)	-1.132*** (0.417)	-1.140*** (0.412)	-0.926** (0.432)
<i>asset productivity</i>	-1.288*** (0.401)	-1.231*** (0.401)	-1.290*** (0.408)	-1.352*** (0.404)	-1.323*** (0.413)
<i>enterprise resource systems</i>	-1.085*** (0.408)	-0.866** (0.415)	-1.069** (0.416)	-1.137*** (0.411)	-0.949** (0.424)
Observations	239	239	239	239	239
Chi-squared	25.23***	32.46***	42.79***	29.92***	51.83***
Degrees of freedom	7	8	8	8	10
Log likelihood	-277.5	-273.9	-268.7	-275.1	-264.2

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes. The intended sample here was 241 of the 381 search chains that actually got initiated by the original searchers responding to our initial survey. However, the actual regressions use only 239 observations, as two observations were completely determined. The dependent variable *chain length* measures the number of steps in the realized search chain in going from an originator to an expert. An ordered logit approach was employed to avoid stringent assumptions about the underlying functional form, while also allowing inclusion of non-completed chains by coding them as being of length “99” (longer than any successful chain; the exact magnitude does not matter since an ordered logit only uses the ordinal ranking of the outcomes).

Table 6. Determinants of the probability of naming an expert in the first step

	(1)	(2)	(3)	(4)	(5)
	Model: Multinomial Logit				
Dependent Variable: <i>reached expert</i>	<i>reached expert</i>				
<i>network distance to expert</i>		-1.138*** (0.354)			-1.050*** (0.394)
<i>tenure</i>			0.184*** (0.045)		0.167*** (0.046)
<i>woman</i>				-1.444* (0.762)	-1.381* (0.794)
<i>self-reported expertise</i>	0.267* (0.146)	0.242 (0.151)	0.219 (0.158)	0.252* (0.147)	0.158 (0.165)
<i>expert in same office</i>	-0.225 (0.492)	-0.811 (0.535)	0.071 (0.527)	-0.266 (0.498)	-0.510 (0.574)
<i>advertisement</i>	-0.011 (0.551)	-0.591 (0.587)	-0.077 (0.581)	0.038 (0.555)	-0.540 (0.620)
<i>asset productivity</i>	0.360 (0.527)	0.160 (0.536)	0.275 (0.563)	0.448 (0.530)	0.227 (0.582)
<i>enterprise resource systems</i>	0.298 (0.552)	-0.126 (0.575)	0.270 (0.590)	0.364 (0.556)	-0.016 (0.612)
Observations	381	381	381	381	381
Chi-squared	525.2	536.5	548.7	529.9	560.8
Degrees of freedom	18	21	21	18	30
Log likelihood	-220.5	-214.8	-208.7	-218.1	-202.6

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Notes. The sample used here was all 381 original searchers, with a multinomial logit regression model used to allow inclusion even of cases where the individual either returned the survey without a name or did not return the survey at all. The dependent variable *reached expert* captures whether or not an original searcher named an expert right away. Similar results, not reported here to conserve space, were obtained when the analysis was repeated by including observations for not just the initiating step but also at other steps in a chain (and using an additional control for the step number).

Table 7. Determinants of characteristics of the next person in a chain

	(1)	(2)	(3)	(4)	(5)	(6)
Model:	Multinomial Logit	Multinomial Logit	Multinomial Logit	Multinomial Logit	Multinomial Logit	Multinomial Logit
Dependent Variable:	<i>next is low network distance to expert</i>	<i>next is low network distance to expert</i>	<i>next is high tenure</i>	<i>next is high tenure</i>	<i>next is male</i>	<i>next is male</i>
<i>network distance to expert</i>		-1.084*** (0.304)		-0.439 (0.286)		-0.140 (0.379)
<i>tenure</i>		0.141*** (0.044)		0.149*** (0.044)		0.036 (0.062)
<i>woman</i>		-1.537*** (0.584)		-0.713 (0.478)		-1.124** (0.497)
<i>advertisement</i>	0.580 (0.445)	0.182 (0.487)	-0.483 (0.443)	-0.729 (0.478)	-1.152** (0.518)	-1.180** (0.554)
<i>asset productivity</i>	0.004 (0.445)	-0.341 (0.482)	0.051 (0.427)	-0.108 (0.450)	1.232* (0.727)	1.250* (0.740)
<i>enterprise resource systems</i>	-0.182 (0.465)	-0.510 (0.501)	-1.017** (0.474)	-1.219** (0.503)	1.501* (0.836)	1.514* (0.850)
Observations	381	381	381	381	381	381
Chi-squared	5.983***	58.04***	9.798***	43.92***	27.97***	46.57***
Degrees of freedom	6	12	6	12	6	12
Log likelihood	-379.7	-353.7	-378.5	-361.4	-329.4	-320.1

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Notes. The sample used here was all 381 original searchers, with a multinomial logit regression model used to allow inclusion even of cases where the individual either returned the survey without a name or did not return the survey at all. The cut-offs for defining the dependent variables *next is low network distance to expert* and *next is high tenure* were based on the respective sample medians. Similar results, not reported here to conserve space, were obtained when the analysis was repeated by including observations for not just the initiating step but also at other steps in a chain (and using an additional control for the step number).

Table 8. Determinants of the remaining chain length and gains from “crossing over”

	(1)	(2)	(3)	(4)	(5)
	Model: Ordered Logit	Ordered Logit	Ordered Logit	Ordered Logit	Ordered Logit
Dependent Variable:	<i>remaining chain length</i>				
<i>decrease in network distance to expert</i>		-2.966*** (0.843)			-2.935*** (0.827)
<i>increase in tenure</i>			-1.614*** (0.410)		-1.614*** (0.428)
<i>cross over from woman to man</i>				-2.433*** (0.778)	-1.969** (0.799)
<i>step</i>	0.311* (0.179)	0.301* (0.181)	0.322* (0.194)	0.300* (0.180)	0.308 (0.195)
<i>network distance to expert</i>	0.503** (0.233)	0.779*** (0.235)	0.495** (0.240)	0.504** (0.245)	0.780*** (0.250)
<i>tenure</i>	-0.109*** (0.027)	-0.104*** (0.026)	-0.143*** (0.029)	-0.108*** (0.027)	-0.136*** (0.028)
<i>woman</i>	0.705** (0.279)	0.685** (0.288)	0.628** (0.292)	2.363*** (0.747)	1.957** (0.773)
<i>self-reported expertise</i>	-0.440*** (0.069)	-0.433*** (0.071)	-0.459*** (0.073)	-0.418*** (0.069)	-0.434*** (0.078)
<i>expert in same office</i>	-0.229 (0.348)	-0.225 (0.354)	-0.270 (0.333)	-0.175 (0.348)	-0.201 (0.335)
<i>advertisement</i>	-1.266** (0.515)	-1.311** (0.516)	-1.484*** (0.537)	-1.326*** (0.512)	-1.565*** (0.538)
<i>asset productivity</i>	-1.599*** (0.483)	-1.631*** (0.483)	-1.797*** (0.489)	-1.564*** (0.487)	-1.814*** (0.487)
<i>enterprise resource systems</i>	-1.488*** (0.488)	-1.546*** (0.485)	-1.637*** (0.481)	-1.481*** (0.488)	-1.671*** (0.473)
Observations	440	440	408	440	408
Chi-squared	103.5***	121.2***	2288***	110.3***	1711***
Degrees of freedom	11	12	11	12	13
Log likelihood	-398.2	-389.0	-360.8	-390.9	-348.7

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes. This table employs all observations (originating or intermediary) corresponding to the 239 chains initiated with an initial survey response and analyzed in Table 5. Since the 142 initial surveys not analyzed in Table 5 are again excluded, the final sample is 440 (i.e., 582 – 142). Observations based on the same survey (see Tables 3a and 3b) were entered with their weights adjusted for duplication. Observations from the same chain were clustered for standard error calculation. The dependent variable *remaining chain length* is defined as the number of realized steps in going from a focal individual (originating searcher or intermediary) to an expert. The variable *decrease in network distance to expert* measures the *network distance to expert* for the searcher *less* the same for the next person she named. The variable *increase in tenure* measures the next person’s tenure *less* the searcher’s tenure. An ordered logit approach was employed to avoid stringent assumptions about the underlying functional form, while also allowing inclusion of non-completed chains by coding *remaining chain length* as “99”.

Appendix 1

Survey questions contained in e-mail surveys to original and intermediary searchers.

[Company name] is conducting a study to find out how [Company name] consultants locate topic experts in [Company name]. Specifically, we are trying to identify experts on [topic]. I would greatly appreciate it if you would respond to the questions below by entering your responses in the spaces provided. All answers will be kept confidential.

Thank you for your assistance,
[name of senior partner sponsoring project]

1. How would you rate your knowledge of [topic] on a scale of 1 to 7? (1=no knowledge, 4= moderate knowledge, 7=expert knowledge).

2. Whom in [Company name] would you contact as an expert on [topic] if you were to do a client case on this topic (name only one person)?

Last Name First Name Office

__ Don't know

__ Would not contact anyone because I have sufficient expertise

3. If you do not know who is an expert on [topic], whom would you suggest as someone who could help identify an expert?

Last Name First Name Office

Appendix II

A comparison with previous small world studies

Kogut and Walker (2001) compare the small-world property in several networks. As described in detail in their article, a network is said to demonstrate the small world property when it has an average path length that is comparable to that in a random graph of the same size (i.e., same number of nodes and edges) but has a clustering coefficient that is much larger than that of the random graph. The following table repeats a similar calculation for the network of consultants in our study, and compares the outcome with that reported for a few other studies on small worlds.

	Path Length (PL)			Clustering Coefficient (CC)			Small
	Actual	Random	Ratio (PLr)	Actual	Random	Ratio (CCr)	
Film actors (Watts-Strogatz, 1998)	3.65	2.99	1.22	0.79	0.0003	2925.93	
Power grid (Watts-Strogatz, 1998)	18.70	12.40	1.51	0.08	0.0050	16.00	
German firms (Kogut-Walker, 2001)	5.64	3.01	1.87	0.84	0.0220	38.18	
German firm owners (Kogut-Walker, 2001)	6.09	5.16	1.18	0.83	0.0070	118.57	
Broadway artists in 1989 (Uzzi-Spiro, 2005)	3.60	2.62	1.37	0.41	0.1820	2.23	
Our network of consultants	3.01	2.25	1.34	0.46	0.0092	49.89	

As these calculations demonstrate, our network does indeed exhibit small-world properties. Specifically, clustering is significantly greater than that expected in a comparable random network (0.46 instead of 0.0092), while average path length is not too different from that in the random network (3.01 instead of 2.25). This leads to a small worlds quotient, or a clustering/path length ratio adjusted for that in a comparable random network, of 37.31 (which signifies a small-world network in line with previous studies).

A word of caution is in order here. Since different studies have used slightly different methodologies in doing the small worlds calculation for their respective networks, the reported measures from different studies are not strictly comparable. For example, of the calculations reported here, only Uzzi and Spiro (2005) adjust for the fact that such networks are typically uni-variate projections of bi-variate networks (where all members of the same team form a fully linked clique), wherein simple calculations can overstate the extent of clustering and understate the true path length when compared with random networks (Uzzi, Amaral and Reed-Tsochas, 2007)

Europe Campus

Boulevard de Constance

77305 Fontainebleau Cedex, France

Tel: +33 (0)1 60 72 40 00

Fax: +33 (0)1 60 74 55 00/01

Asia Campus

1 Ayer Rajah Avenue, Singapore 138676

Tel: +65 67 99 53 88

Fax: +65 67 99 53 99

www.insead.edu

Printed by INSEAD

INSEAD



**The Business School
for the World®**