INSEAD

The Business School for the World®

Faculty & Research Working Paper

The World is Not Small for Everyone: Pathways of Discrimination in Searching for Information in Organizations

Jasjit SINGH
Morten HANSEN
Joel PODOLNY
2008/61/OB/ST

The World is Not Small for Everyone:

Pathways of Discrimination in Searching for Information in Organizations

By

Jasjit Singh*

Morten T. Hansen**

Joel M. Podolny***
**

October 2008

Comments appreciated. Please send to jasjit.singh@insead.edu

* 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 Yale School of Management, 135 Prospect Street, New Haven, CT 06520, USA Fontainebleau 77305 France Ph: +1 203 432 6035 E-Mail: joel.podolny@yale.edu

The World is Not Small for Everyone:

Pathways of Discrimination in Searching for Information in Organizations

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 suggests that people are only a few connections away from the information they seek, but this may not be true for everyone. Some employees may have longer search chains than others in locating experts in an organization—their world may be large and not small. We theorize that two mechanisms—homophily and out-group status—jointly operate to lengthen search. Employees who belong to an out-group by virtue of their gender, tenure and network centrality have less awareness of who knows what in an organization. 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 an out-group and keeping the search chain at the periphery of the organization. To search effectively, employees in out-groups need to engage in heterophilous search behaviors by crossing social status 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 discrimination in the form of employees' unequal pathways to the information they seek, a poorly understood yet important type of organizational inequity in an information economy.

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

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., 1977), 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 that yields short average path lengths has yielded important insights about the properties of the networks as a whole, recent empirical research on the small world phenomenon has paid less attention to individual-level variables that may explain 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 theories cannot explain well why some individuals 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 (Travers and Milgram, 1969; 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).

Yet, understanding why some employees are subjected to arduous search processes for information and how those search processes unfold in organizations is an important, but underexplored, topic in organization theory (Uzzi, Amaral and Reed-Tsochas, 2007). This form of information discrimination, whereby employees in certain social categories confront organization obstacles to finding information, may negatively affect their work performance, with negative consequences for career progressions and pay. This is especially the case in work contexts in which complex tasks require the assimilation of disparate information to be performed well. Such "knowledge work" can be found not only in professional services firms, such as law (Phillips, 2001), investment banking (Podolny, 1993), advertising (Rogan, 2008) and consulting (Hansen and Haas, 2001), but also in functions such as product innovation (Hargadon and Sutton, 1997) and operations (Zander and Kogut, 1995).

To address the paucity of research on individual variation in intra-organizational search processes, 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 out-group status in theories of power and stratification in organization research (e.g., Kanter, 1977; Brass, 1984): employees who belong to out-groups in organizations by virtue of attributes

such as gender, race or tenure may find themselves on the periphery of the information flow in an organization 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 search chain may be longer for out-groups.

To simplify our analysis, we limit our discussion to search for individual topic 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. 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 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. [1] For example, a large global population study using the Internet demonstrated a median path length of five to seven steps 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 may be 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 may give rise to short average path lengths. They proposed that small world networks exhibit two features: individuals are grouped into local cluster 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 why individuals can have short path lengths in networks with overall low density.

[1] Some researchers have also criticized the original study for not taking into account incomplete chains that likely would have increased average path length had they eventually been completed (Kleinfeld, 2002).

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 work 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. Travers and Milgram (1969) hinted at this issue in their original article. A striking feature of their study is not only the average path length of six steps but also the wide distribution of intermediaries required to reach the target, which ranged from 1 to 10 intermediaries for the 64 completed chains in the study.

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, rational-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). 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 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 in search of knowledge 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 a searcher 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 information 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. 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.[2] 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 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 out-group 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

[2] Conceptually, contacting someone who is considered an expert is different from contacting someone who is considered an intermediary, although a searcher may think of them as one and the same.

knows what" in an organization (Wegner, 1987; Wegner at 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 in out-groups may be at a relative disadvantage in locating experts because they may have poorer information about the identity of experts than do in-groups. As a long line of organization research has shown, members of various out-groups 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 out-group positions are 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 or out-group. We define members as having an out-group status to the extent that they have a low network centrality, have short tenure, and whose gender is underrepresented.[3] Which groups constitute out-groups is to some extent empirical and depends on the distribution of employees within a particular organization. For example, only 19.5% of

[3] While gender is a dichotomous variable, tenure and centrality are continuous, suggesting that out-group status is more precisely thought of as a continuous variable (out-groupness).

professionals in our empirical context were women, resulting in women comprising an out-group in our setting, but this may not be the case in other organizations.

Network centrality. The first out-group variable we consider is a searcher's centrality in the organization network. One way of characterizing the degree to which an employee belongs to an out-group in an organization is to assess their position in the organization's social network (Brass, 1984). Centrally placed individuals belong to the core or the in-group while non-central actors belong to the periphery or the out-group (McPherson et al., 2001). Specifically, we consider closeness centrality in a complete network (Freeman, 1978), which in our context is defined as the path through the network that has the lowest number of intermediaries between a searcher and the topic experts.[4] *Network distance* is the converse of closeness centrality: the higher the number of intermediaries between a searcher and the topic experts in the network, the higher the network distance.

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). 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), then it is likely that the focal 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 be passed on by this third person. For example, during a prior project, the third person may have mentioned the name of the expert to the focal employee, who will then know the identity of the expert and can act on this information when searching in the future.

[4] 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.

Intermediaries in an existing network thus funnel information about "who knows what" in an organization on a continuous basis through interactions and joint project work. 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."

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 that 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 out-group 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, the third out-group variable we analyze, 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 important 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 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, these arguments about out-group effects can be summarized in a hypothesis:

> **Hypothesis 1**. Employees who have low network 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.

Using intermediaries to search: the role of homophily

Employees may undertake "social" search by contacting intermediaries and asking them to point them in the direction of experts (Gavetti and Levinthal, 2000). As research on

stratification in organizations has shown, employees 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 who are members of the out-group are likely to contact an intermediary who also is a member of an out-group, while searchers who are a member of an in-group are likely to contact a member of the in-group.

One reason for this homophilous search behavior is familiarity: two colleagues who share one or more similar 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: asking someone for help in finding experts reveals a lack of one's knowledge and exposes oneself to the risk of an unfavorable judgment (cf. 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). 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."

Homophilous search is 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 members of an out-group will contact in-group

members simply because there are few other out-group 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.

Is homophily-based search disadvantageous for out-group members? This hinges on whether intermediaries who are part of the out-group perform worse in their role as an intermediary than those who are part of the in-group. We argue that they will. The reasons for this are the same as those for Hypothesis 1. An intermediary who is part of the out-group is likely to suffer from the same exclusion issues as an original searcher who is part of the out-group. Intermediaries with low network centrality are less likely than intermediaries with a high degree of centrality to identify an 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 longer 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 out-group 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 with out-group vs. in-group members. These arguments about homophilous search can be summarized in two hypotheses:

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

Hypothesis 3. Intermediaries have longer search chains to the extent that they have long network distance, have short tenure, or are in a gender minority.

Crossing boundaries: heterophilous search

We have argued that out-group members are likely to have longer search chains because they have low cognitive awareness of who are the experts (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 members of an out-group). If this is correct, one issue becomes, how can out-group members improve their search for experts in an organization? A simple answer is for out-group members to try to "cross over:" searchers with short tenure can contact intermediaries with longer tenure; female searchers can contact male intermediaries, and searchers 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 potentially negative effects of homophily-based searches for out-group members:

Hypothesis 4. Out-group members who name in-group members as intermediaries shorten the remaining path length from intermediary to expert.

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 in the firm over the years was therefore a mosaic of knowledge that was dispersed across offices and consultants, and some became experts on a particular topic because of the issues they dealt with during their client work. The topic experts in the firm 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 topic experts. Because of the dispersion and unofficial position of topic 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 month, 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. Team members 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 each other 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.

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 expertise 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 $H = \Sigma_i p_i^2$, 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, then this measure is 1.

Figure 2, which depicts the volume of electronic documents and the degree of concentration 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).[5]

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).

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 "hidden" in the organization and 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

[5] The magnitude of expertise on these topics would differ widely by consulting firm. For example, the focal firm did not do much consulting on ERP systems, while this is the main business for other consulting firms.

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 best "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 2).

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 an expert (hereafter called a *core expert*), leading to between 2 and 8 core experts per topic, for a total of 26 core experts (see the last column in Table 2). The cut-off of four or more nominations seems reasonable, as it is difficult to justify labeling an individual a core expert if only three or less knowledgeable individuals surveyed nominated the person. Table 3 lists the demographics and office location of the core 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. Both of these suggest that out-group members 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 in-group.

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"). If the searcher did not know whom to contact as topic expert, we also asked who he or she would suggest 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.)

Chain-based survey. Of the original searchers who responded to the survey, 40 identified one of our pre-identified core experts, and the chain thus stopped, as the original searcher had already reached a core 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 a contact or 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 a core 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 4 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

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).

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

[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).

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 into a sample of 582 observations used for further analysis.[7] We assigned different observations different weights to ensure that duplicate counting of a survey that was entered as multiple observations does not bias the results.[8]

Dependent and independent variables

The first dependent variable, *reached expert*, is an indicator variable for whether or not an original searcher was able to correctly name a core expert in the first step of the chain. Our second dependent variable, *chain length*, measures the number of steps in the search chain starting from any individual (original searcher or an intermediary) to an expert.

We created a binary variable, *woman*, indicating whether an original searcher or an intermediary was a woman (1) or a man (0). We included a variable, *tenure*, denoting 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*, was a more elaborate procedure. 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 work in teams to serve clients, and accordingly a consultant's history of team membership is likely to be an important

[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, in the regression analysis, we assigned weights to each observation (using the "iweight" function in Stata) so that the sum of the fractions for any single survey observation (i.e., a unique individual-topic combination) always summed to 1. For example, if a survey of an individual was entered 4 times as different observations, the weights assigned to each observation for that survey were 0.25, for a sum of 1.

determinant of the relations that the consultant develops 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 core experts identified in our small-world study. The longer ego's path length in the work-flow network to the core experts *prior* to any given search effort, the more intermediaries lie between the experts and ego in the work-flow network. We measure *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 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

[9] We used the three-year cut-off, as it was unlikely that consultants would remember significant proportion of affiliation ties that existed before then. Also, consultants' expertise developed year by year and the expertise from more than three years ago was probably also less relevant and possibly obsolete. We also used the 40-hour criterion because someone who worked fewer than 40 hours on a project would be less likely to develop a meaningful affiliation tie with other team members.

average path lengths.[10] We then 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).[11] This variable can be seen as the converse of closeness centrality (high network distance means low closeness centrality).

Control Variables

We included a dummy 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 core experts (see appendix 1 for the question asked to solicit this information). To account for the possibility that a long geographic distance to the experts may prolong the search, we also included a control variable, *ln(geographic distance to expert + 1)*, which we calculated as the logarithm of one plus the physical distance (in miles) between the office of an employee and that of the nearest expert.[12] 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, but for reasons not directly captured by one of our variables, we also included a control variable, *step*, that measures the sequential position that a respondent occupies in a chain. This variable takes on a value of 1 if the respondent is the original searcher, a value of 2 if the respondent is the first intermediary, and so on.

[10] Kogut and Walker (2001) compared the two measures of small-world properties (clustering and path length) for several networks. Using the same method, our network appears to exhibit small-world properties. See the appendix 2 for detailed calculations of our network and a comparison with small worlds networks reported in previous studies.
[11] This measure thus captures the *minimum* network distance between the focal employee and one of the core experts on the topic in question. We also specified an *average* distance measure between the focal employee and all core experts on the topic, and our results remained the same for both measures.
[12] In analyses not reported in the paper, we also tried using a binary control for whether or not an employee and the closest expert are in the same office, and the findings did not change.

Models

In the analysis of *chain length*, we had a choice between using completed chains only (where the final length is known) or also including non-completed chains. We chose the latter approach in order not to bias the results toward completed chains. Non-completed chains also hold useful information. For example, if one of the chains is observed until the third step, but the intermediary at the third step does not name an expert or a next intermediary, the chain is not completed, yet it still carries useful information because we know that the chain length would have been *at least three*. In order to use information from not just the completed chains but also from incomplete chains, our empirical model for *chain length* uses an interval regression model. This interval regression model can be seen as a generalization of the Tobit model, since it allows the point of right censoring to vary across observations. Specifically, for a successful chain of length x the interval used as the dependent variable is simply the point interval $[x, x]$. However, for an unsuccessful chain that is only observed till step y, the interval is taken as $[y, .]$ to capture the fact that we can still tell that *chain length* could not have been smaller than y.[13]

In our analysis of the dependent variable *reached expert*, we also included non-completed chains in the analysis to control for potential response biases. We estimated a multinomial logit model in which there were three possible outcomes: (i) the individual responded to the survey and named a contact or intermediary who was not a core expert (0); (ii) the individual responded to the survey and named a core expert (1); (iii) or the individual did not return the survey or

[13] In analysis not reported in the paper, we tried two robustness checks. First, we restricted our analysis to completed chains only, and tried estimating an OLS as well as an ordered logit equation to understand drivers of *chain length*. Second, we tried using a two-step model, where a Heckman correction was used to first model whether or not a chain was successful or was left incomplete by a non-response. The main findings regarding drivers of chain length remained essentially the same as those reported in the paper using an interval regression model.

returned it without including a name (2). This approach controls for the potential selection bias of including completed search chains only (Kleinfeld, 2002; Dodds et al., 2003).[14]

RESULTS

Table 5 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 a core expert was 1.89, with a median length of 2.[15] 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 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 the corrected value for chain length one would expect if all surveys were returned to be 2.37.[16]

Descriptive statistics for our main variables are shown in Table 6. Before we test the hypotheses, we show the results from a baseline analysis predicting the length of the chain for the sample of the 381 original searchers. Consistent with our baseline expectations, the results in Table 7 show that searchers who had long network distance, short tenure, and were women had *longer* search chains. The result is robust to including these variables individually (columns 2, 3

[14] In analysis not reported in the paper, we also used an alternative specification by implementing a logit model in which the dependent variable was whether an individual named an expert (1) or not (0), and the main results from this alternate approach were the same.

[15] 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.

[16] 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 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 higher centrality in the work-flow network.[17]

To test hypothesis 1, we analyzed whether an original searcher would identify an expert in the first step. We model with competing risks, such that an original searcher is at risk of naming the expert, naming an intermediary, or not completing the survey, thus taking into account non-responses and incomplete surveys. The results in Table 8 reveal that a long network distance, 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 test the tendency to engage in homophilous search. As shown in Table 9, the dependent variable used in columns 1 and 2 is the network distance of the *next person in a chain*. The results in column 2 reveal that a searcher's network distance positively and significantly predicts the network distance of the next person in the chain. The higher a searcher's network distance (i.e., the lower a person's centrality), the higher the network distance of *the next person*. Interestingly, the tenure and woman variables are also significant in the results displayed in column (2) in Table 9. The longer the tenure of the searcher, the *lower* the network distance of the person he or she picked as an intermediary. In contrast, if the searcher is a woman, the intermediary person she picks has longer network distance than if the searcher had been a man.

Columns (3) and (4) in Table 9 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 tenure of the next person.

[17] In order to examine if there was a particularly strong effect (beyond just an additive effect) of being an out-group member along multiple dimensions, we also tried including interaction variables for tenure, gender and network distance variables. However, none of these interactions turned out to be statistically significant.

Finally, columns (5) and (6) in Table 9 use gender of the *next person* in the chain as the dependent variable, with the reported results supporting the homophily argument that a female respondent is significantly more likely than a male respondent to name a woman as the next person in the search chain. A cross-tabulation of descriptive statistics revealed an overwhelming tendency for same-sex homophily: Male searchers picked another male in 92% of the cases (with the remaining 8% picking a woman). In contrast, 67% of female searchers picked another woman as an intermediary (and 33% 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 10 predict *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 7, where only the 381 original searchers were included, to including all 582 observations, with the interval regression allowing us to include the non-completed chains in the analysis.[18] In this analysis, we conceptualize every observation as an intermediary and analyze the distance from this intermediary to an expert. The results in Table 10 reveal that a long network distance, short tenure, and being a woman all lead to a longer remaining chain length from any given intermediary. The finding is robust to including these variables individually (columns 2, 3 and 4 respectively) as well as to including these three variables together in the model (column 5). These results lend support to hypothesis 3.

[18] In this analysis, we chose to conceptualize an original searcher as an intermediary, assuming that these consultants are also representative of employees that could be asked by someone for advice on who to contact as an expert. This approach allows us to have a sufficiently large sample to test hypothesis 4. In an alternative model, we only included those consultants that were named as intermediaries by the original searchers (n=201, of which only 95 chains were completed). In this model, the coefficients for *network distance to expert*, *tenure* and *woman* continued to be of the expected sign. However, only the coefficient on *tenure* was statistically significant. This lack of significance is most likely due to smaller sample sizes associated with out-group members that were intermediaries. For example, of the 95 (out of 201) intermediaries that did ultimately have a successful chain completion, only six were women.

To test hypothesis 4, we examined whether out-group members that "crossed over" had shorter chain length than those who did not. Table 11 reports three sets of analyses. In the first model (1), we isolated searchers who had a high network distance (set as a path length of 2 to an expert). We then constructed an independent variable, *network distance decrease*, that measures the network distance score for the searcher *less* the network distance score for the intermediary he or she named. The idea is to test whether a searcher who picked someone with lower network distance had shorter chain length. As the result in column (1) in Table 11 reveals, this variable is negative and significant, lending support to hypothesis 4 and the argument that searchers with a high network distance who "cross over" to intermediaries with lower network distance have shorter search chains.

In column (2) in Table 11, we analyzed the effect for tenure. We isolated the sample of searchers who had a short tenure (less than two years). We then created a variable, *increase in tenure*, that measures the intermediary's tenure *less* the searcher's tenure. The idea is that short-tenured consultants who "crossed over" to longer-tenured intermediaries would have shorter search chains. The result reported in column (2) in Table 11 confirms this, lending support to hypothesis 4 (the result is only significant at the 0.1 level; this is likely due to a small sample of 55 observations).

Finally, column (3) in Table 11 examines whether women who named men as intermediaries had shorter chains than women who did not. The estimated coefficient is consistent with the argument that women who "cross over" to naming a man as an intermediary have shorter search chains that women who do not, but the result is not significant. The lack of significance is probably due to the low sample size of 37 observations.

To summarize, the three results reported in Table 11 appear to partially support hypothesis 4: out-groups members who cross over have shorter remaining chain lengths than those who do not. The small sample sizes prevent us from drawing more conclusive evidence.

DISCUSSION

The main finding in this paper is that members of out-groups in an organization—those with low network centrality, short tenure, and women—experienced a form of information discrimination by having longer search chains to reach experts than in-group members. Their world of search was relatively large. Two theoretical mechanisms explain why this was the case. First, out-group members had lower cognitive awareness about "who knows what" in an organization: they fared worse than in-group members in pinpointing an expert in the first step of a search chain. Second, out-group members also selected intermediaries on the basis of homophily, worsening their search process. However, those members of the out-groups that "crossed over" and selected intermediaries from an in-group 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 this search inequality. 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. It also does not apply to settings in which search can be conducted through information systems only, as when employees are able to download all the information they need from a knowledge management system. Our model assumes that some of the search takes places through people. Moreover, 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 women who started a chain, for example, only 12 were completed. Subsequent studies could improve on our study of women in particular by stratifying the initial sample on gender, drawing half the sample of initial searchers among women. We did not do this because we were 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 and the consequences of such a choice.

While our sample size prevented us from dividing the sample into smaller groups for further analysis, our main framework and hypotheses received strong support from the empirical analysis. These results contribute to two lines of research.

First, 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 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, 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 and considerations. It calls for research to move beyond considering rational search strategies only and take into account the rich social fabric surrounding search steps.

Our paper also contributes to the small-world literature by moving the argument to intra-organizational search. To our knowledge, our paper is the first that analyses intra-organizational small-world search processes in a large complex organization, and is also the first that combines data on actual unfolding search chains (step by step) with network structure variables that are

used to predict the direction and length of those chains. This approach should hold considerable promise for future research. While researchers have often collected intra-organizational network data and studied the effect of network properties on task outcomes (e.g., Hansen, 1999), this type of research has stopped short of studying the actual unfolding search processes and has instead postulated how search operates as a causal mechanism (cf. Burt, 1992). How employees 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 network structure affects task outcomes based on the mechanism of search. Building on our approach, researchers who collect intra-organizational network data can add a field experiment of search chains and analyze additional mechanisms. While we focused on one network structure property only (closeness centrality), subsequent studies can study the effects of non-redundancy in ego-centric networks (Burt, 1992). 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.

Our study also contributes to the body of research devoted to understanding discrimination in organizations. Much research has been conducted on various forms of discrimination in organizations, ranging from overt forms to more subtle and implicit forms. The type of information discrimination we studied clearly belongs in the subtle and implicit category. It is an innocuous form: out-group members have longer search chains than do in-group members because they have less cognitive awareness about who knows what and they contact others like themselves (who also have less cognitive awareness), prolonging 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 they belonged to an our-group 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 inequality, in contrast to highly visible outcomes such as promotions.

One possible criticism of this form of discrimination 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, the ability to locate an expert, quickly, likely explained the difference between the ability to assimilate high-quality information and not, leading to different work performances. While we did not empirically analyze these consequences, subsequent studies can link search performances to task outcomes.

The idea that out-group members experience subtle forms of information discrimination in their search for information can open up new lines of research. Our study is an initial attempt at understanding this complex issue. Other studies can extend our model. For example, while out-group members are likely to engage in homophilous search, different organizational contexts may alter this tendency: informal work environments that break down social barriers may encourage out-group members to contact in-group intermediaries and thereby mitigate unequal search processes. Work environments with high degrees of hierarchy and centralization, in contrast, may discourage employees from crossing over social categories in their search. Also, organizations that encourage a wider group of employees, including out-group members, to become experts are likely to reduce the negative impact of homophilous search. In general,

research can study both the conditions that lead to unequal search processes and organizational arrangements that mitigate these tendencies.

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. and 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 and C. McCarty
1982 “Index: an informant-defined experiment in social structure.” Social Forces 61: 99–133.

Baum, J.A.C., A.V. Shipilov and T.J. Rowley
2003 “Where do small worlds come from?” Industrial and Corporate Change, 12: 697-725.

Borgatti, S. and R. Cross
2003 “A relational view of information sharing and learning in social networks.” Management Science, 49: 432-445.

Brandon, D.P. and 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.

Breiger, R.L.
1976. “Career attributes and network structure: A blockmodel study of a biomedical research specialty.” American Sociological Review, 41: 117-135.

Burt, R.S.
1992. Structural Holes. Cambridge, MA: Harvard University Press.

Chatman, J.A. and 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 and W.E. Baker
2003 “The small world of the American corporate elite, 1982-2001.” Strategic Organization, 1: 301-326.

Dodds, P., R. Muhamad, and 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, no. 4 (December 1999): 350-383.

Fleming, Lee, C. King and 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. and D. Levinthal.
2000 “Looking forward and looking backward: cognitive and experiential search.” Administrative Science Quarterly, 45: 113-137.

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. and 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., and 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. and 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, Herminia

1992 “Homophily and differential returns: sex differences in network structure and access in an advertising firm.” Administrative Science Quarterly, 37: 422-447.
1995 “Race, opportunity, and diversity of social circles in managerial networks.” Academy of Management Journal, 38: 673-703.
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. and G. Ahuja
2002 “Something old, something new: a longitudinal study of search behavior and new product introduction.” Academy of Management Journal, 45: 1183-1194.

Killworth, P. and H. Bernard
1978 “Reverse small world experiment.” Social Networks 1: 159–192.
1979 “A pseudomodel of the small world problem.” Social Forces 58: 477-505.

Killworth, P.D., C. McCarty, H. Bernard and M House
2006 “The accuracy of small world chains in social networks.” Social Networks, 28: 85-96.

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. and G. Walker
2001 “The Small World of Germany and the Durability of National Networks.” American Sociological Review, 66: 317-335.

Korte, C. and 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.

Lazarsfeld, P.F. and 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 and 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, and 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. and M. Schwartz
1981 “Interlocking directorates and interest group formation.” American Sociological Review, 46: 851-868.

Mintz, B. and M. Schwartz
1985 The Power Structure of American Business. University of Chicago Press.

Moreland, R., L. Argote and 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. and J. Baron
1997 “Resources and Relationships: Social Networks, Mobility, and Satisfaction in the Workplace.” American Sociological Review, 62: 673-693.

Pool, I. and M. Kochen
1978 “Contacts and influence.” Social Networks 1: 5–51.

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 and 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. and 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. and J. Podolny
1996 “Local Search and the Evolution of Technological Capabilities." Strategic Management Journal, 17.

Szulanski, G.
1996 “Exploring internal stickiness: impediments to the transfer of best practice within the firm.” Strategic Management Journal, 17: 27-43.

Travers, J and S. Milgram
1969 “An experimental study of the small world problem.” Sociometry, 32: 425-443.

Tuma, N. and M. Hannan
1984 Social Dynamics: Models and Methods. Academic Press: New York.

Uzzi, B. and J. Spiro
2005 “Collaboration and Creativity: The Small World Problem.” American Journal of Sociology 111: 447–504.

Uzzi, B., L. AN. Amaral and F. Reed-Tsochas
2007 “Small-world networks and management science research: a review.” European Management Review, 4: 77-91.

Wasserman, S. and K. Faust
1994 Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge.

Watts, D. and 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 and P. Raymond
1991 „Transactive memory in close relationships.“ Journal of Personality and Social Psychology, 61: 923-929.

Zander, U. and B. Kogut
1995 “Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test.” Organization Science, 6: 76-92.

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 2
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 Worlds Coefficient
	Actual	Random	Ratio (PLr)	Actual	Random	Ratio (CCr)	(Q=CCr/PLr)
Film actors (Watts-Strogatz, 1998)	3.65	2.99	1.22	0.79	0.0003	2925.93	2396.85
Power grid (Watts-Strogatz, 1998)	18.70	12.40	1.51	0.08	0.0050	16.00	10.61
German firms (Kogut-Walker, 2001)	5.64	3.01	1.87	0.84	0.0220	38.18	20.38
German firm owners (Kogut-Walker, 2001)	6.09	5.16	1.18	0.83	0.0070	118.57	100.46
Broadway artists in 1989 (Uzzi-Spiro, 2005)	3.60	2.62	1.37	0.41	0.1820	2.23	1.62
Our network of consultants	3.01	2.25	1.34	0.46	0.0092	49.89	37.31

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)

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

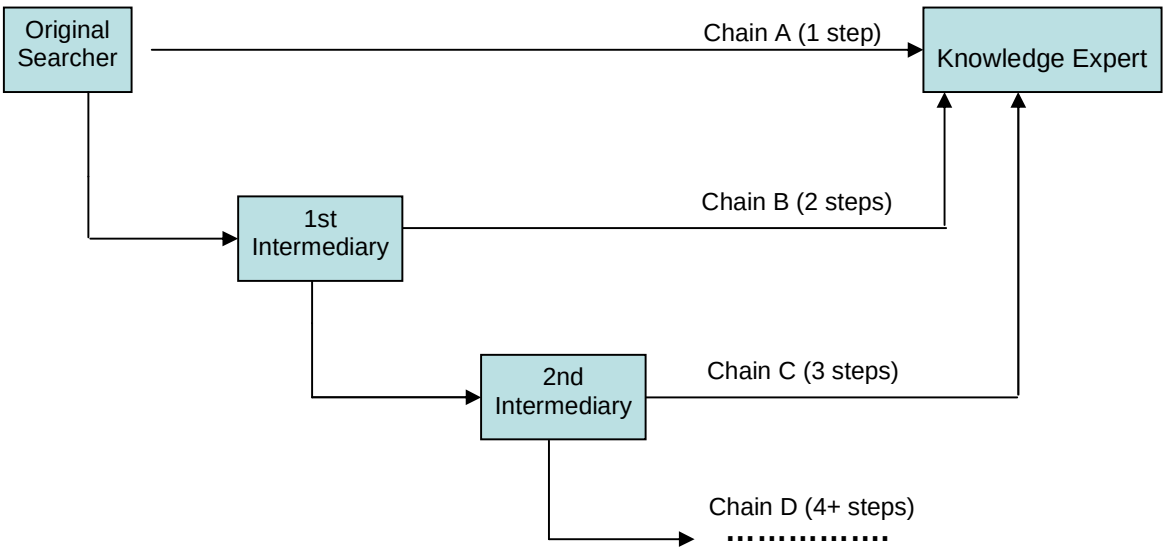

Figure 2. Plot of volume and concentration for 227 topic areas

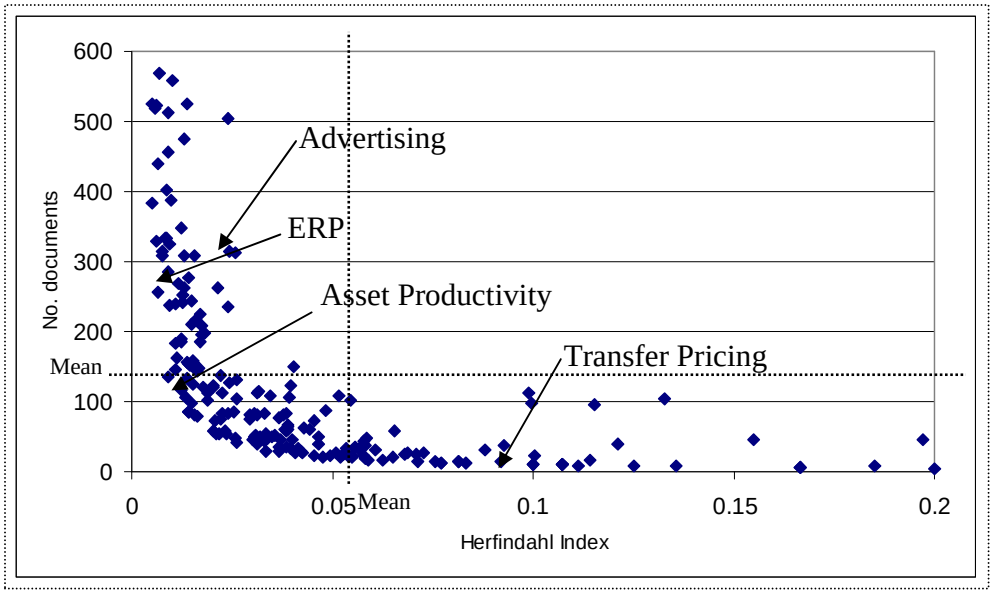

Table 1. Descriptive statistics for the selected four topics

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 2. Results for nomination of core 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 3. Descriptive statistics for the core 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 4. Details of surveys by round, topic and response rates

	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	62	12	6	2	82
Survey response rate	63.3%	78.1%	74.2%	100.0%	67.3%

Table 5. Details of chains by number of steps of search done

Step number	Chains reaching this step	Chains continued to next step	Chains reaching expert at this step	Chains neither reaching expert nor continued
1	381	139	40	202
2	139*	49	46	42
3	49	13	14	22
4	13	2	7	4
5	2	0	0	2
Number of chains			107	272
Mean step for chains			1.89	1.39
Median step for chains			2.00	1.00

*Data missing for two chains, so the remaining columns for the second step only add to 137

Table 6. Descriptive Statistics (N=582 unweighted observations)

		Mean	Std Dev	Min	Max	1	2	3	4	5	6	7	8	9
1	***network distance to expert***	2.02	0.59	1	4	1.00								
2	***tenure***	4.96	4.26	0.15	23.68	-0.20	1.00							
3	***woman***	0.16	0.37	0	1	-0.05	-0.06	1.00						
4	***self-reported expert***	2.26	2.00	0	7	-0.09	0.20	-0.14	1.00					
5	***ln(geographic distance to expert + 1)***	5.66	2.64	0	9.04	0.43	0.02	0.06	-0.14	1.00				
6	***step***	1.48	0.77	1	5	-0.14	0.34	-0.08	0.31	0.00	1.00			
7	***advertisement***	0.25	0.43	0	1	-0.19	0.00	0.10	-0.01	-0.15	0.00	1.00		
8	***asset productivity***	0.24	0.43	0	1	0.00	0.04	-0.04	0.06	-0.14	-0.06	-0.33	1.00	
9	***enterprise resource systems***	0.25	0.43	0	1	-0.01	-0.08	-0.06	-0.05	-0.04	0.03	-0.34	-0.33	1.00

TABLE 7. Determinants of chain length (for original searchers)

	(1)	(2)	(3)	(4)	(5)
Model:	**Interval Reg**	**Interval Reg**	**Interval Reg**	**Interval Reg**	**Interval Reg**
Dep Var:	***chain length***	***chain length***	***chain length***	***chain length***	***chain length***
network distance to expert		0.389***			0.308**
		(0.150)			(0.145)
tenure			-0.081***		-0.073***
			(0.018)		(0.018)
woman				0.483*	0.440*
				(0.253)	(0.238)
self-reported expert	-0.225***	-0.213***	-0.176***	-0.220***	-0.167***
	(0.059)	(0.058)	(0.056)	(0.059)	(0.055)
ln(geographic distance to expert + 1)	0.030	-0.004	0.044	0.029	0.016
	(0.029)	(0.031)	(0.028)	(0.029)	(0.030)
advertisement	-0.360	-0.236	-0.312	-0.380	-0.239
	(0.257)	(0.256)	(0.242)	(0.257)	(0.242)
asset productivity	-0.650***	-0.639***	-0.540**	-0.681***	-0.573**
	(0.248)	(0.243)	(0.235)	(0.249)	(0.233)
enterprise resource systems	-0.479*	-0.375	-0.374	-0.500*	-0.322
	(0.257)	(0.254)	(0.244)	(0.257)	(0.243)
Observations	381	381	381	381	381
Chi-squared	34.4***	41.03***	53.39***	38.24***	60.99***
Degrees of freedom	8	9	9	9	11

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

TABLE 8. Determinants of the probability of an original searcher naming a core expert in the first step

	(1)	(2)	(3)	(4)	(5)
Model:	**Multinomial Logit**	**Multinomial Logit**	**Multinomial Logit**	**Multinomial Logit**	**Multinomial Logit**
Dep Var:	***reached expert***	***reached expert***	***reached expert***	***reached expert***	***reached expert***
network distance to expert		-1.122***			-0.989**
		(0.359)			(0.392)
tenure			0.182***		0.165***
			(0.045)		(0.046)
woman				-1.425*	-1.344*
				(0.763)	(0.791)
resp self assesment	0.198	0.170	0.171	0.184	0.117
	(0.134)	(0.137)	(0.143)	(0.135)	(0.147)
ln(geographic distance to expert + 1)	0.011	0.106	-0.017	0.023	0.077
	(0.066)	(0.074)	(0.073)	(0.068)	(0.081)
advertisement	-0.016	-0.418	-0.121	0.064	-0.366
	(0.577)	(0.597)	(0.608)	(0.584)	(0.631)
asset productivity	0.350	0.254	0.228	0.453	0.258
	(0.542)	(0.548)	(0.582)	(0.547)	(0.598)
enterprise resource systems	0.238	-0.054	0.217	0.323	0.065
	(0.568)	(0.581)	(0.607)	(0.574)	(0.622)
Constant	-2.090***	-0.090	-2.710***	-2.037**	-0.759
	(0.811)	(1.011)	(0.873)	(0.825)	(1.106)
Observations	379	379	379	379	379
Chi-squared	207.4***	218.3***	228.1***	211.9***	239.8***
Degrees of freedom	12	14	14	12	20

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

TABLE 9. Determinants of characteristics of the next person (intermediary) in a chain

	(1)	(2)	(3)	(4)	(5)	(6)
Model:	**OLS**	**OLS**	**OLS**	**OLS**	**Logit**	**Logit**
Dep Var:	***next network distance to expert***	***next network distance to expert***	***next tenure***	***next tenure***	***next is woman***	***next is woman***
network distance to expert		0.486***		-0.128		-0.528
		(0.10)		(0.63)		(0.39)
tenure		-0.0562***		0.288***		-0.0568
		(0.013)		(0.080)		(0.064)
woman		0.328*		0.277		1.648***
		(0.18)		(1.07)		(0.51)
step	-0.260***	-0.102	0.471	0.0191	-0.555	-0.691
	(0.093)	(0.088)	(0.52)	(0.53)	(0.38)	(0.44)
advertisement	-0.356*	-0.186	2.173**	2.060*	0.891*	0.691
	(0.19)	(0.18)	(1.06)	(1.07)	(0.46)	(0.51)
asset productivity	-0.292	-0.170	1.116	0.752	-2.437**	-2.680**
	(0.18)	(0.17)	(1.02)	(1.00)	(1.07)	(1.10)
enterprise resource systems	-0.152	0.0178	-1.665	-1.841*	-2.337**	-2.557**
	(0.18)	(0.17)	(1.04)	(1.04)	(1.07)	(1.11)
Constant	1.814***	0.773**	8.722***	8.172***	-1.108*	0.125
	(0.19)	(0.31)	(1.08)	(1.89)	(0.63)	(1.15)
Observations	240	240	230	230	270	270
R-squared	0.05	0.22	0.07	0.12		
F	2.99***	9.36***	4.00***	4.36***		
Chi-squared					36.01***	48.97***
Degrees of freedom	4	7	4	7	4	7

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

TABLE 10. Determinants of the remaining chain length (from any step in a chain)

	(1)	(2)	(3)	(4)	(5)
Model:	**Interval Reg**	**Interval Reg**	**Interval Reg**	**Interval Reg**	**Interval Reg**
Dep Var:	***chain length***	***chain length***	***chain length***	***chain length***	***chain length***
network distance to expert		0.309***			0.232**
		(0.115)			(0.113)
tenure			-0.072***		-0.066***
			(0.013)		(0.013)
woman				0.402**	0.358*
				(0.194)	(0.185)
step	-0.216**	-0.180*	-0.071	-0.204**	-0.049
	(0.094)	(0.094)	(0.095)	(0.094)	(0.095)
resp self assesment	-0.191***	-0.188***	-0.176***	-0.188***	-0.172***
	(0.041)	(0.040)	(0.039)	(0.040)	(0.039)
ln(geographic distance to expert + 1)	0.045**	0.020	0.053**	0.046**	0.034
	(0.022)	(0.024)	(0.021)	(0.022)	(0.023)
advertisement	-0.418**	-0.356*	-0.442**	-0.455**	-0.427**
	(0.188)	(0.187)	(0.179)	(0.188)	(0.180)
asset productivity	-0.715***	-0.711***	-0.661***	-0.725***	-0.672***
	(0.182)	(0.180)	(0.174)	(0.182)	(0.173)
enterprise resource systems	-0.550***	-0.515***	-0.536***	-0.550***	-0.509***
	(0.187)	(0.186)	(0.179)	(0.187)	(0.179)
Observations	582	582	582	582	582
Chi-squared	72.65***	79.82***	101.4***	77.1***	108.9***
Degrees of freedom	8	9	9	9	11

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

TABLE 11. Gains from crossing over from out-group to in-group

	(1)	(2)	(3)
Model:	**Interval Reg**	**Interval Reg**	**Interval Reg**
Dep Var:	***chain length***	***chain length***	***chain length***
Outgroup sample:	***mindist to expert > 2***	***tenure < 2 years***	***women***
network distance decrease	-0.883***		
	(0.136)		
increase in tenure		-0.059*	
		(0.035)	
cross over from woman to man			-0.409
			(0.521)
step	0.287	-0.543	-0.631**
	(0.494)	(0.343)	(0.302)
network distance to expert	2.422	0.110	-0.326
	(812.646)	(0.346)	(0.370)
tenure	-0.047	-0.638	-0.047
	(0.073)	(0.434)	(0.052)
woman	2.780	0.086	
	(347.546)	(0.505)	
resp self assesment	-0.163*	0.133	-0.228**
	(0.095)	(0.128)	(0.103)
ln(geographic distance to expert + 1)	-0.289**	0.225***	0.046
	(0.132)	(0.066)	(0.046)
advertisement	0.004	0.519	-3.345
	(0.669)	(0.478)	(280.592)
asset productivity	-0.885***	-0.143	-3.570
	(0.313)	(0.523)	(280.592)
enterprise resource systems	-0.716	0.188	-3.147
	(0.591)	(0.523)	(280.592)
Observations	38	55	37
Chi-squared	41.57***	19.3***	36.92***
Degrees of freedom	10	10	10

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

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®