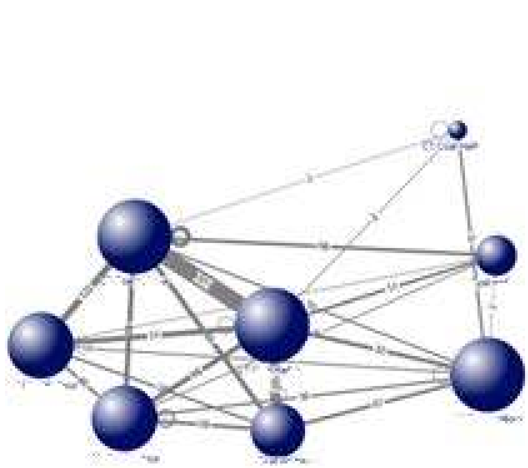
The Impact of intragroup Social Network Topology on Group Performance:

Understanding intra-organizational Knowledge Transfer through a Social Capital Framework

Professor Sean Wise, BA LLB MBA

A research project submitted as a doctoral thesis in fulfillment of the requirements for the Degree of Doctor of Philosophy in Business

December 10, 2012



Abstract

This thesis examines the effects of intragroup social network relations on group performance. Building on prior studies, it views social network topology along structural, relational and cognitive dimensions. Where previous research used a self-reporting questionnaire to gauge these dimensions, this research uses Social Network Analysis (SNA) software to measure e-mail communication logs between group members. The study was conducted in a national travel agency and focused on the social networks of 187 offices, each a subsidiary of the national travel agency. Each office group was tasked similarly and represented a unit of analysis. An analysis of more than 7 million emails was undertaken to generate social network measures for the firm wide network. Subgraphs representing the intraoffice social networks were then generated for each of the 187 travel offices in the greater firm-wide network. NodeXL® software was used to generate group measures representing the dimensions of each office's social network topology. As in prior studies, Centrality, Structural Holes, and Tie Strength (all social network concepts) were used to measure and compare the dimensions of the intragroup social networks. This study contributes by helping to differentiate the concepts of social capital and social network. This research finds the use of email logs to generate SNA more efficient but as effective as prior survey techniques. The study also extends prior work by dynamically examining the tie formation amongst recently hired employees. The study confirms existing views of a curvilinear relationship between social network relations and firm performance. This study finds social network topology a valuable predictor of group performance.

Keywords

Social Network Topology, Social Capital, Intragroup Knowledge Transfer, Performance

Table of Contents

List of Tables

List of Figures

Acknowledgements

It may take a village to raise a child, and it certainly took many people to get this dissertation done. While I can try to thank and acknowledge many, some may be forgotten. Notwithstanding, my gratitude goes out to Professor Robbie Paton and Dr. Robert MacIntosh, two supervisors with their hands firmly on the reins. To my proofreader, soon to be Dr. Eva Nesselroth-Woyzbun, I thank her for knowing the difference between who and whom and there and their. My thanks to Dr. Dave Valerie and Dr. Wendy Culkier, Chair of the Entrepreneurship Department and Vice President: Research and Innovation at Ryerson University, whose support knew no bounds. Also thanks to Thomas Gegenhuber and Milian Miric, my former students, both of whom taught me much about research. Finally, my appreciation to the internet workers on elance.com, none of whom I've met, you counted email frequency, combined hundreds of spreadsheets and made my diagrams pretty; with your help, this research truly became an exercise in global crowdsourcing.

To my wife, who supported me throughout this process and who makes all this worthwhile.

Finally to my family, I may be the first Wise to earn a University degree, but what a degree it is!

Onward and Upwards!

Professor Sean Evan Wise

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature: 

Printed name: Sean Evan Wise

Abbreviations

B2B	Corporate sales
B2C	Retail sales
BETW	Average Betweenness ctr
CLOS	Average Closeness ctr
Cook's D	Cook's distance
DSTA	Days Strong Ties Available
EIG	Average Eigenvector ctr
FTE	Full time equivalents
GLM	General Linear Modelling
GMA	Gross Margin Avg
HF	High Flyer
HOLES	Structural Holes
KPI	Key Performance Indicator
KBV	Knowledge Based View
NREV	Normalized Revenue, Normalized Sales Volume
OLS	Ordinary least-squares
RA	Research Assistant
RBV	Resource Based View
S-D	Service Dominant
SECI	Socialization/Externalization/Combination/Internalization
SNA	Social Network Analysis
TIES	Tie Strength
TSV	Total Sales Volume
TYPE	Office Type

Chapter 1 – Introduction

Introduction to the Research

At the dawn of the twenty-first century, knowledge has quickly become the key asset of commerce (Stewart, 2003). This thesis examines the effects of intragroup social network relations on group performance. To achieve optimal performance, a group attempts to efficiently create, manage, distribute and leverage its knowledge resources (Frey, 2001); by doing so groups drive both innovation and efficiency. Intragroup relationships form a topology over which knowledge can flow.

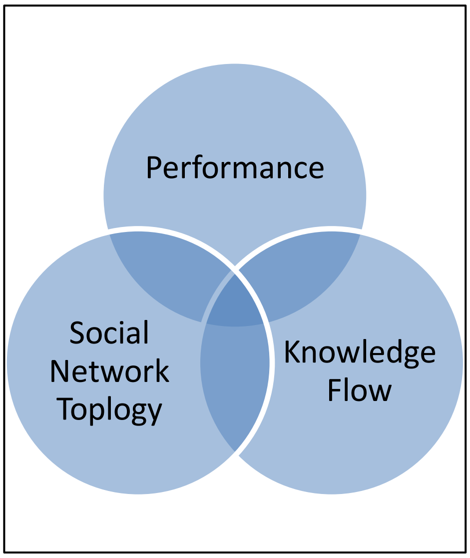


Figure 1: Areas of Research

Network topology is the arrangement of the various elements (links, nodes, etc.) of a computer (Groth and Skandler, 2005; ATIS committee PRQC, 2007) or biological network (Proulx, Promislow and Phillips, 2005). Over the last decade, Social networks have quickly become a paradigm through which a group can be examined (Gulati, 2006). The organizational chart has become less rigid. Twenty-first century groups are seen not as static snapshots but as evolving networks (Hite, 2001).

This research seeks to link a group's social network topology with that group's performance. It presupposes that certain social network topologies facilitate performance through knowledge flow more effectively than others. The research explores the following question: Do high performing groups, undertaking the same task in the same environment, share similar social network topologies, particularly topologies that facilitate knowledge transfer? This research looks to better understand *how social network topology impacts performance through the lens of social capital and knowledge flow*.

1.1 Overview to Research

Firms (collectives of resources geared to a common goal) strive for competitive advantage. In the twenty-first century, such an advantage often comes from a firm's ability to identify and efficiently transfer strategic knowledge between geographically non-proximal locations and arm's length actors (de Pablos, 2006). This research contributes to the social network conversation by augmenting and adding to the knowledge generated by recent works, particularly that of Lechner, Frankenberger and Floyd's 2010 study and Maurer, Bartsch and Ebers' 2011 study, both of which found an inverse curvilinear relationship connecting intragroup social network topology with group performance. Unlike Lechner, Frankenberger and Floyd's 2010 study which focused on strategic initiatives (which often only exist for finite periods of time and have specific explicit goals) as the unit of analysis, this research focuses on *Business as Usual* inside more than 187 groups at a national travel agency. Similarly, as compared to Maurer, Bartsch and Ebers' 2011 study which examined 218 projects in the German engineering industry as the unit of analysis, this research focuses on *Business as Usual* inside more than 187 groups at a national travel agency. Thus, this research contributes by showing the findings to be valid for durable (non-temporary) groups not just strategic initiatives or projects.

Academic literature has long heralded the importance of the knowledge process and the ability to transfer knowledge within groups. In brief, this current project is guided by the following assertions:

A 'firm' can be defined as:

- a collection of productive resources, the disposal of which, between different uses and over time, is determined by administrative decision (Penrose, 1959);

- a set of assets under common ownership and control is equated with ownership (Grossman and Hart, 1986); or

- a pool of learned skills, physical facilities and liquid capital (Chandler, 1962).

Under each and all the above definitions, both the parent company and its 187 subsidiary groups would each be seen as a 'firm'. In this research the terms firm, group and organization are used interchangeably as synonyms.

- A firm seeks to maximize profits (Carroll, 1999).

- Innovation (the process of doing things better) relies on harnessing, leveraging and recombining knowledge stocks (i.e. the knowledge held in a firm's network) (Lawson, 2003).

- Prior research has found it difficult to quantify and empirically measure knowledge flow (Hansen, 2002). Instead, researchers have focused on the antecedents to both (e.g. trust, strength of network ties).

- Social Networks can be described along three network dimensions: structural (e.g. who is near whom); relational (e.g. how strong are the bonds between actors); and cognitive (e.g. how similar are the minds in the network). These dimensions are used as proxies for SNA measures (e.g. centrality represents the structural dimension) (Neergaard, 2005).

- Social Capital can be defined as 'the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition' (Bourdieu, 1983).

- Through the Social Capital of their members, organizations gain knowledge resources (e.g. best practices) that can enhance organizational performance (Wasko and Faraj, 2005).

- Better knowledge flow leads to better performance (Vera, 2003).

- Social Capital facilitates knowledge flow. Social Capital is the lubricant that allows knowledge to flow more easily amongst nodes in a network (Peng, 2009).

This research specifically seeks to investigate the impact that intragroup Social Network Topology has on group performance. It also seeks to explore if measures of Social Network Topology can predict performance. Moreover, this work examines intragroup network topology to determine if each intragroup network dimension (cognitive, relational, and structural) has an inverse curvilinear relationship with overall performance, as suggested by Lechner, Frankenberger and Floyd (2010) in their well-cited *Academy of Management* paper. By exploring the matters above, the author hopes to illustrate how intragroup network structure impacts overall group performance.

1.2 Research Context

This research stands on the shoulders of prior authors exploring knowledge transfer and SNA.

Table 1: Recent Relevant Literatures

Year	Author(s)	Contribution
2004	**Burakova-Lorgnier, Bouzdine-Chameeva and MacGilChrist**	Propose a vision of a network structure from the point of view of knowledge transfer capacity. Find a structure based on strong ties, and thus dense network, assist the transfer of tacit knowledge.
2006	**Ordonez de Pablos**	Develops a conceptual framework for the analysis of knowledge flow between subsidiaries and their parent organization.
2010	**Lechner, Frankenberger and Floyd**	Find (at a group level, inside the group) curvilinear relationships between several SNA variables and group performance.
2011	**Maurer, Bartsch and Ebers**	Find (at project level, across an industry) curvilinear relationships between several SNA variables and group performance.

1.3 Research Questions

While recent works (Lechner, Frankenberger and Floyd, 2010; Maurer, Bartsch and Ebers, 2011) have explored the moderating role that key SNA variables (e.g. Centrality, Tie Strength) have on knowledge transfer, little research has been done to determine how these variables are interrelated and, more importantly, how one can test them dynamically. Nahapiet and Ghoshal (1998) provide a suitable framework of analysis, and it is utilised in this study. The framework consists of three types (or dimensions as they call them) of social capital: structural dimension, cognitive dimension and relational dimension (Fuller, 2006). Note: This is where the distinctions between the concepts of social capital and social network often become blurred.

This research focuses on network characteristics and how they impact performance by driving or hindering knowledge flow. It is this author's hypothesis that some network topologies are better suited to exploration practices, while others are more suited to exploitation practices. This idea was touched upon by Lechner, Frankenberger and Floyd (2010) as they explored the concept of Task Contingencies. Some organizational network structures that yield positive social capital in some task situations can convey social liability in other situations. Imagine a network of creative designers; they form a network structure that is optimal for creative tasks undertaken by the network, but the same topology might not be suited to efficiently execute routine tasks. This is known as the theory of Task Contingency (Donaldson, 2001). Research has polarized task contingency around the degree of task exploration. Prior research shows that exploratory teams (i.e. those teams focused on innovation) complete their projects more quickly if they have a social network structure composed of many strong external ties that are non-redundant. In contrast, teams pursuing tasks that exploit existing expertise (e.g. those teams focused on efficiency) take longer to complete if they have this type of social network structure, mainly because external ties must be maintained but are not needed for the task.

Although this issue was originally addressed by Hansen, Podolny and Pfeffer (2001), recent findings show knowledge transfer mediates between organization members' intra-organizational social capital and organizational performance outcomes of growth and innovation performance (Maurer, Bartsch and Ebers, 2011). Based on this, the research-derived theoretical model is as follows:

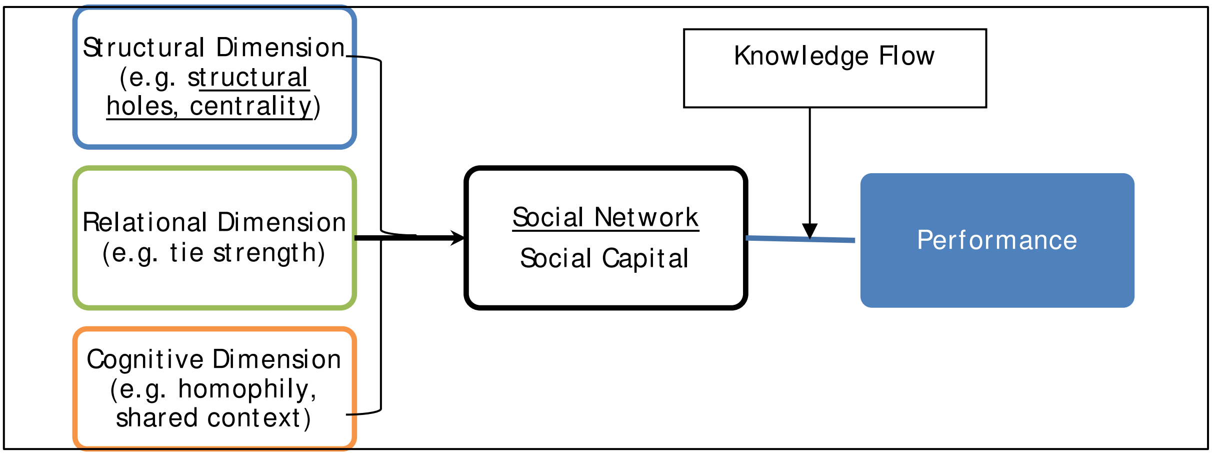


Figure 2: Proposed Research Model

The author seeks to explore how well intragroup social network topology predicts performance and defines the Research Objective as empirically confirming the relationship between an organization's Social Network Topology and organizational Performance. More generally stated as:

Do the Cognitive, Relational and Structural Dimensions of an organization's Social Network have an inverse curvilinear correlation with organizational Performance as predicted?

Additionally, because the data available in this research is time-stamped, this thesis has the opportunity to explore how the dimensions of network topology are formed dynamically, and whether or not such dynamics impact individual performance. This can be examined through the following Research Question:

What is the relationship between the speed at which an individual forms strong ties and that individual's performance?

1.4 Overview of Method

The sample population examined in this research is comprised of all the sales associates of a national travel agency. This organization, which will be called

High Flyer or '**HF**' hereafter, employs over 1800 individuals in Canada and 20,000 full time employees worldwide. More than 80% of **HF** personnel in Canada are exclusively engaged in selling travel products (flights, hotels, car rentals, tours, etc.).

At **HF**, employees are grouped by Office; each Office is staffed with 5 to 8 employees on average (some outliers, e.g. the HF ecommerce office, have 15 full time employees (FTEs). Of the 180+ Offices reviewed, 17 Offices are Corporate (i.e. selling mostly to pre-established business clients via phone and email) and 160+ offices are Retail (i.e. selling mostly to walk-in customers). Corporate teams are assembled into three shared facilities across Canada. Retail teams are located in individual distinct street level (160+) storefronts across Canada. Several Offices were excluded from the sample. The excluded Offices sell wholesale cruise travel and primarily serve as an intermediary with no client contact. Because of the different business models, the cruise offices were removed from the sample.

Ultimately, each member of an Office attempts to maximize individual sales while maintaining a strong margin. **HF** management measures Office performance based on Total Sales Volume and Gross Average Margin. The management's goal is to maximize the performance of every group.

For primary data, this research relied on an HF-provided database containing all **HF** emails sent or received for 2011. More than 7 million email records (To, From, Date, Time) were reviewed, grouped and organized. Any email deemed external (i.e. having a To, From or CC field that lists a non-HF email address) were excluded. Only intragroup emails (e.g. amongst members of the same Office) were examined.

Using an extension to the NodeXL® software package created specifically for this research by the Social Media Foundation and Microsoft, the researcher took an 'x-ray snapshot' of each group's social network. This snapshot is based on underlying data representing a series of social network topology measures, including Tie Strength, Structural Holes and Centrality. To represent the relational dimension of the Social Network, Tie Strength was measured as an indicator of the overall Tie Strength in the group. Centrality and Density (the inverse of Structural

Holes) were measured to represent the structural dimension of the Social Network. Three distinct types of Centrality were measured: Eigenvector, Closeness and Betweenness. A group measure for each was calculated by averaging the individual team members' centrality measures. Initially, ***homophily*** was to be measured to represent the cognitive dimension of the Social Network. Unfortunately, **HF** management instructed the researcher not to undertake any contextual analysis; as this is necessary to measure homophily, homophily measures were abandoned. Alternative measures of the cognitive dimension (e.g. shared background) require HR data for personnel. Access to such data was also denied to the researcher by HF management. For this reason, this research must abandon examining the cognitive dimension. Privacy and confidentiality were cited as the reasons for the denial of access.

Tie Strength, Structural Holes, and Centrality were then correlated, through multiple variant regression and generalized linear modelling, with group performance (Normalized Sales Volume, Gross Margin Average) in order to determine if high-performing groups share similar network measures. Multiple variant regression was used to analyse the findings and to determine how well the model generated deviates from the data provided. In addition to using MVR: GLM techniques, the resulting social graphs were also reviewed using a form of visual analysis.

A summary of variables follows:

Table 2: Social Network Topology Variables

Name	Independent/ Dependent	Acronym
Normalized Revenue	Dependent	nrev
Structural Holes	Independent	holes
Tie Strength	Independent	ties
Average Eigenvector ctr	Independent	eig
Average Closeness ctr	Independent	clos

Average Betweenness ctr	Independent	Betw
Full time equivalents	Independent	Fte
Office Type	Independent (0=corp, 1=retail)	type

To measure onboarding, this research looked to identify and record the moment when a strong tie is formed. Strong Ties were said to exist when two actors exchanged their tenth email. While this number may seem low, one most consider that the colocation of most employees (e.g. proximately amongst officemates) is extremely high, and email use may be limited. Further, one must remember that this number does not reflect the quantity of email sent, only the quantity of intra-organizational (e.g. between office mates) email sent. To determine where to set the bar, the researcher examined the frequency of email distribution, looking to set that bar at a level that would ideally encompass a meaningful set of relationships was the goal. The NodeXl® software package facilitated this visually and the tenth email was selected as a meaningful level.

The earlier that the tenth email was sent, the longer the strong tie bond was in place. An actor who formed a strong tie months before another actor would theoretically benefit from having such a strong tie in place longer. Therefore, this research attempts to explore the correlation between individual performance and the number of days that strong ties were in place.

1.5 Findings

A model predicting performance was generated at the conclusion of this research. This model was then tested to see what extent the model's outputs could be validated by actual data. This model demonstrates the relationships between performance, as represented by normalized sales, and measures of both the structural dimensions and the relational dimension of Social Capital. An inverse curvilinear relationship was found to correlate normalized sales to the number of strong ties present in the network. This coincides with Lechner, Frankenberger and Floyd's findings (2010). An inverse curvilinear relationship was also found to correlate normalized sales to group eigenvector centrality. This, too, furthers Lechner, Frankenberger and Floyd's (2010) theory of the Dark Side of Social Capital (see Chapter 2 for additional discussion), which basically confirms what mothers for time immemorial have taught their children, that 'too much of a good thing can be a bad thing'. For the 180+ offices that comprise the HF national travel agency, top performing firms were found to have high similar intra-team individual eigenvector centrality, few structural holes and strong ties. Similarly, for the 180+ offices that comprise the HF national travel agency, low performing firms were found to have a large number of structural holes, few strong ties and uneven individual eigenvector centrality scores.

Based on limited data, a possible curvilinear relationship was also found to exist between individual performance and onboarding speed (i.e. the speed at which a new agent builds their first strong intragroup ties), but with so few points available it would be inappropriate to rely on these findings.

Table 3: Social Network Topology Findings

	Findings
GLM Analysis	Performance is best measured by Nrev alone. The relational social network measures (log total strong edges and Tie Strength) were highly correlated with performance. Performance had a U relation with log total strong edges. Performance had an inverse U relation with Tie Strength.
Visual Analysis	High performing teams demonstrated mostly of strong tie relationships. High performing teams demonstrated high average eigenvector centrality. High performing teams demonstrated only a few structural holes. Low performing teams seem to lack the ties necessary to facilitate tacit knowledge transfer.

	Low performing teams have strong central actors. Low performing teams have a large number of structural holes.
Onboarding Analysis	A positive relationship between onboarding speed and individual performance was found, although such was based on an extremely limited dataset.

1.6 Strengths of Research

Prior research has been limited mostly to establishing the social network (and related SNA measures) through survey techniques. The possible weaknesses of survey methodology include the recency bias. Further, surveys can be inflexible in that they require the initial study design (the tool and administration of the tool) to remain unchanged throughout the data collection. Further still, with a survey approach the researcher would have to ensure that a large number of the selected sample replied; otherwise the survey's validity may be questioned.

Previous research into Facebook and other social network services, such as Friendster and MySpace, has also been performed using surveys (e.g. Boyd, 2007; Ellison et al., 2006; Stutzman, 2006). While these methods (survey and interview) provide a deep understanding of what individuals are doing and their motivations for doing so, they do not capture large-scale patterns or temporal rhythms exhibited by the collective action of immense numbers of users (Golder et al., 2006). For that one needs to deploy large scale data mining techniques.

Instead of interview or survey, this research uses actual email data to build the social network through communication patterns. The method used herein requires only the consent of management to garner not the perceived social links, but the actual ties created when Actor A speaks with Actor B. Thus, email data observed will tend to be more accurate, as it is based on actual, not perceived, communications in network. In addition, using email data instead of survey data is exponentially faster. In this research, 7 million emails were reviewed, representing all email communications for the 1800+ actors in the HF network. SNA data was generated within a few hours of data exploration. If surveys had been deployed at **HF**, it might have taken weeks, if not months, to gather the data.

Finally, this new method can generate longitudinal information. Survey data is gathered at one moment in time and requires the participants to reflect backwards upon their network. Using email data allows for the generation and a dynamic review of the final SNA data. In the case at hand, survey data would show which new employees had successfully built strong ties as of December 31, 2011. Email data showing the dynamic longitudinal formation of such ties thus offers the researcher a more accurate view of the process of network formation, not simply its output.

1.7 Limitations of Research

The limitations of this research are explored and addressed in greater detail in Chapter 7. In sum, the author acknowledges the following limitations. The researcher focused on only one national firm, in one industry (Travel Services), in one country (Canada). This will limit the generalizability of any findings. The research acknowledges that not all conversations are conducted by email; hence, some network activities (i.e. phone calls and face to face conversations) cannot be traced through this methodology. In fact, there is some evidence that key conversations (especially concerning tacit knowledge) are rarely exchanged over email (Grippa, Zilli, Laubacher and Gloor, 2006). Notwithstanding, there is much more evidence to suggest that email is an effective proxy (Tyler, Wilkinson and Huberman, 2005; Wellman, 2002; Whittaker and Sidner, 1996). The above limitation has a particular impact in the onboarding data, where some high-performing actors showed no strong ties. An alternate explanation, that these actors have strong ties but do not use email, may skew the validity of the findings relating to onboarding speed.

Summary of Chapter 1

This research focuses on how social network characteristics impact performance by driving or hindering knowledge flow. This research assumes some social network topologies are better suited for firms requiring exploration and innovation, while others are more appropriate exploitation and efficiency. This idea was touched upon by Lechner, Frankenberger and Floyd (2010) as they explored the concept of Task Contingencies. Recent findings show that knowledge transfer mediates between an organization member's intra-organizational social capital and organizational performance outcomes of growth and innovation performance (Maurer, Bartsch and Ebers, 2011). This research contributes to such dialogues while attempting to extend existing views on the correlation between group performance and group social network topology and seeking to empirically confirm that social network topology is a strong predictive measure of group performance. This research also attempts to build a critical distinction between social networks and social capital. Finally, this research examines possible correlations between onboarding speed and individual performance.

In this chapter, the researcher provides an overview of this research project. Chapter 2 provides a review of the extant literature as well as a description of the phenomenon and what has been said about it to date. In Chapter 3, the researcher explores the methodologies and methods available and justifies the choice of methods. Chapter 4 focuses on operationalizing the research by determining which measures to collect and how best to collect them. Chapter 4 also sets out the research questions to be tested. Chapter 5 outlines how data was collected, cleaned and transformed prior to analysis. Analysis of the data is covered in Chapter 6. Chapter 7 lists the findings as well as the limitations of this research, ending with a discussion of the contributions made by this research and possible areas to explore in the future.

Chapter 2 — Literature Review and Key Concepts in the Field

Chapter 1 provided an introduction and some background on the general foundations of this research. Next, the researcher explores the literature, both to understand the current landscape and to identify any possible knowledge gaps that this research may be able to address.

Background to Research

Adam Smith (1723-1790) is the father of modern economics. Smithian economics is built around the tenet that for groups (e.g. Firms, Organizations, etc.) to be competitive they must create value (Van de Ven, 1986). In Smithian economics, a group's ultimate aim is to identify and exploit economic opportunities (Van de Ven, 1986). Building upon Smith's work, scholars over the next two centuries determined that, for groups to sustain performance and create value over the long term, they must develop new products and new services in order to pursue new markets and to adapt to new market demands (Brown and Eisenhardt, 1995; Burgelman, 1991; Damanpour, 1992). However, the process of sustaining performance by adapting to market demands is not simple. Such a process is steeped in various social mechanisms by which groups combine and exchange resources as a means of creating value (Moran and Ghoshal, 1996; Nahapiet and Ghoshal, 1998). The rise of the importance of such social mechanisms has reflected a shift not only in market demands but in the very organization of groups themselves.

2.1 The Rise of KBV

Business strategy developed in the 1960s, dominated by economics, focused mainly on firm positioning as the main source of competitiveness (Porter, 1985). These concepts were challenged by the rise of the Resource Based View ('RBV') of the firm. RBV (which was popularised by Barney in 1991) appears earlier in the literature, thanks to both Penrose (1995) and Wernerfelt (1995). Under RBV, a firm (e.g. group, organization) garners competitive advantage through efficient use of resources (Barney, 1991; Teece, Pissano, and Shuen, 1997; Wernerfelt, 1984).

More recently, the paradigm has shifted away from RBV to one in which knowledge is seen as the most important internal resource and the primary source of competitive advantage (Knowledge Based View, 'KBV'). The knowledge-based view of the group argues that knowledge is the resource most necessary for pursuing economic opportunities (Barney, 1991). These shifting market views have rendered services (e.g. financial services) the dominant form of economic employment in the world; for example, service-oriented employment now accounts for 42% of all jobs worldwide (ILO, 2009). In western economies, services play an even more dominating role, accounting for 79.6% of all jobs in the US, 74.5% in the UK, and 69.6% in Canada, according to recent figures (CIA, 2008). The concept of service innovation is rooted squarely in KBV.

While seeking to explain the world of service firms, service dominant logic argues that the services are fundamentally concerned with the application of one critical resource: knowledge (Lusch, 2006; Lusch, Vargo and Tanniru, 2009). This signals the emergence of an economic system based around knowledge and the dominance of a knowledge-based economy in western societies. Consequently, knowledge is now seen as the most important asset among a group's resources for developing new applications and pursuing market opportunities (Moorthy and Polley, 2010; Nonaka and von Krogh, 2009; Zander and Kogut, 1995).

As a result of these paradigm shifts, knowledge creation has come to be viewed as the dominant source of modern competitive advantage (Lyles and Salk, 1996; Moorthy and Polley, 2010; Nonaka, 1994; Nonaka and von Krogh, 2009; Tsai, 2001; Zaheer and Bell, 2005). However, it is important to note that the capacity for knowledge creation is not held solely by individuals (nor is it held within the group itself); but, rather, such capacity resides in the social relationships embedded within the group's network (Burt, 1992). With social relationships as the nexus of knowledge creation, one's social network now represents the most vital of group resources. Thus, it is critical to study the transfer of knowledge (particularly along those social relationships) with respect to the performance of the group in order to gain important insights into the social aspects of organizational design and their impact upon performance. In other words, if one wants to leverage knowledge more effectively, one first needs to understand how knowledge flows across a group's social network.

Under KBV, a firm gains advantage through the purposeful dissemination and creation of knowledge across the organization (Bou-Llusar and Segarra-Cipres, 2006; Grant, 1997). The KBV focuses on the internal relationships of the group, basing its model of group effectiveness more heavily on the knowledge network of a firm rather than on the firm's formal structure. Through this paradigm shift, from RBV to KBV, the challenges of adapting to today's dynamic markets falls on the network and its ability to exchange knowledge in pursuit of economic opportunities. This has not only impacted the dynamics of the organization, but has also become itself the dominant function that the majority of groups perform (services, not goods).

Recent literature confirms these theories, but only for a finite subset of strategic initiatives. Lechner, Frankenberger and Floyd (2010) found curvilinear relationships between several SNA dimensions and group performance. In their concluding their study, they outline some potential future research:

> The intra-organizational social environment exerts significant selection pressures on strategic initiatives. So far, this theoretical proposition has prompted relatively little empirical research on the role of social networks in the success of such groups. The results here confirm the importance of intergroup relations and show that their influence on initiatives' performance is multidimensional and curvilinear. The researchers hope other studies will continue to refine understanding of how networks affect the development of strategic initiatives. (p. 885)

In the same article, the authors Lechner, Frankenberger and Floyd (2010) identify a gap in the literature. This research seeks to address that gap. Lechner, Frankenberger and Floyd (2010) advocate:

> [T]he need to refine [our] understanding of how [group] network relations contribute to group performance. (p. 867)

Further, the following statement from that AOM article informs this research:

> [P]rior work has shown that network features combine to create particular configurations that foster actor performance. (p. 867)

This research then sets out to explore the following foundational question: *Given a set of identical tasks, performed by all groups, is there an optimal social* *network topology?* Is there a certain social network formation (e.g. hub, star-shaped, flat) that will generate optimal performance?

2.2 Theoretical Foundations of the Research

This research seeks to understand how social capital (in particular, social network topology) correlates with performance. The idea that tacit knowledge (i.e. that which cannot be taught explicitly, e.g. how to price travel) is the key to success in the twenty-first century, and that tacit knowledge flows only over strong tie networks, will be discussed. Since it is posited that this tacit knowledge (e.g. best practices for billing) allows for increases in performance, one must first understand the basic concepts at the heart of this research.

2.3 The Firm

Much research to date has focused on the firm as the unit of analysis, that is, a unit of production acting in cooperation towards a common goal (Davis, 1941). At **HF**, each office shares some infrastructure (e.g. IT, HR, Senior Management, etc.), but each office is responsible for generating its own gross sales and gross margin. It is helpful to imagine each Office as interrelated but autonomous. Based on this, each **HF** Office can be seen as a 'firm' under the definition discussed above.

2.3.1 Evolution of the Firm: Theoretical Underpinnings

The aim of a firm is to perform and to fulfil its specified functions (Burgin, 1969). Given this imperative, a firm is forced to utilize available resources in pursuit of economic opportunities; thus, its ability to do so is of vital importance. First, however, the researcher must ask, what is a firm? It is important to understand modern organizational forms, particularly the rise of the organizational paradigm, which sees a group as a network of resources. Moreover, it is important to understand the significance of such networks and their effects on group performance. It is also critical to understand the knowledge processes that take place through a network, because networks ultimately affect the performance of the group. This discussion will begin with a review of 'group' as a management concept.

The dominant approach to strategic management in the last fifteen years has stemmed from the early work of Edith Penrose's book, *The theory of growth of the firm* (1995), which was originally published in 1959. Penrose expanded the foundation for the Resource Based View (RBV) of a group, first popularized by Wernerfelt (1984). Barney (1991) states, in his seminal work, that a group garners competitive advantage by leveraging its own internal resources. Barney has been involved in an on-going debate with Priem and Butler (2001) as to the practical implications of RBV described in his 1991 article. Priem and Butler (2001) argue that the majority of management theories in fact are simply concepts, just short of theories, and thus are lacking in practical relevance. They argue that RBV is an example of yet another management theory which lacks grounding. They argue that RBV is not sufficiently refined to have any practical application. This position was challenged by Barney, who contends that the RBV has all the necessary empirical underpinnings (including some managerial implications) to be considered a legitimate theory. Priem and Butler (2001) counter Barney, claiming that regardless of any debate as to what is considered a theory, RBV is useful but requires a more refined definition. Notwithstanding this debate, both Priem and Butler (2001) and Barney (1991) agree that the value garnered from group resources is dependent on factors outside the Resource Based View.

Fundamentally, the advantage garnered from group resources is not only dependent on those resources but on the group's ability to exploit such (Nelson and Winter, 1982). Thus, even under the RBV of the firm, knowledge is seen as a key resource from which groups garner value. This shift from a Resources Based View (where resources are largely perceived as tangible, material goods) to a Knowledge Based View (where knowledge is seen as the most valuable resource) represents the most significant shift in management thinking since the early twentieth century (Grant, 1997). This new KBV paradigm was built on the earlier work of Zander, Nonaka, Hedlund, Von Krogh, Roos and Spender. Reflecting on Barney's (1986) early work, Grant (1996) suggests that a group is not solely the sum of its resources, but instead it is the transferability of those resources that affects a group's ability to confer a sustainable competitive advantage. Grant (1996) argues that the transferability of resources is particularly important when the resource in question is knowledge and cites the epistemological distinction that is dominant in knowledge management literature—i.e. the difference between *knowing what* and *knowing how*. Grant (1996) argues that, within the KBV paradigm, knowledge assets in themselves do not garner competitive advantage but, rather, as Barney (1986) suggests, the ability to transfer knowledge as needed leads to a competitive advantage.

The shift from RBV to KBV is not a trite distinction based on semantics. Drucker (1988) argues that this paradigm shift (e.g. from RBV to KBV) would change the very nature of organizational forms and the assets those organizations hold dear. Drucker's view is reinforced by the parallel shift in management literature from Goods Dominant Logic (where the primary goal of a group is perceived as the production of a 'good', such as an automobile) to Services Dominant Logic (where the primary goal of a group is perceived as a service generating 'value' from the good; for example, transportation from point A to point B) (Lusch, 2006; Vargo and Lusch, 2008).

2.4 The Firm as a Network

A network is a set of actors connected by a set of ties (Borgatti and Foster, 2003). The actors, often called 'nodes', can be persons, teams, organizations, concepts, etc. According to Borgatti and Foster (2003), ties (i.e. vertices) connect pairs of actors and can be directed (i.e. potentially one-directional, as in giving advice to someone) or undirected (i.e. physically proximate) and can be dichotomous (present or absent, as in whether two people are friends or not) or valued (measured on a scale, as in strength of friendship).

The concept of Networked Organizations (and Organizational Networks) gained popularity in the late 1980s and early 1990s (Borgatti and Foster, 2003). Scholars used the network as an organizational form to describe an entity formed by repetitive exchanges amongst semi-autonomous actors (groups, people) that rely on embedded social relationships (Bradach and Eccles, 1989; Eccles, 1981; Jarillo, 1988; Powell, 1990). According to Borgatti and Foster (2003), the network form of the organization emerged to balance the flexibility of the market organizational form with the predictability of traditional hierarchal organizations forms. The network form of organization features: a flat hierarchy; empowered semi-autonomous workers; and lateral communication paths in knowledge-based industries (Borgatti and Foster, 2003). Further, networks can be seen as defining the actors' environment, creating both opportunities and constraints (Borgatti and Foster, 2003). Drucker (1988) foreshadows the development of such networked organizational forms 20 years ago, while a recent white paper from MIT (Malone, Laubacher and Dellarocas, 2009) argues that the emergent organizational forms create knowledge in concert with each other through their social networks. Consequently, this realm is of great interest to management researchers and has vast implications for modern groups looking to gain competitive advantage from knowledge.

It is important to understand that the social network may prove more important than the formal organizational chart in a discussion of the manner in which an organization impacts its internal flow of knowledge between individuals. A large body of work exists describing the network dynamics of groups and their implications on group performance (Cross, Prusak, Parker and Borgatti, 2001; Cross, Borgatti and Parker, 2002). This literature generally argues that network centrality is related to group performance (i.e. centrality in this sense reflects the position of a group with respect to other groups). Similarly, the literature suggests that individuals who show higher centrality, defined as having a higher number of network ties (Wasserman and Faust, 1994), perform better than those with less centrality. The former is at the macro level (group to group within an industry). The latter is at the micro level (individual to individual within a group).This research focuses on the mezzo (Firm) level and asks at an intragroup level: How do measures (including network centrality) affect overall group performance? This research builds on the discussion of networks and attempts to explore tie strength, tie formation, centrality, position (structural holes) and contribution to overall performance (Burt, 1992).

The literature on formal organizational structures and their implications for group performance is longstanding (Drucker, 1988). Early studies by Pearce and David (1983) and Van de Ven, Delbecq and Koenig (1976) argue that organizational design invariably impacts the overall performance of the group. According to these authors, it is important to consider the dimensions by which organizational design is defined. Mintzberg (1980) synthesized organizational forms into five general models (i.e. Simple Structure, Machine Bureaucracy, Professional Bureaucracy, Divisionalised Form and Adhocracy), which he claims apply without exception to all possible organisational forms. While the discussion following from Mintzberg's work focuses on formal organizational forms, the models themselves suggest that there are underlying social organizational constructs in each organizational form (i.e. groups consist of more than what appears on their organizational chart). In fact, one can see the evolution from the simple structure that was prominent in Smith's day to the modern adhocracy which has defined the post-internet world. However, when referring to a network one is not simply referring to the group's organizational form as an adhocracy. Instead, the term 'network' refers to the larger web of resources and relationships which exist parallel to, but are often separate from, the formal hierarchy.

While the above discussion only scratches the surface of organizational design, it does introduce the fundamental shift in organizational structure which has accompanied the shift towards service businesses. This shift towards services and the more adhocratic organizational structures which came with it have made social networks a more accurate representation of structures within twenty-first century organizations. Thus, the social dimension of networks and the factors associated with interpersonal relationships define the organization and the design of modern groups, and thus are vitally important for modern management research.

The social network dimension of organizational design reflects the knowledge and social resources embedded within personal relationships (Burt, 1992). Podolny and Page (1998) argue that all organizations have their own network form. This form is reflected by social ties rather than the formal structure of the organization. This author strongly concurs with this position. Bienenstock and Bonacich (2003) articulate four organizational network forms (random, scale-free, lattice and bipartite) used to describe networks. While these represent the general shape networks can take, they fail to describe the network in a true organizational sense. More importantly, these forms fail to articulate the substantial shift in organizational design which has brought networks to dominance in modern groups.

2.5 KBV model versus RBV model - Network Paradigms

The Knowledge Based View of the group preceded that of the theory of Dynamic Capabilities and focused attention towards internal group factors rather than environmental factors as the source of a firm's success (Barney, 1991). However, while the Dynamic Capability perspective builds on the importance of knowledge and its adaptation, it is important to first understand the resource-based view, to understand why knowledge and its flow are central to the performance of the group.

KBV theory views knowledge assets as the dominant source of competitive advantage for groups. As discussed earlier, this suggests that the resource central to businesses is not a tangible good but rather knowledge (or, more specifically, a tacit skill which is possessed). Similarly, the tangible assets possessed by a company are not the fundamental source of that company's value proposition; instead it is their ability to utilize those goods in providing value. Comparatively, the Service Dominant (S-D) view proposed by Vargo and Lusch (2004) suggests that tangible assets are knowledge-enabled and do not create value on their own. As such, the assets that are most valued in the service dominant perspective are those specialized skills and competences (tacit knowledge which is internal to the group) which are leveraged in creating value. However, this raises the issue of value creation in a service dominant perspective, which is not simply the value prescribed by the producers (embedded in products) but is co-created with the consumers (based on how much value the consumers derive from it). Given that under the S-D paradigm, value is co-created between separate actors, it is hardly surprising that several scholars (Achrol and Kotler, 2006; Grönroos, 2006; Gummesson, 2006) have likened this interaction to that of network nodes, suggesting that networks play a central role in value creation and exchange under S-D logic.

2.6 Social Capital?

Social capital is a sociological concept which deals with the connections within and between social networks. In *The forms of capital,* Pierre Bourdieu (1983) defines social capital as 'the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition'. Coleman (1990), who was the first to subject the concept to empirical analysis and develop ways of operationalizing it for research purposes, defined social capital as follows:

> Social capital is defined by its function; it is not a single entity but a variety of different entities having characteristics in common: they all consist of some aspect of a social structure, and they facilitate certain actions of individuals who are within the structure (p. 302).

2.7 Social Capital Analysis

Social Network analysis (SNA) studies the patterns of relations amongst individual actors (Wasserman and Faust, 1994). SNA assumes the structure of interacting units (groups in an industry, groups in a group, actors in a group) can lend insights into the nature of these relationships (Farrall, 2004). At the heart of Social Network Analysis is the theory of Social Capital. At the heart of the Theory of Social Capital is the notion of the value of connections (Borgatti and Foster, 2003).

2.8 Group Level Social Capital

Putnam (2000) defines a firm's social capital in terms of broad cross-cutting interconnections among all firm members. The social capital theory emphasizes the possibilities for action that social ties provide the individual, firm or group (Borgatti and Foster, 2003). Social capital studies seek to explain variation in success (i.e. performance or reward) as a function of social ties (Borgatti and Foster, 2003).

In general, more social capital resources lead to greater effectiveness (Guzzo and Shea, 1992; Hackman, 1980). At the firm level, effectiveness is measured by such standards as satisfying external client needs, reaching agreed-upon goals, and being able to come together at some future point to do more work if needed.

It is expected that greater firm social capital resources will make it easier for members' goals and needs to be met; will make the firm more likely to want to come together again in the future; and, ultimately, will allow the firm to reach its goal more easily and with better results (Oh, Labianca and Chung, 2006). Firm social capital is the overall balance of relationships that leads to the maximum flow of group social capital resources (Oh, Labianca and Chung, 2006).

2.9 Dimensions of Social Capital

Like social networks themselves, in order to investigate the concept of social capital one needs a model to operationalize. The three-dimensional model, developed first in organization theory, is one such valuable approach. Social capital can be divided into structural, cognitive, and relational dimensions (Hazleton and Kennan, 2000; Nahapiet and S. Ghoshal, 1998; Tsai and Ghoshal, 1998). These three dimensions identically mirror the dimensions of Social Networks: Structural, Cognitive and Relational (Sheriff, 2012). The structural dimension refers to the information channels that connect individuals and units (Tsai and Ghoshal, 1998). The relational dimension of social capital refers to resources embedded in relationships, e.g. the trust that members develop through intense social interactions (Sheriff, 2012). The relational dimension of social capital is critical for exploitation of knowledge (Tsai and Ghoshal, 1998; Zahra and George, 2002) because trust engenders knowledge transfer (Putnam and Borko, 1997) as opposed to simple information exchange. When actors in a network trust each other, they are often found to be willing to spend the time necessary to ensure that information exchanged is comprehended and can be fully exploited (Sheriff, 2012). The cognitive dimension of social capital refers to the shared meaning and shared understanding that develops among members of the network as they socially interact (Sheriff, 2012). A firm high in the cognitive dimension will often have easier time transferring tacit knowledge, since by definition the actors in the network share mental frameworks and common understandings. All three dimensions are important for the acquisition, comprehension and exploitation of knowledge (Nonaka, 1994).

2.10 The Girders vs. Pipes Debate

In 2003, Borgatti and Foster distinguished between two broad categories of social network theory: Topology and Flow (Farrall, 2004). This debate is often labelled the Girders vs. Pipes debate. Structionalists (girders) hold that an actor's position determines the outcome. Constructionists (pipes) consider that it is the transmission of resources (e.g. the flow of knowledge) along pre-existing social ties that dictates outcome. The girder perspective focuses on social capital while the pipes perspective focuses on the flow of social assets like knowledge over or through the social network. Topology Structionalists discount the actual content of ties while focusing on overall patterns of association (Farrall, 2004), whereas social theorists describe the network structure of social capital on the girders side. Alternatively, flow mechanisms consider network ties as explicit conduits for the flow of social goods. Rogers' (1962) Diffusion of Innovation theory adheres to the pipes perspective, whereas the girders paradigm is best aligned with social capital concepts (e.g. social capital makes up the girders that connect actors). The pipes paradigm is best aligned with social networks (e.g. social networks are the pipes over which knowledge flows).

Perhaps it is neither girders nor pipes alone that provide the optimal theory; perhaps in fact it is both. Much like the Wave-Particle Duality Theory of Light (Greiner, 2001), where light shares attributes of both wave and particle, perhaps a network can simultaneously act both as a girder and as a pipe depending on how the phenomenon is viewed. This is part of the concept of *complementarity*, which says that a phenomenon can be viewed in one way or in another, but not both simultaneously (Chen and Klahr, 1999; Green and Murray, 1989).

Social Capital vs. Social Network; the above debate sheds light on why some research describes the phenomenon as social capital (the strong girders which build within the network) while others describe it as the social network (the pipes over which the knowledge can flow). This researcher prefers the *complementarity* view, that social network forms the pipes over which social capital (a subset of which is tacit knowledge) flows, but equally accepts the duality theory of networks.

2.11 The Dark Side of Social Capital

Lechner, Frankenberger and Floyd (2010) looked at the effect of social network topography on performance. Up until then, most research suggested a strictly positive correlation between social capital measures and performance (e.g. more ties are better). However Lechner, Frankenberger and Floyd theorize that too much social capital could have negative implications. Lechner, Frankenberger and Floyd's (2010) article succinctly summarizes what they call the 'dark side of social capital'.

Whereas most research (Argote, Mcevily, and Reagans (2003) and Burakova-Lorgnier, Bouzdine-Chameeva, and MacGilChrist (2004)) emphasizes the benefits of increased social capital (better relationships, higher trust, more absorptive capacity) Lechner, Frankenberger and Floyd posit that too much social capital could generate negative consequences. Consider the following thought experiment: imagine a group of employees with no Social Capital. Employees would seldom converse, let alone work together. Now imagine a group with too much Social Capital, employees who would chat all day and invest in their relationships at the group, but not actually do much work. Similarly, imagine a social network involving too little trust. Then imagine a social network with too much trust. One posits that, similarly, any benefits from trust are negative at both extremes. Too little trust may undermine knowledge transfer by undermining the confidence of the receiver or dis-incentivising the sender. Too much trust may lead to 'group think'. Lechner, Frankenberger and Floyd (2010) point to this limitation (negative at both extremes) on each of the dimensions of Social Capital, proposing that performance would be optimal at neither end of the spectrum (e.g. neither too much nor too little) but, rather, at some point between the two extremes.

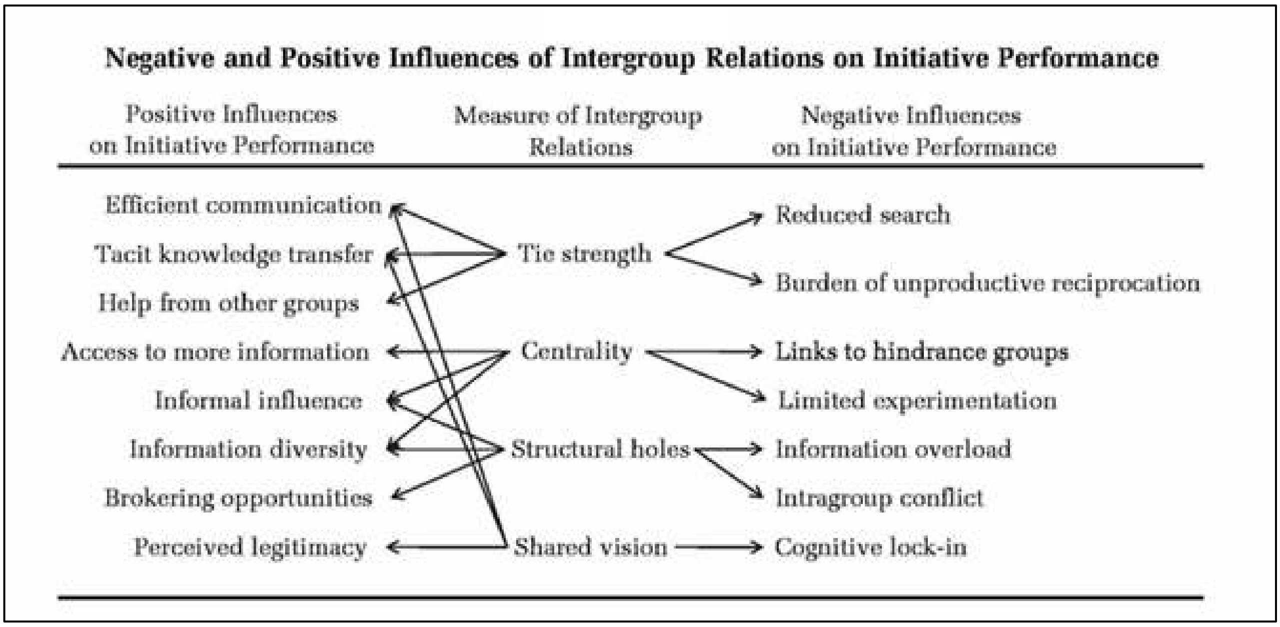


Figure 3: Negative and Positive Influences of Intergroup Relations on Initiative Performances (from Lechner, Frankenberger and Floyd 2010)

The Dark Side of Social Capital, as framed by Lechner, Frankenberger and Floyd (2010), was found consistent in their empirical findings, which generally demonstrated an inverse curvilinear relationship (inverted U) between network dimensions and group performance. Unfortunately, their evidence only found such to be true for two of the three operationalized network dimensions (e.g. Shared Trust, a measure of the cognitive dimension, was found to be linear). This was explained by the authors as simply being the left side of the inverse curvilinear curve of the cognitive dimension. This research extends the findings of Lechner, Frankenberger and Floyd (2010).

2.12 Task Contingency and Social Capital

Some organization network structures that yield positive Social Capital in some task situations convey social liability in other situations. Imagine a network of creative designers, who form a network structure that is optimal for creative tasks undertaken by the network; but such a topology might not be ideally suited to group tasks requiring efficient execution of millions of repetitive routine tasks. This is known as the Theory of Task Contingency. William Richard Scott (1981, p. 114) describes contingency theory in the following manner: ‘The best way to organize depends on the nature of the environment to which the organization must relate’. In *Images of organization*, Gareth Morgan (1997) describes the main ideas underlying the Theory of Task Contingency in a nutshell:

1. Organizations are open systems that need careful management to satisfy and balance internal needs and to adapt to environmental circumstances.

2. There is no one best way of organizing. The appropriate form depends on the kind of task or environment at hand.

3. Management must be concerned, above all else, with achieving alignment and good fit.

4. Different types of organizations are needed in different types of environments.

Research has polarized task contingency around the degree of task exploration. Prior research shows that exploratory teams complete their projects more quickly if they have a social network structure composed of many strong external ties that are non-redundant (Gabbay and Pitts, 2002). In contrast, teams pursuing tasks that exploit existing expertise will take longer to complete the same tasks if they have this type of social network structure, mainly because external ties must be maintained, even though they are not needed for the task. This research proposes that organization network theories of tie strength and structural holes should to be broadened to reflect the effects of task differences, network costs, and difficulties in getting others to help.

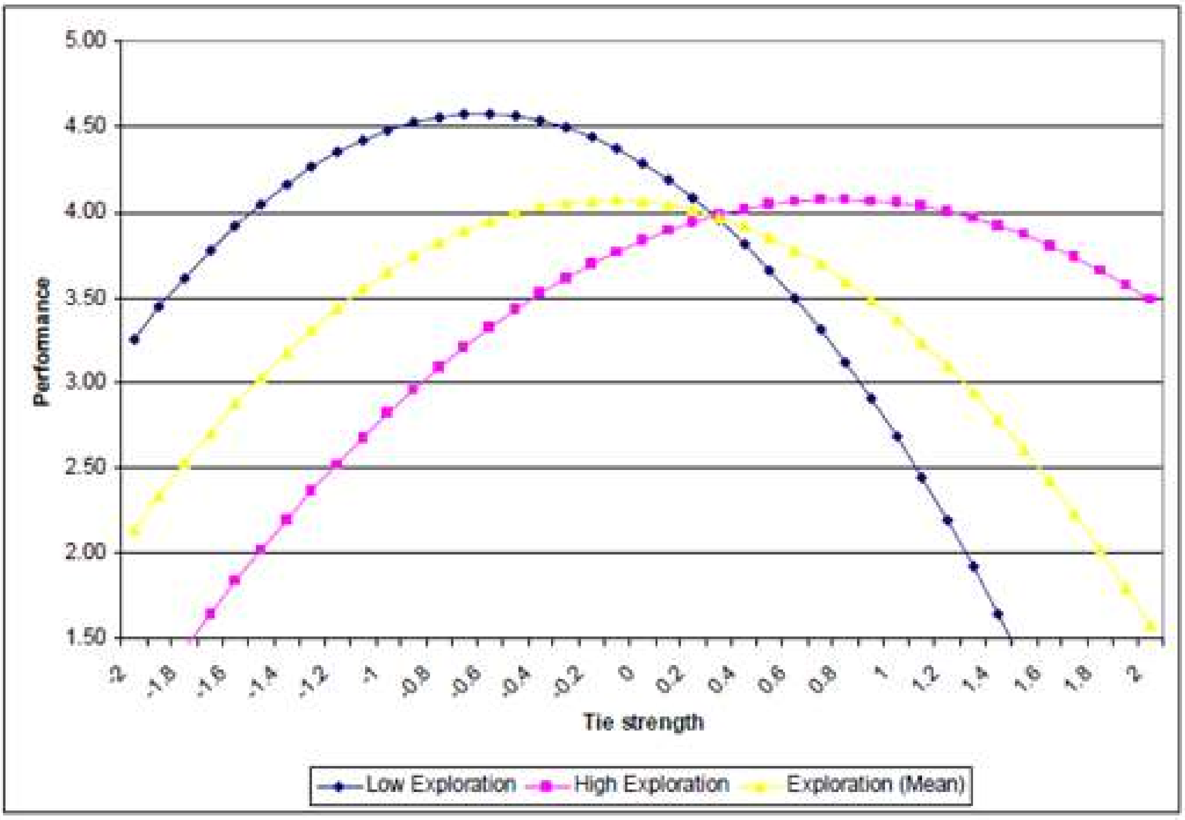


Figure 4: Performance vs. Tie Strength (from Lechner, Frankenberger and Floyd, 2010)

A firm's social network can only be optimized by first examining the level of exploration vs. exploitation that the firm undertakes. Some firms (e.g. an innovative food company) focus on the creation of new knowledge (products, services); these firms can be seen to be primarily explorative. Some firms (e.g. a law firm) focus on efficiencies for competitive advantage, often by taking an innovation and exploiting it for efficiency through diffusion. These firms can be seen to be primarily exploitive. In reality, no firm is strictly exploitive or strictly explorative; all firms undertake some tasks that could be categorized as either. All firms are a mix of explorative vs. exploitive (e.g. even law firms have to create new knowledge, and even innovative food companies must find efficiencies).

Prior research has indicated that exploration moderates relationships between performance and all three dimensions of intergroup social networks (Lechner, Frankenberger and Floyd, 2010). Negative consequences of strong ties and centrality are more pronounced in exploratory initiatives than in 'exploitive' initiatives. Taken to the n^{th} degree, an explorative firm could seek to have a large number of structural holes and a greater diversity in the actor's cognitive background, while an exploitive firm might seek to be less cognitively diverse with stronger and often redundant ties. Seen on a spectrum, three firms (one with a high level of exploration, one with a mean exploration level and one with a low exploration level) will have their optimal points skewed, as can be seen in the following correlation between Performance and Tie Strength.

Figure 4 (from Lechner, Frankenberger and Floyd, 2010) graphs performance vs. tie strength, which is a proxy for the relational dimension. In the graph, three types of initiatives, each with a high, medium and low level of Task Contingency, are examined. In each case, the graph shows an inverse curvilinear relationship. Or, in plain English, too many strong ties have as much negative impact as too few strong ties, regardless of task contingency. Task contingency simply shifts the point of optimal returns.

What is interesting is that performance is optimized in the middle group despite the level of exploration. Also interesting is the fact that the more exploratory the group (i.e. a group with a greater need for innovation), the more tie strength is needed for optimal performance. Firms with a low level of exploration need fewer and less strong ties than groups whose focus is on innovation. This can be explained by the fact that exploratory firms require more ties to more diverse sources of knowledge to drive innovation, while firms focused on exploitation need to defuse innovation more than they need to facilitate its generation.

2.13 Performance

Performance can be defined as *the degree to which any expectation is fulfilled* (Selnes, 1998; Venkatraman and Ramanujam, 1986). In an organizational sense, the expectations to be met are tacitly understood by the shareholders, who appoint the management of that company to fulfil those organizational aims. If the shareholders of an organization are simply seen as investors who buy shares with the expectation that they will rise in value, then their sole interest is financial performance. However, while financial performance may represent the ultimate aim of an organization, it is important that group performance can only be sustained if real competitive advantage is consistently developed. Competitive Advantage occurs when an organization acquires or develops an attribute, or combination of attributes, that allows it to outperform its competitors (Porter, 1985). Thus competitive advantage, when harnessed, leads to performance gains.

Inside each HF Office, a travel agent (actor) fields incoming requests for travel bookings (e.g. flights, hotels, car rental) from either Retail (consumer) clients or Corporate (business) clients. For each request, an agent checks the internal cost of the service being sold, then decides on the gross margin that the agent believes it can acquire without damaging the client relationship. Agents are incentivised to sell as much travel as possible; while gross margin is not incentivised directly, it does significantly contribute to the Total Sales Volume figures. The ability to manage the client's elasticity of demand (i.e. how price sensitive the client is) directly impacts the agent's ability to price effectively and therefore to maximize profit. By definition, this is a tacit knowledge skill. A new employee who masters this skill earlier, or learns such from her team mates quickly, will generate higher profits.

2.13.1 Factors for Firm (Group) Performance

Firm performance is contingent on the group's ability to perceive opportunities and capacity to pursue those opportunities (Van de Ven, 1986). This builds on the fundamental Schumpeterian (1950) concept of creative destruction, whereby new developments degrade the rents appropriated from current business practices.

Under the Schumpeterian theory of economic development, groups require constant innovation and improvement in pursuit of economic opportunities for sustained competitiveness/performance. Schumpeter (1950) outlines the theoretical need for innovation from an organizational perspective. Van de Ven

(1986) later built on this, outlining innovation as *the development and implementation of new ideas by people who over time engage in transactions with others within an institutional order*. Van de Ven (1986) saw innovation as the most crucial mechanism by which groups can ensure their future competiveness. Von Hippel (1988) expanded upon Van de Ven's concept of innovation, identifying two different mechanisms by which innovations allow groups to develop and sustain competitive advantage: 1) allowing groups to develop superior efficiency compared to their competitors; and 2) providing superior value for customers.

2.14 Knowledge

Plato defined knowledge as 'justified true belief'. This early understanding conceptualized knowledge as an ultimate truth which individuals can understand through a complex cognitive process. Later definitions of knowledge evolved to describe an intrinsic understanding of a particular subject which can be applied to specific ends. Sir Francis Bacon aphorized this phenomenon as 'knowledge is power' in his 1597 *Meditations sacrae*. This paradigm represented the pervasive logic of individuals and organizations until the end of the last century: *Knowledge represents the dominant source of sustainable competitive advantage for groups* (Grant, 1997).

A decade ago, the benefit of possessing internal knowledge was not perceived to be as useful as having the capacity to use that knowledge to develop new knowledge resources dynamically (Teece et al., 1997). Throughout his work on knowledge creation processes, Nonaka argues that a group must possess internal knowledge assets to be able to dynamically engage them in the creation of new knowledge and to pursue economic opportunities (Gupta and Govindarajan, 2000; Nonaka, 1994; Nonaka, Toyama and Konno, 2000). From this perspective, the ability to pursue economic opportunities through knowledge assets does not reject the knowledge-based view of the group but, rather, argues that one must possess knowledge to create it. However, if one is interested in the process by which knowledge is created, then the dynamic process of creating knowledge from existing assets is of more interest.

2.14.1 Knowledge Management

More contemporary definitions of knowledge suggest it is:

> that which the researcher comes to believe and value on the basis of the meaningfully organized accumulation of information through experience, communication or inference (Zack, 1999, p. 46).

Dominant organizational paradigms today view knowledge as the most vital of group resources (Seidman, 2011). Even outside the discussion of knowledge-based organizational paradigms, much management literature argues that knowledge is the most relevant strategic resource that groups may possess (Argote and Ingram, 2000; Ipe, 2003; Moorthy and Polley, 2010).

2.14.2 Knowledge Types

Polanyi (1966) creates a valuable distinction between tacit and explicit knowledge. Tacit knowledge is that which an individual possesses internally but may not easily express outwardly. Put another way, explicit knowledge is easy to share (e.g. knowing what a bicycle is), but tacit knowledge is derived from personal experience and is not easily shared (e.g. knowing how to ride a bicycle). More specifically, tacit knowledge entails insights, intuitions and beliefs that are tightly intertwined with personal experience with the knowledge source (Polanyi, 1966). Tacit knowledge is seen as difficult to move between parties, which is both a boon (i.e. competitors cannot easily acquire it) and a bane (i.e. transferring tacit knowledge within a group is just as challenging) (Bou-Llusar and Segarra-Ciprés, 2006).

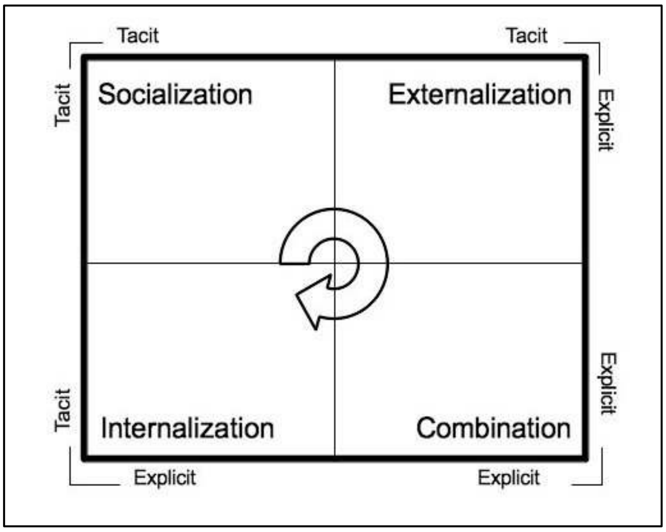


Figure 5: Nonaka's SECI model for Knowledge Transfer

The concept of tacit knowledge is central to any discussion of knowledge transfer and performance. Tacit knowledge is seen as more useful in improving performance but, as mentioned above, tacit knowledge proves much more difficult to transfer (Nonaka, 1994). The challenges associated with transferring tacit knowledge are heavily referred to in the discussion of dynamic capabilities and represent the impetus to develop Dynamic Capabilities (Grant, 1995). However, actual properties of knowledge are not simply binary; instead, they form a spectrum of knowledge. Nissen's text (2005) on knowledge flow offers an excellent illustration of a knowledge hierarchy. Each element of the hierarchy is qualified based on its actionability (i.e. how easily and often does the presence of such lead to actionable steps?) and its abundance (i.e. how prevalent is it?).

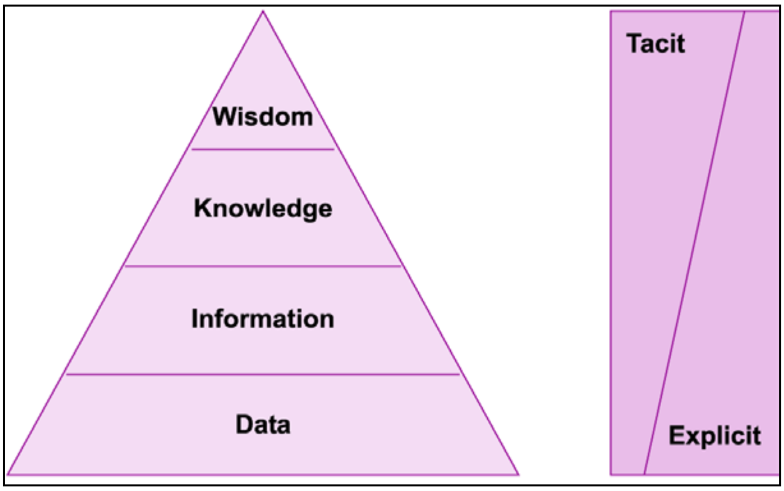


Figure 6: Nissen's hierarchy of Knowledge

Nissen (2005) describes four sub-types of knowledge: Data, Information, Knowledge and Wisdom. The table below summarizes his findings.

Table 4: Knowledge types (from Nissen, 2005)

	Actionability	Abundance	Tacitness	Stickiness
Data	Low	Very high	Low	None
Information	Medium	High	Low	Little
Knowledge	High	Medium	High	Some
Wisdom	Very High	Low	Very high	Extreme

Nissen (2005) outlines different forms of knowledge which vary both in their ability to be transferred and in their ability to be useful. Most, if not all, scholars claim that it is mostly the harnessing of tacit knowledge that leads to innovation and, in turn, to performance (Harlow, 2008; Yang, Brashear and Boles, 2011). The more tacit the knowledge is, however, the harder it is to transfer. Thus stickiness (as defined by Szulanski, 1996) is correlated to utility. However, wisdom (as shown in **Error! Reference source not found.**4) is far more useful (or actionable) than regular knowledge, but it proves even more difficult to transfer. This makes sense since, according to Nissen's hierarchy, wisdom is the most tacit form of knowledge. The benefit of having such actionable knowledge is that when that knowledge is transferred, absorbed, combined and transferred again (i.e. Nonaka's SECI process), the knowledge output and its utility are far greater.

2.15 Knowledge Transfer

Knowledge itself presents a concept that is difficult to measure quantitatively. While it is something that is shared between individuals, knowledge is very specific to the individual and is altered when it is transferred between individuals. Knowledge exchange (transfer) is defined by researchers as the process by which one unit is affected by the experience of another (Argote and Ingram, 2000; Inkpen and Tsang, 2005).

Management literature on the exchange of knowledge has its early roots in the study of the diffusion of innovation. Rogers' (1983) early evidence in this area has proven to be the definitive work on dissemination of innovation published to date. While Rogers' work has not been surpassed in this field, it did lay the theoretical underpinnings for a variety of management literature which followed, namely the body of work on stickiness and knowledge flow. Szulanski's (1996) work on stickiness popularized von Hippel's (1994) concept of 'Sticky Knowledge'. This concept argues that some knowledge is harder to transfer than others. To articulate this concept, Szulanski makes extensive use of Rogers' (1983) early work.

Maurer, Bartsch and Ebers (2011) also find that knowledge transfer mediates between organization members' intra-organizational social capital and organizational performance outcomes of growth and innovation performance. Their work builds on Lechner, Frankenberger and Floyd (2010) by using similar survey techniques. Like Lechner, Frankenberger and Floyd, Maurer, Bartsch and Ebers (2011) also call for further exploration of the influence of intragroup networks on group performance.

On aggregate, the different exchanges of knowledge represent an overall flow (which can be directional) of knowledge across the organization, which is dynamic and changes over time. Whereas knowledge is difficult to quantify, knowledge flow is even more abstract and difficult to measure. Nissen and Levitt (2004) argue that one of the failings of existing knowledge management theories is the lack of a strong cohesive theory in the field of knowledge flow. While there are many works which address the aspects of knowledge flow (Nissen, 2005; Szulanski, 1996; Von Hippel, 1994), there have been few attempts to develop a model which addresses the dynamics of knowledge as it flows (Nissen and Levitt, 2004). The extant literature on the creation of knowledge flow offers a conceptualization of the overall exchange of knowledge across a group, the kind of knowledge exchange that ultimately impacts group innovation.

Scholars agree that the process of continually acquiring new knowledge is necessary for competitive advantage to be sustained (Argote and Ingram, 2000). Within the management literature, the preceding concept was developed parallel to the advancement of dynamic/organizational capability, as laid out by Grant (1995). The dynamic acquisition of knowledge theorized by dynamic/ organizational capability has been established in the management literature, starting with the definitive work by Nonaka (1994). Grant (1995) built upon Nonaka's work, proposing that the development of knowledge creation mechanisms is linked to the performance of the group with respect to its ability to pursue economic opportunities. The learning and knowledge creation process developed by Nonaka and contextualized by Grant was then further developed by Dyer and Nobeoka (2000), who argue that the knowledge creation process is in fact a process of internalizing and combining knowledge to form new knowledge.

A variety of work exists conceptualizing the transfer of knowledge across organizations: Knowledge Transfer (Mowery, Oxley, and Silverman, 1996; Tsai, 2002), Knowledge Sharing (Hansen, 1999; Tsai, 2002), Knowledge Flows (Gupta and Govindarajan, 2000; Schulz, 2001) and Knowledge Acquisition (Darr, Argote, and Epple, 1995; Lyles and Salk, 1996). It is important to consider the conditions that must be present for knowledge to be exchanged successfully (e.g. transferred, shared, flowed, acquired, etc.). This research calls these *knowledge antecedents*. There is a burgeoning body of literature which has attempted to address the exchange of knowledge and the antecedents to facilitating such exchange (Argote and Ingram, 2000; Szulanski and Jensen, 2006). The most relevant for our purposes is the work of Lyles, van Wijk and Jansen (2008) leveraging the earlier work by Inkpen and Tsang (2005), which groups all of the potential antecedents to the transfer of knowledge along three categories based on three sets of underlying characteristics: knowledge characteristics, organizational characteristics and network characteristics.

The literature has also explored the effects of networks on knowledge transfer. Szulanski (1996) argues that higher trust leads to greater transfer of tacit knowledge. Gulati, Nohria and Zaheer (2000) contend that trust reduces search costs associated with the transfer of knowledge. Moreover, considerable literature exists examining social constructs such as centrality and its effects on exposing actors to a wider range of inputs (Burt, 1992). The seminal articles which have emerged are detailed in Table 5 below.

Table 5: Recent Knowledge Transfer Literature

Year	Author(s)	Contribution
1994	**von Hippel**	Puts forth the notion of 'Sticky' information, which suggests that knowledge which proved most useful in stimulating innovation.
1994	**Nonaka**	Popularizes the notion of knowledge creation (a parallel to innovation) and suggests that the knowledge which was most tacit (and useful) is that which contributes to the process of knowledge creation.

Year	Author(s)	Contribution
1995	**Szulanski**	Builds upon von Hippel's concept of 'sticky' information purporting the concept of 'sticky' knowledge as the knowledge which is most useful but difficult to transfer and acquire.
1995	**Zander and Kogut**	Suggests that knowledge and its contextual properties are the most important in facilitating knowledge flow.
1998	**Tsai and Ghoshal**	Underlines the importance of cognitive impediments to knowledge exchange, such as Cultural Distance and Shared Vision.
2000	**Szulanski**	Argues that an important factor in impeding knowledge is knowledge ambiguity (partly linked to knowledge tacitness).
2003	**Argote, Mcevily and Reagans**	Introduces a set of factors which facilitate the flow of knowledge.
2004	**Burakova-Lorgnier, Bouzdine-Chameeva and MacGilChrist**	Proposes a vision of a network structure from the point of view of knowledge transfer capacity. Finds a structure based on strong ties, and thus a dense network with a high level of trust, assists in the transfer of tacit knowledge.
2006	**Ordonez de Pablos**	Develops a conceptual framework for the analysis of knowledge flow between subsidiaries and their parent organization.
2008	**Jansen, van Wijk and Lyles**	Consolidates the existing management literature, articulates the concept of knowledge antecedents and the need for further research into specific antecedents.
2010	**Lechner, Frankenberger and Floyd**	Shows a curvilinear relationship correlating the relational and structural dimensions of networks on performance. Finds that cognitive dimensions have a positive relationship with group performance.

Year	Author(s)	Contribution
2011	**Maurer, Bartsch and Ebers**	Indicates that knowledge transfer mediates between organization members' intra-organizational social capital and organizational performance outcomes of growth and innovation performance.

2.16 Knowledge Transfer Activities

Recent literature (Noethen and Voelpel, 2010) has explored knowledge transfer/knowledge flow as the combination of two sub actions: knowledge seeking and knowledge sharing. The former focuses on the tasks related to determining the knowledge needed and sourcing such. The latter describes the process of transferring the knowledge. A person who engages in a large amount of knowledge seeking would be a receiver who actively seeks knowledge from others to enhance outcomes. A person who engages in a large amount of knowledge sourcing is a sender who shares their knowledge with others. The researcher's definition of knowledge flow includes both knowledge seeking and knowledge sharing. As a result, and to minimize complexity, knowledge flow will be examined only at the group level as including both sub-actions.

As has been mentioned above, knowledge transfer is contingent on a variety of factors. While knowledge specific factors are important, within an organizational context, organization and network specific characteristics also play an important role. The following sections address each of these individually.

2.17 Knowledge Characteristics

The intrinsic characteristics of any given kernel of knowledge are considered to be among the most important factors for such a kernel to be transferred and exchanged (Zander and Kogut, 1995). Tacitness is a prime example of an intrinsic characteristic of knowledge. As mentioned in earlier sections, the degree to which knowledge is tacit is a primary factor in determining how easily the knowledge is transferred (McLaughlin, Paton and Macbeth, 2008; Reed and Defillippi, 1990). Another commonly cited antecedent for knowledge exchange is knowledge ambiguity (Levin and Cross, 2004; Simonin, 1999). The concept of Knowledge Ambiguity emerges from the combined effects of *tacitness, specificity and complexity* of the underlying knowledge and has been long been seen as a hindrance to the transfer of knowledge between groups (intergroup) in alliance literature (Singh, 2005). It has also been suggested that knowledge ambiguity works also on an intragroup basis. As such, Knowledge Ambiguity is a factor which mitigates the dynamic capabilities of individual groups to transfer knowledge (Coff, Coff and Eastvold, 2006).

2.18 Organizational Characteristics

As with the context of knowledge, the very nature of an organization and its design can affect the transfer of knowledge. Organizational characteristics have been generally reduced to two measures in the literature: group size and age (Lyles, van Wijk and Jansen, 2008). While most authors agree that these two antecedents are key to knowledge transfer, the exact effect that each of these factors has on knowledge transfer is still a topic of debate. Both Gupta and Govindarajan (2000) and Laursen and Salter (2006) argue that group size does affect the group's ability to transfer knowledge, while Tsang (2002) found that group size is not a significant factor in knowledge transfer. The effect of organizational age on knowledge transfer is also widely debated, with March and Cyert (1963) taking the position that young groups transfer knowledge better than those that are long established. But March and Cyert (1963) have been challenged by more recent scholars (Gray and Meister, 2004; Yli-Renko, Autio and Sapienza, 2001) who suggest that the age of a group does not have an effect on knowledge transfer at all. Regardless of this contradiction, there are recommendations in the literature which suggest that, to regain the advantage of nascent organizations, established groups should revert to decentralization and disaggregation of their structure (Gupta and Govindarajan, 2000). Gupta and Govindarajan (2000) argue that this process (of decentralization and disaggregation) is in fact a means by which groups induce more open and willing exchange of knowledge. Regardless of any differences in this argument, the discussion of these organizational characteristics suggests that organizational form and design do in fact impact the flow of knowledge directly.

In addition to a group's age and size, a group's 'absorptive capacity' is an organizational attribute which has been found to directly impact the degree to which an organization (or individual) can absorb the knowledge being transferred. Absorptive capacity, conceptualized by Cohen and Levinthal (1990), has been largely positioned as a mezzo (group) level mechanism for internalizing knowledge (Lane, Lyles and Salk, 2001). However, Lane, Lyles and Salk's research suggests that Absorptive Capacity can also be observed at a micro level by focusing on the exchange of knowledge between two individuals. This (micro) notion of absorptive capacity suggests that different actors (groups or individuals) receive knowledge differently. For example, some actors have the ability to absorb all knowledge directed towards them while other actors simply lack this ability. The ability to absorb knowledge thus plays a vital role in internal knowledge flow. Absorptive Capacity has in fact been argued to be the key antecedent to underlying knowledge transfer (Gupta and Govindarajan, 2000; Szulanski, 1996).

2.19 Network Characteristics

As discussed above, social networks represent the outlay of social relationships within groups. In this context, a network is defined *as a pattern of relationships among groups and institutions* (Kogut, 2000). A group's social network can thus be conceptualized as the lattice of pipes through which knowledge and even social capital may flow. These social relationships may extend outside the boundaries of the group but represent the true avenues and channels through which knowledge is exchanged (Inkpen and Tsang, 2005). Fritsch and Kauffeld-Monz (2009) argue that network structures significantly affect a group's ability to transfer and absorb knowledge and that this in turn affects group performance. As such, network characteristics reflect the social faculty of knowledge exchange. Network characteristics can be divided along three dimensions: structural, relational and cognitive. Nahapiet and Ghoshal (1998) argue that each dimension is required to facilitate the combination and exchange of resources embedded within network relationships (i.e. social capital). Nahapiet and Ghoshal's three dimensions are defined as follows:

- **Structural Dimension:** This dimension includes the number of relations within a network; the relative access to information each network actor has (Hansen, 1999; Tsai, 2001); and the centrality of each network actor, as defined by Bonacich (1987) and adopted later by Ahuja (2000), Losada and Heaphy (2004), Powell, Koput and Smith-Doerr (1996) and Tsai (2000). The concept of density refers to the ratio of actual ties to potential ties. A firm with full density would have ties between each actor. The inverse of density is the concept of Structural Holes. When A connects to B and C, but B and C do not connect, then A bridges the structural hole said to exist between B and C.

- **Relational Dimension:** This dimension governs relationships between individual actors in a network. The relational measure of Tie Strength, as defined by Granovetter (1973), is cited as a key relational determinant in facilitating the flow of knowledge between dyads (Hansen, 1999). More recent work extends the Tie Strength concept to an organizational level, arguing that the Tie Strength reinforces trust, which ultimately affects the degree to which knowledge is transferred (Jensen and Szulanski, 2004; Lyles, Lane and Salk, 2001). Other scholars have put forward counterarguments noting that high levels of trust are in fact inhibitors to the exchange of knowledge (e.g. Jensen and Szulanski, 2004; Lyles, van Wijk, and Jansen, 2008; Yli-Renko, Autio and Sapienza, 2001).

- **Cognitive Dimension:** Shared vision and cultural distance are widely accepted as being important cognitive elements which characterize the cognitive dimension of knowledge transfer (Inkpen and Tsang, 2005; Jensen and Szulanski, 2004; Lyles and Salk, 1996; Tsai, 2001). The concept of Homophily is often used as a measure of the cognitive dimension (Breiger, 2004; Louch, 2000; Novak, 2000). Homophily refers to the similarity (in background, experience, culture, training, etc.) between A and B. If A and B are very similar, then the level of homophily is said to be large.

These three dimensions of Social Networks are also widely cited as the three dimensions of Social Capital. This makes sense in that Social Capital is, by definition, a network property and as such it is impacted by the structural, relational and cognitive forces over which it flows.

The properties of individual characteristics and their respective antecedents, as defined by Lyles, van Wijk and Jansen (2008), are shown below in Table below in parallel to a list of facilitators developed by Argote et al. (2003).

Table 6: Antecedents to Knowledge Flow

Antecedents (Lyles, van Wijk and Jansen, 2008)			**Facilitators** (Argote, Mcevily and Reagans, 2003)		**Network Dimensions** (Lechner, Frankenberg and Floyd, 2010)
Knowledge Characteristics		Tacitness Ambiguity	**Knowledge (Context)**	Tacitness Causal Ambiguity	Not Considered
Organizational Characteristics		Firm Size Firm Age Absorptive Capacity	**Unit (Firm)**	Absorptive Capacity Prior Experience	Shared Vision Degree of Exploration
Network Characteristics	Structural	# of Relations Access to Information Centrality	**Relationship (Network)**	Tie Strength Trust	Centrality
	Relational	Strength of ties			Tie Strength
	Cognitive	Shared Vision Cultural Distance			

Argote et al. (2003), Fritsch and Kauffeld-Monz (2009), and Lyles et al. (2008) focus on network characteristics that impact knowledge transfer. Expanding on the previous lists above, Fritsch and Kauffeld-Monz (2009) explore the moderating and mitigating effects of network measures on knowledge transfer, which include but are not limited to frequency of interaction, spatial proximity of network ties and heterogeneity of competencies.

2.20 Knowledge Transfer and Network Topology

Lechner, Frankenberger and Floyd's (2010) paper looked at the effect of network topology on performance. Hansen's (1999) work informed their study, providing evidence that strong ties facilitate the transfer of complex knowledge. Evidence has also been found to suggest that tie strength diminishes opportunistic behaviour between actors (McAllister, 1995). However, Hansen (1999) also finds that relational embeddedness (the closeness of the actors in relation to each other) has a negative effect on performance. The rationale behind this phenomenon is that those who are tied very closely to others may not look beyond their closest ties for knowledge. Based on this prior research it seems likely that network topography plays a significant role in the transfer of knowledge and thus on organizational performance as well.

2.21 Social Capital and Knowledge Transfer

Organization theorists suggest the social capital within organizations is a potentially powerful resource for improving organizational performance (Andrews, 2010). Empirical studies on social capital attempt to quantify social capital's contribution to economic development and show the link between performance and the different social capital dimensions (Tsai and Ghoshal, 1998; Widén-Wulff and Ginman, 2004). Tsai and Ghoshal (1998) made a quantitative study testing the impact of the different social capital dimensions in a multinational company with 15 business units and over 30,000 employees. Tsai and Ghoshal show that each dimension of social capital reinforces the creation of the other dimensions. Tsai and Ghoshal found that creating social capital through these dimensions creates value for the organization (Widén-Wulff and Ginman, 2004).

2.22 Onboarding

Onboarding is the process of acclimating a newly hired employee to an organization (Bradt, 2009). Organizational socialization, i.e. onboarding, is the process through which new employees move from being organizational outsiders to becoming organizational insiders (Bauer and Erdogan 2011). Onboarding refers to the process that helps new employees learn the knowledge, skills and behaviours they need to succeed in their new organizations (Bauer and Erdogan 2011). Further, new employee adjustment is associated with important employee and organizational outcomes, including performance (Bauer 2011). Prior research, including the recent works of Bauer and Erdogan (2011) and Bradt (2009), has relied on surveys to develop these conclusions.

There is a growing dialogue regarding the notion that early and easy knowledge transfer leads to improved performance of new employees (Lawson et al., 2009; Liu, 2011; Vaara, 2012). Alice Snell (2006), a practitioner vice president of the research division of a talent management solution company, writes:

> [A]n effective onboarding process enables new team members to gain access to information, tools and materials needed to perform their function more quickly. Productivity generated by successfully onboarding a new hire sooner will have a direct, positive effect on the overall productivity of the company (p. 32).

Firms that are more successful at rapid onboarding (i.e. bringing newly hired staff to the team) tend to use a relational approach, helping newcomers to rapidly establish a broad network of relationships with co-workers that they can tap to obtain the information they need to become productive (Rollag et al., 2005). Onboarding can be described as the direct bridge between the promise of new employee talent and the attainment of actual productivity (Snell, 2006). The early stages of onboarding are crucial to establishing a lasting bond between employees and the company (Snell, 2006). From this, one concludes that the speed of new strong tie formation would be an important aspect of SNA, yet most prior SNA research has been static, merely confirming the presence or absences of ties. Little work has been done to examine the speed at which those ties are formed, mostly due to the static nature of survey methodology.

2.23 Research Context

Several recent works frame this research. In 2004, Burakova-Lorgnier, Bouzdine-Chameeva and MacGilChrist proposed a vision of a network structure from the point of view of knowledge transfer capacity. They found that a structure based on strong ties, and thus a dense network with a high level of trust, is required to facilitate the transfer of tacit knowledge. De Pablos' 2006 article focused on how intragroup knowledge flow works amongst transnationals (groups with offices in more than one country). De Pablos focused on subsidiaries and parent units of the same group; however, these units had independent governance models and therefore they may be more akin to independently governed groups in an alliance than to intragroup units. De Pablos found that tacitness, cultural distance, social complexity and causal ambiguity all have a negative impact on internal knowledge transfer. De Pablos' research provides a useful model to inform this research. However, de Pablos does not account for a number of network-centric characteristics (e.g. trust, number of ties, Tie Strength, etc.) which may limit the long term impact of her research findings.

In 2007, Wu, Hsu and Yeh published a paper on the determinants of knowledge transfer through a team level analysis. Focusing on sales teams from the travel industry as the target of their empirical sample, their paper reveals results supporting the argument that social capital facilitates knowledge transfer. In 2008, C.T. Butts published 'Social Network Analysis with SNA' in the *Journal of Statistical Software*. This seminal article is a good jumping-on point for the use of software for SNA. Fritsch and Kauffeld-Monz (2009) investigate knowledge transfer in a sample of 16 innovation networks with approximately 300 groups within them. That article supports this research by providing SNA measures and SNA methodology. Fritsch and Kauffeld-Monz (2009) first provide solid descriptions of various mitigating variables on knowledge flow and then a method to explore such both through social network analysis and through survey. Fritsch and Kauffeld-Monz (2009) found:

- Strong ties are more beneficial for tacit knowledge transfer.

- Frequency of interaction leads to more knowledge flow.

- Spatial proximity is not important.

- The presence of network ties within spatial proximity is positively correlated to knowledge flow.

- Network Cohesion is positively correlated to knowledge flow.

This research attempts to port these measures and methods to explore intragroup knowledge transfers amongst a group's internal network (as opposed to Fritsch and Kauffeld-Monz's investigation focusing on inter-group networks).

Wu, Lin et al. (2009) continue the conversation on social networks, performance and knowledge transfer. These authors make the following statements which inform this work:

> A large body of literature on social networks in organizations demonstrates that certain types of network topology are optimal. However, little research leverages the ample data created by people's electronic communications to refine and verify [these] theories. This gap is problematic, because the literature on organizational networks suffers from the same deficits as much of the social network literature: both tend to be focused on small, static networks (p. 1).

Of key importance is (1) the reference to 'static networks'; and (2) their call for further *research to leverage the data created by people's electronic communications to refine and verify theories.* Wu, Lin et al. (2009) find that not only does group level social network topology correlate with group performance, but the authors attribute such to the nodes in a social network. These authors find that an inverted U-shape correlates several SNA measures and performance. This research attempts to operationalize some of Lin et al.'s (2009) measures and methods to further explore intragroup knowledge transfers amongst a group's internal network (as opposed to Fritsch and Kauffeld-Monz's investigation focusing on inter-group networks) through examination of electronic communications.

This research was designed to replicate the object of Lechner, Frankenburger and Floyd's (2010) study (which was an exploration of social capital performance) using actual longitudinal communications rather than self-reported static opinions about communication. Specifically, those authors' use of structural network occurrences as an explanatory variable for the performance of strategic initiatives frames this research's approach. Lechner, Frankenburger and Floyd (2010) find that an inverse curvilinear relationship reflects the relational and structural dimensions of networks and their effect on performance. Lechner, Frankenburger and Floyd (2010) dubs this the Dark Side of the Social Capital Theory. Similarly, their research finds that cognitive dimensions have a positive relationship with group performance. Since Lechner, Frankenburger and Floyd's (2010) publication, several authors have attempted to continue the conversation in the literature, most notably the 2011 work by Maurer, Bartsch and Ebers.

Villena, Revilla and Choi (2011) continue exploring social capital's impact on performance and, like Lechner, Frankenburger and Floyd (2010), find that both too little social capital and too much social capital can undermine performance (e.g. the Dark Side of the Social Capital Theory). Simon and Tellier (2011) explore the longitudinal evolution of networks, while Maurer, Bartsch and Ebers, in their 2011 paper, 'The value of intra-organizational social capital: how it fosters knowledge transfer, innovation performance, and growth', find that knowledge transfer mediates between organization members' intra-organizational social capital and organizational performance outcomes.

In doing so, Maurer, Bartsch and Ebers (2011) pick up from Lechner, Frankenburger and Floyd (2010) by examining 218 projects in the German engineering industry. Maurer, Bartsch and Ebers' findings show that knowledge transfer (conceptualized as the mobilization, assimilation, and use of knowledge resources) mediates between organization members' intra-organizational social capital and organizational performance outcomes of growth and innovation performance. Their findings emphasize the role of knowledge transfer as both a key benefit of social capital and an important driver of the noted organizational performance outcomes (p. 173). With regard to the relational dimension of social capital, Maurer, Bartsch and Ebers find a positive association for tie strength noting that:

> All research should be sensitive to the possibility that different types of organizational settings may display unique relations among dimensions of social capital, knowledge transfer, and performance outcomes (p. 177).

Maurer, Bartsch and Ebers also suggest that it would be a *fruitful avenue* for future research to examine to what extent their findings hold true in other types of intra-organizational settings. The research at the heart of this work is focused on pursuing this *fruitful avenue*.

This research contributes to the social network conversation by augmenting and adding to the knowledge generated by recent works, particularly that of Lechner, Frankenberger and Floyd's 2010 study and Maurer, Bartsch and Ebers' 2011 study, both of which found an inverse curvilinear relationship connecting intragroup social network topology with group performance. Unlike Lechner, Frankenberger and Floyd's 2010 study which focused on strategic initiatives (which often only exist for finite periods of time and have specific explicit goals) as the unit of analysis, this research focuses on *Business as Usual* inside more than 187 groups at a national travel agency. Similarly, as compared to Maurer, Bartsch and Ebers' 2011 study which examined 218 projects in the German engineering industry as the unit of analysis, this research focuses on *Business as Usual* inside more than 187 groups at a national travel agency. Thus, this research contributes by showing the findings to be valid for durable (non-temporary) groups not just strategic initiatives or projects.

At the same time as Maurer, Bartsch and Ebers (2011), another trio of researchers, Yang, Brashear and Boles, examined Social Capital in a selling centre environment (2011), finding that a group's social capital influences selling centre performance through facilitating knowledge transfer and absorption within and across the selling centre. Selling centres in Maurer, Bartsch and Ebers' research all had the same task contingency (i.e. all agents in the selling centre try to sell as much as they can at the best margins they possible). This is similar to the case at the heart of this research whereby **HF**'s offices attempt to maximize sales. Yang, Brashear and Boles define their work as exploratory, suggesting that the propositions they found, need next to be empirically tested. This call for further empirical testing highlights another possible contribution for this research. Yang, Brashear and Boles (2011) state:

> There are many obstacles to conduct empirical research in this area. It requires time-consuming procedures such as snowball interview and network analysis, and is dependent on a high degree of cooperation from selling groups. However, most constructs in our framework have been operationalized by previous studies. Future research can use these measurements to test the proposed conceptual framework (p. 158).

2.24 Models, paradigms, frameworks

In Chapter 4, the researcher reviews relevant material models from recent prior art. From these the researcher developed the following model for this research:

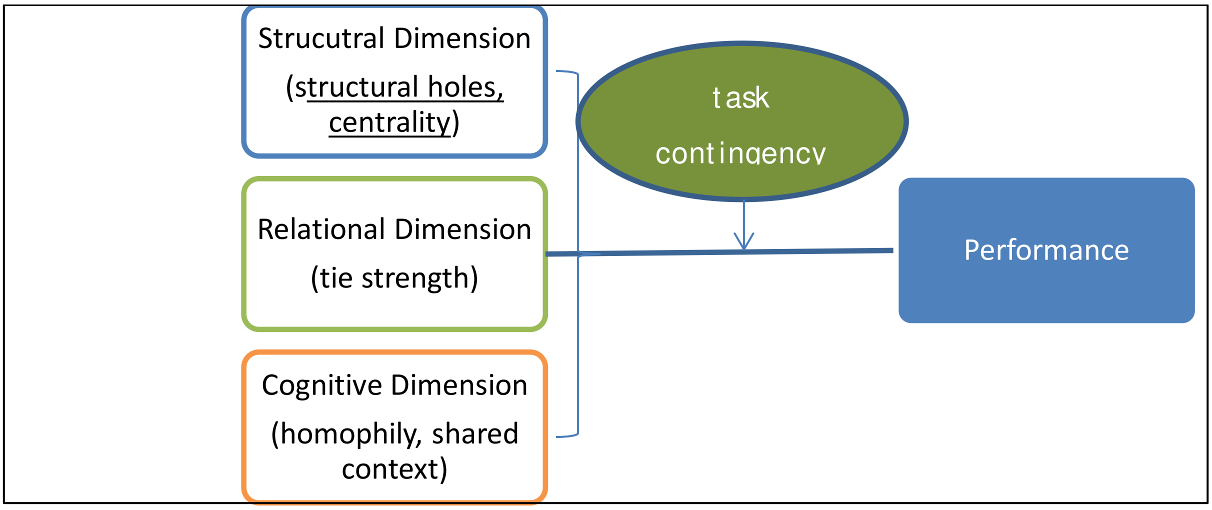


Figure 7: A model for Social Network's impact on Performance

For this research, the level of exploration (i.e. the task contingency level) is held constant for each group since each HF group is undertaking the same task: i.e. *sell the most travel services at the best margin.* By holding the level of task contingency constant, the researcher intends to derive an equation showing the relative impact of each dimension of a social network on group performance. Then the researcher will test the fit of such equation against the observations recorded. In doing so the researcher intends to (1) empirically prove the inverse curvilinear relationship between performance and social network topology measures; and (2) show that with such measures one can predict performance. Through such analysis, the researcher hopes to be able to comment on a wide variety of interesting questions, such as: Which SNA measure best predicts performance? What role does centrality have? Should **HF** Offices have many or few structural holes to maximize performance?

2.25 The Research Questions revisited

From Chapter 1, the research objective was determined to be a theory testing dissertation, empirically confirming the relationship between an organization's Social Network Topology and organizational Performance. This is more generally stated as:

> **Do the Cognitive, Relational and Structural Dimensions of an organization's Social Network have an inverse curvilinear correlation with organizational Performance as predicted?**

This will form the basis for the Research Questions, more thoroughly discussed in Chapter 4.

There is also the following research question regarding onboarding:

> **What, if any, is the correlation between an individual's speed of onboarding and individual's performance?**

2.26 Knowledge Flow and Performance; a Critical Review of Existing Literature

While the theorized correlation between group performance and knowledge flow is well documented in management literature, there is still a need to understand the causal relationships that lead from successful knowledge exchange to performance. There is considerable overlap and disagreement in the literature on this topic. The aim of this research is both to dispel some of those disagreements and to develop further insights into the factors which allow for knowledge processes to influence group performance.

Within the management literature the importance of knowledge processes on group performance is well established (Kogut and Zander, 1992; Lyles and Salk, 1996; Tsai, 2001; Zaheer and Bell, 2005), and regardless of the contradictions as to the mechanisms by which performance is affected (Schreyogg and Kliesch-eberl, 2007; Zaheer and Bell, 2005), it is widely accepted that there is a positive correlation between internal knowledge processes and group performance. The internal knowledge processes which ultimately affect group performance are based on the transfer of knowledge itself, leading to new knowledge creation through dynamically developing their ability to pursue economic opportunities consistently (Eisenhardt, 2002; Nonaka, 1994; Walter and Lechner, 2007; Zack, 1999).

Where this Research Sits

According to Borgatti and Foster's (2003) taxonomy, this research falls within the structuralist social capital paradigm.

Table 7: Typology of research on consequences of network factors (from Borgatti and Foster 2003)

	Social Capital (performance)	Diffusion (social homogeneity)
Structuralist (topology)	Structural capital	Environmental Shaping
Connectionist (flow)	Social Access to Resources	Contagion

This research attempts to join three on-going conversations in academia, including the on-going exploration of the network form of organization; the debate contrasting social capital paradigms in the girders vs. pipes analysis; and the investigation of how organizational network topology correlates with organizational performance. Each will be discussed independently.

Network form Conversation:

The varying organizational forms with which direct human assets are forced to contend, and the specificity of those human assets, should serve to reduce transaction costs and improve efficiency (Fligstein, 1995; Milgrom, 1990; Putterman and Krsozner, 1996). Asset specificity is cultivated as individual human assets tacitly develop knowledge and skills through their individual experiences (Teece 2009). In managing highly specified human assets, transactional costs are significantly heightened as specificity requires high costs for the search and retention of highly specified partners (Rossignoli, 2009). These costs can be directly mitigated through the use of networks where the costs of developing and managing specific assets are mitigated and efficiency is bolstered (Rossignoli, 2009). Over the last decade, organizational structures focused around networks have emerged, which have shaken the traditional notions of formal organizational context (Castells, 2011; Reed et al., 2006; Whitford, 2005).

Pipes vs. Girders Conversation:

In 2003, Borgatti and Foster distinguished between two broad categories of social network theory: Topology and Flow (Farrall, 2004). This debate is often labelled the Girders vs. Pipes debate. Structionalists (girders) believe that an actor's position determines performance outcomes. Constructionists (pipes) believe that it is the transmission (flow) of resources along pre-existing social ties which determines performance outcomes. The girders perspective focuses on social capital while the pipes perspective addresses the flow of social assets (e.g. knowledge). Topology (girders) Structionalists discount the actual content of ties while focusing on overall patterns of association (Farrall, 2004). Social theories describing the network structure of social capital fit on the girders side. Alternatively, flow mechanisms consider network ties as explicit conduits for the flow of social goods.

This research is guided by a theory that it is neither girders nor pipes that provide the optimal theory, but in fact it is both. Much like the Wave-Particle Duality theory of Light (Greiner, 2001) where light shares attributes of both wave and particle, perhaps a social network can act as a girder or as a pipe depending on how the phenomenon is viewed. This is part of the concept of complementarity, which says that a phenomenon can be viewed in one way or in another, but not both simultaneously (Chen and Klahr, 1999).

Topology and Performance Conversation:

As recently as 2010 (Lechner, Frankenberger and Floyd) researchers have called for an exploration of how intragroup topology impacts overall group performance. These researchers have generated a qualitative inductive survey-based work which has led to speculations that balance is required along each of the network dimensions. Lechner, Frankenberger and Floyd (2010) find that, when comparing groups inside an organization, the relational and structural dimensions had an inverted U relationship with group performance (note: they found the cognitive dimension to be simply positively correlated, but if more data points were available to them, perhaps then the cognitive dimension may have also displayed as curvilinear). A year later, Maurer, Bartsch and Ebers (2011) found similar relations.

This research compares a group's intragroup social network dimensions with team performance through deductive empirical data. In doing so, this paper attempts to answer the call posited in recent literature for more empirical examinations of the relationship between a group's network dimensions and overall group performance.

2.27 The Research Gaps

In summary, the gaps this research attempts to address are:

1) Holding task contingency constant across groups, does group social topology (i.e. SNA measures at a group level) have a curvilinear relationship with group performance as predicated?
2) Can software replace survey as a valid method for determining a group's social network, a group's SNA measures?
3) Does the speed of tie formation impact the performance of a new employee?

By exploring the matters above, the author hopes to illustrate how intragroup network structure predicts overall group performance through the transfer of tacit knowledge.

Summary of Chapter 2

Chapter 2 reviewed the background literature that forms the development of the hypotheses. The chapter first examined the evolution of the view of the group (towards Network). It then reviewed organizational knowledge and the knowledge transfer process, examining knowledge transfer under facilitating mechanisms and barriers to transfer, and performance outcomes. Finally, the chapter gave an overview of recent research which serves to frame the gap in the conversation that this research attempts to fill. This research attempts to extend existing views on the correlation between group performance and group social network topology. This research will also attempt to develop and distinguish between the concepts of social networks and social capital.

The third chapter is divided into three main sections. First, the author explores the link between knowledge flow and performance. This is followed by a discussion of how performance can be measured and how this research intends to explore the moderators of performance. Finally, a conceptual model is presented and hypothesises are stated.

Chapter 3 – Methodologies and Methods

In chapter two, the researcher provided a literature review and outlined possible contributions to the field. This third chapter outlines the methodology and is divided into several key sections. First, the author reviews his internal epistemological and ontological stances, discussing the leading paradigms contained within. Next, a discussion follows on how performance can be measured and how this research intends to explore the moderators of performance. Finally, the author presents a conceptual model and states his hypotheses.

Methodology: Epistemological and Ontological Dimensions

The researcher is conducting empirical deductive research, the purpose of which is to test the early findings that the correlative relationship between network dimensions (relational, structural, cognitive) and group performance is inversely curvilinear, or 'U' shaped.

This chapter will first explore the various research paradigms available and then justify the selection of the postpositivist/critical realism paradigm. Based on such, the author will provide justification for the selection of quantitative analysis techniques of social network analysis as well as reasoning for the pursuit of deductive mode research.

3.1 Introduction to Research Paradigms

According to Guba (1990, p. 17), a *paradigm* is a 'basic set of beliefs that guide action'. A *research paradigm* is thus seen to be the beliefs that guide the research process. In his influential work, 'The structure of scientific revolutions', Kuhn (1962) states that a paradigm guides the research efforts and directions of scientific communities, providing a framework into which questions, facts and ideas can be organized

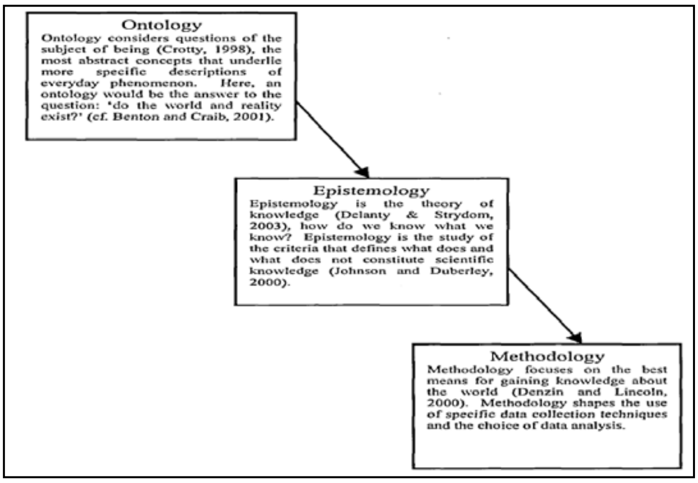


Figure 8: Cascade of Research Paradigms (from Squire 2005)

and evaluated (Kuhn, 1996). According to Kuhn, a research paradigm consists of three inter-related concepts: ontology, epistemology and methodology.

Ontology refers to the nature of reality, Epistemology defines the nature (scope, limits, etc.) of knowledge, and Methodology is the processes by which the researcher searches for knowledge. When one combines these three dimensions of discovery into a paradigm and lays common methodologies upon it, one generates a model such as the one in Figure 9 below, derived from Saunders, Thornhill and Lewis' text 'Research methods for business students' (2006).

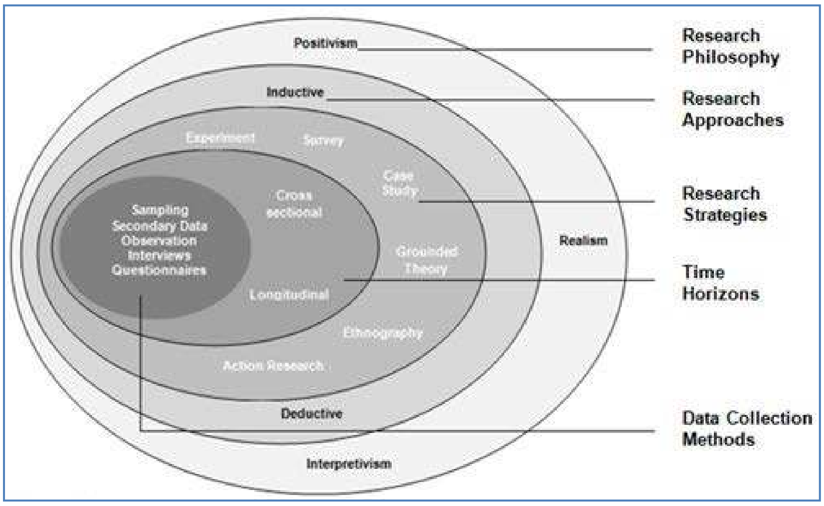


Figure 9: Research process Onion, from (Saunders, Thornhill and Lewis (2006)

Kuhn (1962) calls the social sciences 'pre-paradigmatic', meaning that unlike the physical sciences (chemistry, physics, etc.) where most researchers share or accept commonly held paradigms (e.g. gravity, viscosity), the social sciences (including Management) are not as unified in their research approaches to ontology, epistemology and methodology. This has led to a wide variety of research paradigms. The table below summarizes some of the more prevalent paradigms in management research today. It is beyond the scope of this research to explore these fully, but it will be helpful to touch briefly on each of the key paradigms.

3.2 Exploring Leading Paradigms

Positivist Approach

Positivism searches for causal explanations and fundamental laws. The belief that the natural sciences and the social sciences share common logical and methodological principles is at the heart of positivism (Hughes and Sharrock, 1997). If you believe that the world around us can be objectively observed and analysed, you may be a Positivist. Positivists seek to measure reality through objective, repeatable methods. Positivists tend to first formulate a hypothesis and then craft a research design that focuses on measuring phenomena through objective, arm's length methods which engage the researcher as an independent observer. Positivist methods tend to be quantitative (Neuman, 2000) and are largely orientated towards manipulating and predicting the social world rather than understanding it (Delanty and Strydom, 2003). Positivist studies aim to uncover causal relationships between the objects of interest, so that knowledge can be applied to control or regulate the behaviour of the objects within society (Benton and Craib, 2001).

Post-positivism

Slightly less dogmatic than the absolute nature of positivist researchers, the post-positivist researchers' fundamental tenet is that the world around us is knowable, but through the process of knowing it (exploring it, researching it, and describing it) an observer's bias will temper the findings, meaning that pure objective truth or truth exists but is subject to the interpretation of the finder of such truth.

Post-positivist researchers consider human knowledge to be based not on unchallengeable, rock-solid foundations, but instead rely upon human conjectures. The paradigm of post-positivism rejects the absolute unrelenting nature of Positivist dogma that there is one objective reality, and instead post-positivists suggest that, while determining objective truth is still the end goal, a researcher must consider their own observer bias, participation and impact in determining that truth.

Post-positivists also recognize the fallibility of depending on a single method. Thus, post-positivists support use of critical multiplism, a mixed method research that is an extended version of triangulation (Guba and Lincoln, 1994). Triangulation is a technique whereby methods are selected on the basis of their apparent appropriateness to the research question and will often be combined in an attempt to overcome the bias inherent in single method designs.

One of the most common forms of post-positivism is a philosophy called critical realism. As the research continues to move along the spectrum of epistemology, the researcher starts to take the observer and participants more into account. Unlike positivism, in critical realism meaning is not considered to exist apart from any consciousness, but is constructed through our interactions with reality. Critical realism refers to any position that maintains that there exists an objectively knowable, mind-independent reality, whilst acknowledging the roles of perception and cognition (Bhaskar, 1978).

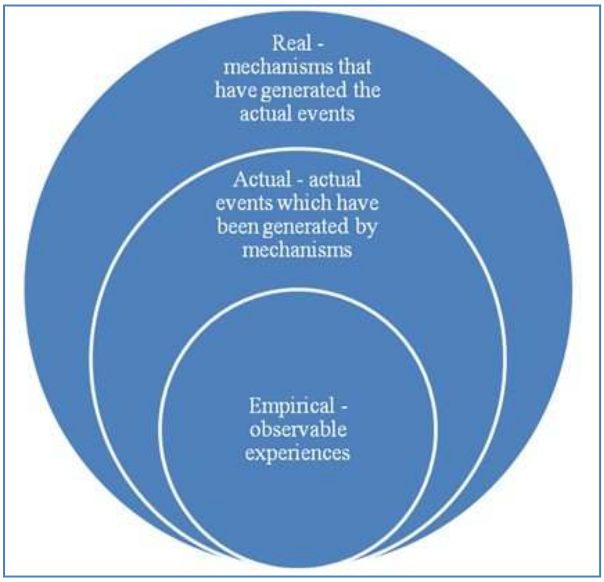


Figure 10: From Mingers and Willcocks (2004)

Critical realism theory states that the theory of knowledge (epistemology) is different from a theory of being (ontology). Critical realism is grounded in the notion that there is a reality which exists independent of its human perception. Critical realists believe that there are unobservable events which cause the observable ones; as such, the social world can be understood only if people understand the structures that generate such unobservable events. This is important in the experimental context because it allows the scientist to distinguish between the event and what causes it. According to this theory, an individual conducting an experiment creates the conditions necessary for the experiment (observable event), but the results are caused by the underlying laws and mechanisms (unobservable events). The realism side of the theory focuses on the existence of real mechanisms which shape events. ‘A central idea of critical realism is that natural and social reality should be understood as an open stratified system of objects with causal powers'. (Morton, 2006, p.2). According to this theory, there are three strata, domains of empirical, actual, and real. The domain of empirical includes observable experiences. The domain of the actual includes actual events which have been generated by mechanisms. Finally, the domain of real includes the mechanisms that have generated the actual events.

Social Constructivism

So far, the paradigms discussed above have focused on objective truths that exist independent of the observer. However, many researchers believe that since all truth is seen through a personal lens, truth cannot be objective. Many researchers believe that the observer is part of what is being observed (Fendt and Kaminska-Labbe, 2011). Under the constructionist paradigm, perception alone is not reality (Gergen, 2003). Constructivists believe that reality is a blend of subjective internal perceptions and external reality. Constructivism is focused not just on the findings of research but on the value underneath such findings, often using inductive logic to facilitate this (Denzin, 1978). Where positivists focus on statistical probabilities to determine truth (Jacking, 1984), social constructionists rely upon theoretical abstraction (Jacobs, 2000).

Interpretivism

Interpretivism is derived from a subjective epistemology that holds that meaning does not exist apart from human consciousness (Crotty, 1998). Rather, meaning is forcibly imposed on the object by the subject (Walsham, 1995; Heshusius, 1996). If one recounts the Zen question, 'If a tree falls in the forest, and there is no one around to hear it, will it make a sound? (Abbott, 2008). One can more easily distinguish between the paradigms. According to Positivists, the tree makes the same sound regardless of the lack of observation. According to Interpretivists (notably, George Berkeley, 1685-1742), there would be no sound, since sound is only created by the vibrations of the falling object colliding with the ear of an observer.

Table 8: Research Paradigms (from Squire, 2005)

Paradigm	Positivism	Post-postivism	Critical Theory *et al*	Interpretivism
Ontology	Naïve Realism: Social world is external to individual cognition and consists of tangible and relatively immutable structure and relationships.	Critical Realism: Social world is external to individual cognition, but it can never be fully understood or comprehended	Historical realism: Reality is shaped by social, political, cultural, ethnic and gender values.	Relativism: Realities are local and relative to the individual or a particular time or culture
Epistomology	Objectivism: Meaning exists apart from the operation of any consciousness. It implies the separation of the subject and object of knowledge so that the observer is uninvolved during the research process.		Constructionism: Meaning comes into existence through our interaction with the realities of the world. There can be no meaning without the mind.	Subjectivism: Meaning is imposed on the object by the subject. Knowledge is generated from the mind without reference to reality.
Methodology	Experiment Simulation Survey Statistics	Experiment Survey Case study	Action Research Feminist Studies Critical Studies Case Study	Ethnography Phenomenological Research Case Study Grounded Theory Heuristic Inquiry

3.3 Choosing a Paradigm

Based on self-awareness and the definitions above, this researcher is neither a pure positivist nor a pure interpretivist. Neither paradigm fits with the researcher's view of the socially constructed world. The intent of this research is to show that intragroup social network topology is predictive of overall group performance and as such behaves in a deterministic manner. As the researcher intends to undertake a deductive approach, it is important that the expected findings can and will be repeatable. Further, recent research (Lechner, Frankenberger and Floyd (2010) and Maurer, Bartsch and Ebers 2011) has provided not only the theory but initial evidence supporting that theory. This researcher now intends to expand those earlier works by empirically examining the impact of network topology on group performance.

Of the many paradigms identified above, this thesis is most consistent with post-positivism. Three factors motivate this choice. First, the thesis is broadly deductive; the researcher identified a series of hypotheses that were derived from existing theory and associated literature. The researcher now wishes to test these hypotheses. Second, post-positivism appears consistent with prevailing paradigms within the field of knowledge transfers (e.g. Schraw's (2006) review of knowledge processes with regards to positivism and post positivism). Third, the researcher recognizes that all research methods are fallible and hopes that the validity of findings is strengthened through a process of triangulation.

In this research, the researcher seeks to explore the causal relationship between network-centric factors and performance. Since a list of such factors can be compiled found in earlier research (see table 5), the focus of this research is not on qualitatively exploring possible relationships but on quantitatively and empirically testing the strength and impact of these factors on performance in order to determine their relative correlations.

This research builds on many early works (Squire, 2005; Szulanski, 1996). While those researchers' works were more exploratory than this current research, the authors seem to share a critical realism stance when it comes to ontology and epistemology.

Finally, long held beliefs and a foundational education in engineering has led this researcher to find that while knowledge is indeed knowable, truth is not absolutely objective, but instead relies upon human conjectures. Thus, the post-positivist/critical realism paradigm suits this researcher and this research best, not only from an ontological and epistemological stance, but also as it allows for the application of triangulation and mixed methods.

The Impact of Critical Realism

Critical realism refers to any position that maintains that an objectively knowable, mind-independent reality exists whilst acknowledging the roles of perception and cognition (Bhaskar, 1978). Critical realism states that there is a reality which exists independent of any human perception thereof. Critical realists believe that there are unobservable events which cause the observable ones, and therefore the social world can be understood only if people understand the structures that generate such unobservable events.

The goal of first exploring one's ontology and epistemology is to ensure a foundation for knowing so that one can determine individually the nature of reality and limits on knowing. Based on the section above, the researcher has determined his own personal ontology/epistemology to be that of critical realism.

Applying the critical realism (post-positivism) view informs the researcher:

- That all data collected is biased by the observer;
- That all data collected may be fallible;
- That in acquiring new knowledge, the best one can do is to explore the causal relationships as seen in the experiment through the researcher's own lens; and
- That whilst empiricism and positivism locate causal relationships at the level of events, Critical Realism locates them at the level of the generative mechanism.

While the researcher agrees with the above, the data at the heart of this work is being collected by software, not through observation, and thus may not be subject to condition 1 above (i.e. software is not biased). However, the view that software is unbiased is naïve at its base. Software of course is written by individuals, and those individuals will build it according to their experiences, goals and biases. Bruno Latour's (1986, 1987) expansive work on the biases of scientific instrumentation and experimentation reminds us that biases are built into instrumentation and laboratory processes (he follows Kuhn). Also see Timothy Lenoir's 'Inscribing science,' which finds that instruments are never neutral because they are born out of epistemological biases that tend to deny or conceal their bias, particularly in empirical research (Lenoir, 1998).

This research builds on the methodology used by Lechner, Frankenberger and Floyd (2010). Lechner, Frankenberger and Floyd (2010) adopt a realist approach by asking the network members to identify the boundaries of the network in order to formulate the domain of the empirical. As with Lechner, Frankenberger and Floyd (2010), the domain of real here includes the unobservable mechanisms generating the performance results.

3.4 Choosing the Research Methodology

Research Considerations

From Punch (1998), the researcher derives the following topics to explore when contemplating the adoption of a particular method:

1. **Research Questions.** What exactly are you trying to find out? Focus on the 'exactly', as this can lead you in either a qualitative or quantitative direction.

2. **Viewpoint**. Are you interested in making standardized and systematic comparisons, or do you really want to study this phenomenon or situation in detail?

3. **The Literature.** How have other researchers dealt with this topic? To what extent do you wish to align your own research with standard approaches to the topic?

4. **Practical Considerations.** Issues of time, money, availability of samples and data, familiarity with the subject under study, access to situations, gaining cooperation.

5. **Knowledge pay-off.** Will you learn more about this topic using quantitative or qualitative forms of research? Which approach will produce more useful knowledge? Which will do more good?

6. **Style.** Some people prefer one approach over another. This may involve paradigm and philosophical issues or different images about what a good piece of research looks like.

For this research, leveraging Punch (1998):

1. As the author is trying to test how network dimensions impact group performance, a quantitative approach seems most applicable.

2. The researcher is interested in making standardized and systematic comparisons.

3. Other researchers have dealt with this topic in a conceptual manner. These researchers 'discovered' the sources of knowledge friction and this research now wishes to 'explore' them further, focusing on the relationship between those friction sources and performance.

4. Focusing on how to improve performance by addressing the most significant factors regarding knowledge flow is a topic that should elicit sufficient responses to allow for valid data gathering and statistical relevance.

5. As noted, the qualitative work has already been done to some extent. Prior research calls now for more empirical quantitative work.

6. As noted above, the researcher's undergraduate training in Engineering has tilted his perspective towards the positivist side of the epistemological spectrum.

Punch's key questions confirm that the researcher and research are most suited to a Post-Positivist approach.

3.5 Research Mode

There are two main modes of research: inductive and deductive (Buckley, 1976). A deductive mode involves testing theory. An inductive mode aims to generate theory based on fact-finding activities. In deductive mode, the researcher creates a hypothesis, *a priori*, and his research then goes on to prove or disprove that hypothesis. The goal of deductive research is to move from specific facts to generalizable phenomena; consequently deductive researchers often gravitate towards the positivist paradigm. The goal of inductive research is to generate theory based on specific facts. Prior to the data gathering no substantive hypotheses are created *a priori.* The metaphor of explorers vs. exploiters may be helpful. Explorers know not what they seek, but seek it anyway, sailing to faraway lands to gather facts. Exploiters already know the land exists but seek to better understand the land, to map it better and understand how those facts interrelate. In this metaphor, explores are inductive while exploiters are deductive.

For this research, prior authors have undertaken the exploration at an intragroup level, and now the author seeks to better understand how network topology impacts group performance. Based on this, the deductive mode is appropriate for theory testing the prior suggested correlations between the dimensions of social capital and performance.

3.6 Research Design and Methods

According to Churchill (1979), research design provides an overall guide for the collection and analysis of data of a study. The importance of research design stems from its role as a critical link between the theory and argument that informed the research and the empirical data collected (Nachmias and Nachmias, 1981).

A choice of research design 'reflects decisions about the priority being given to a range of dimensions of the research process' (Bryman and Bell, 2007, p. 40). This will have considerable influence on lower-level methodological procedures such as sampling and statistical packages. Research Design is therefore a blueprint that enables researchers to find answers to the questions being studied for any research project. Along with the clear research plan it provides, constraints and ethical issues that a study will inevitably encounter must also be taken into account (Saunders, Thornhill, and Lewis, 2006).

As the researcher wishes to explore group performance, data gained through observational research techniques (based on the researcher's observations) will be needed on the group. One must be careful when utilizing this method and work to minimize and mitigate any issues that might arise from observer bias. The researcher must also be aware and ensure that the observations are both reliable and generalizable. Since the values recorded through observation will be objective and stable, the impact of issues of bias, reliability and generalization will be minimal.

3.7 Measuring Social Networks

As discussed in Chapter 2's literature review, firm structure and form are evolving. Where once a firm's structure was dictated by its organizational chart, in today's world firms are more accurately visualized as a non-hierarchical network of resources (people, technology, knowledge) and relationships. Social Network Analysis (SNA) is a research method used to highlight the relationships between people (Mead, 2001). SNA is often used to describe the relationship, examine information flows, and analyse patterns that develop between individuals and organizations (Wasserman and Faust, 1994). The result is a visual representation similar to the one presented in Figure 11.

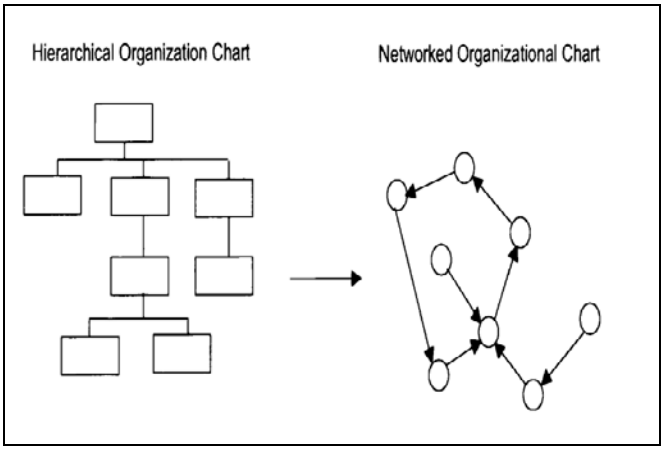


Figure 11: Network vs. Hierarchy

SNA can be used to map knowledge flows and measure relationships between actors in a network (Liebowitz, 2004). SNA provides a perspective not only on how embedded are actors in a network, but also on how a structure emerges from the interactions of actors in the network. One type of SNA approach advocates collecting information about each actor's ties with all other actors in a network (Hanneman, 2001), whereas another method uses a snowball technique by identifying key actors, gathering information on their relationships and then about the subsequent relationships with an expanding set of actors. A third method would be to use "egocentric" methods (Liebowitz, 2004), with the selection of certain individuals as focal nodes, and analysing their immediate relationships.

As illustrated, traditional organizational charts tend to focus on command and control, whereas SNA shows the group based on defined relationships and practical deployment. SNA can help illustrate the true informal networks behind a group's success. According to Huberman and Hogg (1995), these informal networks coexist within the formal structure of the organization and can be used to solve problems more efficiently.

In general, there are two main approaches to SNA. The first explores the firm as a whole, and is aptly named the Whole Network approach. The other is called the Ego Network approach. The Whole Network approach looks at relationships between individuals within the group as a gestalt, while the ego network approach focuses on a particular individual and his or her relations. Since the researcher is focused on the firm performance, it is more productive to pursue a whole network approach. The typical barrier to whole networks analysis for surveys is that whole networks require almost 100% survey participation in order to be valid, which can be extremely difficult to achieve.

3.8 Measuring Firm Performance

To understand what constitutes successful firm performance, it is important to underline the Smithian tenet of maximizing available group resources in economic exchange. Notwithstanding, a discussion of group performance must not be limited to static performance (e.g. Share price on January 1); rather, it must be extended to sustainable, comparable and objective metrics (to differentiate these from short term financial gains/losses caused by anomalies) when describing group performance for any practical value to be gained from it.

For the case at hand, **HF** Senior Management regularly collects and monitors two key performance measures:

- Total Sales Volume ➔ the total annual travel services sold by all members of that group.

- Gross Margin Average ➔ the average of the individual gross margins on travel services sold by all members of that group.

The researcher acknowledges there is a large number of performance measures listed in prior research (table 5). Notwithstanding, the researcher sees no reason to deviate from the Performance measures used by **HF** Senior Management. From this, the proposed model would be:

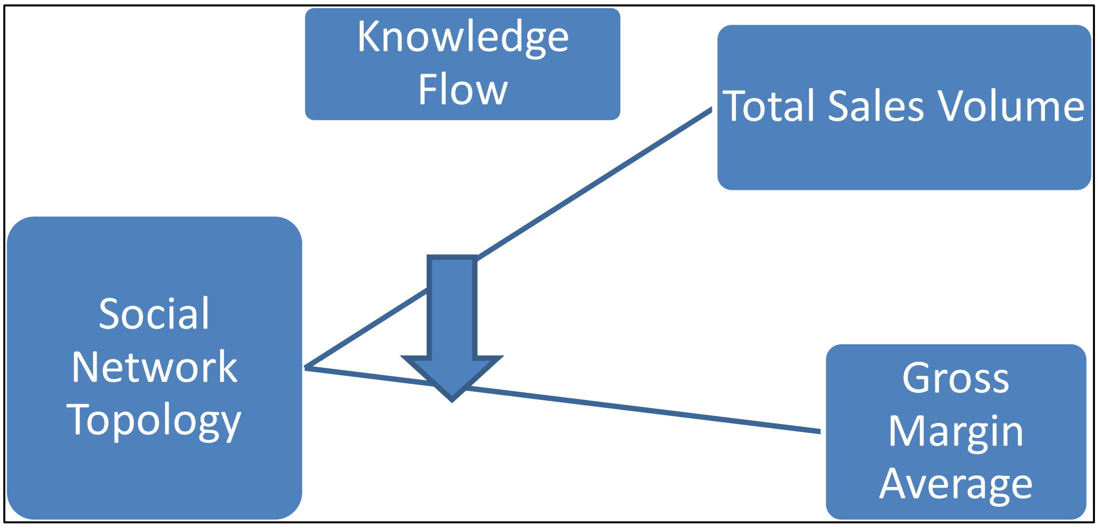


Figure 12: The Research Model

3.9 Measuring Network Topology

Social network data can be viewed as a social relational system characterized by a set of actors and their social ties (Wasserman and Faust, 1994, p. 89). Social network analysis seeks to understand the network structure by description, visualization, and (statistical) modelling. Social network data consist of various elements.

A social network is a very simple concept; it is a set of actors (or points, or nodes, or agents) who may have relationships with one another. Networks can have few or many actors and may support various kinds of relations between pairs of actors. Network analysis is a fundamental approach to the study of social structure (Wellman, 1983). It is typically undertaken through either Statistical Modelling or Visual Analysis.

3.10 Multi Variant Analysis

A substantial amount of information is needed to describe even small social networks. According to Hanneman (2005), managing this data to reveal patterns of social structure can be tedious and complicated. All of the tasks of social network methods are made easier by using tools from mathematics. For the manipulation of network data and the calculation of indices describing networks, it is most useful to record information through mathematics (Hanneman, 2005).

The most direct way to research a social structure is to analyse the patterns of ties which link its members (Wellman, 1983). In doing so, the researcher hopes to concentrate on studying how the pattern of ties in a network provides significant opportunities and constraints, because it affects the access of people and institutions to such resources as information, wealth, and power. (Wellman, 1983).

For statistical modelling, the researcher will first use software to generate measures for Social Network Topology (based on email records provide by **HF**). This generates the social graph for each office along with each office's SNA measures (e.g. centrality) then the researcher will explore the possible correlations between those SNA measures and Performance. This will be done through the standard, well-accepted statistical technique of multiple regression analysis. In this technique a number of possible independent variables, e.g. the SNA measures of Network Topology, are tested for possible correlations with suitable measures of performance as the dependent variable.

In a typical example, the correlation of activity across multiple possible independent variables with 180+ observations may be checked by multiple regression analysis. This may find that Performance correlates with some combination of these SNA variables. The statistical information which is usually provided in reporting such a correlation consists of ***n***, the number of observations, ***r***, the multiple correlation coefficient, ***r2***, which is a measure of the explained variance, and ***s***, the standard deviation. The statistical significance of the correlation equation and of each independent variable is also given in the form of a ***p*** value or as an *F* statistic from which a ***p*** value can be readily determined. Multiple Regression is a well-accepted form of statistical modelling (Topliss, 1972).

3.11 Visual Analysis

From the early days of SNA, images of networks have been used both to develop structural insights and to communicate those insights to others (Freeman, 2000). Social networks are inherently visual in nature. Visual analytic tools and techniques have been used in social network analysis (Shen, 2008).

The use of visual images is common in many branches of science, and such images are important for progress in various fields (Arnheim, 1970; Freeman, 2000; Klovdahl, 1981; Koestler, 1964; Taylor, 1971; Tufte, 1983; Tukey, 1972). Historian Alfred Crosby has gone much further, proposing that visualization is one of only two factors that are responsible for the explosive development of all of modern science, the other being measurement (Freeman, 2000).

Visualizations of social networks have been used to aid SNA from the beginning (Freeman, 2000). The visualization of networks is important because it is a natural way to communicate connectivity, allowing for fast pattern recognition by humans. However, there are great challenges when visualizing networks by hand (Di Battista, 1999); thus the rise of SNA software. Two distinct display forms have been used to visually construct network images, one based on points and lines and the other on matrices. In most point and line displays the points (nodes) represent social actors and the lines (vertices) represent connections among the actors. In matrix displays the rows and columns both represent social actors, and numbers or symbols in the cells show the social connections linking those actors. The overwhelming majority of network images have involved the use of points and lines (Freeman, 2000).

3.12 Social Network Software

SNA software tools are not just for scientists anymore. Moderators, administrators and other community experts also have a stake in learning more about the structural dynamics of their interactions. The emergent challenge for designers and educators is to build easy-to-learn interfaces that enable these SNA users to discover community patterns and individual roles they might not otherwise see (Bonsignore, 2009).

Social network analysis has emerged as a powerful method for understanding the importance of relationships in networks. However, interactive exploration of networks is currently challenging because (1) it is difficult to find patterns and comprehend the structure of networks with many nodes and links; and (2) current systems often consist of a medley of statistical methods and produce overwhelming visual output, which leaves many analysts uncertain about how to explore in an orderly manner (Perer and Shneiderman, 2006).

The earliest use of computational procedures in producing point and line diagrams focused on the problem of determining locations for the actors (points). Bock and Husain (1952) and Proctor (1953) were the first to report using computational procedures to aid in placing points. They both used factor analysis but produced very different kinds of images (Freeman, 2000).

In the 1970s, Alba (1972) worked with Gutmann and Kadushin to develop an early program (SOCK) that, along with a Stromberg-Datagraphics 4060 plotter, could produce point and line graphics automatically. The program was intended to serve as a general-purpose network analysis and image-producing device (Freeman, 2000). Over the following decades more than two dozen software packages have been developed for SNA.

The advent of the World Wide Web in the mid-1990s revolutionized opportunities for network imaging (Freeman, 2000). In practice, a network visualization of a domain can be messy, particularly when the network is large. Visualizations are useful to leverage the powerful perceptual abilities of humans, but overlapping links and illegible labels of nodes often undermine this approach (Perer and Shneiderman, 2006). Existing SNA software tools often involve extensive pre-processing or intensive programming skills that can challenge practitioners and students alike.

At present there are more than two dozen software applications that can visualize SNA. A thorough review thereof is beyond the scope of this work, but the author recommends Mark Huisman's authoritative article on the topic, 'Software for social network analysis' (2005). In the last decade analytical tools have improved greatly and many SNA software applications have come to market. NodeXL® is one such software application.

NodeXL® is an open-source template for Microsoft Excel that integrates a library of common network metrics and graph layout algorithms within the familiar spreadsheet format, offering a low-barrier-to-entry framework for teaching and learning SNA (Bonsignore, 2009). NodeXL® was chosen as the software package for this research for a multitude of reasons, including but not limited to the following facts: NodeXL® is free and open-sourced; NodeXL® is relatively easy to use; NodeXL® scales for large data sets; NodeXL® facilitates the visualization of SNA through email usage; and this researcher has direct access to the developers of the NodeXL® project.

3.13 Email for SNA

As corporations grow, knowledge becomes dispersed and communication and coordination become increasingly challenging (Ackerman, Pipek and Wulf, 2003). While face to face communication is not always possible, social computing (e.g. email, blogs, twitter) tools are highly accessible, uniquely positioning them to provide collaborative enterprise-solutions (Stecher et al., 2009). Over the last decade SNA has received a major boost from the ability to use email data mining and software to generate the network image. Prior to these developments, the practice of SNA was manual and iterative (i.e. you had to ask each person about his or her relationships). Through the practice of email data mining, SNA has become much less expensive and time consuming.

Email requires an inherent social network and this can be leveraged to visualize connections (Nardi et al., 2002). According to Tyler, Wilkinson and Huberman (2005) email is a strong tool for discovering the community structure of organizations as email has become the predominant means of communication in our information society. Email pervades business, social and technical exchanges, and as such it has been established as an indicator of collaboration and knowledge exchange (Wellman, 2002; Whittaker and Sidner, 1996). Bulkley and Van Alstyne also agree with this proposition but go a step further in their 2008 paper. They suggest that email may actually be a better tool than network surveys when conducting SNA; stating:

> [N]etwork surveys can provide reasonable measures of general communication tendencies, studies of informant inaccuracy have demonstrated that self-reporting (as done with network surveys) become increasingly unreliable for capturing details of interactions. (Bulkley and Van Alstyne, 2008, p. 5).

Bulkley and Van Alstyne (2008) report that the most prolific communicators over email were the most prolific communicators across all media. In fact, measured email activity was found to be directly correlated with self-reported estimates (reliable here, because of the small numbers of nodes) ($p < 0.01$).

An additional argument supporting the use email vs. survey comes from the snapshot vs. dynamic view of networks. Surveys, used in many SNA papers, gather information at a single point in time. But social networks form over time, growing stronger or weaker as the levels of interaction impact tie strength. Watts (2001) discusses the process of network formation through a dynamic paradigm, where self-interested individuals can form and sever links. Jackson (with Watts, 2002) delved deeper into dynamic network formation in their study, finding:

> The payoff to an individual from an economic or social activity depends on the network of connections among individuals. Over time individuals form and sever links connecting themselves to other individuals based on the improvement that the resulting network offers them relative to the current network (p. 265).

This view was originally established by Skyrms and Pemantle (2001), who, publishing in the same journal and issue as Watts (2001), write, 'modelling network structure as dynamic increases the realism (of the result) without rendering the problem analysis intractable'.

When an email is sent, more than just the text is sent. Email architecture also contains specific time and date data which is added at the time when email is sent. This allows researchers to monitor or review tie strength longitudinally, something not available to researchers using traditional survey methods. By this logic, email affords researchers a shield against the recency bias. In this research, the researcher is concerned with the social networks at **HF** during the period of Jan 1, 2011, through December 31, 2011. A survey administered on or around December 31st would indicate the relationships in the network only at that moment in time. A travel agent who left the group on December 1st would not be included in the surveyed sample, and thus may not show up in the network. Unlike surveys, SNA uses objective longitudinal data and, in this example, would include the departing travel agent, thereby providing a fuller understanding of the network.

As part of this research, the researcher accessed the email communication logs of the groups under review. **HF** provided the research with an email database of 7 million. This represents all incoming and outgoing email from the **HF** servers, from which a subset was generated that represented only intragroup communication. From this subset an edge list (i.e. who speaks to whom) was generated. From the edge list SNA data was gleaned (e.g. centrality, density, etc.). NodeXL® then leveraged the edge list to draw a social graph representing the social network topology of that group. This process was then repeated for each of the **HF** Offices.

While the researcher agrees with the authors' claims (Tyler, Wilkinson and Huberman (2005); Wellman, 2002; Whittaker and Sidner, 1996) above that email is an appropriate tool to examine group structure, the researcher does acknowledge that a recent study by MIT may provide a valid counterpoint. In their recent article, 'E-mail may not reflect the social network' (2006), authors Grippa, Zilli, Laubacher and Gloor suggest that one must be cautious before adopting holus bolus the use of email to determine how work truly gets done. The authors remind readers that face to face interactions are still the most efficient way to transfer tacit knowledge. They suggest that, in groups where co-location of personnel predominates, those actors may opt for more synchronous forms of communication (phone, instant message and face to face). Notwithstanding even those cases where face to face communication is available, email may be used to arrange for such, thus furthering the concept that email records are a proxy for relationships. This is important for this research, since HF intraoffice groups are all collocated.

The researcher also agrees that the proportion of communication that is email, as compared to other forms (phone, online chat, face to face), will fall as the percentage of co-located actors rises. Notwithstanding, email continues to be an appropriate tool for measuring relationships and knowledge flow. According to Bulkley and Van Alstyne in their 2006 Sunbelt Conference paper, 'Our analyses provided significant evidence supporting the interpenetration of email measures as proxies for more general communication patterns even though email use in any organization is context specific'. (Bulkley and Van Alstyne, 2004; Rice and Steinfeld, 1994). This view is also supported by several papers from leading authors, including Wellman (2002) and Whittaker and Sidner (1996), who find that email is a strong indicator for levels of collaboration and knowledge exchange, even if email is not the tool being used directly for such collaboration and knowledge exchange.

3.14 Analysis

This research uses a multi-method approach to analysis, cross-correlating the results obtained through different analysis techniques. Triangulation is defined as 'the combination of methodologies in the study of the same phenomenon' (Denzin, 1978; p. 291). The effectiveness of triangulation relies on the premise that the weaknesses of any single method will be compensated by the balancing strengths of the other method (Jick, 1979). This assumes that the weaknesses of individual methods are discrete rather than overlapping and that the strengths of two or more methods are complementary, to the extent that weaknesses are offset. However, from the discussion above, it is apparent that any research method chosen will have inherent flaws, and the choice of that method will limit the conclusions that can be drawn. It is therefore 'essential to obtain corroborating evidence from a variety of methods' (Scandura and Williams, 2000, p. 1249).

Given that this research is formulated within the post-positivistic paradigm, methodological triangulation is appropriate. Post-positivist researchers generally recognize that any single method may be fallible, and therefore triangulation offers the opportunity to compensate for specific limitations. Using multiple methods helps fill the gaps left by any given method and provides an important cross-check of individual analyses (Connidis, 1983).

It should be noted that true methodological triangulation is not being offered herein. To do so, data would need to be drawn from multiple sources and by multiple methods. This is not the case for this research. This research uses the same data but undertakes two different types of analysis (GLM and Visual). While the research acknowledges this is not a true triangulation of methods, it should provide additional insight. In summary, once the data was gathered, mathematical and visual analysis was undertaken.

3.15 Ethical Considerations

Ethical approval was granted by the University of Glasgow (where the researcher studies) and Ryerson University (where the researcher lectures). Approval was sought and granted for the use of the above described SNA email mining technique. No primary data was removed or copied from the server. The SNA software parsed the large volume of emails on the server, mapped out the SNA for the group and generated the SNA measures necessary for this research; any impact on the participants was deemed to be minimal. Further, the researcher has no role with HF, limiting the possibility of negative ramifications for HF employees.

Summary of Chapter 3

This chapter presented an overview of research paradigms, locating this current project within post-positivist research methods. Of the many paradigms identified, this research falls within the parameters of post-positivism and is informed by a critical realism view of epistemology, which is consistent with the prevailing paradigms within the field of knowledge transfer. The research is based on multiple methods; thus, the researcher seeks to strengthen the validity of the findings through a process of triangulation. The following chapter will detail the processes by which the methods will be operationalized, and the models upon which this research is based are discussed further.

Chapter 4 – Operationalization

This chapter is divided into five main sections. First, the researcher explores the link between knowledge flow and performance, and how performance can be measured. Next, the author explores the moderators of performance. Next, the researcher presents a conceptual model and the hypotheses to which such a model might lead. Then, the researcher sets out the method for moving from the theory behind the phenomenon to operationalizing the research. Finally, general definitions for Performance and for measuring social network topology are explored.

Overview of method

The sample population examined in this research is comprised of the sales associates of a nationwide travel agency. This organization, which will be called '**HF**,' employs over 1800 individuals in Canada and more than 20,000 worldwide. More than 90% of **HF** Canadian personnel are engaged in selling travel products (flights, hotels, car rentals, tours, etc.). These employees are grouped by Office, each of which is staffed by 5 to 15 travel agents. About 10% of the offices are designated as Corporate (selling mostly to pre-established business clients via phone and email), while 170+ offices are categorised as Retail (selling mostly to walk-in customers). Corporate teams are assembled in three colocation offices across Canada. Retail teams are located in individual group based street level storefronts. There are also two offices that focus on wholesaling cruise travel, primarily serving the other offices rather than the clients; because of this difference in business model, the two cruise offices were removed from the sample.

Each member of an Office attempts to maximize sales while maintaining a strong margin. **HF** management judges Office performance based only on Total Sales Volume. It is management's goal to maximize the performance of every Office. **HF** produced a database containing all **HF** emails for 2011. More than 7 million records (To, From, Date, Time) were reviewed, grouped and organized. Only intragroup emails (e.g. email amongst members of the same Office) were examined; all others were excluded.

Using an extension to the NodeXL® software package created for this research by the Social Media Foundation and Microsoft, the researcher took an ‘x-ray snapshot’ of each group’s social network. This was done by first creating an edge list (showing who spoke to whom, when, how often, etc.) based on intragroup email communications from January 1, 2011 to December 31, 2011. This snapshot is based on underlying data representing a series of social network topology measures, including Tie Strength, Structural Holes and Centrality.

To represent the cognitive dimension of Social Networks, homophily was measured as an indicator of cognitive distance and shared vision. To represent the relational dimension of Social Networks, Tie Strength was measured as an indicator of the overall Tie Strength in the group. To represent the structural dimensions of Social Networks, Centrality and Density (the inverse of structural holes) was measured. At **HF’s** request, email content was not reviewed. This effectively undermined any attempt to measure homophily; as a result, the cognitive dimension could not be explored.

Tie Strength, Density, and Centrality were then correlated with group performance (Normalized Sales Volume) in order to determine if high-performing groups share similar network measures. Multiple Variant Regression techniques were then used to analyse the findings and to determine if an equation optimizing the social network measures could be generated, followed by a visual analysis.

4.0 The Research Questions

According to Sarantakos (1998), the research methods will depend not only on the methodology of the researcher but also on the research questions. From Chapter 1, the following research objective was determined to be a theory testing, and empirically confirming the relationship between an organization’s Social Network Topology and organizational Performance. This can be more generally stated as:

Do the Cognitive, Relational and Structural Dimensions of an organization's Social Network have an inverse curvilinear correlation with organizational Performance as predicted?

From this the following research questions can now be extended based on Chapter 3:

- **Does a group's level of Tie Strength have an inverse curvilinear correlation with the Group's Performance?**
- **Does a group's level of Structural Holes have an inverse curvilinear correlation with the Group's Performance?**
- **Does a group's degree of Centrality have an inverse curvilinear correlation with the Group's Performance?**

And

- **What, if any, is the relationship between the number of days a strong tie is place and an individual's performance?**

4.1 Key Concepts that underlie this work.

From an organizational perspective, new economic opportunities are always emerging, and existing opportunities may fade away. To perceive and pursue these opportunities, groups need to create new knowledge consistently and disseminate it widely throughout their organization (Kusunoki, Nonaka and Nagata, 1998). Given that the creation of new knowledge stems from the combination of existing knowledge from a variety of sources, the ability to transfer knowledge is vitally important. This research sits at the crux of three overlapping areas of study:

- Social Capital;
- Knowledge Transfer; and
- Performance.

Li and Zhu (2009), in their work on the Influence Mechanism of Social Capital to Informal Knowledge Transfer, propose a theoretical model which links social capital, knowledge transfer and performance. They suggest that to improve effective informal knowledge transfer, one must improve across one or more of the three dimensions of social capital.

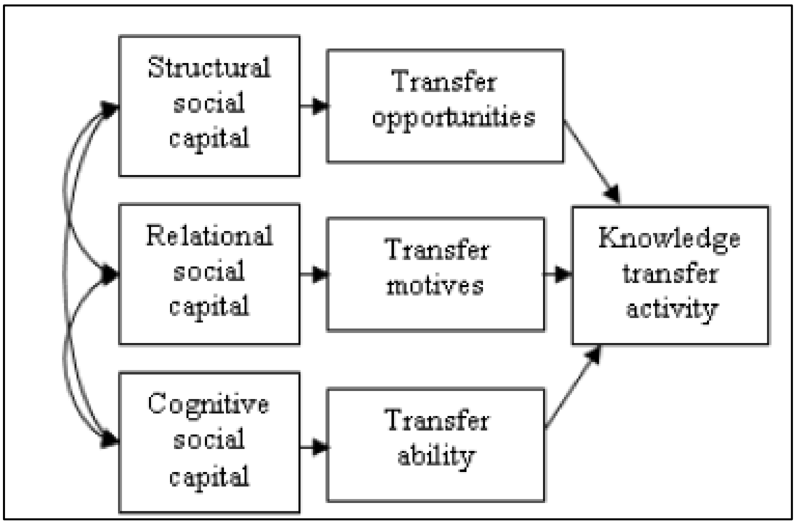


Figure 13: Model of Social Capital's impact to Knowledge Transfer (from Li and Zhu, 2009)

A large body of academic literature has argued that knowledge transfer directly impacts the performance of the group; while performance can be determined directly, however, knowledge flow is more difficult to measure. Knowledge flow is determined by several factors, including the type of knowledge (e.g. tacit or explicit), the absorptive capacity of the receiver, and the network over which that knowledge flows (Lin et al., 2011).

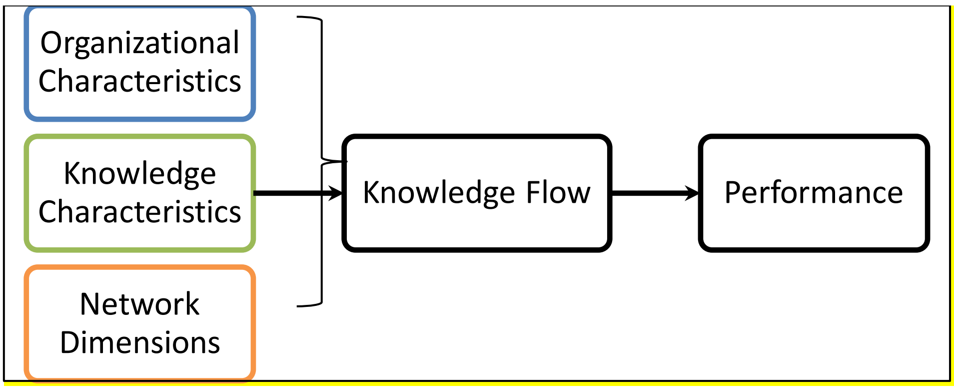


Figure 14: Knowledge Flow model (from Lin et al., 2011)

As much has already been written on both the impact of knowledge type and the value of knowledge flow (Kamhawi, 2010; Mu, 2008; Sturdy, 2009) this research examines the relatively less explored concept of how the Network Dimensions (e.g. topology) impact Performance.

4.2 Unit of Analysis

In social network analysis, one has many possible units to analyse. In all cases the vertex (edge) represents the relationship and the nodes are the parties privy to that relationship. Nodes can be groups (Schweitzer, 2009), strategic units (Lechner, Frankenberger and Floyd (2010), departments (Marin and Wellman, 2011) or individuals (Borgatti, 2005). Looking at the adjacent figure, it is empirically irrelevant whether Nodes 1,2,3,4,5,6 are offices, firms, groups, teams, or individuals. Yet it is contextually important to be mindful of the outcomes based on the level of analysis.

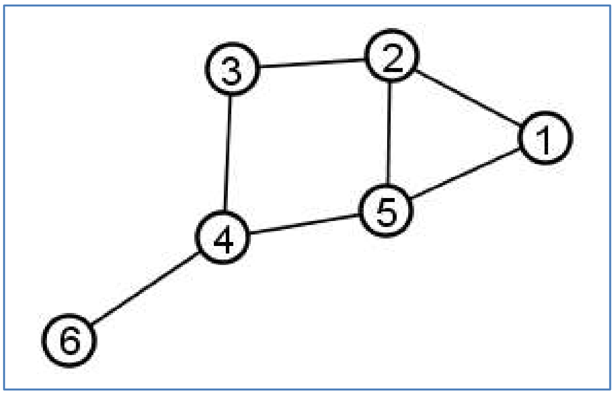


Figure 15: A typical Social Graph

This research examines independent units of production, which are referred to internally as Offices but herein contain 5-15 individuals, each selling travel services. Each Group is ranked based on annual gross sales generated. While each Group is not a legal entity (e.g. subsidiary), each group acts (from a cost and revenue perspective) as an independent unit. Even though Groups share common infrastructure (e.g. billing, HR, IT), a case can be made that each Group can be treated as an independent unit for analysis for the purpose of this research.

4.3 Prior Models

There are a number of useful models in the existing literature. From de Pablos (2006), the researcher acquires the following model:

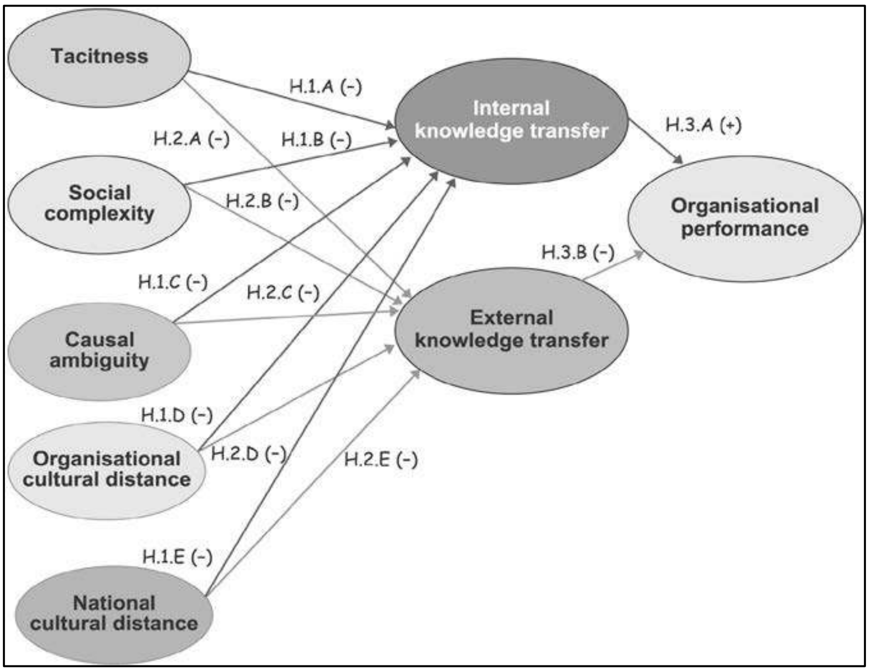


Figure 16: Model of Social Capital's impact on Performance (from de Pablos, 2006)

From Oh, Labianca, and Chung's (2006) paper on Multilevel Model of Group Social Capital, the researcher applies the following model of Social Capital:

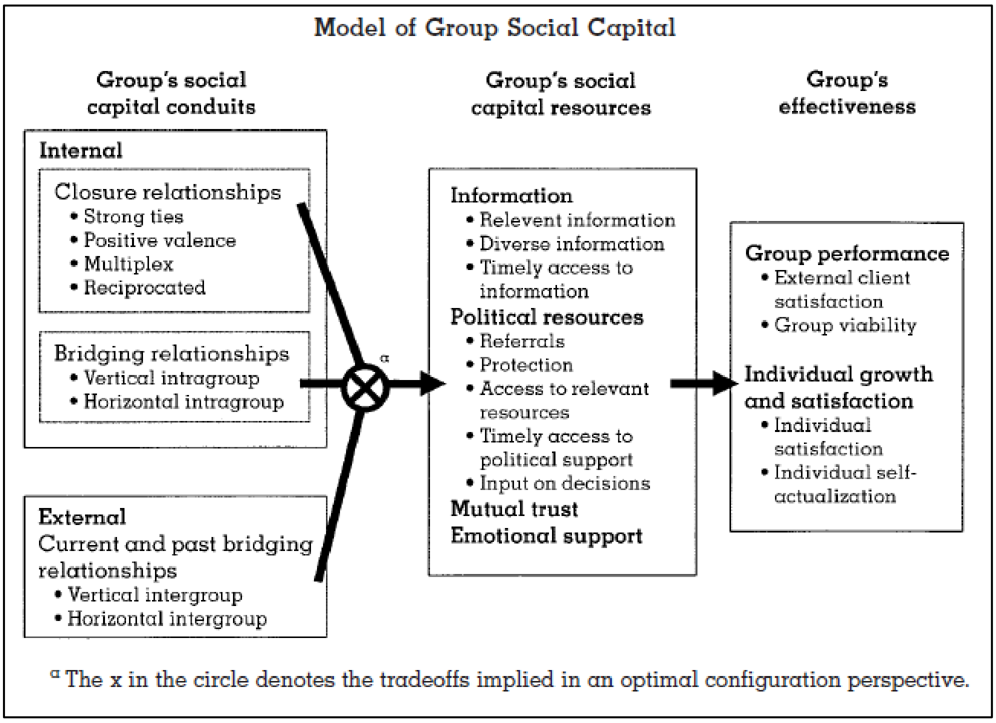


Figure 17: Multilevel Model of Group Social Capital (from Oh, Labianca and Chung, 2006)

From Mu (2008) the researcher learns that social capital, especially that which is rich in trust-based ties:

1. develops between nodes through an interaction process;
2. accelerates knowledge flow; and
3. acts as an informal governance mechanism between nodes.

Weak ties help groups to build initial relationships, and strong ties help groups to acquire higher-quality and fine-grained knowledge.

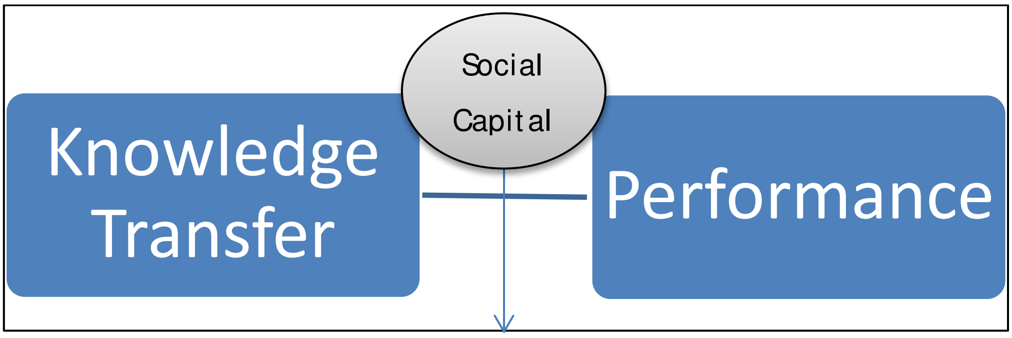


Figure 18: Model of Social Capital's impact (from Mu, 2008)

Pearson, Carr and Carr, in their 2008 paper, ‘Toward a theory of familiness: a social capital perspective’, amalgamate early research (Leana and Van Buren, 1999; Nahapiet and Ghoshal, 1998; Tsai and Ghoshal, 1998) to generate a model linking the dimensions of social capital (and social networks) to capabilities:

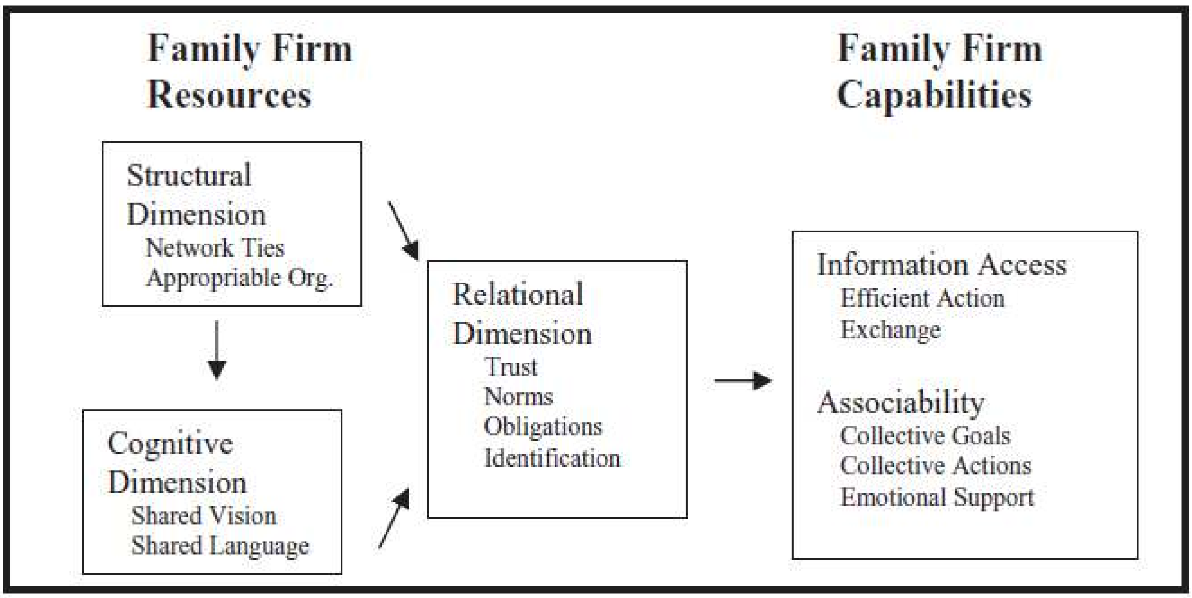


Figure 19: Model for Group Social Capital (from Pearson, Carr, and Carr, 2008)

Li and Zhu (2009) proposed the following model; in doing so, those authors err in categorizing Tie Strength under the Structural dimension, whereas the majority of research classifies Tie Strength as an indicator of the Relational Dimension of the social network (e.g. Lechner, Frankenberger and Floyd (2010):

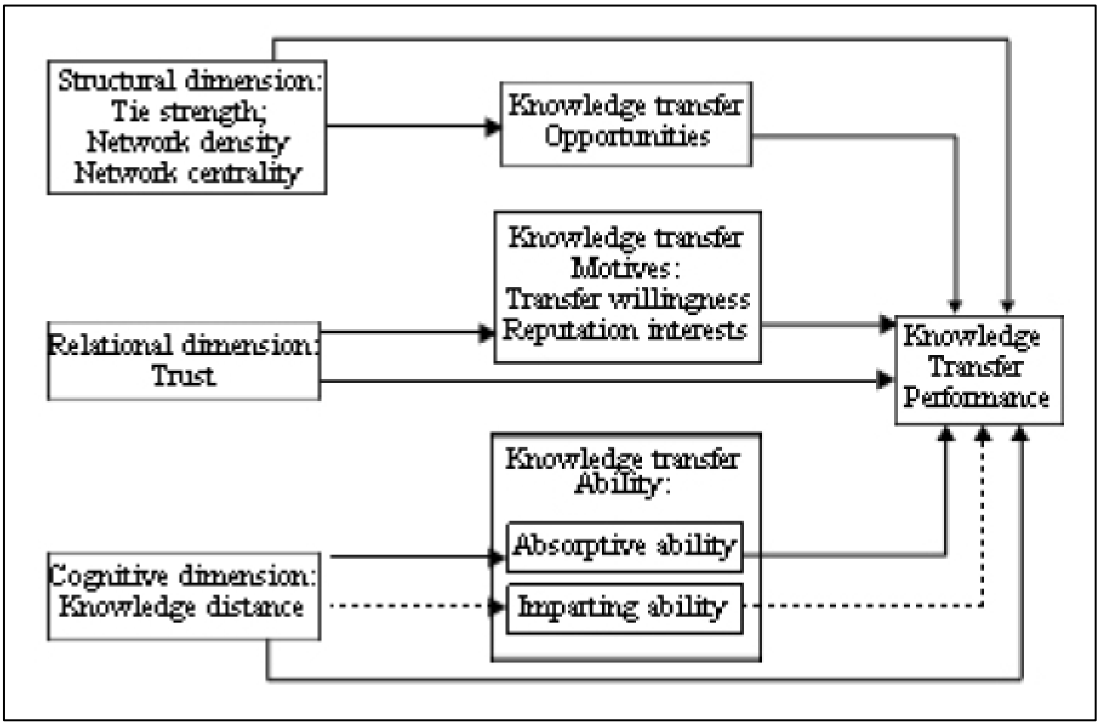


Figure 20: Model of Social Capital's impact to Knowledge Transfer (from Li and Zhu, 2009)

More recent findings have found that knowledge transfer mediates between organization members' intra-organizational social capital and organizational performance outcomes of growth and innovation performance (Maurer, Bartsch and Ebers, 2011).

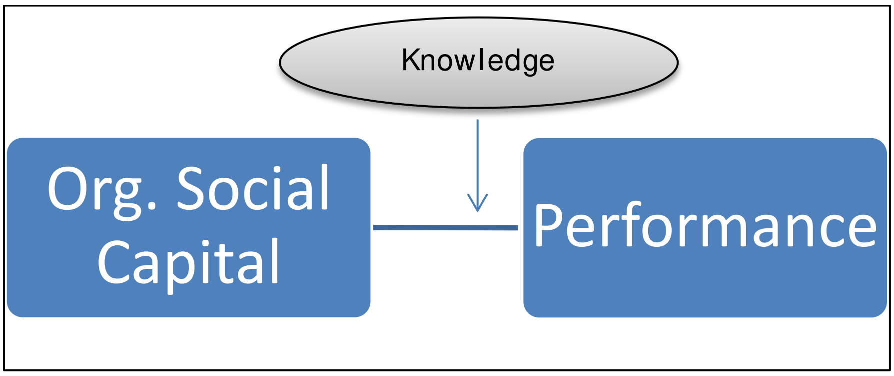


Figure 21: Knowledge Flow model (from Maurer, Bartsch and Ebers, 2011)

While these prior models offer insight, the next step is to aggregate the conceptual model, previously introduced above, into a manner which relates the structure of the group's social network to organizational performance.

4.4 The Model for this Research

The aforementioned models provide a valuable introduction into our particular logic, yet they focus heavily on the knowledge processes external to the group. As this research concentrates primarily on the internal knowledge processes which impact the group's ability to transfer knowledge (internally) and perform as reflected by the network structure of the group, the model adopted by this research looks directly at the effects of network topography on group performance. The conceptual model which forms the logical framework for how network topology impacts organizational performance is given below. This model outlines the causal relationships between the different factors, identified in management literature, which are known to impact the transformation of knowledge into performance.

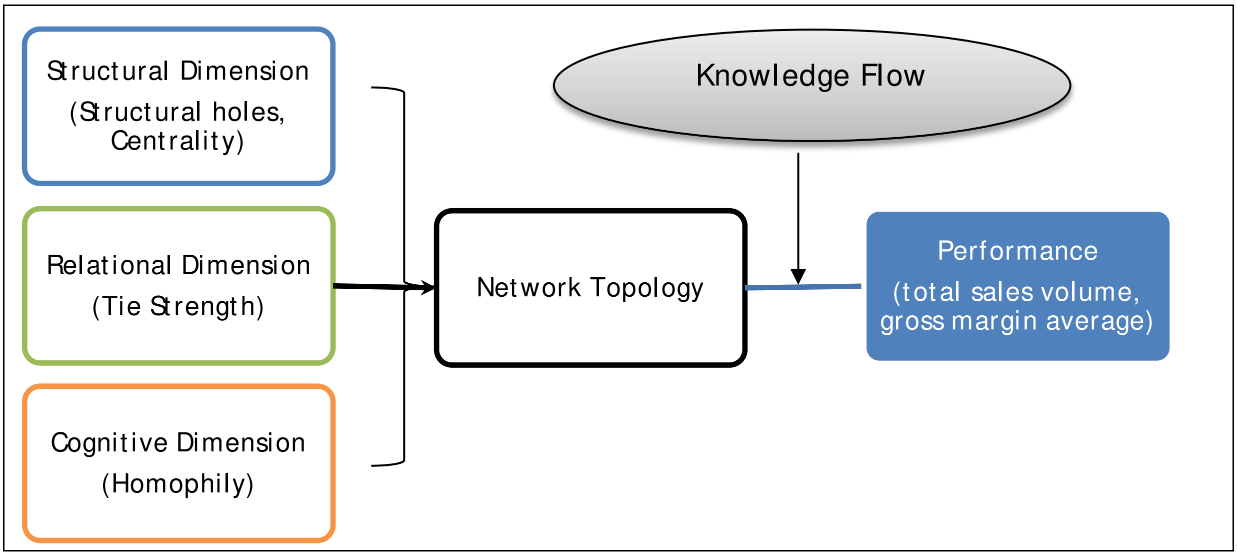


Figure 22: The Proposed Research Model

Since the task (and task contingency) was held constant across groups (e.g. sell the most travel and the best margin), the researcher need not control for industry or country. Instead, the control variables were:

- Is the group Corporate? Yes = 1, No = 0

- As not all groups have the same number of travel agents, the number of full time staff.

It is clear from this model that any knowledge-specific variables have been omitted. However, this is warranted, as one of the underlying assumptions of this research is the notion that a similar nature of knowledge exists in similar groups. That is to say those groups in a given industry (i.e. Travel sales) would have comparable tactility and complexity of knowledge being transferred across their organizations and thus can be omitted.

4.5 Measuring Performance

In its broadest definition, performance reflects the degree to which an outcome has met expectations. In the context of a modern corporation, the degree to which the group performs reflects the degree to which management's execution has met the expectations set by the shareholders. As such, the shareholders' expectations are dependent on the group's resources and its potential to exploit those resources to economic ends. Given that expectations and group performance are highly contextual and dependent on the group's specific resources, it sometimes proves difficult to compare groups that possess different resources. Similarly, the nature of the group's business and the sector in which it operates define the group's expectations. In the case of HF, all groups being measured have similar context, task contingency and resources.

Performance is the dependent variable in this research. To measure performance, the researcher first collected the Total Sales per Person and individual Gross Margin data from the HF Senior Management. Then the researcher aggregated to the data to generate Total Sales Volume and Average Gross Margin for each Office. Using Total Sales Volume, all groups were ranked 1 through 180. Finally, using standard techniques, a Normalized Revenue Per Group was generated as Nrev. Gross Margin is a percentage and thus is not required to be normalized. Instead, Gross Margin was averaged for each group; this was dubbed Gross Margin Avg.

4.6 Measuring Network Topology

Social Network Topology is the study of qualitative properties of Social Networks. To date, the literature has focused on three key dimensions: Cognitive, Relational and Structural (Haythornthwaite, 1996; Lechner, Frankenberger and Floyd, 2010; Maurer, Bartsch and Ebers, 2010; Uzzi, 1997). Historically, these dimensions have been converted to measures that empirically explore the social network topology using social network measures, including Centrality, Density, Structural Holes, and Tie Strength. These measures and their application to this research are presented in the following sections.

4.7 Centrality (a measure of the structural dimension)

Centrality is the concept of being 'in the thick of things'. Centrality has been used in social network analysis to determine the degree to which a given actor is 'important' within a network. Several measures have been derived from this definition of centrality, degree of centrality, closeness centrality, betweenness centrality, eigenvector centrality, information centrality (Ni, Sugimoto and Jiang, 2010). These metrics have been used in recent studies as a means of quantifying the flow across a network (Borgatti, 2005). The different measures of centrality reflect slightly different network phenomena; however, each measure of centrality allows us to perceive how 'central' given actors may be within a network. Three centralities are of interest:

- **Eigenvector Centrality** is the measure of the influence of a node on the network. Thus it is the influence that any one group member (Travel Agent) can have on the group (Office). A node with high eigenvector centrality will be able to strongly influence other members of that group.

- **Closeness Centrality** determines the distance between the nodes. In mapping social graphs there is a natural distance between pairs of nodes. This distance (farness) is defined by the length of the shortest path to connect them. The distance of a node is calculated as the sum of all the shortest paths. Closeness Centrality is the inverse of Farness. It is often regarded as a measure of how long it would take to spread information along the shortest paths. Since this research focuses on the performance benefits resulting from the spread of tacit knowledge (e.g. best practices), Closeness Centrality is an insightful dimension in this research.

- **Betweenness Centrality** refers to the extent to which a node (representing an actor) lies between other nodes in the network. This measure takes into account the connectivity of the node's neighbours, giving a higher value for nodes which bridge clusters. Betweenness centrality may be the most appropriate value to measure in this research, as it reflects the number of people with whom a person is connecting indirectly through their direct links.

All three measures of centrality tend to be defined as an individual measure and not as a group measure. Most centrality metrics are calculated on ego networks (not whole networks, as with this research) to generate the centrality of the individual. At the turn of the century, a triad of authors created algorithms and methods to calculate group-wide centrality based on individual centrality. They did this by looking at a subset of Whole Networks called Weighted Networks. A weighted network is a whole network in which ties are not just either present or absent, but have some form of weight attached to them. The weight represents the tie strength (relationship) between the actors connected by it. Opsahl, Agneessens and Skvoretz's (2010) approach makes it possible to gauge Group level Centrality through a hybrid methodology which combines the various metrics used in the current research. Thus, Group Centrality was calculated by first aggregating each Group's individual centrality measures and then dividing by the number of individuals in the group. This yields the following measures:

- Avg. Eigenvector Centr.
- Avg. Closeness Centr.
- Avg. Betweenness Centr.

A group with higher Avg. Centrality will be seen as less decentralized (Hui, 2008). Decentralization facilitates innovation better than exploitation (Sahay, 2011). **HF** Groups are mostly focused on exploitation, getting the most from the assets on hand. Based on this, one would predict that groups with higher average centrality would have an easier time facilitating knowledge flow and would be in turn better able to drive higher performance.

4.8 Structural Holes (a measure of the structural dimension)

Network cohesion is a structural measure of a social network which reflects the degree of redundancy occurring within a group. That is to say, the number of redundant ties (paths between actors) within a network represents network cohesion (Burt, 1992). If a network is cohesive, then it can better tolerate actor defection. Network cohesion (sometimes called network redundancy) has the potential to affect the knowledge processes of a group (Fritsch and Kauffeld-Monz, 2009) and, as such, it is of interest to knowledge transfer.

Network cohesion is a metric reflective of the entire network and thus must be calculated on a group-wide level. This has been done in past studies through empirical survey-based social network analysis (Burt, 1992). The measurement of network cohesion also allows one to account for structural holes occurring in the network. Thus, through the use of this metric, it will be possible to identify the presence and frequency of structural holes within a group. Structural holes are disconnections between nodes in a social network (Ahuja, 2000). The theory of network cohesion is often operationalized as either Density or Structural Holes.

Density is the inverse of Structural Holes (Zaheer and Soda, 2009) meaning a group with 100% density will have no structural holes. To generate a measure of Structural Holes, the researcher calculated: 1/Density. Structural holes can lead to the arrival of non-redundant knowledge to the network (Rodan, 2010); however, with too many structural holes, it will be difficult to diffuse innovations (e.g. best practices) throughout the group. The researcher predicts that top performing groups will have fewer holes than low performing groups. Thus, the inverse is true—top performing groups should have higher Density than lower performing groups.

4.9 Tie Strength (a measure of the relational dimension)

A social network is made up of actors. These actors have relationships. These relationships are described as 'ties'. Ties are often naturally associated with a strength that differentiates them from each other. **Tie strength** has been operationalized as weight. In a social network, the weight of a tie is generally a function of duration, emotional intensity, intimacy, and exchange of services (Granovetter, 1973). Barrat, Barthelemy and Pastor-Satorras (2004) generalize degree centrality to weighted networks by taking the sum of weights instead of the number of ties, while Brandes (2001) and Newman (2001) utilize Dijkstra's (1959) algorithm of shortest paths for generalizing closeness centrality and betweenness centrality to weighted networks, respectively.

While there are methods to quantify the number of connections in a network, it is also important to consider the strength with which those connections interact. The degree to which two ties have a 'strong' or 'weak' bond cannot be wholly attributed to the frequency with which they interact. Frequency of contact may have some correlation with tie strength, but it cannot serve as an all-encompassing substitute for tie strength. One can easily imagine a relation where there is great frequency of contact but little tie strength; for example, the relationship between the researcher and the coffee barista who works at Starbucks and provides the researcher with his daily dose of caffeine. They interact daily and frequently, but there is no emotional intensity, intimacy or duration, so despite its frequency, the tie strength between the barista and the researcher would likely be weak. Notwithstanding this matter, recent literature has accepted this limitation and forged ahead using frequency (i.e. edge weight) as a proxy for Tie Strength (Pepe, 2011). The reasoning lies in the following correlation: If A and B have a strong relationship, one where tacit knowledge transfers through face to face interactions, some coordination is still required (i.e. to set up the face to face meeting). Based on this, it seems reasonable to conclude that if A and B have a high frequency of email, they may have a strong relationship; based on the inverse, if A and B have a weak tie relationship, they would be unlikely to have a high frequency of interaction.

Tie Strength indeed reflects an established, working relationship, it thus would seem logical to imply that trust exists within that relationship (i.e. how could two people have a strong relationship without trust?). This research focuses on intragroup ties and small groups, to generate a measure of Tie Strength, researchers first calculated the potential number of ties (relationships) amongst each group. This was based on the well accepted formula (Yuan, 2010):

$$\text{Max ties possible} = \frac{n(n-1)}{2}$$

Thus, a group with 5 FTEs has the potential for 10 relationships, some of which may be STRONG and others WEAK.

Contextually, each Office is made up of 6-16 geographically proximate travel agents. Typically an Office is less than 300 square feet. Thus, all members of that Office are in close proximity. One would expect that such close proximately facilitates low-cost face to face communication. In fact, one can assume that face to face communication would be the dominant form of communication in the Office, with email being used mostly to arrange meetings asynchronously or to share explicit information (e.g. a new incentive plan). Thus, in this specific context, the researcher expects email to be deployed only weakly for communication; this pattern is consistent for all Offices at **HF**, which allows for valid comparisons. For this reason a low frequency (10 emails per year) was set as the threshold for strong ties. From this logic, a weak tie was defined as having less than 10 edges (i.e. instances of communication) over one year. As for setting the bar at 10 edges, one must remember that this number does not reflect the quantity of email sent, only the quantity of intra-organizational (e.g. between office mates) mail sent. To determine where to set the bar, the researcher examined the frequency of email distribution, looking to set that bar at a level that would ideally encompass a meaningful set of relationships. The NodeXl® software package facilitated this visually and the tenth email was selected at the point when a tie went from weak to strong.

A strong tie is thus defined herein as any tie with an edge weight of 10 or greater and a tie with higher edge weight is deemed stronger than the tie based on less frequent contact. The final step was calculating a measure of Tie Strength derived by dividing the number of actual Strong Ties by the Number of Potential Strong Ties.

$$\text{Group level Tie Strength} = \frac{(\#\text{ of Strong Ties})}{(\#\text{ max potential strong ties})}$$

Examining the concept in practice:

- Ct Central has 6 FTEs during 2011.
- The maximum number of potential strong ties is calculated to be 15.
- The actual number of strong ties (those with edge weights greater than 10) detected was 13.
- The Tie Strength measure for HF Central = 13/15 = 0.87

Strong ties are required for tacit knowledge to flow (Hansen, 1999; Levin and Cross, 2004; Li and Zhu, 2009; Nie, 2010). Therefore, **HF** Offices with more strong ties will be better able to share best practices. The researcher predicts that groups lacking strong ties (over which that can flow) will not be able to gain efficiencies from best practices since these practices will be harder to disseminate. Since **HF**'s task contingency is mostly exploitive in nature, one predicts top groups will be dominated by a majority of strong ties.

4.10 Homophily (a measure of cognitive dimension)

Shared vision and mutual values facilitate a common understanding (Tsai and Ghoshal, 1998). Shared vision and systems promote mutual understanding amongst actors and may provide a crucial bonding mechanism that helps actors integrate knowledge (Tsai and Ghoshal, 1998). It is important that network actors have a shared vision; otherwise a 'lack of shared vision' may arise as a barrier to the transfer of knowledge. Shared vision refers to the clarity and coherence with which all network actors understand and embrace their organizational goals. A disparity in vision between network actors can impede the exchange of knowledge and ultimately impede performance (Inkpen and Tsang, 2005). This dimension aims to assess the degree to which all network actors understand what their organizational (high level) goals are in an explicit sense. As such, this barrier reflects the cohesion of vision and goals at the most basic level. In order to measure shared vision, a survey would need to be deployed, but would be counter to the goals of this research. Shared vision looks for consistency and congruency on group level goals and strategies. Cultural distance looks for consistency and congruency between actors in a group. This is the fundamental concept of Homophily.

The notion of homophily is well-known in network analysis. Homophily assumes that similar nodes are more likely to be linked together. It is based on Social Identity Theory (Pratt, 2001), which acknowledges that it is in our nature to be drawn to those who are like ourselves (Brass, 1995). As a result, like seeks like, and like works more efficiently with like. For instance, two engineers in Silicon Valley who graduated from the same school in Bangalore may have relatively short cultural distance, and as result they would have a greater ability to transfer knowledge. As cultural distance grows (say between an engineer and a graphic designer) the flow of knowledge may become more difficult. Cultural distance increases the cost of entry and hampers a group's ability to transfer core competences (Palich and Gomez-Mejia, 1999).

In some papers (e.g. Cillo, 2005; Nooteboom, Vanhaverbeke, Duysters, Gilsing and Van Den Oord, 2006) cultural distance is referred to as 'Cognitive Distance,' which is then is defined as the discrepancy in the frames of reference between two or more people involved in the exchange of knowledge manifested in the different cognitive focuses, such as perspectives, norms of conduct and more technical capabilities). In both cases these (Cognitive Distance and Cultural Distance) measure the cognitive dimensions of network topology.

Homophily refers to the tendency for people to interact more with their own kind, whether by preference or induced by opportunity constraints (McPherson, Smith-Lovin and Cook, 1987), as defined by such individual characteristics as race, gender, educational class, organizational unit and so on. More recently, organizational research on homophily has focused on its effects on group and individual performance outcomes (e.g. Reagans and Zuckerman, 2001).

On the positive side, interacting exclusively with similar parties is thought to be efficient to the extent that similarity:

(a) facilitates transmission of tacit knowledge (Cross, Borgatti and Parker, 2001, p. 229);

(b) simplifies coordination (Ancona and Caldwell, 1992; O'Reilly, Caldwell and Barnett, 1989); and

(c) avoids potential conflicts (Pelled, Eisenhardt and Xin, 1999; Pfeffer, 1983).

On the negative side, limiting communication among similar parties prevents a group from reaping the benefits of diversity and promotes us-vs.-them thinking (Krackhardt and Stern, 1988).

Cross, Borgatti and Parker (2001) lay the groundwork with their finding that homophily facilitates transmission of tacit knowledge. At **HF**, little exploration is required. Inside agents attempt to maximize and exploit opportunities to boost performance. In the case of **HF**, the transfer of best practices (e.g. tacit knowledge flow) is at the heart of such exchanges. The ease transfer should be enhanced when Offices are homogenous. Therefore one would expect **HF** Offices with high homophily to be high performers. Tacit knowledge flow requires shared mental constructs and less cognitive distance (Clark, 2011). So the research expects that Offices with higher Homophily will be able to transfer best practices (e.g. tacit knowledge) more easily.

To calculate homophily, one may examine the homogeneity of communications. If wider mental constructs and more variable language are deemed to be present, the network will be said to be low in homophily and high in cognitive distance. By using DICTION software (http://www.dictionsoftware.com/) it is possible to ascertain the degree to which individuals communicate and interact using the same language. This is possible because individuals exhibiting greater homophily will tend to communicate in similar ways. As a result, the difference in language (used in email) between various individuals in the network should indicate to what degree they exhibit shared vision. Notwithstanding the value of the above, the senior management of **HF** requested that the researcher not access the content of the 7 million emails sent in 2011. Alternative methods of measuring a group's level of homophily would require access to relevant human resource information (e.g. age, sex, academic background, culture, country of origin, etc.).

Unfortunately, **HF** did not wish to provide the researcher with this data. Hence, the researcher was not able to measure Cultural Distance nor Homophily, two widely accepted measures of the Cognitive Dimension of Social Networks. Because neither of the accepted measures of the Cognitive Dimension of Social Networks was available to the researcher, the Cognitive Dimension was excluded from this research.

4.11 Onboarding Speed

Firms that are more successful at rapid onboarding tend to use a relational approach, helping newcomers to rapidly establish a broad network of relationships with co-workers that they can tap to obtain the information they need to become productive (Rollag et al., 2005). Researchers have articulated that social ties have the potential to facilitate the flow of all kinds of resources within teams, which correspondingly determines the success of those teams (Balkundi and Harrison, 2006). In addition to exploring the relationship between performance and social network topology, this research additionally examines the concept of dynamic tie formation. Knowledge sharing and application are widely recognized as the key determinants of team performance (Choi et al., 2010; Janhonen and Johanson, 2011).

Tacit knowledge travels over strong ties (Hansen, 1999; Levin and Cross, 2004; Li and Zhu, 2009; Nie, 2010). Tacit knowledge enhances performance (e.g. as best practices are shared, individual performance grows). The researcher suggests that those Offices who are able to form strong ties faster will be able to benefit from the ability to transfer tacit knowledge earlier. This is particularly informative with regard to onboarding. During Onboarding, new nodes (FTEs) are added to the network. Those nodes form ties with the other members of the Office. The quicker those ties become strong, the quicker tacit knowledge can flow.

Earlier, it was decided that a tie is deemed strong upon the tenth interaction (i.e. email) between the nodes on the vertex. The primary data (i.e. edge list) provided by **HF** includes all dated communications over the 365-day period from January 1, 2011, to December 31, 2011. To measure onboarding, the following steps were undertaken:

- All emails sent by and sent to the new employee are aggregated.
- These emails are sorted by date sent.
- The date of the tenth intra-Office interaction is noted.
- That data is converted to a number (e.g. Jan 1st is #1, Jan 2nd is #2, Dec 31st is #365).
- That number is subtracted from 365.
- This yields the number of days that such strong ties existed during 2011.
- This value was deemed the measure of Tie Formation.

This approach was leveraged from the earlier work of Kenis and Knoke (2002) in 'How organizational field networks shape inter-organizational tie-formation rates'.

4.12 Control Variables

Size

Many studies have included size as an organizational factor that may impact knowledge flow. Most studies which use size as a control variable find a positive relationship between size and knowledge flow (Gupta and Govindarajan, 2000; Laursen and Salter, 2005). It should be noted that other studies have not found size to have a positive impact on knowledge flow. Tsang (2002) finds size and knowledge flow to have no correlation, while Makino and Delios (1996) find size to have a negative impact on knowledge flow. To calculate size, the researcher will count the number of Sales Agents (**FTEs**) in each office.

Corporate Status

While all Offices undertake the same sales goals, the Offices do not all share the same client approach. Approximately 15% of **HF** Offices are Corporate, which means that FTEs book commercial travel for pre-existing B2B customers. The remaining offices are Retail; at these offices, FTEs book retail travel for those who walk into the retail storefront. With this in mind, the researcher expects that Corporate Offices will have larger gross sales and larger margins. Some of the reasons for this are as follows:

- HF Corporate have a greater Elasticity of Demand and are thus not as price sensitive, allowing Agents to increase the margins.
- HF Corporate are not paying their travel with personal funds. This may make them even less Price Sensitive.
- HF Corporate offices have fewer clients, but more repeat business. Retail clients tend to book less frequently than corporate clients. This makes it harder for Agents to know the client's sensitivity to margin.

Based on the above, the researcher predicts that, based on RANK, more Corporate offices will be in the Top 10 than Retail offices. Similarly, the research predicts that based on RANK, more retail offices will appear among the 10 lowest performing groups.

Degree of Exploration (a control variable)

Degree of Exploration is a measure that represents the level of exploration involved in the proper execution of a task. Every task has both explorative elements and exploitive elements. For example, designing a new form of brake pads is very explorative (i.e. innovative tasks that require novelty), while painting a house might be seen as primarily exploitive (i.e. routine tasks that involve scale). One can argue that all professions (e.g. house painter, lawyer, travel agent) are made up of explorative and exploitive tasks, which when normalized (over a large sample) become constant (e.g. each house painter needs 10% creativity and 90% efficiency to be a high performer). Similarly, two pop stars trying to break into the music business would have similar task contingency. For pop stars performance is measured in overall record sales, which is dependent upon having success with both explorative tasks (e.g. writing songs) and exploitive tasks (e.g. touring and singing the same songs).

At **HF**, all the Offices (are filled with sales agents who are pursuing the same performance goals (i.e. Total Sales Volume, Average Gross Margin) and the same task contingency; meaning that a member of Office 1 has the exact same job as a member of Office 99. In all cases, selling travel requires the same explorative/explorative balance of tasks.

For this reason, the researcher is confident that all groups have the same task contingency.

4.13 Summary of variables considered

The range of variables considered in this research and their academic foundation is outlined in the following figure. Performance measures are the dependent variables. Structural and Relational dimensions are the independent attribute variables. Number of Employees, Task Contingency, etc. are control variables.

Table 9: Variable Summary

	Definition and Source	Measure	Corr. With KT	Authors
Performance	maximizing available group resources in economic exchange	**Gross Sales Volume**	+	Raymond Van Wijk, Justin J. P. Jansen, Marjorie A. Lyles, 2008
		Gross Margin	+	Raymond Van Wijk, Justin J. P. Jansen, Marjorie A. Lyles, 2008
Organizational Controls	contextual factors relating the group itself	**Size**	+	Dhanaraj et al., 2004 Gupta and Govindarajan, 2000 Laursen and Salter, 2006
			no effect	Tsang, 2002
			-	Makino and Delios, 1996
		Decentralization	+	Gupta and Govindarajan, 2000
Network Characteristics	**structural dimension: place in network**	**# of relations / ties**	+	Gupta and Govindarajan, 2000
		centrality	+	Tsai, 2001; Ahuja, 2000
		Network Cohesion	+	Fritsch and Kauffeld-Monz, 2009
	relational dimension: nature of relations	**tie strength**	+	Argote, Reagans and McEvily, 2003
	cognitive dimension: shared understanding	**shared vision/ system**	+	Inkpen and Tsang, 2005 Lane et al., 2001

4.14 Variables used

Based on the above, the following variables will be collected:

Table 10: Variables Used

Name	Independent/Dependent	Acronym
Normalized Revenue	Dependent	nrev
Structural Holes	Independent	holes
Tie Strength	Independent	ties
Average Eigenvector ctr	Independent	eig
Average Closeness ctr	Independent	clos
Average Betweenness ctr	Independent	Betw
Full time equivalents	Control	Fte
Office Type	Control (0=corp, 1=retail)	type

4.15 Population/sampling

For this research the population and sample are the same. Over 7 million edges were collected based on all incoming and outgoing **HF** mail during 2011. The researcher thus had all intragroup emails for all Offices at **HF** (including Central). As our sample equals our population, the sampling error would be 0.

4.16 Analysis

Data handling will be described in Chapter 5. The primary data for this research is derived from the edge list provided. This list, which includes each piece of mail sent through the **HF** servers over 2011, records the sender and receiver for each email. Each instance of sender and receiver is an 'edge' and duplicates are counted as 'edge weight'.

For example: If Mr. A emailed Mrs. B 100 times over 2011, the data would show 100 instances of A → B, represented as A → B with an edge weight of 100.

Data Analysis will be handled in two distinct rounds. Round 1 will see the weighted edge list for each Office (group) entered into NodeXL as a standalone network. This will generate SNA metrics for that Office. This will be repeated 188 times (once per Office) and generate SNA metrics (Density, Centrality, Tie Strength, etc. for all Offices. Round 2 of Data Analysis will leverage the data sheet output from Round 1. Standard quantitative analysis will then be deployed. Alternative analysis (e.g. structural equation modelling, visual analysis) will also be deployed. These methods will be explained in Chapter 5.

4.17 Pilot Study

The researcher conducted a full pilot study with a manufacturing company prior to launching this study. The company used in the pilot study manufactures various types of ribbon and finishing for high performance outerwear. Employees work in Toronto, Canada, or Buffalo, USA. In total the researcher collected email from the group's email servers dating January 1, 2011, to December 31, 2011. In total, 101 email accounts were tapped and more than 1.1 million emails were reviewed.

The result, shown in the figure below, reveals four cliques around which most staff were strongly associated. The social graph also identified several prominent structural holes. The researcher then discussed these findings with management. During the post hoc interview, the following interesting confirmatory facts were found.

- Most employees work in one of four general areas: Manufacturing, Admin, Sales or R&D. These groups matched the four cliques identified visually.

- The tightest (largest average centrality) group is the manufacturing team. Management suggested this was because each member of that team was hired through internal recruitment (e.g. on direct employee recommendation). During the visit, it was interesting to note that the entire manufacturing team comes from the same ethnic background, lives in the same sub-community, and all display large homophily as a result.

- Structural holes identified visually were confirmed. Management confirmed that there were holes in their network; for example, sales reps had little contact with R&D staff.

Based on these results, no changes were implemented post pilot.

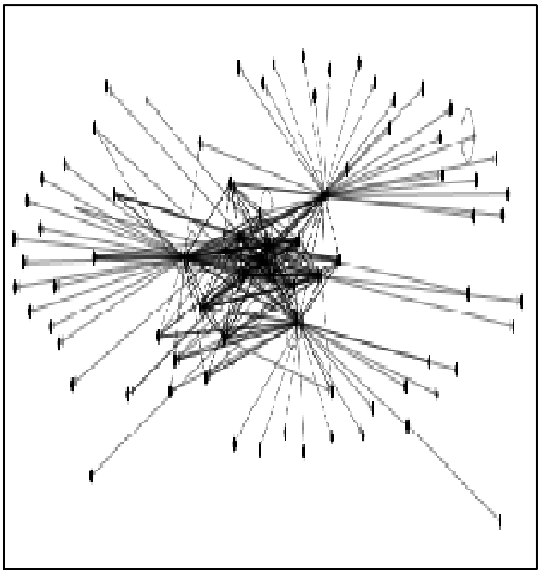


Figure 23: Social Graph of Pilot group

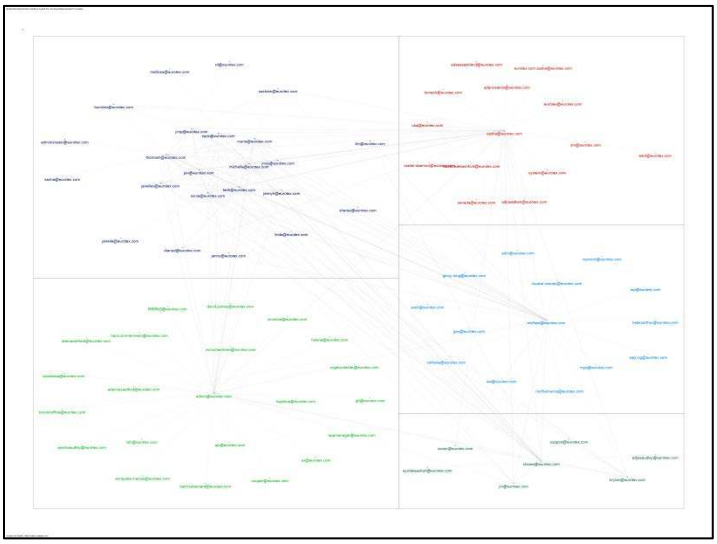


Figure 24: Social Network of Pilot group

4.18 Limits to testing/analysis

These research results cannot be generalized beyond the particular context in which they were generated. As with Inkpen and Tsang (2005), it is likely that different types of organizational settings may display unique relations among the various dimensions of social capital, knowledge transfer, and performance outcomes.

Summary of Chapter 4

This chapter explored the link between knowledge flow and performance. The researcher discussed a variety of social network measures and how the author intends to explore the moderators of performance. Finally, the chapter presented a conceptual model and RQs which can be derived from the model. In this chapter, the principles introduced in the literature were contextualized within this research and the groundwork for the methodology was laid; however, Chapter 5 will further articulate the specific methodology adopted by this literature review. The following chapter explores the issues, choices and background behind this study's methodology; additionally, it sets the definitions for the measures used and outlines how those measures will be collected.

Chapter 5 – Data

Introduction

Now that the phenomenon of interest (i.e. how social capital predicts group performance) has been converted to research questions and the research measures operationalized, it is time for data collection. In this chapter, the researcher will describe the specific methodology adopted for this research. The issues, choices and background behind this study's methodology will be justified and measures will be defined and explained as they are operationalized.

5.1 Measures

The following variables were used in this research:

Table 11: Variables Used - Measures

Name	Independent/Dependent	Acronym
Normalized Revenue	Dependent	nrev
Structural Holes	Independent	holes
Tie Strength	Independent	ties
Average Eigenvector ctr	Independent	eig
Average Closeness ctr	Independent	clos
Average Betweenness ctr	Independent	Betw
Full time equivalents	Control	Fte
Office Type	Control (0=corp, 1=retail)	type

5.2 Data Collection Process

The data were processed in four steps:

1. Organizing the data
   a. Received data from IT provider.
   b. 250MB of edge lists representing 7 million emails.
2. Filtering the data
   a. First filtered to exclude all mail which does not have an **HF** URL email address in both the SENDER and RECEIVER columns of the edge list. This left only the internal mail of **HF** which is mail sent from an **HF** FTE to another **HF** FTE.
   b. Duplicates were rolled up into a Weighted Edge List.
3. Obtaining the SNA metrics.
   a. A NodeXL workbook was created for each Office (group). Placed in each were Edges where both the SENDER and RECIEVER were from the same Office.
   b. This created subgraphs for each Office and generated SNA metrics for that Office.
4. Aggregating the results
   a. Several offices were removed from the Data sheet. These offices sold only cruise travel, and mostly sold such on behalf of other offices. The researcher concluded that this task was not identical to the task being carried out by other **HF** Offices that sell retail travel through storefronts to walk-in clients, or by other **HF** Offices that sell corporate travel to pre-established business clients.
   b. SNA metrics for 188 Officers were created and aggregated into a data summary sheet.

Using Research Assistants and Elance.com

Three research assistants ('RAs') were hired to assist the researcher with the coding, filtering and amalgamation of the raw data. Each was given a zipped file with the raw data and written instructions regarding how to assemble, clean and organize the data. To address the issues concerning interoperate viability, the researcher compared two KPIs from the RAs' output, specifically Number of FTEs and Number of Offices Found. Their results were highly congruent with two of the three assistants generating identical outputs. Research assistants B and C then went on to assist with assembling the edges and vertices into standalone workbooks (one per office). The researcher loaded the edges into NodeXL and calculated group metrics. Using Research Assistants to code large amounts of data is not unusual; what was unusual was the method of recruitment of research assistants. The researcher used www.elance.com to hire assistants, as the service offers affordable support. These assistants were tasked with the mechanical routines (e.g. counting vertices). RAs were also leveraged during the Analysis portion of this research for similar purposes.

5.3 SNA Data and Graphs

The raw data provided by the IT providers of **HF**, included:

- 7 million emails
- Individual Annual Gross Revenue
- Individual Average Gross Margin

An edge list was created based on the emails provided. These edges were collected by Office and subgraphs (one for each group) were generated. For each subgraph (Office) NodeXL generated:

- **Density** ➔ The density of a graph G = (V,E) measures how many edges are in set E compared to the maximum possible number of edges between vertices in set V. If a group has complete density (i.e. 1) there are no structural holes.

- **Eigenvector Centrality** ➔ The measure of influence a Vertex (node) has on the network. This measure of Centrality takes into account not just the number of connections, but the number of important connections.

- **Closeness Centrality** ➔ This measure of Centrality represents how fast information spreads amongst the network sequentially. Closeness is calculated as the inverse of Farness. The farness of a node is defined as the sum of its distances to all other nodes and its closeness is defined as the inverse of the farness (Sabidussi, 1966).

- **Betweenness Centrality** ➔ The measure of the control of a human on the communication between other humans in a social network (Freeman and Linton, 1977).

- **Average Geodesic Distance** ➔ Geodesic Distance is the shortest distance between two vertices in a graph. The measure calculates the number of edges in the shortest path connecting them as defined by Bouttier, Di Francesco and Guitter (2003). This measure is the average for the entire Office.

The researcher also obtained a visual subgraph for each **HF** Office, similar to those shown below. The spheres are vertices which each represents an FTE. The relative size of the sphere visualizes centrality (e.g. big spheres have more influence in the group than small spheres). By comparing the size of spheres relatively, one can see how the centrality is balanced across the networks. The thickness of the edges (lines connecting Vertices) represents the edge weight (number of emails sent). The number on the edge represents its weight. Only edges of ten or more are displayed. Total Strong Edges is calculated as the sum of the numbers on the edges in the subgraph.

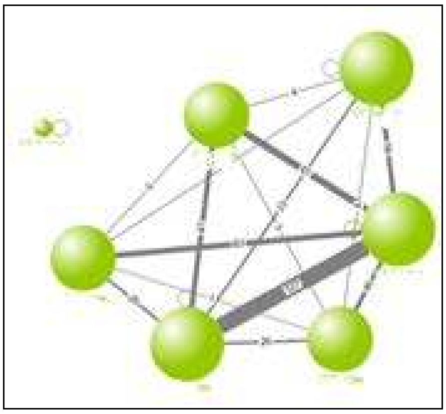


Figure 25: Subgraph for HF Central Office

For a full catalogue of office Subgraphs please see Appendices B and C.

5.4 Creating Secondary Measures from Primary Data

Leveraging the literature and the primary data, the data were transformed in the following ways:

- **Total Strong Edges** (a strong edge was one with an edge weight of 10 or more), any vertices connected by ≤10 emails, were dropped and deemed 'non-strong ties'. This was necessary since, at Edge Weight of > 1, all edges would be deemed strong (i.e. the researcher had to set the bar at 10 or more to see any difference). Thus Total Strong Edges is the aggregate of the strong edge weights of the group.

- The researcher aggregated Individual Annual Gross Revenue to the Office level ➔ **Total Sales Vol.**

- The researcher counted the number of FTEs contributing to Total Sales Vol ➔ **# FTEs.**

- The researcher normalized the Total Sales Volume ➔ **Normalized Sales Vol.**

- The researcher aggregated Individual Average Gross Margin to the Office level and divided by #FTEs ➔ **Avg Gross Margin.**

- The researcher divided 1 by Density to generate → Structural Holes (the inverse of Density).

- The researcher took $\frac{(\#FTEs-1)*\#FTEs}{2}$ → **Potential Edges.**

- For **Ties** → Total Strong Ties / Potential Edges.

- **Avg group centralities** → calculated based on aggregating each group members individual centrality and dividing by the number of FTEs.

5.5 Aggregate Data

All primary and secondary data was aggregated to a large data matrix which can be found in Appendix A: Data Sheet. The first 25 rows of such are included below for illustrative purposes.

Table 12: Sample 25 Rows from Data Sheet

	FTE	RANK	Gross Revenue(TSV)	Normalized Rev	Rev/FTE	Gross Margin Avg.	Density	Structural Holes	Avg. Geo Distance	Avg. Eigenvector Centr.	Avg. Betweenness Centr.	Avg. Closeness Centr.	Total Strong Edges	Potential edges	Ties
E-Commerce	14	1	$19,668,280.01	$4.85	$1,404,877.14	6.94%	0.128	7.800	0.645	0.091	0.000	0.114	29	91	0.32
CT Pemberton	6	2	$14,069,885.00	$3.47	$2,344,980.83	14.93%	0.933	1.071	0.833	0.167	0.000	0.200	32	15	2.13
CT Central	6	3	$11,647,103.92	$2.87	$1,941,183.99	14.18%	0.600	1.667	1.040	0.200	0.600	0.197	13	15	0.87
CT Coal Harbour	10	4	$11,129,319.04	$2.74	$1,112,931.90	16.35%	0.768	1.302	1.031	0.125	0.625	0.124	50	45	1.11
CT Dundas	9	5	$10,786,718.20	$2.66	$1,198,524.24	13.81%	0.690	1.448	0.811	0.143	0.000	0.171	33	36	0.92
CT Bay St	8	6	$10,509,500.08	$2.59	$1,313,687.51	10.52%	0.767	1.304	0.000	0.167	0.333	0.179	26	28	0.93
CT Mission	6	7	$9,955,509.58	$2.45	$1,659,251.60	14.42%	0.650	1.538	1.040	0.200	0.600	0.199	16	15	1.07
Upper Canada Mall	6	8	$9,755,775.97	$2.40	$1,625,962.66	11.39%	0.357	2.800	1.081	0.143	0.714	0.133	16	15	1.07
CT Kensington	5	9	$9,632,926.33	$2.37	$1,926,585.27	12.81%	0.800	1.250	1.000	0.167	0.500	0.170	28	10	2.80
CT Marine	8	10	$9,563,512.72	$2.36	$1,195,439.09	13.52%	0.733	1.364	1.056	0.167	0.667	0.163	26	28	0.93
CT King	6	11	$9,539,009.55	$2.35	$1,589,834.93	18.03%	0.700	1.429	0.889	0.167	0.167	0.189	21	15	1.40
CT Burrard	7	12	$8,617,291.45	$2.12	$1,231,041.64	14.73%	0.667	1.500	0.800	0.200	0.000	0.250	23	21	1.10
Coquitlam	9	13	$7,629,396.94	$1.88	$847,710.77	10.89%	0.267	3.750	1.231	0.167	1.000	0.137	10	36	0.28
CT City Hall	8	14	$7,319,635.60	$1.80	$914,954.45	18.12%	1.000	1.000	0.800	0.200	0.000	0.250	22	28	0.79

Similar information at the individual level was obtained from the raw data for the 8 new hires which joined **HF** prior to the start of 2011. A tie was deemed strong once the frequency of 10 was reached and the date of this benchmark was recorded, and the remaining days calculated (defined as Days Strong Ties Available, DSTA). Theoretically, the earlier this happened, the longer each agent had the ability to access to tacit knowledge.

Since tacit knowledge includes best practices for selling travel, one hypothesis would be:

H_{ob} Larger DSTA (i.e. more days with strong ties in place) the larger Performance.

For the onboarding data, the following was found:

Table 13: Onboarding Data

Agent	2011 SV	2011 AGM	office	tie str	# of edges	in degree	out degree	tenth tie date	day #	DSTA
MA	$1,281,126.71	19.30%	ct central	strong	15	14	14	07/03/2011	66	299
GH	$1,383,797.97	17.57%	ct central	strong	19	18	21	01/03/2011	60	305
SP	$1,341,491.13	12.30%	ct dundas	strong	61	27	31	06/03/2011	65	300
KC	$932,329.64	13.39%	ct city hall	very strong	491	34	39	12/05/2011	132	233
MC	$1,048,394.62	15.93%	ct city hall	very strong	216	26	18	19/04/2011	110	255
NP	$1,016,079.91	18.47%	ct king	weak	6	12	13	n/a	0	0
HF	$1,249,760.30	23.00%	ct king	weak	5	24	24	n/a	0	0
DW	$755,218.96	19.98%	ct king	strong	10	21	31	27/05/2011	147	61

5.5 Analysis

From this data, multiple variant regressions will be carried out alongside a visual analysis. While such findings will not be sufficient to generate an inter-measure validity, it still may prove helpful to analyse the data using three separate and independent analyses and to compare and contrast the results of these.

Summary of Chapter 5

This chapter operationalized the specific methodology adopted for this research. The researcher explored the issues, choices and background behind this study's methodology as well as set the definitions for the measures to be used alongside an outline of how those measures will be collected.

Chapter 6 – Analysis

Introduction

At the conclusion of Chapter 5, a sample data sheet was presented. This research focuses on exploring the relationship between Performance (measured as Normalized revenue, Gross Margin, Rank, etc.) and multiple social network dimensions (e.g. Centrality, Tie Strength, etc.). Given that performance and most of the social network measures are continuous variables, and given that the researcher's epistemology, ontology and methodology guide the researcher down a positivist/critical realist path of quantitative deduction, a quantitative analysis is deemed most appropriate.

Prior literature informs the researcher that performance should be inversely curvilinear with relational and structural measure of social capital (Lechner, Frankenberger and Floyd (2010). Multivariate regression analysis will be pursued to establish the relationship between independent, continuous and dependant variables. The unified approach provided by General Linear Modelling (GLMs) is the starting point for the analysis.

Why GLM Regression? The researcher, being a critical realist, first wanted to start the analysis with a general approach. Traditional statistics approaches are based on teaching a number of disparate tests, for example, t-tests, ANOVA, MANOVA, MANCOVA, Mann-Whitney, Kruskal-Wallis, Friedman, etc. This form of statistical education is not ideal for a number of reasons:

1. It does not give a theoretically-unified method for data analysis;
2. It does not allow appropriate tests to be easily identified;
3. It assumes experimental designs and random selection;
4. One cannot easily add extra variables to the statistical tests; and
5. It does not provide a simple path.

The Generalized Linear Model (GLM) refers to a family of statistical models that extend the linear parametric methods (e.g. OLS), regression and analysis of variance, to data types where the response variable is discrete, skewed and/or nonlinearly related to the explanatory variables (Hutcheson and Sofroniou, 2006). GLMs are univariate models which predict the behaviour of one particular variable (Hutcheson and Moutinho 2012), in this case, Performance. GLMs (proposed by Nelder and Wedderburn (1972)) represent a family of statistical techniques that can be used to analyse a wide variety of research problems. They are sufficiently general to be applicable to much social science data and provide a comprehensive set of analytical tools (Hutcheson and Moutinho, 2011). GLMs enable building of descriptive and predictive models that are sufficiently general to be applicable to much social science data. GLMs can be used to model data collected from survey and experimental studies and can replace many of the more traditional hypothesis tests that are still in use (Hutcheson, Moutinho, 2012).

Of particular importance is the unified theoretical framework that GLMs offer, as this enables certain 'economies of scale' to be realised that allow a whole range of data to be analysed using similar techniques (Hutcheson, Moutinho, 2012). The use of the techniques will be described using a modelling procedure whereby a particular variable can be modelled (or predicted) using information about other variables. As theory suggests, the relationship being examined herein is curvilinear (not linear); as a result some transformation may be required prior to undertaking the GLM regression approach of ordinary least squares. A continuous response variable (one that can legitimately be described using the mean) can be modelled using ordinary least-squares (OLS) regression (Hutcheson and Moutinho (2012). Other advantages of using GLM techniques include:

- Adjustment for correlations between rating factors - traditional 'one-way' analysis is biased by such correlations.
- Multivariate methods allow for investigations into interaction effects.
- Produces statistics to allow testing of significance of rating factors, parameter estimates and model goodness of fit.
- Does not rely on subjectively selecting LDFs (and hence, possibly inserting bias into the results).

But analysis through GLM is not without its disadvantages which may include:

- In theory, any distribution function could be used as an assumption. In practice, the number of error distribution assumptions available is somewhat restricted.

- To the extent that distribution assumptions are inaccurate, the GLM will produce biased estimates. In these cases, the measuring statistics will be biased as well (if the bias is great, it will be obvious).

- Mathematics behind GLM are difficult to explain to most.

6.1 Performance may be predicted by measures of Social Network Topology.

To run a GLM, one only needs to identify the variable to model and the data that is going to be used to model it. If the variable being predicted is numeric, the GLM model is OLS (Hutcheson, Moutinho, 2012). For more on the advantages and disadvantages of using GLMs, please see Hutcheson, Moutinho (2012).

The researcher followed these steps to perform the GLM analysis: [1]

1. Identify the independent and dependant variables.

2. Input the data into statistics software.

3. Derive the regression equation/model.

4. Calculate and interpret the coefficient of determination.

5. Test the significance of the regression model.

6. Test the significance of the regression coefficients.

7. Test the independent variables for collinearity.

8. Plot the residuals against the value of ***y*** generated by the regression equation (i.e. test the observed data vs. the outputs of the regression model/equation).

9. Test the residuals for randomness.

[1] Derived from three sources: Introduction to business research, Vol 3. S. 5/25, Edinburgh Business School (2010); Hutcheson and Moutinho, The SAGE dictionary of quantitative management research (2011, p. 132); Hutcheson and Sofroniou, The multivariate social scientist (2006).

10. Revise the regression model by adding or deleting variables.

11. Repeat the analysis with revised model until *best fit* model derived.

12. Write up results.

Following GLM, visual analysis was explored in order to give colour and context to the findings. From the early days of SNA, images of networks have been used both to develop structural insights and to communicate those insights to others (Freeman, 2000). Social networks are inherently visual in nature. Visual analytic tools and techniques have been used in social network analysis (Shen, 2008). The use of visual images is common in many branches of science, and such images are important for progress in the various fields (Arnheim, 1970; Beliën and Leenders, 1996; Freeman, 2000; Koestler, 1964; Klovdahl, 1981; Taylor, 1971; Tufte, 1983; Tukey, 1972). Historian Alfred Crosby (1997) has gone much further. Crosby has proposed that visualization is one of only two factors that are responsible for the explosive development of all of modern science; the other factor is measurement (Freeman, 2000).

Two distinct display forms have been used in the literature to construct network images, one based on points and lines and the other based on matrices. In most point and line displays the points (nodes) represent social actors and the lines (vertices) represent connections among the actors. In matrix displays the rows and columns both represent social actors and numbers or symbols in the cells show the social connections linking those actors. The overwhelming majority of network images have involved the use of points and lines and are adopted for this research (Freeman, 2000). Visualizations of social networks have been used to aid SNA from the beginning (Freeman, 2000). The visualization of networks is important because it is a natural way to communicate connectivity and allows for fast pattern recognition by humans. However, there are great challenges when visualizing networks by hand (Di Battista, 1999).

Following GLM and Visual analysis of the general SNA data matrix, the researcher will deploy simple regression to explore the onboarding data to answer the following research questions (see Chapter 4 for how these were formulated):

What is the relationship between a group's Social Network Topology and that group's Performance?

- What is the relationship between **Tie Strength** and Performance?
- What is the relationship between **Centrality** and Performance?
- What is the relationship between **Structural Holes** and Performance?

What is the correlation between onboarding speed and individual performance?

Transformed to Hypothesis, the research questions would appear as:

- H1: In the case of **HF**, the relationship between **Tie Strength** and Performance is inversely curvilinear.

E.g. as tie strength grows, performance improves, but after an optimal point is passed, additional tie strength undermines performance (perhaps because maintaining such strong ties can be exhausting).

- H2: In the case of **HF**, the relationship between **Centrality** and Performance is inversely curvilinear.

E.g. as the level of average centrality grows, performance improves, but after an optimal point is passed, additional centrality undermines performance (perhaps because all parties are deemed central, knowledge search may take longer).

- H3: In the case of **HF**, the relationship between **Structural Holes** and Performance is inversely curvilinear.

E.g. as the number of structural holes grows, performance improves, but after an optimal point is passed, additional structural holes undermine performance (these holes add non-redundant ties, but need to be bridged for knowledge to flow between actors on either side of the hole).

- H4: There is a positive correlation between onboarding speed and individual performance.

E.g. The faster the onboarding speed (i.e. shorter time to form a strong tie bond to be formed), the better the performance.

These hypotheses can be expressed as follows:

1) **What variable(s) best represent performance?**

2) **How is performance impacted by social network?**

3) **What structural measures were significantly associated with performance?**

    a. **What is the relationship between performance and centrality?**

    b. **What is the relationship between performance and structural holes?**

4) **What relational measures were significantly associated with performance?**

    a. **What is the relationship between performance and strength of ties?**

Based on prior research, the researcher predicts that high-performing teams will not only share similar social network topologies, but because of the task contingency (mainly exploitive) involved, one would expect high-performing **HF** Offices to:

- Have a majority of strong tie relationships, to facilitate tacit knowledge to flow.

- Include Groups where Centrality is shared equally amongst members (little hierarchy).

- Have few structural holes.

6.2 Data Variables

The researcher is exploring the significant SNA factors which impact performance. In this case our independent variable **IV** is Normalized Performance.

Table 14: Data Variables – IV Normalized Performance

Name	**Independent/Dependent**	**Acronym**
Normalized Revenue	Dependent	nrev
Structural Holes	Independent	holes
Tie Strength	Independent	ties
Average Eigenvector ctr	Independent	eig
Average Closeness ctr	Independent	clos
Average Betweenness ctr	Independent	Betw
Full time equivalents	Control	Fte
Office Type	Control (0=corp, 1=retail)	type

6.3 Data Manipulation

Imputation

Before getting to the modelling stage, many data manipulations need to be done. One significant item to address is missing data. Density and structural holes had over 40 missing points out of 187 observations and needed to be imputed. The researcher chose imputation over simply omitting the offices with missing data from the analysis. Otherwise, the researcher would have lost 20% of the group data sets.

There were also five missing points within each of the following SNA measures:

- average eigenvector centrality,
- average betweenness centrality,
- average closeness centrality, and

- total strong edges.

Using the original data matrix with missing variables, the researcher used the R package Amelia to impute the missing values. The advantage of Amelia is that it combines the comparative speed and ease-of-use of our algorithm with the power of multiple imputation (King, 2001). The researcher chose imputation over simply leaving Offices with missing data out of the analysis in order to maximize the value of the data. The alternative, dropping Offices without complete data, would have led to a loss of more than 20% of the data sets, and that was deemed unacceptable.

There were only a few missing points within density, average geo distance, average eigenvector central4ity, average betweenness centrality, average closeness centrality, and total strong edges. To address this, simple imputation was applied on these variables by using the mean value for the missing points. There were 30 missing points within structural holes, and multiple imputation method (by using multiple imputation modelling) was employed.

Outliers

An examination of outliers was undertaken by the researcher. To this the researcher generating the following boxplots (Figure 26 on the following page), which clearly indicates that the eCommerce Group is a potential outlier.

After much deliberation, the researcher decided to leave the outlier in the data, as it was judged appropriate when one examined the data in context. The eCommerce office is made up of 15 FTEs, as opposed to the typical six to eight. This most likely contributes largely to the Group's extraordinary performance and in turn may have led to its outlier status.

However, this is expected as eCommerce Group is a hybrid group selling retail services, but only through the internet. This makes eCommerce Group's clients similar to those serviced by retail offices but gives eCommerce Group the efficiencies associated with corporate offices which receive most of their requests electronically.

Notwithstanding the foregoing, leaving eCommerce Group in the data to be analysed may have the most impact on tie strength, since that is based on the number of potential edges (relationships) activated, which in turn is based on FTE. E.g. a Group of six FTEs has 15 potential edges, while a Group of 15 has 105. Thus, by leaving in the eCommerce data, one would expect large variance in Tie Strength, potential number of ties and number of Strong Edges.

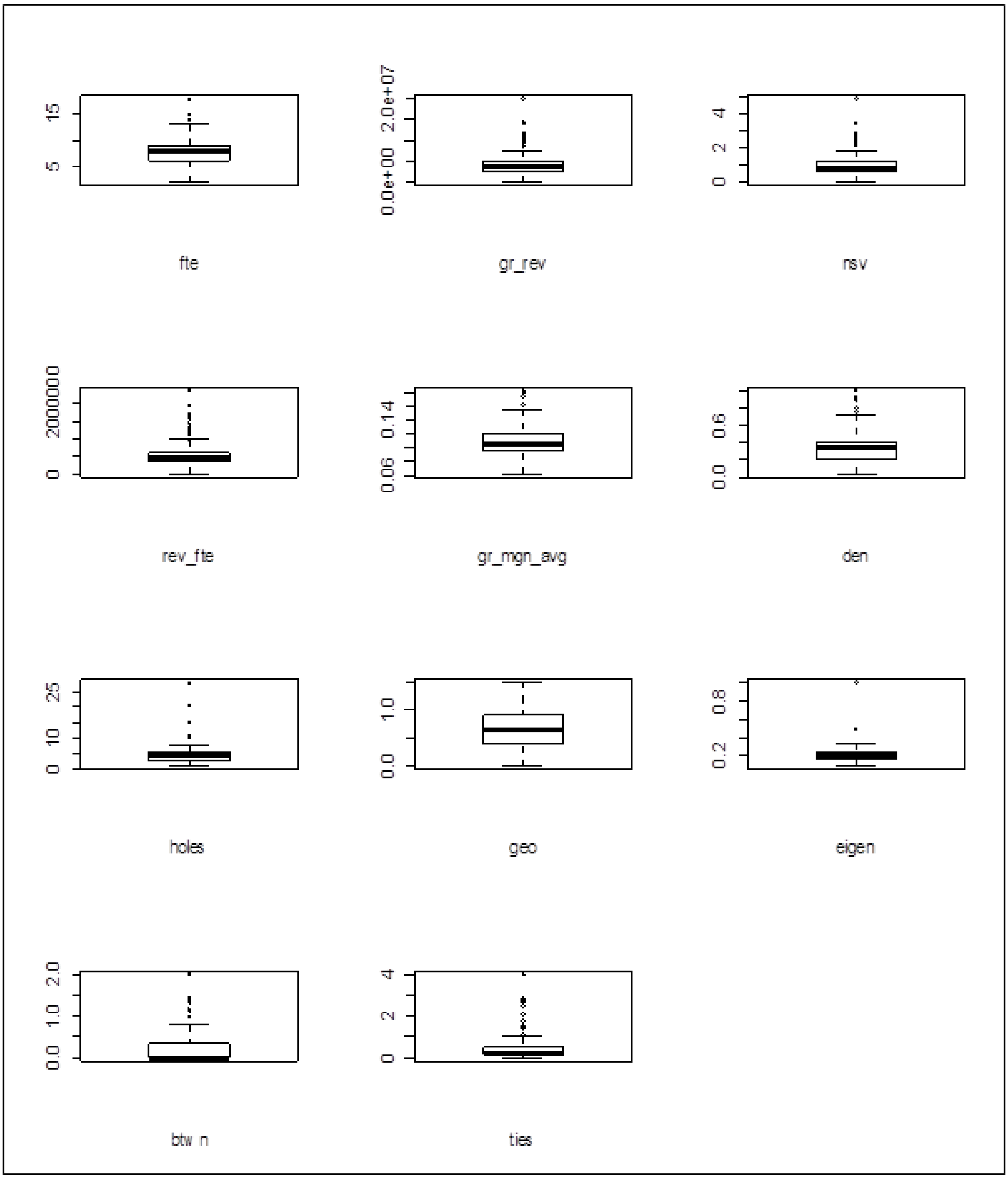


Figure 26: Checking for Data Outliers using Box Plots

6.4 Measuring Performance

For Performance **HF** provided had many measures (e.g. Rank, Gross Margin, Nrev). TSV and normalized Rev (Nrev) are the same (TSV divided by Nrev was a constant). Nrev is highly correlated with TSV/FTE (correlation coefficient 0.78) and not correlated with Gross Margin Avg. (correlation coefficient 0.04). Therefore, Nrev and Margin Avg. could be chosen to represent performance, with FTE included as a covariate in the multiple regression analysis. In such cases, where you have two possible dependant variables it may be insightful to create a scale from these measures, effectively amalgamating them. For this research, this is not only insightful, but also necessary as GLM:OLS only allows for one dependant variable to be analysed at a time (Hutcheson and Moutinho, 2012). This raises the issue of what is the best measure of Performance, Nrev, GMA or a scale dubbed 'performance'.

Table 15: Performance, Nrev, GMA

		Nrev	Gross Margin Avg	Performance
Nrev	Pearson Correlation	1	.199**	.999**
	Sig. (2-tailed)		.006	.000
	N	187	187	187
Gross Margin Avg	Pearson Correlation	.199**	1	.231**
	Sig. (2-tailed)	.006		.002
	N	187	187	187
Performance	Pearson Correlation	.999**	.231**	1
	Sig. (2-tailed)	.000	.002	
	N	187	187	187

From this the researcher learns that the Performance scale (made up of Nrev + Gross Margin Avg) is 99.9% correlated Nrev. This suggests that simply using Nrev as the measure of Performance may be sufficient. This makes sense as a travel agent's gross revenue is based on both the quantity of travel booked and the quality of the margin/profit built into each such booking (i.e. GMA). An agent with better margins (i.e. higher GMA) who sells the same number of trips will have a higher Nrev. To confirm this, Factor Analysis was conducted.

Factor Analysis is a form of statistical interpretation which provides insight into the relation amongst similar variables. Factor Analysis can be used to reduce a large number of variables to a smaller number of variables while minimizing the negative impact of doing so (Hutcheson and Moutinho, 2011). Factor Analysis was deployed at the conclusion of model development to confirm that only Nrev needed to be included in the regression model used to predict performance at **HF**.

A few additional notes about the Measures

By reviewing the data provided by the research with a critical lens, the researcher found:

Normalized Sales Volume (Nrev) was calculated by dividing each Group's total sales volume (TSV) by the mean, a well-accepted approach. TSV and normalized Rev (Nrev) are the same (TSV divided by Nrev was a constant). Because the focus of this research is on the Group level of analysis, the researcher chose not to focus on performance measures of the individual (e.g. Nrev/FTE). The researcher has concerns about centrality. With the average Group size being six to eight FTEs, the researcher speculates that centrality may have less impact on performance at **HF** than in other instances, the consideration being the relative lack of deviation (e.g. spread) available. This makes sense to the researcher since a small group limits the number of potential edges that can be formed (six people can only have five edges).

Density and structural holes are inverse to each other. After imputation, only one could be used in the multivariate regression model, and that was determined to be density, because density had a better distribution. However, most prior literature focus on structural holes. So this will need to be revisited in the final analysis. It is also worth while noting that both Density and Structural Holes have upper limits. Density is based on number of connections divided by number of possible connections. Structural holes are limited by the number of nodes in the graph (i.e. if you have three unconnected nodes, the most holes would be n-1). Therefore the upper limit of Density would be 100% and the upward limit of structural holes would be n – 1, where n is the number of FTEs/nodes.

Total strong edges divided by potential edges equals *ties*. The researcher notes, that because strength of ties was equivalent to strong edges / potential ties, it is highly correlated with both strong edges and potential ties. Therefore, it should not be included as an independent variable in the multivariate regression along with strong edges and potential ties. Most of the dependent variables and independent variables are not normally distributed; therefore, they need some form of transformation.

A note about dropping offices:

The researcher choose to drop the Cruise Offices from the data as these Offices do not sell travel in the same manner as **Retail or Corporate offices.** Cruise Office sales are mainly resales to other Offices. The researcher also considered dropping **HF**BT offices (as these are hybrid offices which sell corporate travel to small business through retail outlets) but, in the end, decided to retain them because they are hybrids and may offer insight.

A note about Coding:

While most of the variables herein are continuous, Office Type is not. Office Type (**type**) is an unordered categorical scale measurement with only two discrete values: Corporate (B2B sales) and Retail (B2C sales). An unordered categorical scale of measurement is achieved when the data are recorded as categories which have no meaningful order. In order to include type into the analysis, it must first be dummy coded. To achieve this a value of 0 is assigned to corporate groups and a value of 1 is assigned to retail offices (B2C).

6.5 Descriptive statistics after data imputation

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
nsv	Normalized Rev	187	1.0000000	0.6117057	0.0013250	4.8450446
margin_avg	Gross Margin Avg	187	0.1086568	0.0200384	0.0632005	0.1811672
Density		187	0.2936827	0.2437317	0	1.0000000
hole	Structual Holes	187	5.1555647	4.3223876	-1.9864017	28.0000000
distance	Avg Geo Distance	187	0.6227333	0.3632668	0	1.4814810
eigen	Avg Eigenvector Centr	187	0.2344673	0.1174872	0.0909091	1.0000000
between	Avg Betweenness Centr	187	0.2137079	0.3195605	0	2.0000000
close	Avg Closeness Centr	187	0.2584936	0.1978167	0	1.0000000
strong	Total Strong Edges	187	9.9679144	7.7743054	1.0000000	50.0000000
potential	Potential edges	187	32.0000000	21.2251119	1.0000000	153.0000000
ties	Strength of Ties	187	0.4609956	0.5754630	0	4.0000000

Figure 27: Descriptive statistics after data imputation

The researcher reviewed the descriptive statistics (above) to identify any outliers or any other abnormal data issues that need to be addressed. This chart flagged Structural Holes, Total Strong Edges and Potential Ties as having high variance. As suspected, the presence of the eCommerce Group data displays as a large variance in the number of potential ties (see above for more). Notwithstanding this factor, the general statistics (number, mean, standard deviation, variance) showed that the data set was complete and without issue.

6.6 Analysis with General Linear Modelling

Generalized Linear Models (GLMs) enable descriptive and predictive models to be built that are sufficiently general to be applicable to much social science data. GLMs can be used to model data collected from survey and experimental studies and can replace many of the more traditional hypothesis tests that are still in common use (Hutcheson and Moutinho, 2012).

Of particular importance is the unified theoretical framework that GLMs offer, as this enables certain 'economies of scale' to be realised that allow a whole range of data to be analysed using similar techniques (Hutcheson and Moutinho, 2012).[2] The use of the techniques will be described using a modelling

[2] Note: A more detailed account of model diagnostics across the wide range of models available within the GLM framework can be found in McCullagh and Nelder (1989), in which the authors give a detailed discussion of the techniques available and illustrate the advantages of moving beyond the traditional parametric model with Normal errors (Hutcheson and Moutinho (2012)).

procedure whereby a particular variable can be modelled (or predicted) using information about other variables. A continuous response variable (a variable that can be described using the mean) can be modelled using ordinary least-squares (OLS) regression (Hutcheson and Moutinho, 2012) and written as follows:

Performance may be predicted by measures of Social Network Topology.

$$Y = Fn(x)$$

Whereas:

- Y is Nrev, and the DV is continuous; it represents Performance.
- X represents Social Network Measures (e.g. Tie Strength, Centrality, etc.) and Controls (e.g. FTE, Type).
- Since Y is continuous, the researcher uses GLM:OLS through R and R cmdr.
- Theory suggests a curvilinear relationship between Performance and each of the Social Network measures.

Checking GLM Assumptions

For a regression model to be valid:

1. The sample needs to be representative of the population for the inference prediction.
2. The error is a random variable with a mean of zero conditional on the explanatory variables.
3. The independent variables are measured with no error.
4. The predictors are linearly independent (i.e. it is not possible to express any predictor as a linear combination of the others).
5. The errors are uncorrelated, that is, the variance–covariance matrix of the errors is diagonal and each non-zero element is the variance of the error.

In this data, the variance of the error is constant across observations and there are sufficient conditions for the least-squares estimator to possess desirable properties; in particular, these assumptions imply that the parameter estimates will be unbiased, consistent, and efficient in the class of linear unbiased estimators (Gauss, 1809). There are five different factors on which the efficacy of a regression model is determined.

Those factors are Normality, Linearity, Homoscedasticity, Independence and Model specification. If these five conditions are not met, transformation may be required. Transformations can help evolve data into a form that meets these key tests. An assessment of the normality of data is a prerequisite for many statistical tests, as normal data is an underlying assumption. In order to leverage GLM through Ordinary Least Squares (OLS) the data must met this assumptions. To test for normality, the researcher leverages two approaches: a **Tests of Normality** table and the **Normal Q-Q Plots**. If data proves to be abnormal, then some transformations may need to be deployed before multi-variant regression can be carried out.

For the numerical Tests of Normality, the researcher chose Kolmogorov-Smirnov and Shapiro-Wilk as key tests from prior literature (Bellieni, 2009; Unterseher, 2011).

Table 16: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
FTE	.122	187	.000	.972	187	.001
Performance	.151	187	.000	.811	187	.000
Strong Edges	.185	187	.000	.827	187	.000
potential edges	.185	187	.000	.867	187	.000
Str of Ties	.232	187	.000	.648	187	.000
Avg Geo Dist	.101	187	.000	.956	187	.000
Avg Eigen Centr	.255	187	.000	.677	187	.000
Avg B/w Centr	.283	187	.000	.712	187	.000
Avg Close Centr	.160	187	.000	.881	187	.000
Str Holes	.243	187	.000	.759	187	.000
Density	.195	187	.000	.904	187	.000
Gross Margin Avg	.091	187	.001	.942	187	.000
Nrev	.147	187	.000	.812	187	.000

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

The Kolmogorov–Smirnov statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution, or, alternatively, between the empirical distribution functions of two samples. The null distribution of this statistic is calculated under the null hypothesis that the samples are drawn from the same distribution (in the two-sample case) or that the sample is drawn from the reference distribution (in the one-sample case). In each case, the distributions considered under the null hypothesis are continuous distributions but are otherwise unrestricted. In the special case of testing for the normality of the distribution, samples are standardized and compared with a standard normal distribution. This is equivalent to setting the mean and variance of the reference distribution equal to the sample estimates, and it is known that using these to define the specific reference distribution changes the null distribution of the test statistic.

Since the probability associated with the test of normality of < 0.001 is less than or equal to the level of significance (0.01), the researcher rejects the null hypothesis and concludes that all but Gross Margin Average appear to be not normally distributed. The researcher conducted log transformations which were necessary to transform this data into a form that would test positive for normality.

Yet various studies have found that, even in this corrected form, the Kolmogorov–Smirnov statistic is less powerful for testing normality than the Shapiro–Wilk test. In particular, R.B. D'Agostino makes a very strong statement in *Goodness-of-fit techniques* (1986): 'The Kolmogorov-Smirnov test is only a historical curiosity. It should never be used'. For that reason the researcher pursued a second test for normality, the Shapiro-Wilk Test.

If the Significance value of the Shapiro-Wilk Test is greater than 0.05, then the data is normal. If it is below 0.05, then the data significantly deviate from a normal distribution. Based on the Shapiro-Wilk Test, NONE of the variables seem to follow the normal distribution.

Next, the researcher applied the graphical approach to test for normality and generated **Normal Q-Q Plots**. In order to determine normality graphically the researcher uses the output of a normal Q-Q Plot. If the data are normally distributed, then the data points will be close to the diagonal line.

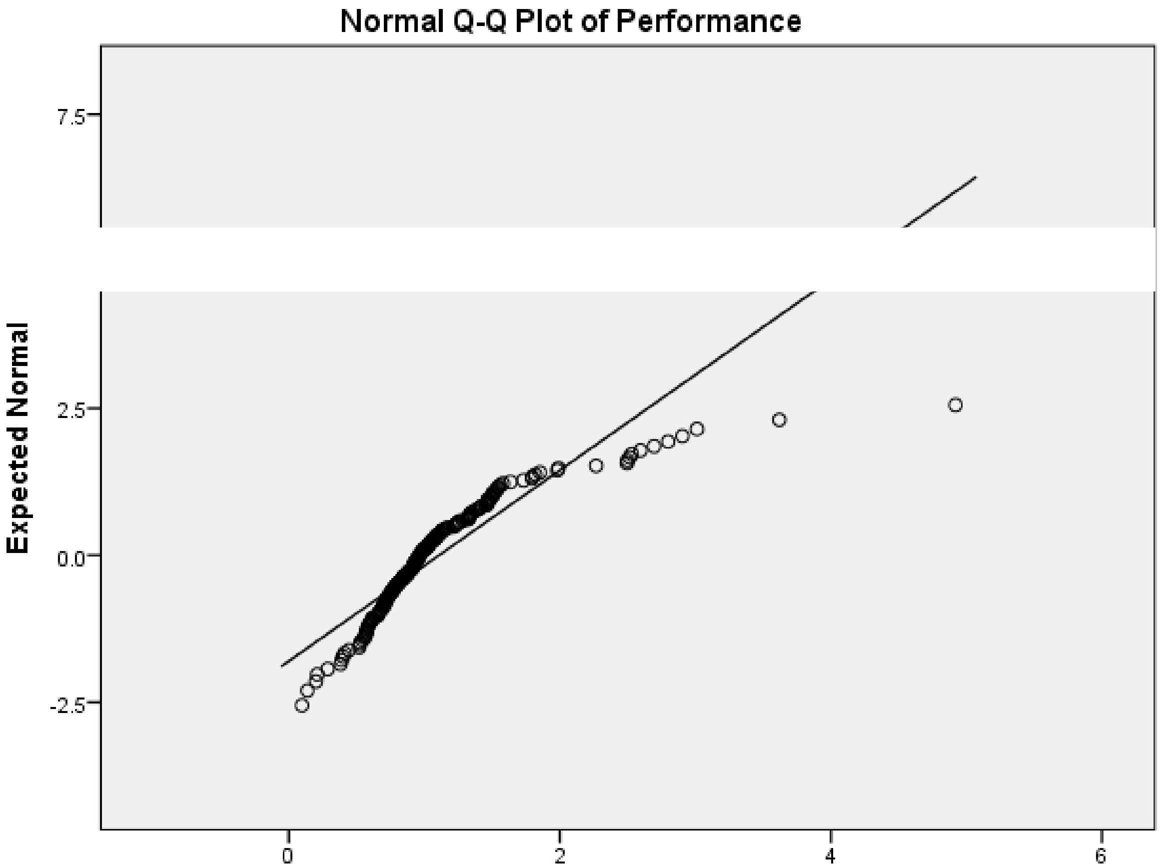


Figure 28: Normal Q-Q Plot of Performance

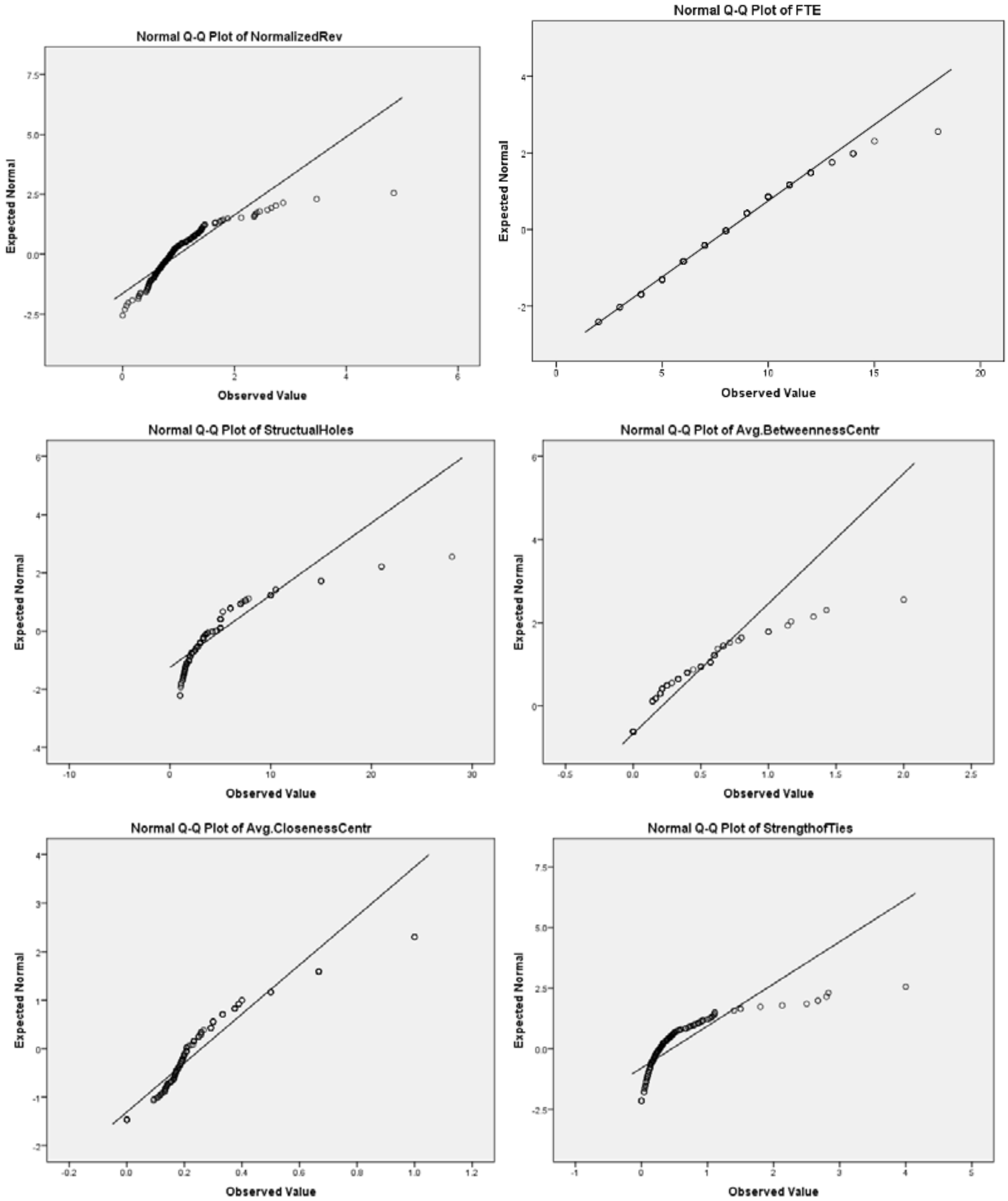


Figure 28.2: Normal Q-Q Plots

In Normal Q-Q plots a diagonal line is drawn representing the expected values for normal distribution. If the actual distribution (the dots) follows that diagonal line, then normality can be concluded. Based on the above plots, however, the researcher concludes few, if any, IV are normally distributed. This is entirely consistent with the results of the Shapiro-Wilk Test (which found no variables follow the normal distribution) and the Kolmogorov–Smirnov test. When the normal distribution is not readily seen, one may perform transformations with the hope that undertaking such will transform the data into normally distributed values. However, there are some limitations on transformation (see below).

6.7 Transformation

From earlier analysis, most of the variables are not normally distributed; therefore they need some form of transformation. However, there are some limitations on transformation (see below). For transformation, the researcher undertook Box-Cox power transformation (Sakia, 1992) to determine the transformation. Transforming data means performing the same mathematical operation on each piece of original data. Some transformation examples from daily life are currency exchange rates and converting Celsius values into Fahrenheit.

The Box-Cox Transformation searches for a value of lambda such that the transformation may correct for non-normality. In order to find the optimal and closely competing lambda values, the Box-Cox Transformation modifies the original data using the equation below for W_i (a standardized transformed variable). It then calculates the standard deviation of the variable W. The goal is to find the value of lambda that minimizes the standard deviation of W.

$$W_i = \frac{(Y_i^{\lambda} - 1)}{\lambda\ G^{\lambda-1}} \quad \text{when } \lambda \neq 0$$

$$W_i = G\ In(Y_i) \quad \text{when } \lambda = 0$$

where: Y_i: Original data
G: Geometric mean of all the data
λ: Lambda values

The lambda value (λ) for the transformation of the dependent variable was determined to be: $\lambda = 0.5$. This was deduced from the following graph for the log-likelihood of the transformation.

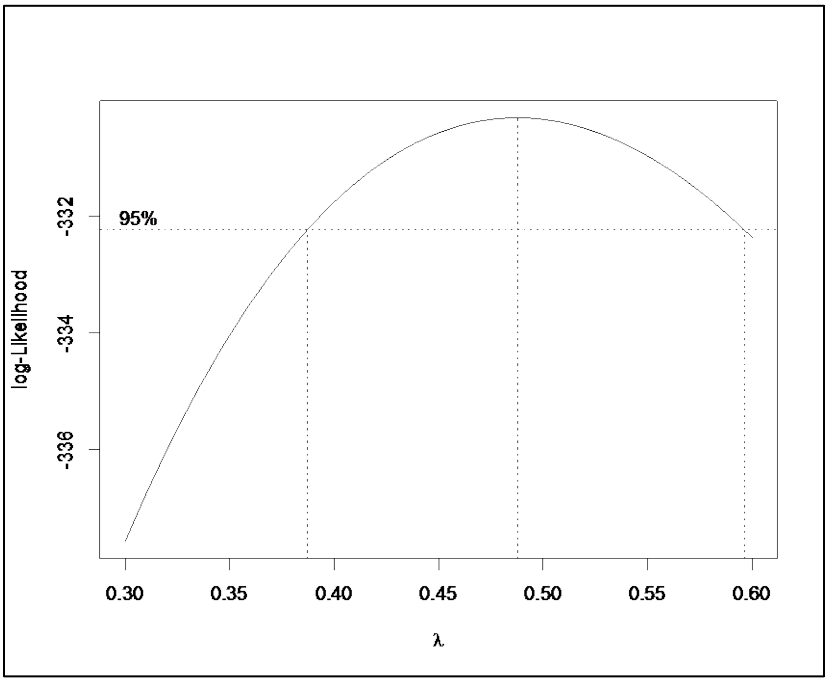


Figure 29: Estimating λ for the Box Cox Procedure

Revisiting The Tests for Normality and remembering if the skewness and kurtosis statistics fall between -2 and +2, it's generally considered acceptable to assume that the data is normally distributed. One can see that the log of the various measures does push skewness into the -2 to +2 range, thus driving the data to normality.

Table 17: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
log of Avg Eigen centr	.187	187	.000	.917	187	.000
log of Total Strong Edges	.062	187	.074	.988	187	.128
log of Potential Edges	.150	187	.000	.923	187	.000

a. Lilliefors Significance Correction

Thus, by transforming Avg Eigenvector Centrality, Total Strong Edges and Potential Edges the researcher are able to find a data set that is normally distributed. When logs generated negative values, the values must first undergo Reflection. Reflection is computed by subtracting all of the values for a variable from one plus the absolute value of maximum value for the variable. The absolute value of maximum of Average Eigenvector Centrality is 1, so the Reflective value will be (1+1) – Eigen. This yields all positive numbers between 1 and 2.

Details of data transformation:

1. Distributions of dependent variables (Nrev and gross margin avg.) look OK, no need to transform; distributions of independent variables do not look very OK.

2. For avg eigenvector centrality, total strong edges and potential edges that all had positive values; therefore one can run log transformation.

3. For density, avg geo distance, avg betweenness centrality and avg closeness centrality that had multiple zero values, it is hard to find a mechanism to reasonably assign a different non-zero value (e.g. 0.00001) to each of the zero values, therefore no log transformation can be made. Note: Square root transformation could be run, but the variables with square root transformation did not have better normality. Therefore these variables in their original forms will be used in the regression.

Other data handling

Density and structural holes are inverse with each other. After imputation, only one could be used in the multiple regression model, and it must be density because density had fewer missing points. Strength of ties was equivalent to the ratio of strong edges to potential ties. It is highly correlated with both strong edges and potential ties (with correlation coefficient of 0.62 and -0.56, respectively). Therefore, to avoid collinearity it should not be included as an independent variable in the multiple regression along with strong edges and potential ties. It should be included, however, in the factor analysis.

6.8 LOWESS (Locally Weighted Scatterplot Smoothing)

When there is suspicion that the relationship is not completely linear, LOWESS (locally weighted scatterplot smoothing) is helpful in suggesting what form to use for fitting the polynomial terms of the regressors (such as squares or cubic terms). It provides a smooth fit from the dependent to the independent variable. By checking the smoothing plot we can see what form the relationship may take, and a parametric model can then be chosen to approximate this form.

The following LOWESS plots showed some curvilinear patterns in most of the social network variables. In these curvilinear patterns the number of 'turns' was 1 or 2, which corresponds to the quadratic term and the cubic term, respectively.

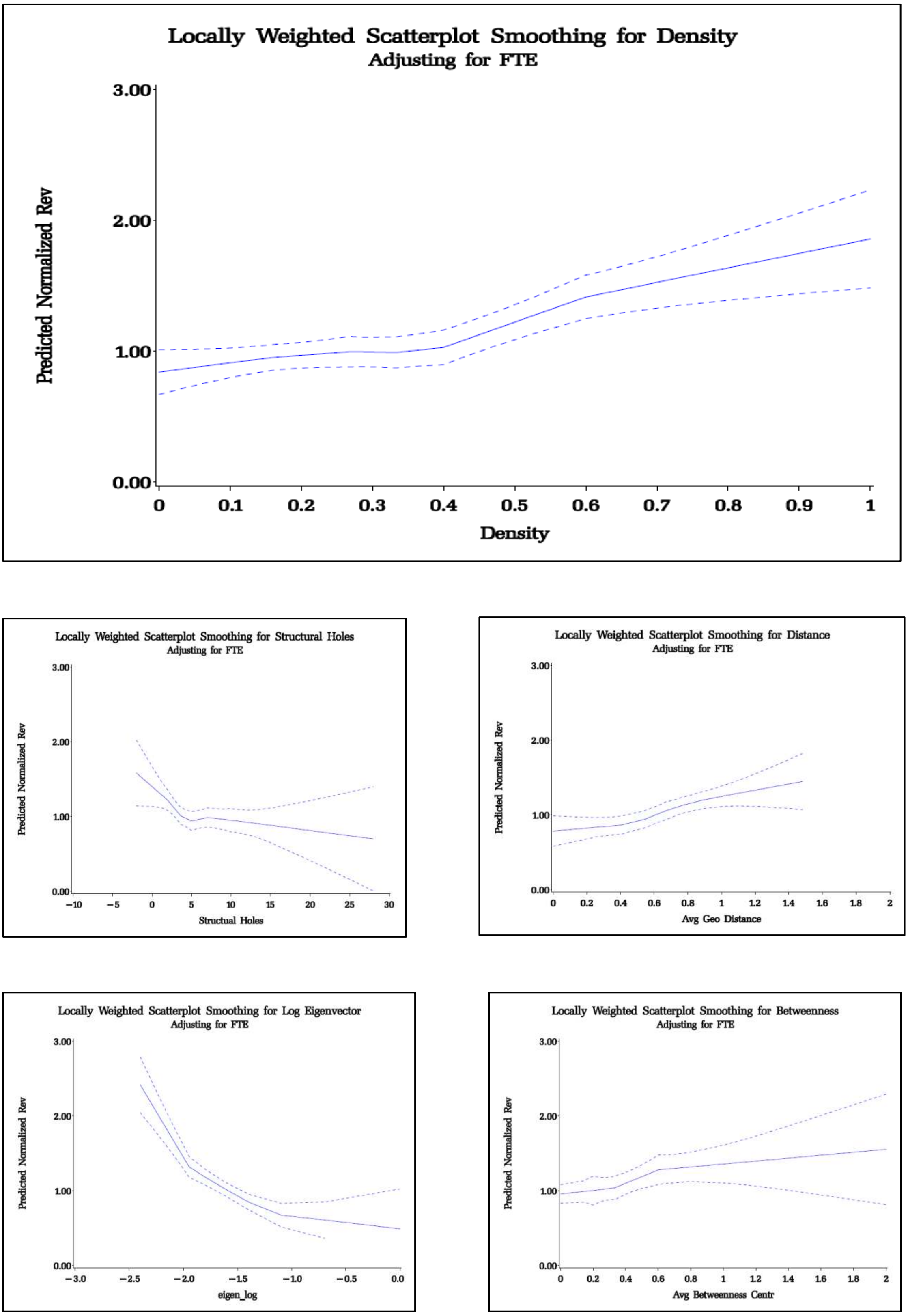


Figure 30: Locally Weighted Scatterplot Smoothing

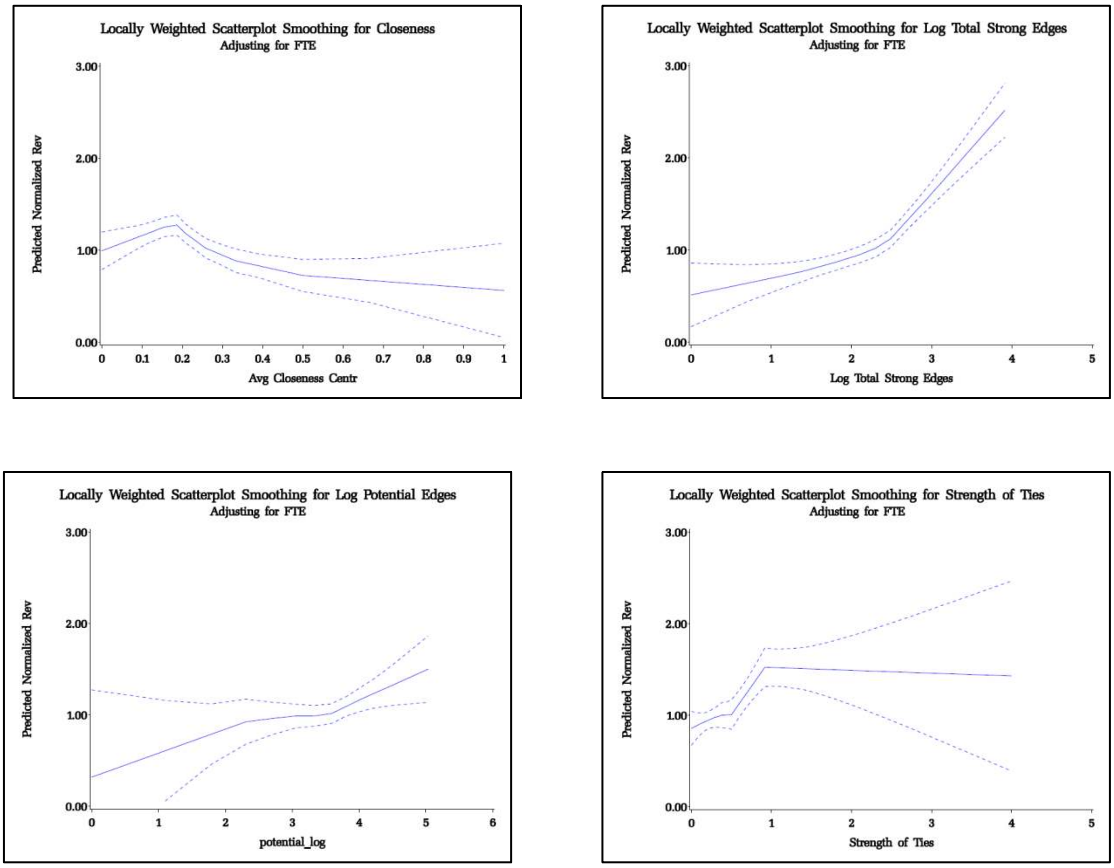


Figure 31: Locally Weighted Scatterplot Smoothing

The LOWESS curves above suggest that in this data a potential curvilinear relationship involved some quadratic term or cubic term. Therefore, the square term and the cubic term of each social network variable were created, and the set of polynomial terms (linear, quadratic, and cubic) in each social network variable was tested against performance.

Polynomial regression for each social network variable

In these polynomial regressions, none of the social network variables was correlated with gross margin avg (all the p-values >0.05), when controlling for FTE and Corporate. Therefore, gross margin avg is not a representative of performance in terms of the relationship between performance and social network. The only representative of performance in this regard is Nrev. This confirms that only Nrev is necessary to generate a valid, well-fitting prediction model. Note: None of the potential edges terms was correlated with Nrev since all the p-values >0.05.

Table 18: Social Network Variable Polynomial Regression

Social Network Variable		Nrev	
		Parameter Est.	p-value
Density	Self	0.54	0.001
Structural holes	Self	-0.02	0.017
Geo Distance	Self	0.41	<0.0001
Eigenvector	Log	0.81	0.017
	Square of log	0.53	<0.0001
Betweenness	Self	0.34	0.002
Closeness	Self	2.23	0.008
	Square	-7.85	0.001
	Cubic	5.74	0.001
Total strong edges	Log	-0.25	0.140
	Square of log	0.15	0.001
Tie Strength	Self	0.97	<0.0001
	Square	-0.23	<0.0001

The results of the polynomial regressions show that:

- **density, structural holes, distance, and betweenness had a linear regression;**

- **log eigenvector, log total strong edges, and tie strength had a quadratic term involved; and**

- **closeness had a cubic term involved.**

These will be taken into account when the multiple regression model is developed.

In the table, a positive sign of a parameter estimate for a quadratic term corresponds to the U shape in the LOWESS plots, whereas a negative sign of a parameter estimate for a quadratic term corresponds to the inverse U shape in the LOWESS plots.

From this, the researcher confirms Lechner, Frankenberger and Floyd's (2010) findings that an inverse curvilinear (inverse U shape) relationship exists between Nrev (representing Performance) and both Centrality (representing the Structural dimension of **HF**'s social network); and the square of Tie Strength (representing the Relational dimension of **HF**'s social network).

Multiple polynomial regression for the set of social network variables

1. Model selection

A stepwise selection process was run to choose a best fit to predict Nrev. The criterion p-value to enter was set as 0.20, and the criterion p-value to stay was set as 0.05. The selected variables were log total strong edges, square of log eigenvector, and square of log total strong edges, controlling for FTE and Corp.

2. Final multiple polynomial regression

The selected variables were put into a final multiple polynomial regression model. To minimize the problem of multi-collinearity, one resolution is to 'orthogonalize' the vectors in the regression through the method of Gentleman-Givens transformations:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	40.638596648	8.1277193295	50.80	<.0001
Error	181	28.959603775	0.1599978109		
Corrected Total	186	69.598200423			

Root MSE	0.3999972636
R-Square	0.5839029803

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.36493783958832	0.1843782722	1.98	0.0493
FTE	1	0.04600260804126	0.0131990798	3.49	0.0006
Corp	1	1.00642993499137	0.1232389944	8.17	<.0001
Log Total Strong Edges	1	-0.50770699840839	0.1830063048	-2.77	0.0061
Square of Log Total Strong Edges	1	0.14878486388649	0.0406592624	3.66	0.0003
Square of Log Eigenvectors	1	0.20160391725911	0.0629753402	3.20	0.0016

Figure 32: Polynomial Regression Model for HF

The p-values are 0.006 for log total strong edges and 0.0003 for its square term, and is 0.002 for square of log eigenvector, indicating that **between log transformed total strong edges and Nrev, both a linear relationship and a quadratic relationship existed** (e.g. the final regression equation will have Strong Edges represented twice)**. For log transformed eigenvector, when controlling for other covariates, there was only a quadratic relationship**. Also worth noting: both FTE and Corp were significant positive predictors of Nrev.

The multiple regression model can be written as:

Nrev = 0.365 + (0.046)*FTE + (1.006)*Corp – (0.508)*Log_Strong + (0.202)*(Log_Eigenvector)(Log_Eigenvector) + (0.149)*(Log_Strong)(Log_Strong)

Or cleaned up:

$$Nrev = 0.365 + 0.046FTE + 1.006Corp - (0.508)Log\_Strong + 0.202Log\_Eigenvector^2 + 0.149Log\_Strong^2$$

6.9 Model diagnostics

A residual is generally a quantity left over at the end of a process. Residuals and statistical errors are not the same thing. The *error* of a sample is the deviation of the sample from the (unobservable) *true* function value, while the *residual* of a sample is the difference between the sample and the *estimated* function value. In GLM, the difference between observed values and the value generated by the regression equation is known as the *residual*. If a residual is small, the values generated by the equation are very close to those observed. A large residual indicates there is deviation between observed and predicted values. If the general linear model developed is appropriate, it is reasonable to expect the residuals to exhibit properties that agree with the stated assumptions (e.g. Normality, Linearity, Homoscedasticity, etc.).

According to a forthcoming Quantitative Analysis text by Hutcheson and Moutinho (2012):

> Regression models for continuous data (including ANOVA and ANCOVA) assume a number of things about the data and the relationships between the variables. These include the assumptions of linearity, Normality (of the model residuals), constant variance and the absence of influential outliers. In OLS regression, statistical inference is weakened when data depart from these assumptions. Even when used for solely descriptive purposes the analysis is improved if the statistical assumptions are met, since a better model-fit is usually obtained (p. 1).

A comprehensive description of checking assumptions in regression models is provided in Hutcheson and Sofroniou (1999). A minimal set of four diagnostic graphs taken from this source will be deployed enabling basic checks to be made on the assumptions of linearity, over and under dispersion, normally-distributed residuals and outliers (Hutcheson and Moutinho, 2012).

First, a goodness of fit test is run to test the strength of the model. The significance of individual and groups of variables in a multiple OLS regression model can be calculated by comparing the deviance statistics (RSS) for nested models (Hutcheson and Moutinho, 2012). Doing so allows one to compare two models. In this research, the first is a full model (e.g. with many independent variables) called the larger model, the second a model 'less full' (e.g. with less independent variables used) is called the nested model:

$$RSSdi\_ = (RSSp) ▯ (RSSp+q)$$

Where RSS is the measure of deviance, p is the smaller, nested model, and p + q is the larger model. Consider the following ANOVA table:

Table 19: Analysis of Deviance Table (Type II tests) of GLM model 1

	Sum Sq	Df	F value	Pr(>F)
FTE	0.0128	1	0.0802	0.77747
Avg.Betweenness.Centr	0.1915	1	1.1978	0.27560
Avg.Closeness.Centr.	0.0076	1	0.0473	0.82810
Avg.Eigenvector.Centr	0.0105	1	0.0656	0.79817
Avg.Geo.Distance	0.0003	1	0.0019	0.96525
Density	0.5960	1	3.7281	0.05547
Structual.Holes	0.0022	1	0.0138	0.90670
Strength.of.Ties	0.0468	1	0.2926	0.58942
Total.Strong.Edges	3.6636	1	22.9159	4.154e-06 ***
Potential.edges	0.0920	1	0.5752	0.44944
Office.Type	9.9043	1	61.9517	7.712e-13 ***
Residuals	23.0214	144		

Response: Performance

Sign if codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '. 0.1 ' ' 1

Next, graphics can be leveraged to check for violations:

Table 20: Violation checking

Testing For:	Use:
linearity	Scatterplot showing actual values of Y values against fitted values (those predicted from the model).
over/ under dispersion	Scatterplot of predicted values against residuals.
normality	Histogram of residuals.
outliers	Cook's distance against observation number.

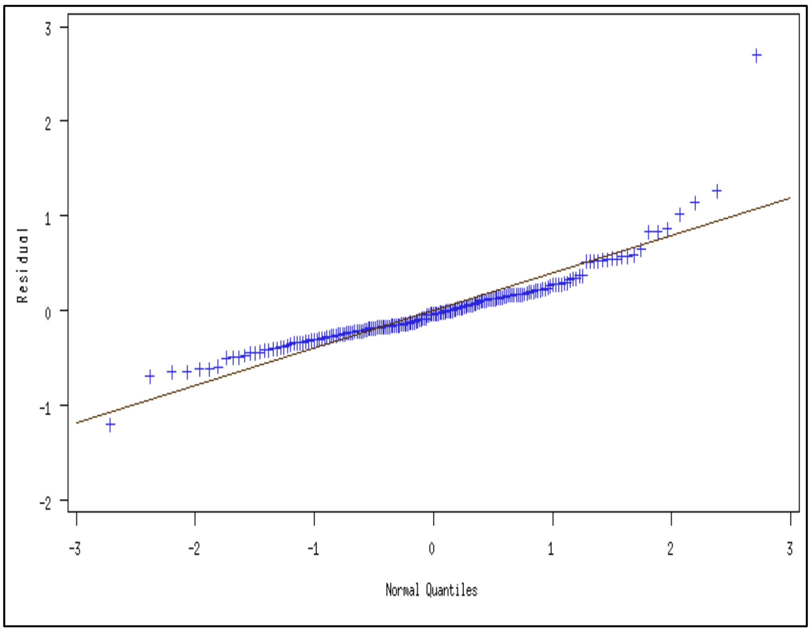
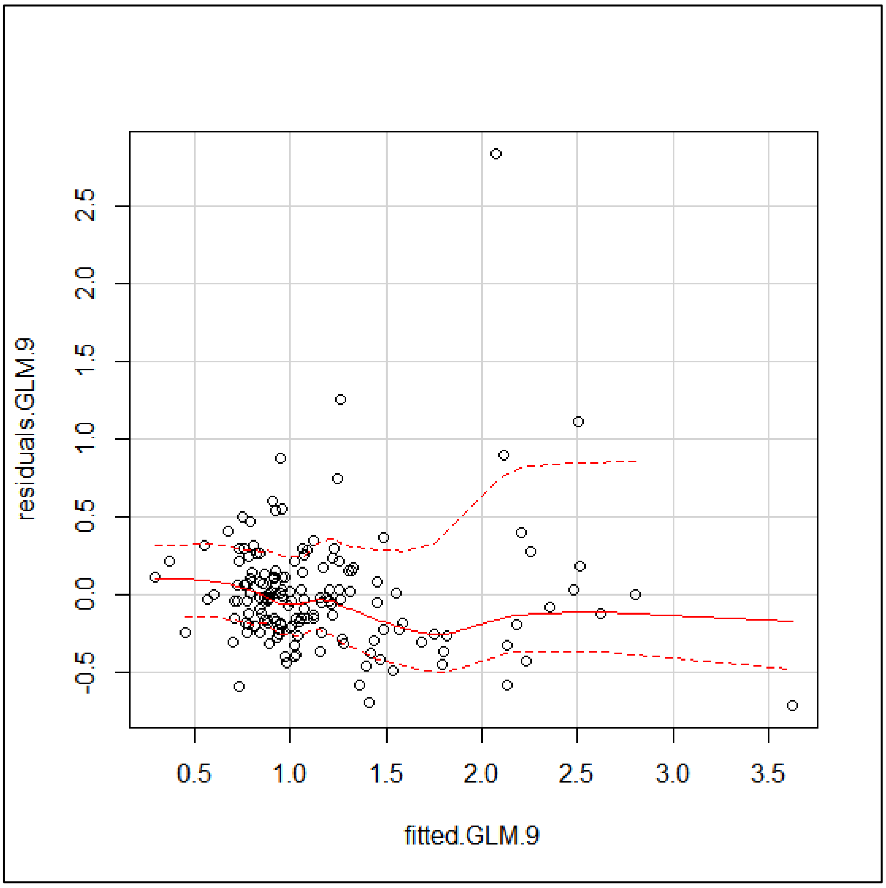


Figure 33: Residuals

First using R cmdr, the researcher calculated and added to the data set the following:

- Fitted values
- Residuals
- Cook's distances

Then graphs were drawn using the graphical functions of Rcmdr.



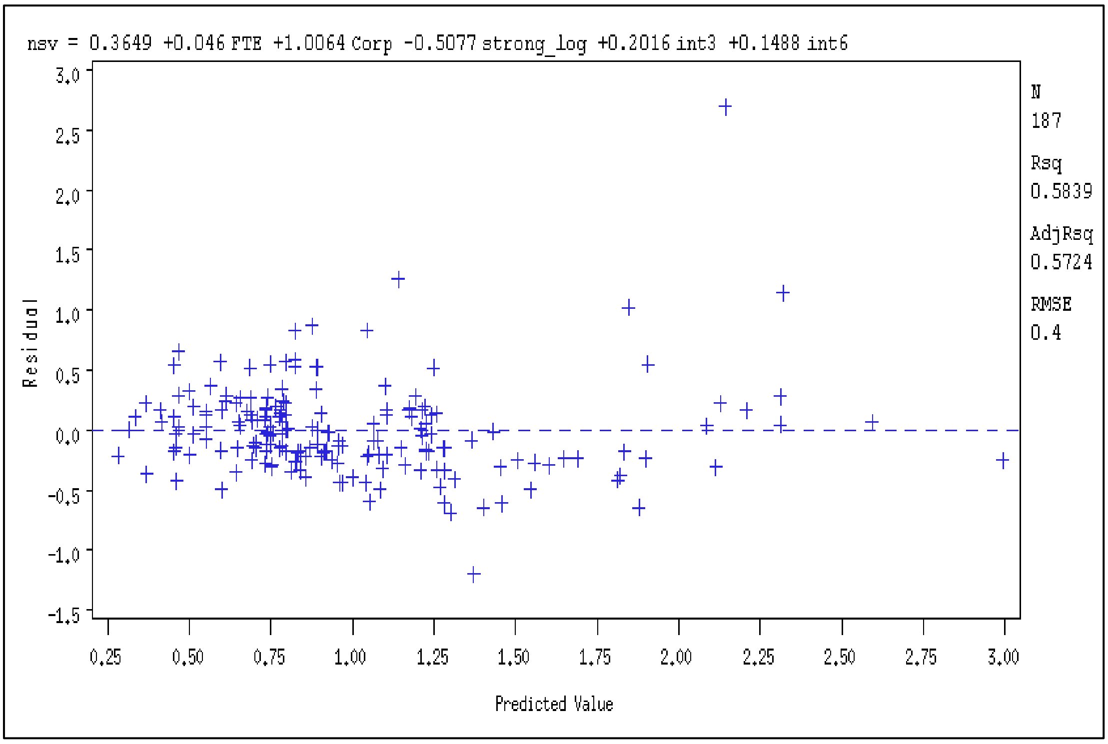


Figure 34: Residuals

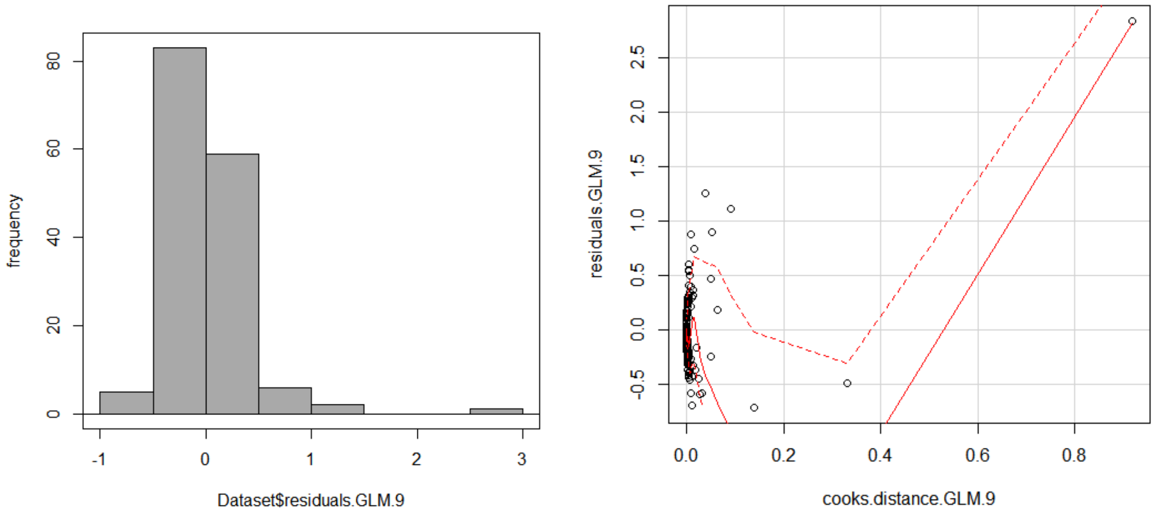


Figure 35: Residuals

When the four key assumptions for the model are met, the graphs should have the following properties:

- The predicted values of the response variable should show a linear relationship when plotted against the observed values of the response variable.
- The residuals should show no obvious pattern when plotted against the predicted values.
- The residuals should be roughly Normally distributed.
- There should be no extreme outliers.

As the graphs above demonstrate, all assumptions are met and based on such, the model fit is found to be acceptable.

6.10 GLM Findings

After using the Box-Cox procedure to transform the data to normality, an ordinary least squares method was followed through stepwise selection. The following model was determined to be the best fit for the data at hand.

$Nrev = 0.365 + 0.046FTE + 1.006type - 0.508Log\_strong + 0.202Log\_Eigen^2 + 0.149*Log\_strong^2$

This multivariate regression model shows that when the social network measures are considered together, only log transformed total strong edges (linear, quadratic) and log transformed eigenvector (quadratic) significantly predicted performance.

Put more generally, in the case of **HF** teams, Performance can be predicated using only the number of strong tie relationships amongst the group and by measuring how central the average travel agent is in the network. This could be interpreted as meaning:

- Performance can be increased to a greater extent if an office is able to increase Tie Strength rather than average eigenvector centrality; and/or

- Adding another FTE would not have as significant an impact as changing a retail store to a corporate office (although it would be significantly easier).

6.11 Interpretation of the GLM model

With the p-value of the F-test <.0001, the regression model is statistically significant. The R-Squared is 0.5839, meaning 58.39% of the variation in Nrev can be explained by the predictors in the model. The R-Squared Adjusted is often used to summarize the fit, as it takes into account the number of predictor variables in the model. In this model, the adjusted R-Squared indicates that about 57.24% of the variation in Nrev is accounted for by the model.

The coefficients for each of the predictive variables indicates the amount of change one could expect in Nrev, given a one-unit change in the value of that variable, and given that all other variables in the model are held constant. For example, for every 1-point increase in FTE, we would expect a 0.046 increase in Nrev, assuming all other variables in the models are held constant.

6.12 Visual Analysis

Following GLM, a visual inspection of the subgraphs was conducted. The NodeXL software generates subgraphs for each Office. Each subgraph has the following elements:

Vertices:

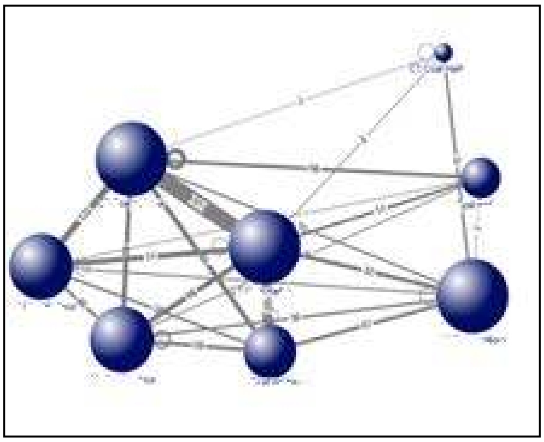


Figure 36: Subgraph of high performing office

Circles/spheres of colour. Each represents a FTE. The relative size of the vertices represents their Centrality in the group (i.e. five same-sized spheres would mean centrality was shared equally amongst all members, e.g. all members are influential; one large sphere and five small spheres indicates that one member of the team is more influential than the others. Often this is the team captain.)

Edges:

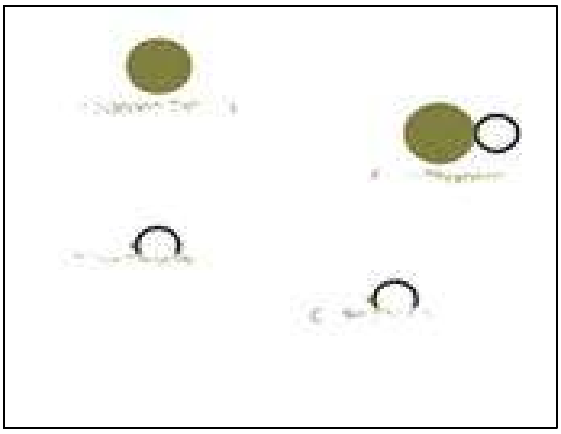


Figure 37: Subgraph of low performing office

These represent relationship connections. The thickness of the edge and the number label indicate the relative strength of the relationship (i.e. if the edge between A and B was thick and labelled 32, but the edge between A and C was thinner and labelled 2, then this would mean the relationship between A:B is 16x stronger than between A:C).

Social graphs for the top performing groups are listed in Appendix B. Social graphs for low performing groups are listed in Appendix C. Based on reviewing such, the researcher observes the following:

Top 10 Groups (T10G):

- Very Dense
- Few Structural Holes
- Centrality spread out over 50% of the vertices
- Have multiple influencers
- A majority of relationships are 'strong ties'

Bottom 10 Groups (B10G):

- Have a large number of structural holes
- Are not dense
- Have many weakly connected members
- Have isolated members
- Have few or no 'strong ties'
- Have little centrality or influence over each other

6.13 Analysing the Onboarding Data

Simple bivariate analysis was conducted in the following manner:

1. Descriptive statistics and normality checking
2. Regression diagnostics
   a. Scatter plot
   b. Test on normality of residuals
   c. Test on homogeneity of variance of residuals

d. Test on randomness of residuals

1) Plot of the residuals against the observation order

2) Runs test (Wald-Wolfowitz test) for randomness

e. Test for outliers: Cook's distance

3. Regression

a. Equation of regression

b. Coefficient of determination (R-square)

c. Residual calculation

d. Significance of the regression model (F-test)

4. Partial correlation between performance and days with strong ties

5. Findings of the onboarding GLM model

Onboarding Variables

- The IV in the onboarding data is Onboarding Speed, being the number of days that strong ties were available (DSTA).

- The DV in the onboarding data is individual Performance (PERFORMA).

Employee	Gross Revenue	D.S.T.A.
MA	1,281,126.71	299
GH	1,383,797.97	305
SP	1,341,491.13	300
KC	932,329.64	233
MC	1,048,394.62	255
NP	1,016,079.91	0
HF	1,249,760.30	0
DW	755,218.96	61

Figure 38: Onboarding Data Matrix

A two-sided probability value of <0.05 was considered statistically significant. All statistical analyses were performed using the SAS software.

1. Descriptive statistics and normality checking for Onboarding

In the onboarding study, the dependent variable is performance, and the independent variable is days with strong ties. With regards to Performance:

```
Variable:  PERFORMA  (Gross Revenue)

                              Moments

N                            8    Sum Weights                 8
Mean                 1126024.9    Sum Observations   9008199.24
Std Deviation       222122.531    Variance           4.93384E10
Skewness            -0.4743411    Kurtosis           -0.9733514
Uncorrected SS      1.04888E13    Corrected SS       3.45369E11
Coeff Variation     19.7262539    Std Error Mean     78532.1741


              Basic Statistical Measures

    Location                    Variability

Mean      1126025     Std Deviation            222123
Median    1149077     Variance             4.93384E10
Mode            .     Range                    628579
                      Interquartile Range      337104


           Tests for Location: Mu0=0

Test           -Statistic-    -----p Value------

Student's t    t  14.33839    Pr > |t|    <.0001
Sign           M         4    Pr >= |M|   0.0078
Signed Rank    S        18    Pr >= |S|   0.0078


                   Tests for Normality

Test                  --Statistic---    -----p Value------

Shapiro-Wilk          W      0.93261    Pr < W      0.5401
Kolmogorov-Smirnov    D     0.211257    Pr > D     >0.1500
Cramer-von Mises      W-Sq    0.0473    Pr > W-Sq  >0.2500
Anderson-Darling      A-Sq  0.287314    Pr > A-Sq  >0.2500
```

Figure 39: Onboarding – Performance Variable

Based on the above tests for normality, the variable of performance is normally distributed. Mean ± STD is 1126025 ± 222123.

With regards to Days with strong ties:

```
Variable:  DSTA  (Days with Strong Ties)

                            Moments

N                          8    Sum Weights                8
Mean                 181.625    Sum Observations        1453
Std Deviation     137.092187    Variance          18794.2679
Skewness          -0.5953713    Kurtosis          -1.9473982
Uncorrected SS        395461    Corrected SS      131559.875
Coeff Variation   75.4809016    Std Error Mean    48.4694077


              Basic Statistical Measures

    Location                    Variability

Mean     181.6250     Std Deviation           137.09219
Median   244.0000     Variance                    18794
Mode       0.0000     Range                   305.00000
                      Interquartile Range     269.00000


           Tests for Location: Mu0=0

Test           -Statistic-    -----p Value------

Student's t    t  3.747209    Pr > |t|    0.0072
Sign           M         3    Pr >= |M|   0.0313
Signed Rank    S      10.5    Pr >= |S|   0.0313


                   Tests for Normality

Test                  --Statistic---    -----p Value------

Shapiro-Wilk          W     0.789842    Pr < W      0.0223
Kolmogorov-Smirnov    D     0.271076    Pr > D      0.0826
Cramer-von Mises      W-Sq  0.127328    Pr > W-Sq   0.0394
Anderson-Darling      A-Sq  0.745824    Pr > A-Sq   0.0314
```

Figure 40: Onboarding - Days with Strong Ties variable

Based on the above tests for normality, the variable of days with strong ties available (DSTA) is normally distributed. Mean ± STD is 181.6 ± 137.1.

2. Regression diagnostics

a. Scatter plot

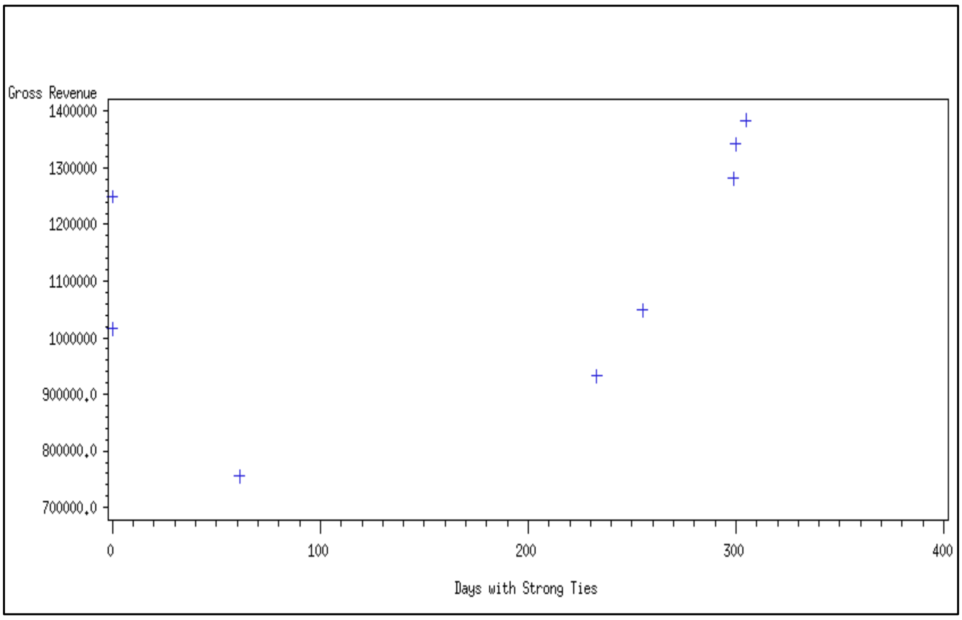


Figure 41: Scatter plot – Days with Strong Ties

While the scatterplot above does not suggest the presence of outliers, the data displayed do seem to demonstrate a curvilinear relationship between gross revenue and onboarding speed. This makes sense because the longer such ties are in place, the longer the agent has access to tacit knowledge. However, the relationship is not simply positive, as high performance can also be seen in agents that have never had access to strong ties. In order for OLS regression to be valid the data must meet the assumptions of normality, linearity, etc. As a result, some transformation may be required. The scatter plot above clearly indicates a quadratic pattern. Therefore, it is not a linear relationship, and any derived regression model must include a quadratic term.

b. Test on normality of residuals

The plot of normal quartiles against residuals based on the specified regression model demonstrated that the residuals are close to a normal distribution, meaning the normality assumption is satisfied.

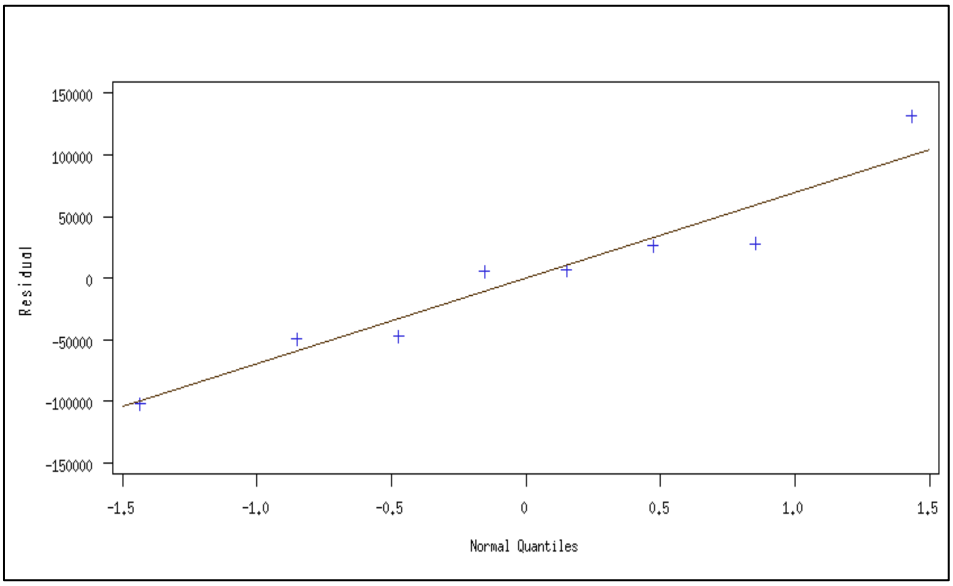


Figure 42: Normal Quartiles Against Residuals

c. Test on homogeneity of variance of residuals

If the model is well-fitted, there should be no pattern to the residuals plotted against the fitted values. The plots below showed that there is no obvious pattern, and the residuals seem to have equal variation around line 0, indicating the model is well-fitted and the equal variance assumption seems acceptable. The White test, which tests the null hypothesis that the variance of the residuals is homogenous, returned a p-value of 0.50, re-confirming the homogeneity of residual variance.

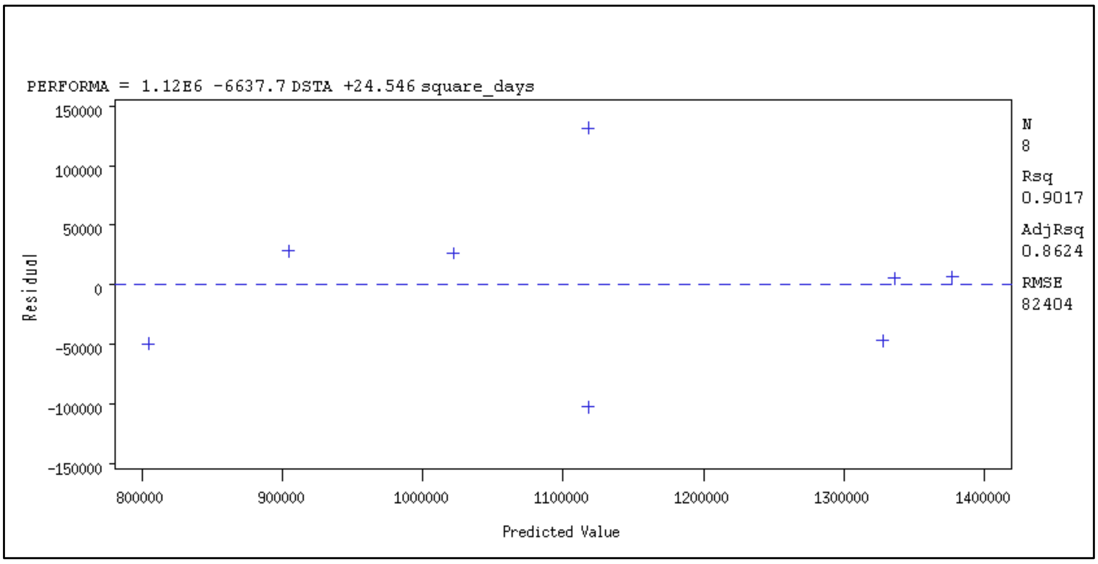


Figure 43: Homogeneity of Variance of Residuals

d. Test on randomness of residuals

1) Plot the residuals against the observation order to detect order trend.

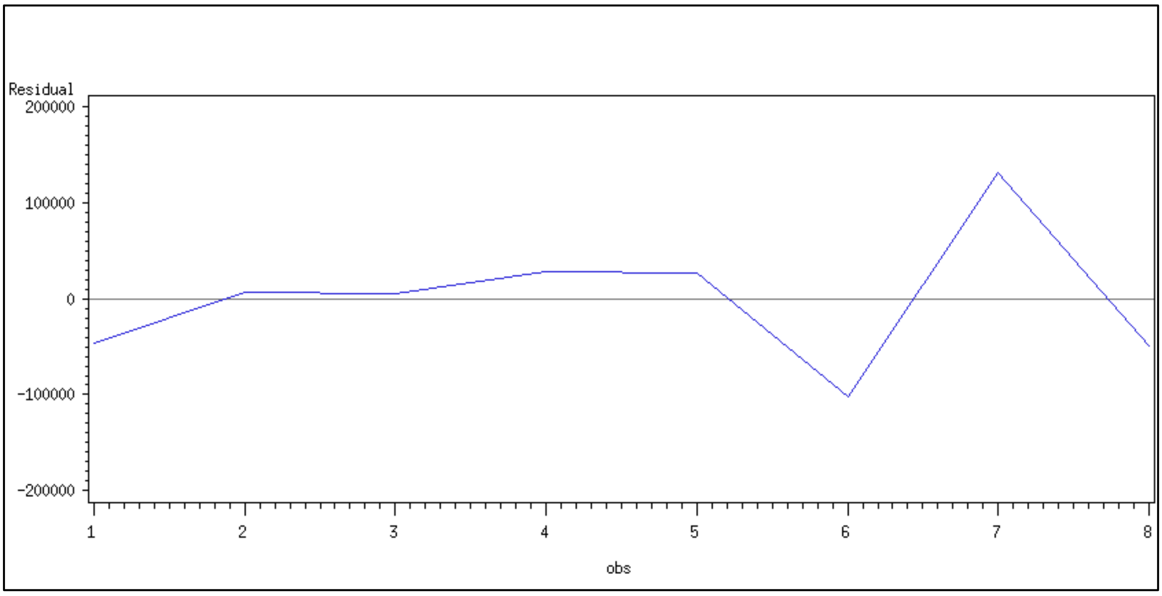


Figure 44: Randomness of Residuals

There is no obvious order in residual plot and so the residuals do not seem to have correlations.

2) The Runs test (Wald-Wolfowitz test) showed that the data are random (p=0.8368):

Wald-Wolfowitz Z	Pr > $\lvert Z \rvert$
-0.20597	0.8368

Figure 45: Wald-Wolfowitz

e. Test for outliers

Cook's distance (e.g. Cook's D) is a measure of change in the parameter estimates when the observation is deleted. The higher the value, the more influential the observation. The critical values are listed in the table below.

According to the table, the appropriate criterion of Cook's D for this study is greater than 2, and the Cook's D of all the observations are below 2, so one is able to conclude that no observation is an outlier.

Alpha = .05, Cook's D Critical Values

	Sample Size									
Predictors	n=5	10	15	20	25	50	100	200	400	800
1	8.323333	1.905	1.196	0.909	0.747	0.425	0.248	0.146	0.085	0.005
2	41.03	2.366	130.3	0.942	0.753	0.406	0.23	0.132	0.076	0.043
3		3.027	1.409	0.968	0.755	0.39	0.216	0.123	0.07	0.04
4		4.256	1.559	1.009	0.767	0.379	0.207	0.115	0.065	0.037
5		7.1	1.771	1.068	0.787	0.373	0.199	0.11	0.062	0.036
6		16.74	2.088	1.147	0.817	0.37	0.193	0.106	0.06	0.034
7		96.45	2.557	1.25	0.856	0.369	0.19	0.104	0.057	0.032
8			3.387	1.386	0.905	0.37	0.187	0.101	0.056	0.031
9			4.956	1.569	0.966	0.373	0.185	0.099	0.055	0.031
10			8.68	1.82	1.041	0.377	0.184	0.097	0.053	0.03
15				9.995	1.882	0.413	0.183	0.093	0.05	0.028
20					11.16	0.476	0.187	0.092	0.049	0.026
40						2.167	0.236	0.096	0.047	0.025
80							0.818	0.123	0.05	0.024

Figure 46: Cook's Distance

Obs	NAME	PERFORMA	residual_ student	cookd
1	MA	1281127	-0.67138	0.05921
2	GH	1383798	0.10022	0.00176
3	SP	1341491	0.07847	0.00084
4	KC	932329.6	0.45589	0.05443
5	MC	1048395	0.38106	0.01863
6	NP	1016080	-1.67472	0.77140
7	FC	1249760	2.15636	1.27891
8	DW	755219.0	-0.81580	0.18822

Figure 47: Outliers

On the other hand, the studentized residuals (seen above) can also be used to detect possible outliers; any observations with absolute values greater than 2 should be considered as outliers. Using this criterion, subject **HF** can be considered as an outlier. However, it is not an extreme outlier at all. Considering the small sample size, this observation should not be excluded from the analysis.

3. Regression

a. Equation of regression

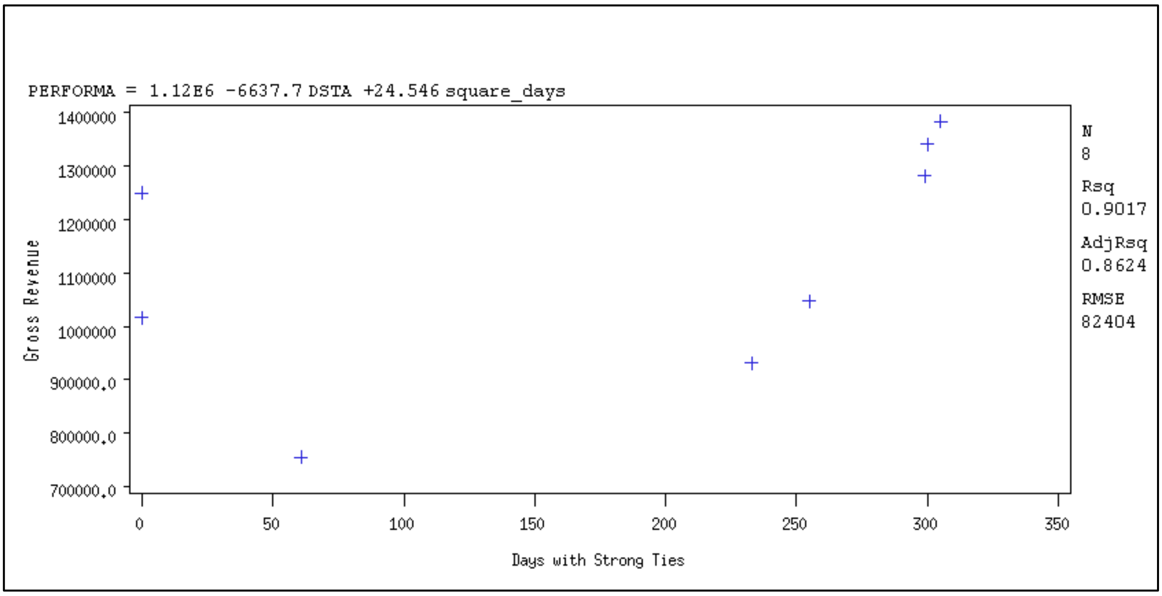


Figure 48: Regression

$$\text{PERFORMA} = 1.12*10^6 - 6637.7\ \text{DSTA} + 24.546\ (\text{DSTA})^2$$

Interpretation the regression equation: DSTA has a quadratic relationship with PERFORMA. When DSTA=0, PERFORMA equals $1.12*10^6$. The lowest value of PERFORMA is 671260, when DSTA = 135.21.

b. Coefficient of determination (R-square)

The coefficient of Determination, R^2, measures the percentage of the variation in the DV that is explained by its relationship with the IV. The coefficient of determination (R-square) of the model is 0.9017. The adjusted R-square is

0.8624. Both are high, indicating approximately 90% of the variability of PERFORMA is explained by DSTA.

c. Residual calculation

Obs	NAME	PERFORMA	residual
1	MA	1281127	-46857.23
2	GH	1383798	6685.79
3	SP	1341491	5441.88
4	KC	932329.6	28113.87
5	MC	1048395	26683.31
6	NP	1016080	-102151.10
7	FC	1249760	131529.29
8	DW	755219.0	-49445.81

Figure 49: Regression - Residual

d. Significance of the regression model (F-test)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	3.114169E11	1.557085E11	22.93	0.0030
Error	5	33951991983	6790398397		
Corrected Total	7	3.453689E11			

Root MSE	82404	R-Square	0.9017
Dependent Mean	1126025	Adj R-Sq	0.8624
Coeff Var	7.31812		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	1118231	55407	20.18	<.0001
DSTA	Days with Strong Ties	1	-6637.73536	1288.01604	-5.15	0.0036
square_days		1	24.54599	4.19761	5.85	0.0021

Figure 50: Regression - F-test

The p-value for the F-test is 0.003, indicating the model is well fit.

4. Partial correlation between performance and days with strong ties

Partial correlation testing is checking the correlation between two variables while controlling for some covariates. In the onboarding regression model, when checking the correlation between PERFORMA and DSTA, $(DSTA)^2$ will be controlled for. When checking the correlation between PERFORMA and $(DSTA)^2$, DSTA will be controlled for.

Pearson Partial Correlation Coefficients, N = 8
Prob > |r| under H0: Partial Rho=0

	PERFORMA	DSTA
PERFORMA Gross Revenue	1.00000	-0.91737 0.0036
DSTA Days with Strong Ties	-0.91737 0.0036	1.00000

Figure 51: Correlation Coefficients

Pearson Partial Correlation Coefficients, N = 8
Prob > |r| under H0: Partial Rho=0

	PERFORMA	square_days
PERFORMA Gross Revenue	1.00000	0.93404 0.0021
square_days Square of DSTA	0.93404 0.0021	1.00000

Figure 52: Correlation Coefficients

In both cases, the correlation coefficients are very high, with the absolute values >0.9, indicating high correlations.

5. Findings of the onboarding GLM model

DSTA has a quadratic relationship with PERFORMA. The linear and quadratic terms of DSTA can explain 90% of the variability of PERFORMA.

6.14 Summary of Findings

6.14.1 Findings on Performance and Social Network Topology:

After using the Box-Cox procedure to transform the data to normality, an ordinary least squares method was followed using stepwise selection. The following model was determined to be the best fit for the data at hand.

$$\mathbf{Nrev = 0.365 + 0.046FTE + 1.006type - 0.508Log\_strong + 0.202Log\_eigen^2 + 0.149*Log\_strong^2}$$

This multivariate regression model shows that when the social network measures are considered together, only log transformed total strong edges (linear, quadratic) and log transformed eigenvector (quadratic) significantly predicted performance. Put more generally, in the case of **HF** teams, performance can be predicated using the number of strong tie relationships amongst the group and by measuring the centrality of the average travel agent in the network, which could be interpreted as meaning:

- performance can be increased to a greater extent if an office is able to increase Tie Strength rather than average eigenvector centrality.
- adding another FTE will not have as significant an impact as changing a retail store to a corporate office (although it would be significantly easier).

6.14.2 Findings on Interpretation of the GLM model:

With the p-value of the F-test <.0001, the regression model is statistically significant. The R-Squared is 0.5839, meaning 58.39% of the variation in Nrev can be explained by the predictors in the model. The R-Squared Adjusted is often used to summarize the fit, as it takes into account the number of predictor variables in the model. In this model, the adjusted R-Squared indicates that about 57.24% of the variation in Nrev is accounted for by the model. The coefficients for each of the predictive variables indicates the amount of change one could expect in Nrev, given a one-unit change in the value of that variable and given that all other variables in the model are held constant. For example, for every one increase in FTE, we would expect a 0.046 increase in Nrev, assuming all other variables in the models are held constant.

6.14.3 Findings on Onboarding and Performance:

With limited data, it is hard to say with any degree of certainty, but at **HF** there appears to be a curvilinear relationship between an individual's performance and the speed at which they build the strong tie relationships necessary for tacit knowledge transfer.

$$\textbf{PERFORMA} = 1.12*106 - 6637.7\ \textbf{DSTA} + 24.546\ (\textbf{DSTA})^2$$

DSTA has a quadratic relationship with PERFORMA. The linear and quadratic terms of DSTA can explain 90% of the variability of PERFORMA. The faster someone builds strong ties, the higher their individual performance, at least based on a small sample.

Summary of Chapter 6

In chapter 6, the researcher analysed the data through a variety of methods, most notably GLM and visual analysis. In the end, a quadratic equation was generated indicating that group normalized revenue can be predicted based on the number of employees, the number of strong relationships in that group, the presence of structural holes and average closeness centrality (being the inverse of the sum of distances between members in the group).

Chapter 7 – Conclusions

This thesis aims to enrich our understanding of the role of social networks in firm performance. In the introduction it was noted that recent studies (notably Lechner, Frankenberger and Floyd, 2010; and Maurer, Bartsch and Ebers, 2011) explore the moderating role that social networks have on knowledge transfer within organizations. The links between knowledge transfer and performance have been explored at length in prior research (Argote, Mcevily and Reagans, 2004; Maurer, Bartsch and Ebers, 2011; Zander and Kogut, 1995). Therefore, this thesis seeks to extend the findings of prior studies (Bulkley and Van Alstyne, 2004; Rice and Steinfeld, 1994; Wellman, 2002; and Whittaker and Sidner, 1996) by considering how SNA variables (e.g. Centrality and Tie Strength) are interrelated and whether there is a temporal dimension to their relationship to each other or to firm performance. Hence, in summary, the objective of this research is to theory test the proposed relationship between group's Social Network Topology and that group's Performance.

7.1 Research Questions Revisited

Using data gathered from organizational units within a single organization, performance was based on group normalized annual sales revenue (**Nrev**). Each unit of analysis comprised a small co-located team repeatedly executing highly similar tasks. Social Network topology was assessed by examining e-mail traffic between members of these teams. In the extant literature, measurement of SNA occurs along three dimensions: cognitive, relational, and structural. In this study, the relational dimension was operationalized as Tie Strength and the structural dimension was operationalized as Centrality and Structural Holes. The cognitive dimension was not incorporated here since it would have required access to the content of e-mail exchanges and this was considered too challenging in terms of the ethical approval of the study. Therefore, it was possible, with the data gathered in this study, to test the following hypotheses:

1. What is the relationship between **Tie Strength** and **Performance**?

2. What is the relationship between **Centrality** and **Performance**?

3. What is the relationship between **Structural Holes** and **Performance**?

Additional the following questions were also explored, specifically:

4. What measures best represent Performance?

5. What, if any, is the relation between **onboarding speed** and individual **performance**?

The literature to date has suggested that performance is correlated to the dimensions of social capital through an inverse curvilinear relationship (e.g. Lechner, Frankenberger, and Floyd; 2010); this research confirmed those findings.

7.2 Contributions to Theory

The analysis of e-mail traffic within teams allows this research to tease out differences between the theoretical concepts of social capital and social network. The former constitutes the social lubricant that facilitates knowledge transfer, while the latter can be defined as the framework over which the former flows. This is an important and oft overlooked difference, since social networks are easier to operationalize than social capital but lead to similar results. Thus, social network measures may be used as a proxy when trying to facilitate performance driven by knowledge transfer. This research contributes to an understanding of how social capital flows over/through the social network, thereby facilitating knowledge transfer, which in turn influences performance.

Unlike Grippa, Zilli, Laubacher and Gloor's 2006 paper, 'E-mail may not reflect the social network', but similar to the findings by Tyler, Wilkinson and Huberman (2005), this research finds that email can be used as a reliable proxy for community structure within organizations. The researcher agrees with prior authors' claims (Tyler, Wilkinson and Huberman, 2005; Wellman, 2002; Whittaker and Sidner, 1996) that email is an appropriate tool to examine group structure. This author does acknowledge that a recent study by MIT may provide a valid counterpoint. In their 2006 article, 'E-mail may not reflect the social network', authors Grippa, Zilli, Laubacher and Gloor suggest that one must be cautious before adopting holus bolus the use of email to determine how work truly gets done. Those authors remind readers that face to face interactions were still the most efficient way to transfer tacit knowledge. They suggest that in groups where co-location of personnel predominates (e.g. the case of **HF**), those actors may opt for more synchronous forms of communication (phone, instant message and face to face). Yet even in those cases where face to face communication is available, email may be used in this fashion, further supporting the concept that email records are a proxy for relationships.

The researcher concedes that the proportion of communication that is email, as compared to other forms (e.g. phone, online chat, face to face), will fall as the percentage of co-located actors rises. Similarly, one expects a drop in email use where actors are in close proximity (as in **HF**). At HF the percentage of actors in close proximity is 100%. Notwithstanding, email continues to be an appropriate tool for measuring relationships and knowledge flow inside of **HF**. For, in the case of **HF**, all network actors share office space and are rarely more than two metres apart. Notwithstanding the likelihood that face to face interactions dominate social interactions amongst the actors and thus are still vital to organizational performance, any group that does not exchange at least ten communications via email in a 365 day period is likely filled with structural holes or weak ties, since even if key knowledge is shared face to face email would still be used for organizational purposes is described above.

The idea that email is appropriate in such settings was first proposed by Bulkley and Van Alstyne in their 2006 Sunbelt Conference paper. In that paper, the authors find significant evidence supporting the interpenetration of email measures as proxies for more general communication patterns even though email use in any organization is context-specific (Bulkley and Van Alstyne 2006; Rice, 1994). This view is also widely supported by several papers from leading authors, including Wellman (2002) and Whittaker and Sidner (1996), who find that email is a strong indicator for levels of collaboration and knowledge exchange, even if email is not the actual tool for such collaboration and knowledge exchange.

Maurer, Bartsch and Ebers' findings (2011) showed that knowledge transfer mediates between intra-organizational social capital and project performance. Similarly, this research found a correlation between the relational and structural dimension and organizational performance. This research showed that, in the case of **HF**, measures of relational dimensions (e.g. tie strength) have a greater impact on performance than do measures of structural dimensions (e.g. centrality). However, the author recommends caution against reading too much into this finding, as it is possible that centrality's impact is greatly mitigated by the small number of actors in each Office (i.e. typically six or less). After all, in such a small group, particularly working in close proximity, centrality may be less important than strong ties. Whereas in larger projects, such as those examined by Lechner, Frankenberger, and Floyd, centrality's role in predicating performance may be greater.

This research contributes to the social network conversation by augmenting and adding to the knowledge generated by recent works, particularly that of Lechner, Frankenberger and Floyd's 2010 study and Maurer, Bartsch and Ebers' 2011 study, both of which found an inverse curvilinear relationship connecting intragroup social network topology with group performance. Unlike Lechner, Frankenberger and Floyd's 2010 study which focused on strategic initiatives (which often only exist for finite periods of time and have specific explicit goals) as the unit of analysis, this research focuses on *Business as Usual* inside more than 187 groups at a national travel agency. Similarly, as compared to Maurer, Bartsch and Ebers' 2011 study which examined 218 projects in the German engineering industry as the unit of analysis, this research focuses on *Business as Usual* inside more than 187 groups at a national travel agency. Thus, this research contributes by showing the findings to be valid for durable (non-temporary) groups not just strategic initiatives or projects.

Empirical results from prior research yielded conflicting correlations, including positive (Florin, Lubatkin, and Schulze, 2003; Subramaniam and Youndt, 2005), insignificant (Batjargal, 2003; Lee, Lee and Pennings, 2001) and negative associations (Edelman, Bresnen, Newell et al., 2004; Gargiulo, 1993), between measures of social capital and dimensions of organizational performance. Lechner, Frankenberger, and Floyd's findings (2010) brilliantly reconcile these apparent inconsistences with their "Dark Side of Social Capital" Theory. This research extends Lechner, Frankenberger, and Floyd's findings (2010) of the inversely curvilinear correlation between Performance and several dimensions of Social

Networks by extending the level of analysis from strategic initiatives to groups (in this case Offices, each of which acts as an independent unit of production). In so doing, this research suggests that it is possible for SNA to be framed in performance terms and to suggest ‘best fit’ topologies.

This research empirically contributes to the concept that social network topology can be a predictive measurement of group performance. Extant non-empirical research has had limited ability to make such an association due to methodological constraints; however, researchers have called for the investigation of the predictive capacity.

This research also enriches the understanding that social network topologies may be leveraged to enhance organizational goals with certain social network topologies being identified as being more effective. To date, although called for, few empirical studies have pursued exploring such a link.

This research also demonstrates that SNA can be achieved in a cost-effective and accurate manner by adopting methodology leveraging software tools. To date, most SNA has been hampered by expensive, time-intensive surveys. This prior survey-based research is both costly and non-dynamic. This research illustrates that, through software methodology, SNA has the great potential to deliver ‘affordable’ and ‘dynamic’ SNA research outcomes.

This research leverages a SNA software-driven methodology rather than a survey-driven procedure, and in doing so affords researchers access to dynamic SNA measures as a proxy for social capital. Prior research was limited to static analysis (e.g. conducted through a survey at the end of the period as to the shape of the social network). Unlike earlier research efforts, the technique used herein can detect WHEN strong ties were formed, something the survey method cannot. By accessing dynamic longitudinal data on tie formation, this research affords a means of investigating the impact of topological factors over time; for example, this research focused on the correlation between individual performance and the speed at which that individual is able to form strong intragroup ties.

7.3 Answering the Research Questions

A firm's social network can only be optimized by first examining the level of exploration vs. exploitation that the firm undertakes. Some firms (e.g. an innovative food company or a Pharmaceutical firm) focus on the creation of new knowledge (products, services); these firms can be seen to be primarily explorative. Some firms (e.g. a large scale manufacturing firm) focus on efficiencies for competitive advantage, often by taking an innovation and exploiting it for efficiency through diffusion. These firms can be seen to be primarily exploitive. In reality, no firm is strictly exploitive or explorative; all firms undertake some tasks that could be categorized as either. All firms are a mix of explorative vs. exploitive (e.g. even a large scale manufacturer) have to create new knowledge, even innovative food companies must find efficiencies).From prior research, exploration moderates relationships between performance and all three dimensions of intergroup social networks. Negative consequences of strong ties and centrality are more pronounced in exploratory initiatives than in 'exploitive' initiatives. Taken to the nth degree, an explorative firm would seek to have a large number of structural holes and a greater diversity in the actor's cognitive background, while an exploitive firm would seek to be less cognitively diverse, with stronger and often redundant ties.

Task Contingency dictates which form optimizes performance. A firm whose task contingency is innovation centric may be better served by a social network containing structural holes (to access diverse information), more weak ties and evenly spread centrality, while a firm with an exploitive focus, might benefit from less ties overall, but from strong intragroup ties (to facilitate knowledge transfer). HF groups sell travel repeatedly. Little new knowledge creation may be needed, but in order to exploit best practices, tacit knowledge must flow. Best practices in turn drive efficiency, which in turn drives performance. Thus, one predicts top performing HF groups should have a majority of strong ties internally, have a balanced centrality, and have few structural holes.

Hypothesis 1: What is the relationship between **Tie Strength** and **Performance?**

As far back as Burt (1992), there have been attempts to explore the impact of **tie strength** on overall **performance**. This investigation demonstrated that a group's relational dimension, as represented by **Tie Strength**, does correlate with group **Performance**. In fact, both were found share an inverse curvilinear relationship. Similar results were also found in prior research.

Earlier theory (Lechner, Frankenberger and Floyd, 2010) suggests that the more exploratory the group (i.e. a group with a greater need for innovation), the more tie diversity is needed for optimal performance. Firms with a high level of exploration need fewer and less redundant ties than groups whose focus is on innovation. This can be explained by the fact that exploratory firms require more ties to more diverse sources of knowledge to drive innovation, while firms focused on exploitation need to diffuse innovation more than they need to facilitate its generation.

HF groups are more exploitive in nature than explorative (e.g. HF groups have limited need to innovate, but instead drive performance through tacit knowledge transfer). From this, the researcher predicted that high-performing HF groups would have mostly strong ties. This was found to be the case. In this research, top performing firms had many strong and deep intragroup ties. This can be explained by looking at the level of analysis. Within HF groups, best practices need (strong ties) to be shared easily. Tacit knowledge transfer requires strong intragroup ties. However, strong ties between HF groups (i.e. intergroup ties) would be burdensome to maintain. Thus the strong ties within the top performing HF groups facilitate the sharing of tacit knowledge which in turn drives those groups' optimal performance.

Hypothesis 2: What is the relationship between **Centrality** and **Performance?**

There have been attempts to explore the impact of **centrality** on overall **performance** (Burt, 1992). **Centrality** is the concept of being 'in the thick of things'. **Centrality** has been used in social network analysis to determine the degree to which a given actor is 'important' within a network. Several measures (degree of centrality, closeness centrality, betweenness centrality, eigenvector centrality, information centrality (Ni, Sugimoto and Jiang, 2010)) used in prior studies as the means of quantifying the flow across a network (Borgatti, 2005).

Based on prior studies, one can conclude that a group with higher Centrality will be seen as less decentralized (Hui, 2008). Decentralization facilitates innovation better than exploitation (Sahay, 2011). For **HF** Groups, mostly focused on exploitation, getting the most from the assets on hand is key. Based on this, one would predict that groups with higher average centrality would have an easier time facilitating knowledge flow and in turn would be in turn better able to drive higher performance. This investigation demonstrated that a group's structural dimension, as represented by Average Group **Centrality**, does correlate with group **Performance**. In fact, both share an inverse curvilinear relationship.

Only Eigenvector Centrality was found to have an impact on **HF** group performance. Eigenvector Centrality reflects the influence individual actors have on the network. In high performing networks one might expect high Eigenvector Centrality.

Hypothesis 3: What is the relationship between **Structural Holes** and **Performance?**

Network cohesion is a structural measure of a social network; it reflects the degree of redundancy occurring within a group. That is to say, the number of redundant ties (paths between actors) within a network represents network cohesion (Burt, 1992). If a network is cohesive, then it could better tolerate actor defection. Network cohesion (sometimes called network redundancy) has the potential to affect the knowledge processes of a group (Fritsch and Kauffeld-Monz, 2009) and, as such, it is of interest to knowledge transfer.

Network cohesion is a metric reflective of the entire network and thus must be calculated as a group-wide measure. This has been done in past studies through empirical survey based social network analysis (Burt, 1992). The measurement of network cohesion also allows us to account for **structural holes** occurring in the network. The Theory of Network Cohesion is often operationalized as either Density or **Structural Holes**. Burt (1992) and others have explored the impact of position (**structural holes**) on overall **performance. Structural holes** can lead to the arrival of non-redundant knowledge to the network (Rodan, 2010); however, with too many **structural holes**, it becomes difficult to defuse innovations throughout a group. Thus, through the use of this metric, it was possible to identify the presence and frequency of **structural holes** within a group.

The researcher predicted that top performing groups would have fewer structural holes as compared to low performing groups. This was found to be true. Amongst the top ten performing **HF** groups, no more than one structural hole per group was found. Similarly, as expected the bottom ten performing **HF** groups, were found to be riddled with **structural holes**. This investigation demonstrated that a group's structural dimension, as represented by **Structural Holes**, does correlate with group Performance. In fact, both share an inverse curvilinear relationship.

Hypothesis 4: What measures best represent **Performance?**

Group performance is contingent on the group's ability to perceive opportunities and capacity to pursue those opportunities (Van de Ven, 1986). Von Hippel (1988) expanded upon Van de Ven's concept of innovation by identifying two different mechanisms by which innovations allow groups to develop and sustain competitive advantage: 1) developing superior efficiency compared to their competitors; and 2) providing superior value for customers. This investigation found that **normalized group revenue** is sufficient to represent group **Performance**. **Rank, Gross Margin, Rev/FTE** are not required to make a valid predictive model. This can be reconciled based on the fact that **Rank, Gross Margin,** and **Rev/FTE** are all contributors to over group revenue, i.e. the higher the **group revenue,** the higher the **Rank;** better **Gross Margins** lead to better **group revenue**; and **Rev/FTE** when aggregated and multiplied by FTE yields a **group revenue**.

This research found that Performance (nrev) can be predicted validly with only:

a. a measure of the structural dimension (Eigenvector Centrality);

b. a measure of the relational dimension (Tie Strength);

c. the number of employees in group and

d. the type of office it is (e.g. corporate vs. retail)

Hypothesis 5: What, if any, is the relation between **onboarding speed** and individual **Performance?**

Researchers have articulated that social ties have the potential to facilitate the flow of all kinds of resources within teams, which correspondingly determines the success of those teams (Balkundi and Harrison, 2006). Firms, that are more successful at rapid onboarding, tend to use a relational approach. This helps newcomers to rapidly establish a broad network of relationships with co-workers to obtain the information they need to become productive (Rollag et al., 2005). In addition to exploring the relationship between performance and social network topology, this research additionally examines the concept of dynamic tie formation. Knowledge sharing and application are widely recognized as the key determinants of team performance (Choi et al., 2010; Janhonen and Johanson, 2011).

To facilitate knowledge sharing, strong ties need to be formed since tacit knowledge travels over strong ties (Hansen, 1999; Levin, 2004; Li and Zhu, 2009; Nie, 2010), Tacit knowledge enhances performance (e.g. as best practices are shared, individual performance grows). This is particularly informative with regard to onboarding. During Onboarding, new nodes (FTEs) are added to the network. Those nodes form ties with the other members of the Office. The quicker those ties become strong, the quicker tacit knowledge can flow. The researcher predicted that those newly hired agents who are able to form strong ties faster, will better benefit from ability to transfer tacit knowledge earlier, and this in turn will drive the performance of the agents.. Put more succinctly, the researcher predicted a strong positive correlation between individual **performance** and **onboarding speed.** This investigation indeed found a positive relationship between

onboarding speed and individual **performance**, although this was found based on very limited data.

Extrapolating from the Research Question findings:

- Only groups with tie strength significant enough to facilitate tacit knowledge transfer had high performance.
- Top performing groups had strong ties throughout out the network, although each had one or two visible structural holes. Low performing groups had low Density. Low performing offices' social networks had many structural holes throughout their network. These may have potentially undermined the formation of strong ties, which in turn may have impeded tacit knowledge transfer and performance.
- Individual agents who quickly formed strong ties faster performed better, sooner.

Table 21: Summary of Research Question Findings

Results	Findings
What is the relationship between Centrality and Performance?	Inverse Curvilinear
What is the relationship between Centrality and Performance?	Inverse Curvilinear
What is the relationship between Structural Holes and Performance?	Curvilinear
What, if any, is the relation between onboarding speed and individual performance?	Positive

7.4 The Developed Model for Predicting Performance

Predicting Group Performance from Social Network Topology

After using the Box-Cox procedure to transform the data to normality, an ordinary least squares method was followed through stepwise selection. The following model was determined to be the best fit for the data at hand:

$$\boxed{\begin{gathered}\mathbf{Nrev = 0.365 + 0.046FTE + 1.006type - 0.508Log\_strong} \\ \mathbf{+ 0.202Log\_eigen^2 + 0.149{*}Log\_strong^2}\end{gathered}}$$

With the p-value of the F-test <.0001, the regression model is statistically significant. The R-Squared is 0.5839, meaning 58.39% of the variation in Nrev can be explained by the predictors in the model. The R-Squared Adjusted is often used to summarize the model fit as it takes into account the number of predictor variables in the model. In this model, the adjusted R-Squared indicates that about 57.24% of the variation in Nrev is accounted for by the variables in the model.

The coefficients for each of the predictive variables indicate the amount of change one could expect in Nrev, given a one-unit change in the value of that variable, and given that all other variables in the model are held constant. For example, for every one increase in FTE, we would expect a 0.046 increase in Nrev, assuming all other variables in the models are held constant. This multivariate regression model shows that when the social network measures are considered together, only log transformed total strong edges (linear, quadratic) and log transformed Eigenvector Centrality (quadratic) significantly predicted performance. Put more generally, in the case of **HF** teams, once one knows the Office type (ie. Retail or Corporate) and the # of FTEs, then Performance can be predicated using only the number of strong tie relationships amongst the group and by measuring how central the average travel agent in the network. This could be interpreted as meaning:

- Performance can be increased to a greater extent if an office is able to increase Tie Strength rather than average eigenvector centrality.
- Adding another FTE will not have as much impact as changing a retail store to a corporate office (although it would be significantly easier).

Revisiting the Model

In Chapter 4, the researcher reviews relevant material models from recent prior art. From these the researcher developed the following model for this research:

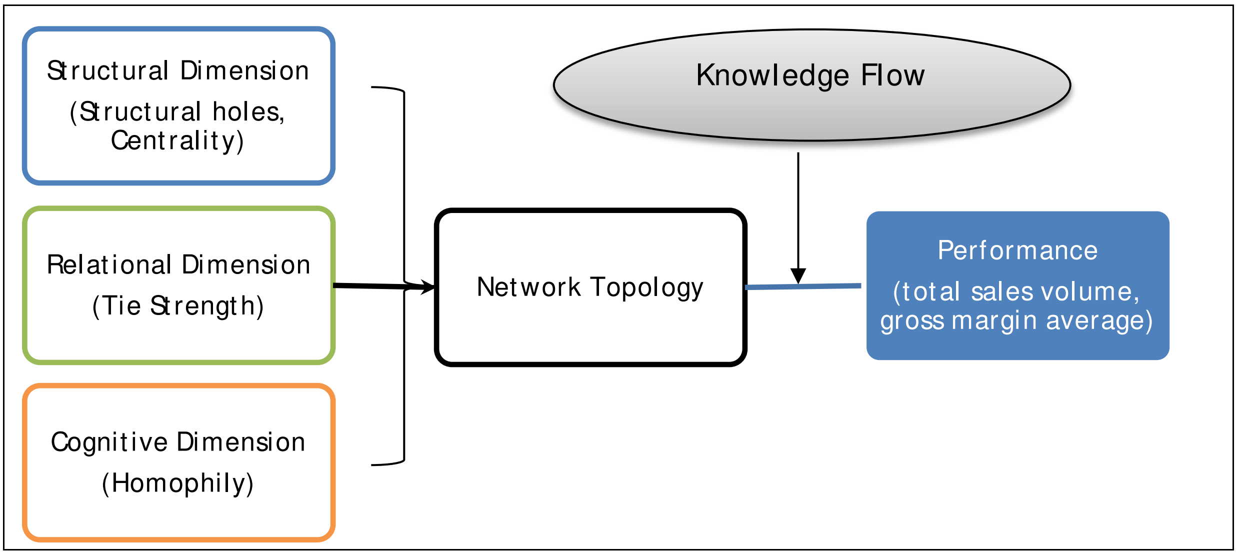


Figure 53: The Research Model

This research concentrates primarily on the internal knowledge processes which impact the group's ability to transfer knowledge (internally) and perform as reflected by the network structure of the group, the model adopted by this research looks directly at the effects of network topography on group performance. The conceptual model which forms the logical framework for how network topology impacts organizational performance is given above. This model outlines the causal relationships between the different factors, identified in management literature, which are known to impact the transformation of knowledge into performance. The model below summarizes the findings of this research visually.

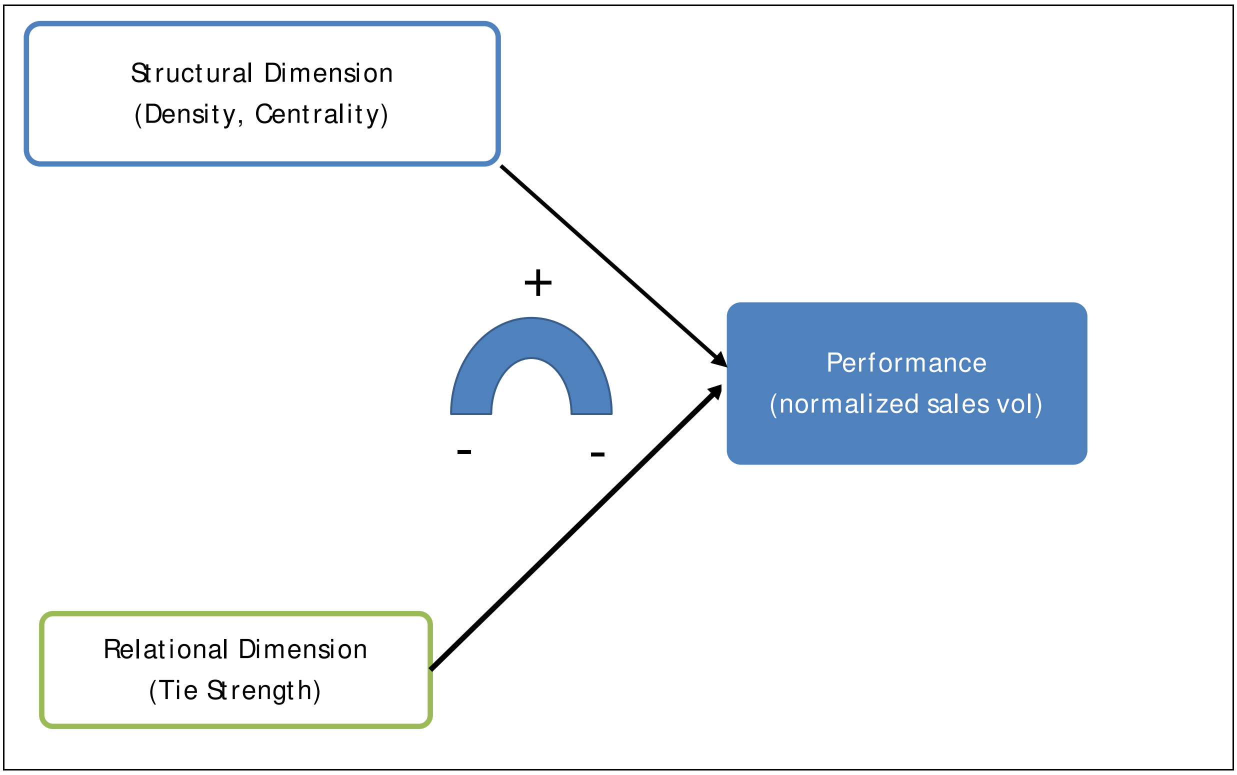


Figure 54: The Research Model showing the Inverse Curvilinear Relationship found

Predicting Individual Performance from Onboarding Speed

With limited data, it is hard to say with any degree of certainty, but at **HF** there appears to be a curvilinear relationship between an individual's performance and the speed at which they build the strong tie relationships necessary for tacit knowledge transfer.

$$\textbf{PERFORMA = 1.12*106 – 6637.7 DSTA + 24.546 (DSTA)}^2$$

DSTA has a quadratic relationship with PERFORMA. The linear and quadratic terms of DSTA can explain 90% of the variability of PERFORMA. Thus the data seems to indicate that the faster someone built strong ties the higher their individual performance, at least on a small sample.

The table below summarizes the findings across the various sets of independent analysis:

Table 22: Independent Analysis Summary

	Findings
GLM Analysis	Performance is best measured by Nrev alone. The structural social network measures (density, avg geo distance, avg. betweenness, log eigenvector, and closeness) are all significantly associated with performance. The relational social network measures (log total strong edges and Tie Strength) were highly correlated with performance. Performance had a U relation with log total strong edges. Performance had an inverse U relation with Tie Strength.
Visual Analysis	High performing Offices had mostly strong tie relationships. High performing Offices had high average eigenvector centrality and centrality was evenly distributed (all members had almost identical). High performing Offices had only a few structural holes. Low performing Offices seem to lack the ties necessary to facilitate tacit knowledge transfer. Low performing Offices unevenly distributed centrality (typically one or two actors with high eigenvector centrality, and the remaining actors will little centrality). Low performing Offices have too many structural holes.
Onboarding Analysis	A positive relationship between onboarding speed and individual performance was found, although this was found based on very limited data.

7.5 Implications for Practice

Management of HF's primary goal is to increase shareholder value by optimizing performance. This research empirical informs that goal by providing insight into the optimal HF office size, structure and social topology. Top performing groups were seen as sharing social network topology, as were underperforming groups. HF management may now consider management intervention to address underperforming groups' topology.

Do more agents make more money?

As Offices add more staff, Performance is enhanced. This is likely because more potential ties become available for knowledge flow. There may be a cap to this, after which too many staff becomes a problem

Why was only eigenvector centrality found to be predicative?

This research explored three measures of centrality. Centrality is the concept of being 'in the thick of things'. Centrality has been used in social network analysis to determine the degree to which a given actor is 'important' within a network. Several measures have been derived from this definition of centrality including: degree of centrality, closeness centrality, betweenness centrality, eigenvector centrality and information centrality (Ni, Sugimoto and Jiang, 2010). The different measures of centrality reflect slightly different network phenomena; however, each measure of centrality allows us to perceive how 'central' given actors may be within a network. Three centralities were explicitly explored in this interest:

- **Eigenvector Centrality** is the measure of the influence of a node on the network. Thus, it is the influence that any one group member (Travel Agent) can have on the group (Office). A node with high eigenvector centrality will be able to strongly influence other members of that group. A group with high average eigenvector centrality would see many nodes able to influence the network.

- **Closeness Centrality** determines the distance between the nodes. In mapping social graphs there is a natural distance between pairs of nodes. This distance (farness) is defined by the length of the shortest path to connect them. The distance of a node is calculated as the sum of all the shortest paths. Closeness Centrality is the inverse of Farness. It is often regarded as a measure of how long it would take to spread information along the shortest paths. This research focuses on the performance benefits resulting from the spread of tacit knowledge (e.g.

best practices). The researcher predicts that this measure of centrality would be predictive. However, this was not found to be the case. Closeness Centrality played no role in the final model. Perhaps the small size of the groups (i.e. six or less), the large number of redundant ties and the close proximity of actors in the networks lead to the lack of impact from Closeness Centrality.

- **Betweenness Centrality** refers to the extent to which a node (representing an actor) lies between other nodes in the network. This measure takes into account the connectivity of the node's neighbours, giving a higher value for nodes which bridge clusters. Betweenness centrality reflects the number of people with whom a person is connecting indirectly through their direct links. As tie redundancy rises, this measure may become less meaningful.

In this research only Eigenvector Centrality was proven a predictor of group performance. One possible explanation might be that eigenvector centrality deals with the ability to influence others in the network. An actor with high eigenvector centrality will be able to strongly influence other members of that group. High performing teams had a high group average of eigenvector centrality, which when examined visually appears to represent a group in which most members have high eigenvector centrality. As for why neither of the other forms of centrality were found to be predictive, one possible explanation might relate to the size of the network. Most groups have six or less actors; this limits the amount of alternative paths knowledge might take and thus may undermine the impact of closeness centrality. Similarly, betweenness centrality reflects the number of people with whom a person is connecting indirectly through their direct links. In most teams direct links are sufficient, indirect paths are rarely needed due to the small network size.

Why were Strong Ties present in the final model?

There were several possible measures for the relational dimension. Potential ties, Number of Strong Ties, and the ratio of the two (Tie Strength) were available to the predictive model, but only the actual number of strong ties was found to have significant impact on the model. Again, the issue of small network size may be at play but the researcher prefers an alternative view. The number of potential ties is directly related to the number of FTEs in the network. FTEs are included directly in the regression model, thereby possibly undermining the need to include potential ties.

The ratio of potential ties to strong ties was also not found material. Perhaps this indicates it is less a feature of how many strong ties are formed from the potential number that could be formed and more a feature of the absolute number of strong ties over which tacit knowledge travels.

Why were Structural Holes not in the final model?

Structural holes were not found to be predictive of group performance. This seems odd as the visual analysis clearly demonstrates that top performing offices have one or no holes and low performing offices are riddled with structural holes. Structural holes tend to confer a network with access to non-redundant information. This novel information is key to forming new knowledge. But at **HF** innovation (the appliance of creativity to generate new knowledge) seems to have less impact than ensuring current knowledge is fully exploited based on HF's task contingency.

This is not of great concern to the researcher, as the structural dimension of social capital is measured by both structural holes and centrality, and as shown above centrality plays a large role in predicting performance.

Why does only Gross Revenue matter when it comes to Performance?

This investigation found that normalized group revenue is sufficient to predict group Performance. Rank, Gross Margin and Rev/FTE were not required to make a valid predictive model. A model was adopted from theory testing and found to be consistent with prior research (e.g. the inverse curvilinear relationship between the relational and structural dimensions of social capital). This investigation found a positive relationship between onboarding speed and individual performance, although this was found based on very limited data.

Is there an optimal HF Topology?

The top performing offices were all found to be strong tie networks with no more than one structural hole. This makes sense as high-performing offices were found to be closed dense interpersonal networks, necessary for tacit knowledge to be facilitated and for commercial norms to develop. Strong relational capital was also found to be high in these offices. A simple visual comparison between the top ten performing teams and the bottom ten performing teams says it all. High performing teams were found to be closed, densely populated networks. It was also shown, albeit weakly, that a correlation may exist between the ability to form strong ties and individual performance?

7.6 Generalization of Findings

This research is a theory testing dissertation. It focused on providing empirical evidence of the inverse relationship between social network topology and group performance. This work confirmed such. While there are sufficient data points to generalize the impact of social network topology on **HF** group performance, the use of one firm (HF) for analysis limits the ability to generate more generalized findings. Notwithstanding, one can imagine similar results occurring at firms with similar task contingency (e.g. car rental agencies; real estate firms). All groups in this study worked in Canada for the same multinational corporation. As a result, these findings cannot be widely generalized.

Further, the task contingency of **HF** is very exploitive in nature (as opposed to innovative) with FTEs repeating the same task many hundreds of times over a one-year period. Based on this, tacit knowledge (e.g. best practices) would likely need to travel efficiently over a strong tie network to generate increases in performance. Thus one would expect high performing firms to have many (but not all) strong intragroup tie relationships. Further, high performing groups would need to have high average eigenvector centrality (i.e. no one actor is more important than another).

One might conclude that other groups undertaking repetitive homogenous jobs with highly exploitive task contingencies (e.g. car salesmen, tax auditors, real estate agents, etc.) might share similar social network topologies to that of **HF** groups.

7.7 Limitations of Findings

All research involves trade-offs and compromises (Mackert, 2009). One often trades accuracy for limitations. By tightly focusing one's research, one may generate a deeper understanding of phenomena but limit the generalizability of those results. In summary, the author acknowledges the following limitations:

- The research focused on only one group (**HF**), one industry (Travel Services), and one task (selling travel). This extremely limits the findings to that explicit case. One cannot apply the findings to all travel agencies in Canada, let alone other countries or industries. Using only one corporation for data may limit broader generalizations. While 180+ offices were examined, all work for the same parent corporation. This too limits our ability to draw broader generalizations from the work.

- Relational measures (e.g. Tie Strength) were calculated using the frequency of email communications. But not all conversations are conducted by email; hence some network activities (i.e. phone calls and face to face conversations) cannot be traced through this methodology. Further, since tacit knowledge transfer may be the driving force, this issue is exacerbated. Most tacit knowledge transfers best through face to face conversations (Alexander, 2012; Dinur, 2011; Lin et al., 2011; Wu et al., 2008). In the case of **HF** most agents are in extremely close proximity to their teammates because most offices have six to eight agents in a space smaller than 1000 square feet. It thus seems likely that email would be used less frequently. Yet email still serves as a proxy for relationships (e.g. agent 1 emails agent 2 a hundred times in a year; it is unlikely that no relationship exists).

- Lechner, Frankenberger and Floyd (2010) and other researchers have used perception (i.e. with whom *do* you speak?) not actual (with whom *did* you speak?), communications as this research does. In doing so, perception based studies may suffer from recency effects as well as incorrect subjective opinions as to the frequency. While this research avoids both such issues, it does so at a cost. Communications counted objectively by the software lack subjective context, and in doing so all communications are treated as equal, which may impact the results.

- No cognitive measures were allowed by **HF**. Originally, the researcher wanted to test all three dimensions of social capital. The Cognitive dimension, which is often operationalized as cognitive similarity, shared norms and/or homophily, provides insight into the ease with which knowledge can be transferred cognitively. Prior research has shown actors who share norms and terms of reference are able to more easily transfer tacit knowledge. While those extremely diverse groups may have difficulty doing so. In the case at hand, agents act under the same task contingency (i.e. undertaking the selling of travel for maximum profit). This may suggest that a **HF** group with high homophily and shared norms would have an easier time transferring best practices (and other forms of tacit knowledge). It is disappointing that this researcher was not able to explore such, and it limits the researcher's ability to confirm that the inverse curvilinear relationship exists across all three dimensions of the social network.

- Another potentially limiting issue is office type. Of the top ten **HF** offices, eight offices are corporate. Of the bottom ten, nine offices are retail. It is also worthy to note that the only potential outlier in the data comes from a retail e-commerce office (b2c) that acts like a corporate office (b2b). It might be interesting, in the future, to run the data as two distinct sets: corporate and retail. One argument for not doing separate analysis is that for the office type, if split, the corp variable offices (e.g. was this a B2B or B2C office) comprise less than 10% of the total records. This would be a small sample, and the analysis arising from that might therefore be of limited value.

7.8 Future extensions

The research provides empirical evidence of the inverse curvilinear relationship between group performance and social network topology. As with Inkpen and Tsang (2005), it is likely that different types of organizational settings may display unique relations among the various dimensions of social capital, knowledge transfer, and performance outcome), it would be a fruitful avenue for future research to examine to what extent our findings hold in other types of intra-organizational settings. For instance, would these results appear in other settings with similar task contingencies (e.g. A Car Rental Franchise)? More research on the interrelationship between social network topology and performance is warranted. Specifically, it would be interesting to extend this research in any of the following manners:

1. Do these findings apply to **HF** groups outside of Canada?

2. Do these findings apply to other travel agencies?

3. Do these findings hold true in other industries?

4. Each **HF** group is nested and has its own social relations with 'head office', the same head office that decides which resources go to which team (e.g. new agents, sales leads). Interfirm network position could thus impact group performance because of this resource allocation (e.g. teams well connected to head office may get the best leads). Further research to explore what role interfirm network position plays on group performance would be fruitful.

5. More data related to onboarding speed and performance could be pursued to further evidence and explore any potential correlations. e.g. Could onboarding speed ever lead to negative results?

Summary of Chapter 7

This thesis examines the effects of intragroup social network relations on group performance. Building on prior studies, social network topology was viewed along structural, relational and cognitive dimensions. Where previous research used a self-reporting questionnaire approach to generate SNA measures, this research uses Social Network Analysis Software to leverage e-mail communication logs to produce SNA measurers.

This study was conducted on a national travel agency where the social networks of 180+ offices were examined. Each office was tasked similarly and represented a unit of analysis. An analysis of more than 7 million emails was used to generate social network measures for the firm wide network. Subgraphs, representing the intraoffice social networks, were generated for each office. NodeXL® software was used to generate group measures representing the dimensions of each office's network topology. As in prior literature, Centrality, Structural Holes, and Tie Strength were used to measure and compare the dimensions of the intragroup networks. This study confirms empirically existing findings of an inverse curvilinear relationship correlating social network topology and firm performance (i.e. The Dark Side of Social Capital Theory). This study also extends prior research on new employee socialization (e.g. onboarding) by dynamically examining the tie formation amongst recently hired employees, finding a positive correlation between an individual's onboarding speed and their performance.

Appendix A: Data Sheet

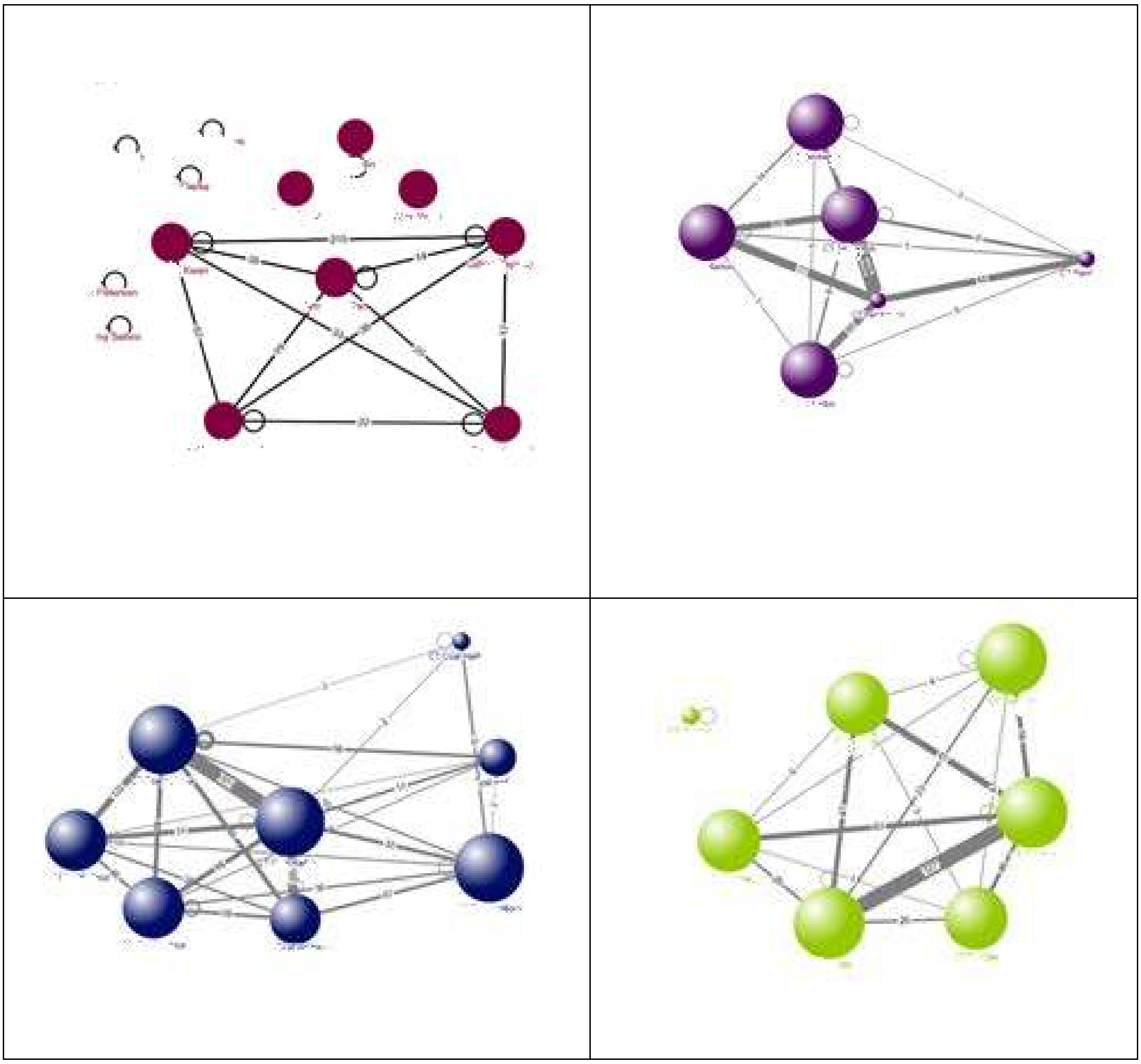
FTE	Office Type	RANK	Gross Revenue(TSV)	Normalized Rev	Rev/FTE	Gross Margin Avg.	Density	Structual Holes	Avg. Geo Distance	Avg. Eigenvector Centr.	Avg. Betweenness Centr.	Avg. Closeness Centr.	Total Strong Edges	Potential edges	Strength of Ties
14	retail	1	$19,668,280.01	$4.85	$1,404,877.14	6.94%	0.128	7.800	0.645	0.091	0.000	0.114	29	91	0.32
6	corp	2	$14,069,885.00	$3.47	$2,344,980.83	14.93%	0.933	1.071	0.833	0.167	0.000	0.200	32	15	2.13
6	corp	3	$11,647,103.92	$2.87	$1,941,183.99	14.18%	0.600	1.667	1.040	0.200	0.600	0.197	13	15	0.87
10	corp	4	$11,129,319.04	$2.74	$1,112,931.90	16.35%	0.768	1.302	1.031	0.125	0.625	0.124	50	45	1.11
9	corp	5	$10,786,718.20	$2.66	$1,198,524.24	13.81%	0.690	1.448	0.811	0.143	0.000	0.171	33	36	0.92
8	corp	6	$10,509,500.08	$2.59	$1,313,687.51	10.52%	0.767	1.304	0.000	0.167	0.333	0.179	26	28	0.93
6	corp	7	$9,955,509.58	$2.45	$1,659,251.60	14.42%	0.650	1.538	1.040	0.200	0.600	0.199	16	15	1.07
6	retail	8	$9,755,775.97	$2.40	$1,625,962.66	11.39%	0.357	2.800	1.081	0.143	0.714	0.133	16	15	1.07
5	corp	9	$9,632,926.33	$2.37	$1,926,585.27	12.81%	0.800	1.250	1.000	0.167	0.500	0.170	28	10	2.80
8	corp	10	$9,563,512.72	$2.36	$1,195,439.09	13.52%	0.733	1.364	1.056	0.167	0.667	0.163	26	28	0.93
6	corp	11	$9,539,009.55	$2.35	$1,589,834.93	18.03%	0.700	1.429	0.889	0.167	0.167	0.189	21	15	1.40
7	corp	12	$8,617,291.45	$2.12	$1,231,041.64	14.73%	0.667	1.500	0.800	0.200	0.000	0.250	23	21	1.10
9	retail	13	$7,629,396.94	$1.88	$847,710.77	10.89%	0.267	3.750	1.231	0.167	1.000	0.137	10	36	0.28
8	corp	14	$7,319,635.60	$1.80	$914,954.45	18.12%	1.000	1.000	0.800	0.200	0.000	0.250	22	28	0.79
13	retail	15	$7,173,466.15	$1.77	$551,805.09	8.23%	0.381	2.625	1.278	0.167	1.333	0.134	11	78	0.14
9	retail	16	$7,076,703.52	$1.74	$786,300.39	7.97%	0.143	7.000	0.545	0.200	0.000	0.300	7	36	0.19
5	corp	17	$6,737,288.69	$1.66	$1,347,457.74	13.62%	0.650	1.538	1.040	0.200	0.600	0.199	18	10	1.80
9	retail	18	$6,712,903.00	$1.65	$745,878.11	8.14%								36	0.00
4	corp	19	$6,702,369.52	$1.65	$1,675,592.38	14.40%	0.650	1.538	1.040	0.200	0.600	0.199	17	6	2.83
8	retail	20	$5,974,770.24	$1.47	$746,846.28	8.86%	0.333	3.000	1.429	0.143	2.000	0.104	14	28	0.50
14	retail	21	$5,958,367.34	$1.47	$425,597.67	10.97%			0.000	0.200	0.000	0.000	5	91	0.05
6	corp	22	$5,926,266.63	$1.46	$987,711.11	17.46%								15	0.00
10	corp	23	$5,852,168.02	$1.44	$585,216.80	10.41%	0.500	2.000	1.000	0.250	0.500	0.258	9	45	0.20
8	retail	24	$5,773,395.95	$1.42	$721,674.49	8.82%	0.500	2.000	0.824	0.200	0.200	0.233	11	28	0.39
12	retail	25	$5,764,728.49	$1.42	$480,394.04	10.46%	0.333	3.000	0.778	0.167	0.167	0.194	12	66	0.18
9	retail	26	$5,758,192.79	$1.42	$639,799.20	11.21%	0.476	2.100	0.741	0.143	0.000	0.179	24	36	0.67
12	retail	27	$5,753,412.76	$1.42	$479,451.06	8.81%			0.000	0.250	0.000	0.000	4	66	0.06
9	retail	28	$5,736,982.07	$1.41	$637,442.45	13.07%	0.905	1.105	0.939	0.143	0.286	0.154	37	36	1.03
8	retail	29	$5,721,373.76	$1.41	$715,171.72	9.60%	0.200	5.000	0.727	0.200	0.200	0.233	6	28	0.21
8	retail	30	$5,682,962.70	$1.40	$710,370.34	9.64%	0.333	3.000	0.750	0.100	0.000	0.120	31	28	1.11
9	retail	31	$5,656,932.46	$1.39	$628,548.05	10.34%	0.286	3.500	1.037	0.143	0.571	0.132	15	36	0.42
8	retail	32	$5,612,676.10	$1.38	$701,584.51	8.89%	0.800	1.250	1.000	0.167	0.500	0.171	21	28	0.75
10	retail	33	$5,551,810.18	$1.37	$555,181.02	9.90%	0.143	7.000	1.250	0.250	1.000	0.208	4	45	0.09
9	retail	34	$5,521,058.23	$1.36	$613,450.91	9.90%	0.619	1.615	0.944	0.167	0.333	0.179	20	36	0.56
8	retail	35	$5,506,478.00	$1.36	$688,309.75	10.92%	0.267	3.750	0.941	0.200	0.400	0.207	6	28	0.21
11	retail	36	$5,486,824.96	$1.35	$498,802.27	10.98%	0.238	4.200	0.778	0.167	0.167	0.194	12	55	0.22
8	retail	37	$5,434,722.99	$1.34	$679,340.37	12.89%	0.333	3.000	1.243	0.143	1.143	0.116	13	28	0.46
12	retail	38	$5,329,602.86	$1.31	$444,133.57	8.46%	0.286	3.500	0.857	0.125	0.250	0.133	20	66	0.30
9	retail	39	$5,269,958.46	$1.30	$585,550.94	11.12%			0.000	0.250	0.000	0.000	4	36	0.11
11	retail	40	$5,236,915.72	$1.29	$476,083.25	8.65%	0.357	2.800	0.741	0.143	0.000	0.179	26	55	0.47
9	retail	41	$5,234,543.74	$1.29	$581,615.97	7.67%	0.667	1.500	0.769	0.167	0.000	0.208	17	36	0.47
4	corp	42	$5,210,135.59	$1.28	$1,302,533.90	14.00%	0.500	2.000	0.889	0.333	0.333	0.389	5	6	0.83
10	retail	43	$5,176,875.96	$1.28	$517,687.60	12.55%	0.190	5.250	1.481	0.143	1.429	0.093	11	45	0.24
11	retail	44	$5,152,714.73	$1.27	$468,428.61	8.31%	0.143	7.000	0.500	0.167	0.000	0.250	8	55	0.15
8	retail	45	$5,088,188.07	$1.25	$636,023.51	8.60%	0.714	1.400	0.811	0.143	0.000	0.171	31	28	1.11
15	retail	46	$5,022,524.57	$1.24	$334,834.97	9.94%	0.278	3.597	0.690	0.111	0.000	0.139	22	105	0.21
12	retail	47	$4,988,919.66	$1.23	$415,743.31	10.39%			0.000	0.250	0.000	0.000	4	66	0.06

FTE	Office Type	RANK	Gross Revenue(TSV)	Normalized Rev	Rev/FTE	Gross Margin Avg.	Density	Structual Holes	Avg. Geo Distance	Avg. Eigenvector Centr.	Avg. Betweenness Centr.	Avg. Closeness Centr.	Total Strong Edges	Potential edges	Strength of Ties
11	retail	48	$4,982,582.45	$1.23	$452,962.04	9.81%	0.067	15.000	0.250	0.167	0.000	0.333	8	55	0.15
10	retail	49	$4,930,875.44	$1.21	$493,087.54	12.28%	0.286	3.500	1.037	0.143	0.571	0.132	10	45	0.22
12	retail	50	$4,918,626.68	$1.21	$409,885.56	11.93%	0.238	4.200	0.778	0.167	0.167	0.194	13	66	0.20
5	retail	51	$4,883,739.48	$1.20	$976,747.90	11.65%			0.000	0.200	0.000	0.000	5	10	0.50
13	retail	52	$4,748,638.84	$1.17	$365,279.91	11.07%	0.133	7.500	0.667	0.167	0.167	0.194	9	78	0.12
9	retail	53	$4,742,942.69	$1.17	$526,993.63	8.74%	0.100	10.000	0.500	0.500	0.000	1.000	2	36	0.06
7	retail	54	$4,662,483.06	$1.15	$666,069.01	10.70%	0.222	4.500	1.128	0.111	0.778	0.095	19	21	0.90
12	retail	55	$4,607,527.74	$1.14	$383,960.65	9.53%	0.095	10.500	0.615	0.143	0.143	0.167	9	66	0.14
9	retail	56	$4,604,505.99	$1.13	$511,611.78	10.12%	0.214	4.667	0.632	0.143	0.000	0.190	16	36	0.44
7	retail	57	$4,586,621.16	$1.13	$655,231.59	9.70%	0.200	5.000	1.059	0.200	0.600	0.187	7	21	0.33
6	retail	58	$4,539,890.45	$1.12	$756,648.41	12.43%	0.167	6.000	0.400	0.333	0.000	0.667	4	15	0.27
9	retail	59	$4,535,434.42	$1.12	$503,937.16	9.55%	0.467	2.143	1.000	0.167	0.500	0.167	11	36	0.31
10	retail	60	$4,328,858.81	$1.07	$432,885.88	10.32%	0.190	5.250	0.842	0.143	0.286	0.148	11	45	0.24
8	retail	61	$4,285,405.01	$1.06	$535,675.63	8.92%	0.524	1.909	1.027	0.143	0.571	0.140	16	28	0.57
9	retail	62	$4,274,362.57	$1.05	$474,929.17	8.64%	0.306	3.273	0.974	0.111	0.444	0.109	19	36	0.53
7	retail	63	$4,272,333.49	$1.05	$610,333.36	10.89%	0.429	2.333	0.815	0.143	0.143	0.164	18	21	0.86
9	retail	64	$4,231,049.45	$1.04	$470,116.61	8.83%	0.200	5.000	1.059	0.200	0.600	0.187	9	36	0.25
9	retail	65	$4,152,833.85	$1.02	$461,425.98	10.00%	0.067	15.000	0.333	0.250	0.000	0.500	3	36	0.08
10	retail	66	$4,134,058.43	$1.02	$413,405.84	9.82%			0.000	0.250	0.000	0.000	4	45	0.09
6	retail	67	$4,120,627.64	$1.02	$686,771.27	8.77%	0.600	1.667	1.120	0.200	0.800	0.185	7	15	0.47
5	retail	68	$4,056,035.58	$1.00	$811,207.12	9.10%	0.667	1.500	0.769	0.167	0.000	0.208	25	10	2.50
12	retail	69	$4,027,403.80	$0.99	$335,616.98	11.59%	0.300	3.333	0.545	0.200	0.000	0.300	11	66	0.17
6	retail	70	$4,023,437.80	$0.99	$670,572.97	9.57%	0.333	3.000	0.889	0.333	0.333	0.389	5	15	0.33
9	retail	71	$3,936,349.85	$0.97	$437,372.21	11.34%	0.333	3.000	0.778	0.167	0.167	0.194	11	36	0.31
6	retail	72	$3,906,716.10	$0.96	$651,119.35	11.35%	0.400	2.500	0.941	0.200	0.400	0.207	9	15	0.60
10	retail	73	$3,879,132.39	$0.96	$387,913.24	10.87%	0.200	5.000	0.800	0.250	0.250	0.292	5	45	0.11
8	retail	74	$3,877,460.34	$0.96	$484,682.54	13.02%	0.500	2.000	0.600	0.250	0.000	0.375	5	28	0.18
18	retail	75	$3,841,280.79	$0.95	$213,404.49	9.23%	0.095	10.500	0.727	0.200	0.200	0.233	6	153	0.04
9	retail	76	$3,786,382.17	$0.93	$420,709.13	11.62%	0.524	1.909	1.056	0.167	0.667	0.163	21	36	0.58
5	retail	77	$3,781,267.75	$0.93	$756,253.55	9.23%	0.333	3.000	0.800	0.250	0.250	0.292	4	10	0.40
7	retail	78	$3,773,265.39	$0.93	$539,037.91	10.04%	0.100	10.000	0.333	0.250	0.000	0.500	4	21	0.19
7	retail	79	$3,769,034.86	$0.93	$538,433.55	15.60%	0.267	3.750	0.941	0.200	0.400	0.207	8	21	0.38
9	retail	80	$3,752,193.36	$0.92	$416,910.37	10.63%	0.100	10.000	0.333	0.250	0.000	0.500	6	36	0.17
10	retail	81	$3,712,728.83	$0.91	$371,272.88	9.98%	0.100	10.000	0.333	0.250	0.000	0.500	5	45	0.11
14	retail	82	$3,698,201.85	$0.91	$264,157.28	12.31%	0.267	3.750	0.889	0.167	0.333	0.172	12	91	0.13
10	retail	83	$3,680,446.91	$0.91	$368,044.69	8.57%	0.100	10.000	0.286	0.200	0.000	0.400	4	45	0.09
7	retail	84	$3,678,313.12	$0.91	$525,473.30	13.01%	0.700	1.429	1.040	0.200	0.600	0.199	15	21	0.71
3	retail	85	$3,660,711.36	$0.90	$1,220,237.12	15.26%	0.400	2.500	0.941	0.200	0.400	0.207	8	3	2.67
11	retail	86	$3,645,826.70	$0.90	$331,438.79	9.37%	0.200	5.000	0.889	0.333	0.333	0.389	3	55	0.05
9	retail	87	$3,638,291.62	$0.90	$404,254.62	10.04%	0.200	5.000	0.545	0.200	0.000	0.300	7	36	0.19
11	retail	88	$3,630,013.36	$0.89	$330,001.21	8.90%	0.067	15.000	0.250	0.167	0.000	0.333	8	55	0.15
8	retail	89	$3,628,340.81	$0.89	$453,542.60	12.78%	0.300	3.333	0.545	0.200	0.000	0.300	11	28	0.39
8	retail	90	$3,591,090.36	$0.88	$448,886.30	9.24%								28	0.00
6	retail	91	$3,549,754.96	$0.87	$591,625.83	11.81%	0.139	7.200	0.737	0.143	0.143	0.167	13	15	0.87
10	retail	92	$3,547,704.17	$0.87	$354,770.42	9.50%	0.067	15.000	0.462	0.143	0.000	0.214	10	45	0.22
5	retail	93	$3,526,773.01	$0.87	$705,354.60	9.20%	0.400	2.500	0.667	0.167	0.000	0.222	15	10	1.50
9	retail	94	$3,511,410.52	$0.86	$390,156.72	9.38%			0.000	0.333	0.000	0.000	3	36	0.08
9	retail	95	$3,509,600.60	$0.86	$389,955.62	11.98%	0.381	2.625	0.923	0.167	0.333	0.178	16	36	0.44
7	retail	96	$3,462,252.16	$0.85	$494,607.45	10.17%	0.533	1.875	1.222	0.167	1.167	0.141	15	21	0.71
8	retail	97	$3,451,760.02	$0.85	$431,470.00	9.84%			0.000	1.000	0.000	0.000	1	28	0.04

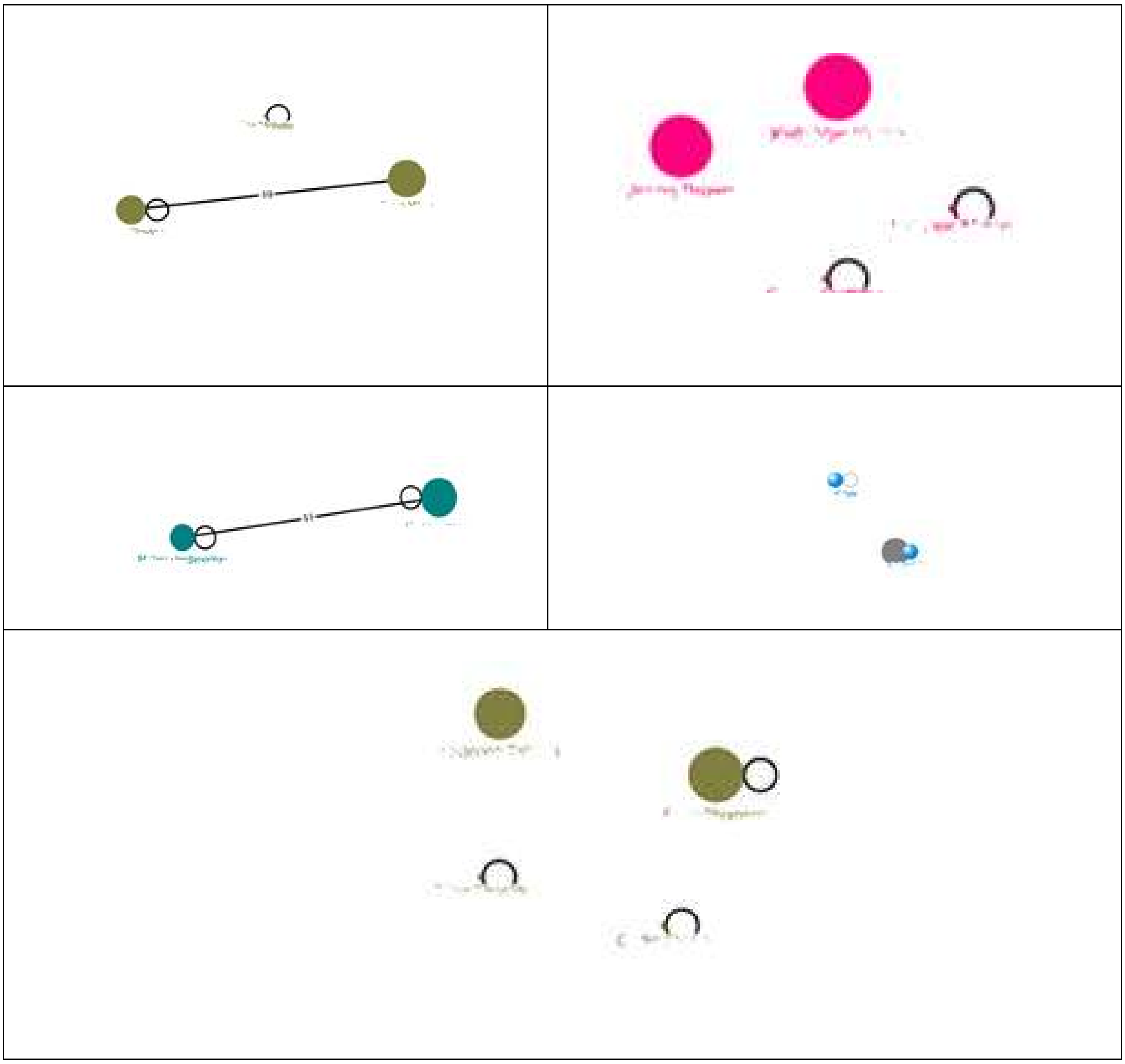
FTE	Office Type	RANK	Gross Revenue(TSV)	Normalized Rev	Rev/FTE	Gross Margin Avg.	Density	Structual Holes	Avg. Geo Distance	Avg. Eigenvector Centr.	Avg. Betweenness Centr.	Avg. Closeness Centr.	Total Strong Edges	Potential edges	Strength of Ties
6	retail	98	$3,445,670.64	$0.85	$574,278.44	9.88%	0.167	6.000	0.333	0.250	0.000	0.500	4	15	0.27
8	retail	99	$3,434,352.54	$0.85	$429,294.07	8.54%	0.357	2.800	0.714	0.125	0.000	0.156	22	28	0.79
6	retail	100	$3,430,703.63	$0.85	$571,783.94	11.47%	0.600	1.667	1.120	0.200	0.800	0.185	10	15	0.67
8	retail	101	$3,412,779.23	$0.84	$426,597.40	10.29%	0.036	28.000	0.250	0.167	0.000	0.333	8	28	0.29
7	retail	102	$3,362,929.56	$0.83	$480,418.51	11.97%	0.500	2.000	0.600	0.250	0.000	0.375	9	21	0.43
9	retail	103	$3,340,776.22	$0.82	$371,197.36	11.06%	0.267	3.750	0.889	0.167	0.333	0.172	10	36	0.28
8	retail	104	$3,335,468.37	$0.82	$416,933.55	10.67%	0.067	15.000	0.250	0.167	0.000	0.333	7	28	0.25
6	retail	105	$3,326,314.43	$0.82	$554,385.74	9.78%			0.000	0.200	0.000	0.000	5	15	0.33
7	retail	106	$3,325,276.04	$0.82	$475,039.43	10.79%	0.067	15.000	0.400	0.333	0.000	0.667	5	21	0.24
8	retail	107	$3,322,047.87	$0.82	$415,255.98	9.15%	0.400	2.500	1.000	0.250	0.500	0.258	6	28	0.21
6	retail	108	$3,268,316.82	$0.81	$544,719.47	11.72%	0.200	5.000	0.545	0.200	0.000	0.300	11	15	0.73
7	retail	109	$3,238,174.93	$0.80	$462,596.42	12.40%	0.300	3.333	0.545	0.200	0.000	0.300	8	21	0.38
12	retail	110	$3,235,032.29	$0.80	$269,586.02	12.12%	0.095	10.500	0.615	0.143	0.143	0.167	8	66	0.12
8	retail	111	$3,226,643.62	$0.79	$403,330.45	9.45%	0.500	2.000	0.600	0.250	0.000	0.375	8	28	0.29
9	retail	112	$3,169,835.09	$0.78	$352,203.90	9.78%	0.200	5.000	0.800	0.250	0.250	0.292	4	36	0.11
8	retail	113	$3,169,402.97	$0.78	$396,175.37	12.24%	0.400	2.500	0.941	0.200	0.400	0.207	11	28	0.39
11	retail	114	$3,115,509.94	$0.77	$283,228.18	11.10%	0.048	21.000	0.250	0.167	0.000	0.333	7	55	0.13
10	retail	115	$3,108,947.30	$0.77	$310,894.73	14.33%			0.000	0.333	0.000	0.000	3	45	0.07
8	retail	116	$3,106,943.65	$0.77	$388,367.96	8.07%			0.000	0.333	0.000	0.000	3	28	0.11
4	retail	117	$3,072,774.57	$0.76	$768,193.64	11.12%	1.000	1.000	0.667	0.333	0.333	0.500	9	6	1.50
6	retail	118	$3,068,160.55	$0.76	$511,360.09	8.75%	0.100	10.000	0.400	0.333	0.000	0.667	4	15	0.27
9	retail	119	$3,049,575.31	$0.75	$338,841.70	12.37%	0.400	2.500	0.941	0.200	0.400	0.207	10	36	0.28
7	retail	120	$3,031,983.95	$0.75	$433,140.56	14.83%	0.143	7.000	0.545	0.200	0.000	0.300	7	21	0.33
10	retail	121	$3,001,678.77	$0.74	$300,167.88	6.32%	0.133	7.500	0.727	0.200	0.200	0.233	6	45	0.13
5	retail	122	$2,957,159.42	$0.73	$591,431.88	11.59%	0.500	2.000	0.824	0.200	0.200	0.233	11	10	1.10
8	retail	123	$2,956,009.18	$0.73	$369,501.15	13.85%	0.600	1.667	0.706	0.200	0.000	0.267	12	28	0.43
9	retail	124	$2,928,486.17	$0.72	$325,387.35	12.18%	0.048	21.000	0.286	0.200	0.000	0.400	5	36	0.14
6	retail	125	$2,908,066.39	$0.72	$484,677.73	8.84%	0.300	3.333	0.545	0.200	0.000	0.300	7	15	0.47
7	retail	126	$2,890,894.57	$0.71	$412,984.94	12.59%	0.333	3.000	0.800	0.250	0.250	0.292	7	21	0.33
7	retail	127	$2,868,815.54	$0.71	$409,830.79	11.88%	0.167	6.000	0.400	0.333	0.000	0.667	4	21	0.19
9	retail	128	$2,851,165.93	$0.70	$316,796.21	7.77%			0.000	0.250	0.000	0.000	4	36	0.11
8	retail	129	$2,831,613.68	$0.70	$353,951.71	9.84%			0.000	0.500	0.000	0.000	2	28	0.07
9	retail	130	$2,826,917.91	$0.70	$314,101.99	12.24%	0.300	3.333	0.600	0.250	0.000	0.375	6	36	0.17
7	retail	131	$2,810,930.69	$0.69	$401,561.53	12.14%	0.167	6.000	0.333	0.250	0.000	0.500	4	21	0.19
10	retail	132	$2,796,635.39	$0.69	$279,663.54	8.82%	0.200	5.000	1.059	0.200	0.600	0.187	8	45	0.18
7	retail	133	$2,752,802.87	$0.68	$393,257.55	10.63%	0.167	6.000	0.400	0.333	0.000	0.667	3	21	0.14
9	retail	134	$2,752,681.97	$0.68	$305,853.55	11.00%	0.200	5.000	0.545	0.200	0.000	0.300	9	36	0.25
7	retail	135	$2,729,624.36	$0.67	$389,946.34	10.06%	0.400	2.500	1.077	0.167	0.667	0.154	10	21	0.48
12	retail	136	$2,720,001.28	$0.67	$226,666.77	10.14%	0.467	2.143	1.000	0.167	0.500	0.165	15	66	0.23
11	retail	137	$2,711,385.84	$0.67	$246,489.62	9.22%	0.100	10.000	0.333	0.250	0.000	0.500	4	55	0.07
8	retail	138	$2,643,214.39	$0.65	$330,401.80	9.98%	0.200	5.000	0.727	0.200	0.200	0.233	7	28	0.25
7	retail	139	$2,633,239.69	$0.65	$376,177.10	9.13%	0.200	5.000	0.727	0.200	0.200	0.233	7	21	0.33
6	retail	140	$2,599,886.33	$0.64	$433,314.39	9.78%	0.067	15.000	0.250	0.167	0.000	0.333	6	15	0.40
8	retail	141	$2,599,308.62	$0.64	$324,913.58	12.27%	0.200	5.000	0.727	0.200	0.200	0.233	7	28	0.25
9	retail	142	$2,563,454.36	$0.63	$284,828.26	11.66%			0.000	0.250	0.000	0.000	4	36	0.11
7	retail	143	$2,558,136.05	$0.63	$365,448.01	10.14%	0.300	3.333	0.545	0.200	0.000	0.300	6	21	0.29
7	retail	144	$2,498,181.12	$0.62	$356,883.02	9.01%	0.524	1.909	1.027	0.143	0.571	0.140	22	21	1.05
10	retail	145	$2,496,618.66	$0.62	$249,661.87	9.68%			0.000	0.250	0.000	0.000	4	45	0.09
9	retail	146	$2,483,442.23	$0.61	$275,938.03	10.31%	0.143	7.000	0.500	0.167	0.000	0.250	7	36	0.19
9	retail	147	$2,475,076.09	$0.61	$275,008.45	10.98%	0.167	6.000	0.333	0.250	0.000	0.500	5	36	0.14

FTE	Office Type	RANK	Gross Revenue(TSV)	Normalized Rev	Rev/FTE	Gross Margin Avg.	Density	Structual Holes	Avg. Geo Distance	Avg. Eigenvector Centr.	Avg. Betweenness Centr.	Avg. Closeness Centr.	Total Strong Edges	Potential edges	Strength of Ties
10	retail	148	$2,432,081.05	$0.60	$243,208.11	11.80%	0.133	7.500	0.400	0.167	0.000	0.667	6	45	0.13
8	retail	149	$2,404,091.40	$0.59	$300,511.43	10.72%	0.500	2.000	1.250	0.250	1.000	0.208	4	28	0.14
10	retail	150	$2,385,149.27	$0.59	$238,514.93	10.56%	0.900	1.111	0.880	0.200	0.200	0.230	16	45	0.36
4	retail	151	$2,377,127.22	$0.59	$594,281.81	10.70%			0.000	0.500	0.000	0.000	2	6	0.33
5	retail	152	$2,363,445.71	$0.58	$472,689.14	14.76%			0.000	0.500	0.000	0.000	2	10	0.20
7	retail	153	$2,354,026.74	$0.58	$336,289.53	10.51%	0.167	6.000	0.400	0.333	0.000	0.667	3	21	0.14
7	retail	154	$2,309,726.63	$0.57	$329,960.95	8.90%	0.300	3.333	0.545	0.200	0.000	0.300	10	21	0.48
8	retail	155	$2,303,415.36	$0.57	$287,926.92	10.01%	0.167	6.000	0.400	0.333	0.000	0.667	2	28	0.07
9	retail	156	$2,285,001.17	$0.56	$253,889.02	10.34%			0.000	0.250	0.000	0.000	5	36	0.14
6	retail	157	$2,269,490.57	$0.56	$378,248.43	10.13%	0.167	6.000	0.400	0.333	0.000	0.667	5	15	0.33
8	retail	158	$2,258,015.23	$0.56	$282,251.90	12.23%	0.133	7.500	0.727	0.200	0.200	0.233	6	28	0.21
5	retail	159	$2,254,737.51	$0.56	$450,947.50	14.14%	0.400	2.500	0.941	0.200	0.400	0.207	8	10	0.80
11	retail	160	$2,149,153.41	$0.53	$195,377.58	10.74%	0.100	10.000	0.286	0.200	0.000	0.400	7	55	0.13
11	retail	161	$2,112,552.40	$0.52	$192,050.22	9.83%	0.095	10.500	0.727	0.200	0.200	0.233	5	55	0.09
11	retail	162	$2,040,225.39	$0.50	$185,475.04	11.60%			0.000	0.250	0.000	0.000	4	55	0.07
6	retail	163	$2,020,877.77	$0.50	$336,812.96	11.24%	0.500	2.000	0.600	0.250	0.000	0.375	10	15	0.67
4	retail	164	$1,972,646.53	$0.49	$493,161.63	11.15%	0.333	3.000	0.400	0.333	0.000	0.667	3	6	0.50
6	retail	165	$1,971,070.79	$0.49	$328,511.80	10.75%	0.333	3.000	0.889	0.333	0.333	0.389	4	15	0.27
7	retail	166	$1,961,712.40	$0.48	$280,244.63	9.57%	0.167	6.000	0.400	0.333	0.000	0.667	4	21	0.19
7	retail	167	$1,905,531.75	$0.47	$272,218.82	10.23%	0.300	3.333	0.545	0.200	0.000	0.300	9	21	0.43
5	retail	168	$1,901,356.03	$0.47	$380,271.21	9.13%	0.300	3.333	0.600	0.250	0.000	0.375	5	10	0.50
9	retail	169	$1,888,431.48	$0.47	$209,825.72	11.41%	0.333	3.000	0.800	0.250	0.250	0.292	8	36	0.22
6	retail	170	$1,874,465.41	$0.46	$312,410.90	12.64%	0.067	15.000	0.250	0.167	0.000	0.333	6	15	0.40
14	retail	171	$1,870,401.01	$0.46	$133,600.07	10.99%								91	0.00
8	retail	172	$1,858,426.07	$0.46	$232,303.26	10.04%			0.000	1.000	0.000	0.000	1	28	0.04
6	retail	173	$1,819,481.08	$0.45	$303,246.85	9.43%	0.300	3.333	1.059	0.200	0.600	0.187	8	15	0.53
3	retail	174	$1,801,652.72	$0.44	$600,550.91	13.80%	1.000	1.000	0.667	0.333	0.000	0.500	8	3	2.67
5	retail	175	$1,797,551.77	$0.44	$359,510.35	8.66%	0.333	3.000	0.500	0.500	0.000	1.000	1	10	0.10
6	retail	176	$1,730,295.71	$0.43	$288,382.62	9.41%	0.167	6.000	0.333	0.250	0.000	0.500	5	15	0.33
6	retail	177	$1,711,937.22	$0.42	$285,322.87	9.82%			0.000	0.500	0.000	0.000	2	15	0.13
5	retail	178	$1,279,867.83	$0.32	$255,973.57	9.13%	1.000	1.000	0.500	0.500	0.000	1.000	3	10	0.30
6	retail	179	$1,278,706.09	$0.31	$213,117.68	13.38%			0.000	0.500	0.000	0.000	2	15	0.13
2	retail	180	$1,176,990.52	$0.29	$588,495.26	9.85%								1	0.00
9	retail	181	$1,175,116.31	$0.29	$130,568.48	9.20%			0.000	0.333	0.000	0.000	3	36	0.08
6	retail	182	$1,121,558.68	$0.28	$186,926.45	12.09%	0.333	3.000	0.400	0.333	0.000	0.667	5	15	0.33
4	corp	183	$709,960.07	$0.17	$177,490.02	11.63%			0.000	0.500	0.000	0.000	2	6	0.33
8	retail	184	$402,632.02	$0.10	$50,329.00	10.92%			0.000	0.333	0.000	0.000	3	28	0.11
2	retail	185	$283,486.72	$0.07	$141,743.36	12.99%	0.333	3.000	0.400	0.333	0.000	0.667	4	1	4.00
5	retail	186	$172,104.43	$0.04	$34,420.89	9.71%	0.067	15.000	0.400	0.333	0.000	0.667	3	10	0.30
3	retail	187	$5,378.87	$0.00	$1,792.96	9.71%			0.000	0.333	0.000	0.000	3	3	1.00

Appendix B: Subgraphs of High Performing Offices



Appendix C: Subgraphs of Low Performing Offices



References

Achrol, R.S. and Kotler, P. (2006). The service-dominant logic for marketing: a critique. In R. Lusch and S. Vargo, *The Service-dominant logic of marketing: dialog, debate, and directions*. Armonk, NY: M.E. Sharpe, pp. 320–33.

Ackerman, M.S., Pipek, V. and Wulf, V. (2003). *Sharing expertise: beyond knowledge management*. MIT Press.

Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: a longitudinal study. *Administrative Science Quarterly*, 45, 425-455.

Ahuja, G. (2000). The duality of collaboration: Inducements and opportunities in the formation of interfirm linkages. *Strategic Management Journal,* 21, 317-343.

Alba, R. (1972). SOCK. *Behavioural Science,* 17, 326-327.

Alexander, B. and Dijst, M. (2012). Professional workers @ work: importance of work activities for electronic and face-to-face communications in the Netherlands. *Transportation*, 39 (5), 919-940.

Ancona, D.G. and Caldwell, D.F. (1992). Bridging the boundary: external process and performance in organizational teams. *Administrative Science Quarterly,* 37(4), 634-665.

Argote, L. and Ingram, P. (2000). Knowledge transfer: a basis for competitive advantage in firms. *Organizational Behaviour and Human Decision Processes*, 82(1), 150-169.

Argote, L., Mcevily, B. and Reagans, R. (2003). Managing knowledge in organizations: an integrative framework and review of emerging themes. *Management Science*, 49(4), 571-582.

Arnheim, R. (1970). *Visual thinking*. London: Faber.

ATIS Committee PRQC. 'Network topology'. ATIS Telecom Glossary 2007. *Alliance for Telecommunications Industry Solutions*.

Balkundi, P. and Harrison, D. (2006). Ties, leaders, and time in teams: Strong inference about network structure's effects on team viability and performance. *Academy of Management Journal*, 49(1), 49–68.

Barney, J. (1986). Strategic factor markets: expectations, luck, and business strategy. *Management Science*, 32(10), 1231-1241.

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99-120.

Barrat, A., Barthelemy, M. and Pastor-Satorras, R. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11), 3747–3752.

Batjargal, B. (2003). Social capital and entrepreneurial performance in Russia: a longitudinal study. *Organization Studies*, 24, 535-556.

Bauer, T.N. and Erdogan, B. (2011). Organizational socialization: the effective onboarding of new employees. In S. Zedeck, ed. *APA handbook of industrial and organizational psychology, Vol 3: Maintaining, expanding, and contracting the organization*, *APA handbooks in psychology*. Washington, DC: American Psychological Association, pp. 51–64.

Beliën, S. and Leenders, R. (1996). Comparison of visualization techniques and packages (version 2.0). Available at: SARA Visualization Center, http://www.sara.nl/Rik/REPORT.update/Report.html (Accessed 17 January 1996).

Bellieni, C.V., Cordelli, D.M., Marchi, S., et al. (2007). Sensorial saturation for neonatal analgesia. *The Clinical Journal of Pain,* 23(3), 219–221.

Benton, T. and Craib, I. (2001). *Philosophy of social science*. New York: Palgrave.

Bhaskar, R. (1978). *A realistic theory of science*. Sussex: Harvester Press.

Bienenstock, E.J. and Bonacich, P. (2003). Balancing efficiency and vulnerability in social networks. In R. Brieger, K. Carley and P. Pattison, eds. *Dynamic social network modeling and analysis: workshop summary and papers*. Washington, DC: National Academies Press.

Bonacich, P. (1987). Power and centrality: a family of measures. *American Journal of Sociolog*y 92(5), 1170–1182.

Bonsignore, E. M., Dunne, C., Rotman, D., Smith, M., Capone, T., Hansen, D. L. and B. Shneiderman. (2009). First steps to NetViz Nirvana: evaluating social network analysis with NodeXL., SIN '09. Proceedings of the Fourth International Conference on Communities and Technologies. Vancouver: IEEE Computer Society Press.

Borgatti, S. (2005). Centrality and network flow. *Social Networks*, *27*(1), 55-71.

Borgatti, S.P. and Foster, P.C. (2003). The network paradigm in organizational research. *Journal of Management,* 29(6), 991-1013.

Bou-Llusar, J.C. and Segarra-Ciprés, M. (2006). Strategic knowledge transfer and its implications for competitive advantage: an integrative conceptual framework. *Journal of Knowledge Management*, 10(4), 100-112.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel, ed. *Soziale Ungleichheiten (Soziale Welt, Sonderheft 2*), Göttingen: Otto Schwartz, p. 249.

Bouttier, J.; Di Francesco, P. and Guitter, E. (2003). Geodesic distance in planar graphs. *Nuclear Physics* B 663(3), 535–567.

Boyd, D. (2007). None of this is real: networked participation in Friendster. In J. Karaganis, editor, *Structures of Participation in Digital Culture*. New York : Social Science Research Council, Chapter 8.

Bradach, J.L., and Eccles, R.G. (1989). Price authority and trust: from ideal types to plural forms. *American Review of Sociology*, 15, 97–118.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, *25*, 163-177.

Bradt, G.B., and Vonnegut, M. (2009). Onboarding: How to get your new employees up to speed in half the time. Hoboken, NJ: Wiley.

Brass, D. (1995). A social network perspective on human resources management. *Research in Personnel and Human Resource Management*, 13, 39–79.

Breiger, R. L. (2004). The analysis of social networks. In M.A. Hardy and A. Bryman, eds. *Handbook of Data Analysis*, London, U.K.: Sage Publications, pp. 505–526.

Brown, S.L. and Eisenhardt, K.M. (1995). Product development: past research, present findings, and future directions. *The Academy of Management Review*, 20(2), 343-378.

Bryman, A. and Bell, E. (2007). *Business research methods*. New York: Oxford University Press.

Buckley, P. and Casson, M. (1976). *The future of multinational enterprise*. London: Macmillan.

Bulkley, N. and Van Alstyne, M. (2008). Why information should influence productivity. In Manuel Castells, ed. *The network society: a global perspective*. Cheltenham, U.K. and Northampton, MA: Edward Elgar Publishers, pp. 145-173.

Burakova-Lorgnier, M., Bouzdine-Chameeva, T. and MacGilchrist, R. (2004). Network structure as a predictor of tacit knowledge transfer capacity. Proceedings for the 2004 UKAIS conference, 10-12 May, Glasgow, UK.

Burgelman, R.A. (1991). Intraorganizational ecology of strategy making and organizational adaptation: theory and field research. *Organization Science*, 2(3), 239-263.

Burgin, G.H. (1969) On playing two-person zero-sum games against nonminimax players. *IEEE Transactions Systems Science and Cybernetics*, 5(4), 369-370.

Burt, R.S. (1992). *Structural holes: the social structure of competition*. Cambridge, MA: Harvard University Press.

Butts, C.T. (2008), Social network analysis with SNA, *Journal of Statistical Software*, 24(6).

Carroll, A.B. (1999). Corporate social responsibility: evolution of a definitional construct. *Business and Society*, 38(3), 268–295.

Castells, M. (2011). A network theory of power. *International Journal of Communication*. In M. Castells, ed. *The rise of network society, the information age: economy, society and culture*. Cambridge, MA, and Oxford: Blackwell.

Chandler, A. (1962). *Strategy and structure: chapters in the history of the industrial enterprise*. Cambridge, MA: MIT Press.

Chen, Z. and Klahr, D. (1999). All other things being equal: acquisition of the control of variables strategy. *Child Development*, 70, 1098–1120.

Choi, H.S., Lee, H. and Yoo, Y. (2011). The impact of information technology and transactive memory systems on knowledge sharing, application, and team performance: a field study. *MIS Quarterly*, 34 (4), 855–870.

Churchill, J.G. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing*, 16, 64-73.

CIA.0. (2008). The world factbook: GDP - COMPOSITION BY SECTOR. Available at: https://www.cia.gov/library/publications/the-world-factbook/fields/2012.html (Accessed 8 November 2012).

Cillo, P. (2005). Fostering market knowledge use in innovation: the role of internal brokers. *European Management Journal*, 23(4), 404-412.

Clark, L. (2011). Seeing the social capital in agricultural innovation systems: using SNA to visualise bonding and bridging ties in rural communities. *Knowledge Management for Development Journal*, 6(3), 206-218.

Coff, R.W., Coff, D.C. and Eastvold, R. (2006). The knowledge-leveraging paradox: how to achieve scale without making knowledge imitable. *Academy of Management Review*, 31(2), 452-465.

Cohen, W.M. and Levinthal, D.A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly,* 35, pp. 118-152.

Coleman, J. (1990). *Foundations of social theory.* Cambridge, MA: Belknap Harvard.

Connidis, I. (1983). Integrating qualitative and quantitative methods in survey research on aging: an assessment. *Qualitative Sociology*, 6(4), 334-352.

Cross, R., Borgatti, S.P. and Parker, A. (2001). Beyond answers: dimensions of the advice network. *Social Networks*, 23(3), 215–235.

Cross, R., Borgatti, S.P. and Parker, A. (2002). Making invisible work visible: using social network analysis to support strategic collaboration. *California Management Review*, 44(2), 25-47.

Cross, R., Prusak, L., Parker, A., and Borgatti, S.P. (2001). Supporting knowledge creation and sharing in social networks. *Organizational Dynamics*, 30(2), 100-120.

Crotty, M. (1998). *The foundations of social research*. Crows Nest, NSW: Allen & Unwin.

D'Agostino, R.B. and Stephend, M.A. (1986). *Goodness-of-fit techniques.* New York: Marcel Dekker.

Damanpour, F. (1992). Organizational size and innovation. *Organization Science*, 13(3), 375-402.

Darr, E.D., Argote, L. and Epple, D. (1995). The acquisition, transfer, and depreciation of knowledge in service organizations: productivity in franchises. *Management Science*, 41(11), 1750-1762.

Davis, A., Gardner, B.B. and Gardner, M.R. (1941). *Deep south*. Chicago: The University of Chicago Press.

Delanty, G. and Strydom, P. (2003*). Philosophies of social science: the classic and contemporary readings*. Maidenhead, UK: Open University Press.

Denzin, N.K. (1978). *The research act: a theoretical introduction to sociological methods*. New York: McGraw-Hill.

de Pablos, P.O. (2006). Transnational corporations and strategic challenges. *The Learning Organization*, 13(6), 554-559. doi: 10.1108/09696470610705433.

Dhanaraj, C., Lyles, M., Steensma, H.K. and Tihanyi, L. (2004). Managing tacit and explicit knowledge transfer in IJVs: the role of relational embeddedness and the impact on performance*. Journal of International Business Studies*, 35, 428–443.

Di Battista, G., Eades, P., Tamassia, R. and Tollis, I.G. (1999). *Graph drawing: algorithms for the visualization of graphs*. Upper Saddle River, NJ: Prentice Hall.

Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1, 269–271.

Dinur, A. (2011). Tacit knowledge taxonomy and transfer: case-based research. *Journal of Behavioral and Applied Management*, 12(3), 246-281.

Donaldson, L. (2001). The contingency theory of organizations. Thousand Oaks, CA: Sage Publications.

Drucker, P. (1988). The coming of the new organization. *Harvard Business Review*, 66(1), 45-53.

Dyer, J.H., and Nobeoka, K. (2000). Creating and managing a high-performance knowledge-sharing network: the Toyota case. *Strategic Management Journal*, 21(3), 345-367.

Eccles, R.G. 1981. The quasifirm in the construction industry. *Journal of Economic Behaviours and Organization*, 2, 335–357.

Edelman, L.F., Bresnan, M., Newell, S., Scarborough, H. and Swan, J. (2004). The benefits and pitfalls of social capital: empirical evidence from two organizations in the United Kingdom. *British Journal of Management* 15(S1), 59–69.

Eisenhardt, K. and Santos, F.M. (2002). Knowledge-based view: a new theory of strategy. In A. Pettigrew, H. Thomas and R. Whittington, eds. *Handbook of strategy and management*. London: Sage.

Ellison, N., Steinfeld, C. and Lampe, C. (2006). Spatially bounded online social networks and social capital: the role of Facebook. In Proceedings of the Annual Conference of the International Communication Association, June 19-23, 2006, Dresden, Germany.

Farrall, K. (2004). Web graph analysis in perspective, v 1.0. Available at: http://farrall.org/papers/webgraph_as_content.html (Accessed 14 June 2011).

Fendt, J. and Kaminska-Labbe, R. (2011). Relevance and creativity through design-driven action research: introducing pragmatic adequacy. *European Management Journal*, 29(3), 217-233.

Fligstein, N. and Freeland, R. (1995). Theoretical and comparative perspectives on corporate organization. *Annual Review of Sociology*, 21, 21-43.

Florin, J., Lubatkin, M. and Schulze, W. (2003). A social capital model of high growth ventures. *Academy of Management Journal*, 46(3), 374-384.

Freeman, L.C. (2000). Visualizing social networks. *Social Structure*, 1(1).

Frey, R.S. (2001). Knowledge management, proposal development, and small businesses. *Journal of Management Development*, 20(1), 38-54.

Fritsch, M. and Kauffeld-Monz, M. (2009). The impact of network structure on knowledge transfer: an application of social network analysis in the context of regional innovation networks. *Annals of Regional Science*, 44(1), 21-38.

Fuller, S. (2006). In search of sociological foundations for the project of humanity: the new sociological imagination. *History of the Human Sciences*, 22(2), April 2009, 138-145.

Gabbay, M.J., and Pitts, A.M. (2002). A new approach to abstract syntax with variable binding. *Formal Aspects of Computing*, 13, 341–363.

Gargiulo, M. (1993). Two-step leverage: managing constraint in organizational politics. *Administrative Science Quarterly*, 39, 1-19.

Gauss, C.F. (1963). *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientum*. Translated by Davis, C.H. New York: Dover (Original work published 1809).

Gergen, Kenneth J. (2003). *An invitation to social constructionism.* Thousand Oaks: Sage.

Golder, S.A., Wilkinson, D. and Huberman, B. (2006). Rhythms of social interaction: messaging within a massive online network. 3rd International Conference on Communities and Technologies, East Lansing, MI. June 28-30, 2007.

Granovetter, M.S. (1973). The strength of weak ties. *The American Journal of Sociology*, 78(6), 1360-1380.

Grant, R.M. (1995). Prospering in dynamically competitive environments: organizational capability as knowledge integration. *Organization Science*, 7(4), 375-387.

Grant, R.M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17, 109-122.

Grant, R.M. (1997). The knowledge-based view of the firm: Implications for management practice. *Long Range Planning*, 30(3), 450-454. doi: 10.1016/S0024-6301(97)00025-3.

Gray, P.H. and Meister, D.B. (2004). Knowledge sourcing effectiveness. *Management Science*, 50(6), 821-834.

Green, A.R. and Murray, T.K. (1989). A simple intravenous infusion method in rodents for determining the potency of anticonvulsants acting through GABAergic mechanisms. *Journal of Pharmacy and Pharmacology*, 41, 879-880.

Greiner, W. (2001). *Quantum mechanics: an introduction*. New York: Springer.

Grippa, F., Zilli, A., Laubacher, R. and Gloor, P.A. (2006). E-mail may not reflect the social network. In Proceedings of the North American Association for Computational Social and Organizational Science Conference, pp. 1-6.

Grönroos, C. (2006). What can a service logic offer marketing theory? In R.F. Lusch and S.L. Vargo, eds. *The service-dominant logic of marketing: dialog, debate, and directions*. Armonk, NY: M.E. Sharpe, pp. 354–64.

Grossman, S., Hart, O., 1980. The costs and benefits of ownership: a theory of vertical and lateral integration. *Bell Journal of Economics*, 11, 42-64.

Groth, D. and Skandier, T. (2005). Network+ study guide, 4th edn. Sybex, Inc.

Guba, E. (1990). *The paradigm dialog*. London: Sage.

Guba, E.G. and Lincoln, Y.S. (1994). Paradigmatic controversies, contradictions, and emerging confluences. In N.K. Denzin and Y.S. Lincoln, eds. *Handbook of qualitative research*. Thousand Oaks, CA: Sage, pp. 105-117.

Gulati, R., Nohria, N. and Zaheer, A. (2000). Strategic networks. *Strategic Management Journal*, 215, 203-215.

Gulati, S. (2006). Application of New Technologies. In S. Glen and P. Moule, eds. *e-Learning in Nursing*. Basingstoke: Palgrave, pp. 20-37.

Gummesson, E. (2006). Many-to-many marketing as grand theory: a Nordic school contribution. In R.F. Lusch and S.L. Vargo, eds. *The service–dominant logic of marketing: dialog, debate, and directions*. Armonk, NY: M.E. Sharpe, pp. 339–53.

Gupta, A.K. and Govindarajan, V. (2000). Knowledge flows within multinational corporations. *Strategic Management Journal*, 21(4), 473-496.

Guzzo, R.A. and Shea, G.P. (1992). Group performance and intergroup relations in organizations. In M.D. Dunnette and L.M. Hough, eds. *Handbook of industrial and organizational psychology*. Palo Alto, CA: Consulting Psychologists Press, pp. 269-313.

Hackman, J. and Oldham, G. (1980). *Work redesign*. Reading, MA: Addison-Wesley.

Hanneman, R.A. (2001). *Introduction to social network methods*. Available at: http://faculty.ucr.edu/~hanneman/SOC157/TEXT/TextIndex.html (Accessed 31 August 2012).

Hanneman, R. A. and Riddle, M. (2005). Introduction to social network methods. Riverside: University of California.

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

Hansen, M.T. (2002). Knowledge networks. *Organization Science*, *6*(1), 76-248.

Hansen, M.T., Podolny, J.M. and Pfeffer, J. (2001). So many ties, so little time: a task contingency perspective on corporate social capital. In S.M. Gabbay and R.Th.A.J. Leenders, eds. *Social Capital of Organizations*, *Volume 18*. Bingley: Emerald Group, pp. 21-57.

Harlow, H. (2008). The effect of tacit knowledge on firm performance. *Journal of Knowledge Management*, 12(1), 148-163.

Haythornthwaite, C. (1996). Social network analysis: an approach and technique for the study of information exchange. *Library and information science research*, 18(4), 323-342.

Hazleton, V. and Kennan, W. (2000). Social capital: reconceptualising the bottom line*, Corporate Communications* 5(2), 81–6.

Heshusius, L. and Ballard, K., eds. (1996). *From positivism to interpretivism and beyond*. New York: Teachers College Press.

Hite, J.M. and Hesterly, W.S. (2001). The evolution of firm networks: from emergence to early growth of the firm. *Strategic Management Journal*, 22(3), 275–286.

Huberman, B.A. and Hogg, T. (1995). Distributed computation as an economic system. *The Journal of Economic Perspectives*, 9(1), 141-152.

Hughes, J. and Sharrock, W. (1997). *The philosophy of social research*. London: Longman.

Hui, P. and Crowcroft, J. (2008). Human mobility models and opportunistic communication system design. *Royal Society Philosophical Transactions*, 366, 1872.

Huisman, M. and Van Duijn, M.A.J. (2005). Software for social network analysis. In P.J. Carrington, J. Scott, and S. Wasserman, eds. *Models and methods in social network analysis*. New York: Cambridge University Press, pp. 270–316.

Hutcheson, G. and Moutinho, L. (2008). *Statistical modelling for management*. London: Sage Publications.

Hutcheson, G. and Sofroniou, N. (1999). *The multivariate social scientist*. London: Sage Publications.

ILO. (2009). Global Employment Trends: January 2009. Geneva. Available at: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---comm/documents/publication/wcms_101461.pdf (Accessed 30 August 2012).

Inkpen, A.C. and Tsang, E.W. (2005). Social capital, networks, and knowledge transfer. *Academy of Management Review*, *30*(1), 146-165.

Ipe, M. (2003). Knowledge sharing in organizations: a conceptual framework. *Human Resource Development Review*, 2(4), 337-359.

Janhonen, M. and Johanson, J. (2011). Role of knowledge conversion and social networks in team performance. *International Journal of Information Management*, 31(3), 217–225.

Jarillo, J.C. 1988. On strategic networks. *Strategic Management Journal*, 9(1), 31–39.

Jensen, R. and Szulanski, G. (2004). Stickiness and the adaptation of organizational practices in cross-border knowledge transfers. *Journal of International Business Studies*, *35*(6), 508-523.

Jick, T.D. (1979). Mixing qualitative and quantitative methods: triangulation in action. *Administrative Science Quarterly*, 24, 602-611.

Kamhawi, E. M. (2010). The three tiers architecture of knowledge flow and management activities. *Information and Organization,* 20(3-4), 169-186.

Kenis, P. and Knoke, D. (2002) How organizational field networks shape interorganizational tie-formation rates. T*he Academy of Management Review*, 27(2), 275-293.

Klovdahl, A.S. (1981). A note on images of networks. *Social Networks*. 3, 197-214.

Koestler, A. (1964). *The act of creation*. New York: Macmillan.

Kogut, B. (2000). The network as knowledge: generative rules and the emergence of structure. *Strategic Management Journal*, 21(3), 405-425.

Krackhardt, D. and Stern, R.N. (1988). Informal networks and organizational crises—An experimental simulation. *Social Psychology Quarterly,* 51(2), 123-140.

Kuhn, T.S. (1996). *The structure of scientific revolutions*. 3rd edn. Chicago and London: University of Chicago Press.

Kusunoki, K., Nonaka, I. and Nagata, A. (1998). Organizational capabilities in product development of Japanese firms: a conceptual framework and empirical findings. *Organization Science*, 9(6), 699-718.

Lane, P.J., Lyles, M.A. and Salk, J.E. (2001). Absorptive capacity, learning, and performance in international joint ventures. *Strategic Management Journal*, *22*(12), 1139-1161.

Latour, B. (1987). *Science in action: how to follow scientists and engineers through society.* Cambridge, MA: Harvard University Press.

Laursen, K. and Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal*, 27(2), 131-150.

Lawson, B. Developments in managing innovation, knowledge and e-business (2003) Knowledge and Information Technology Management. Human and Social Perspectives, Idea Group Publishing, Hershey, P 1-13.

Lawson, B., Petersen, K.J., Cousins, P.B. and Handfield, R.B. (2009). Knowledge sharing in interorganizational product development teams: the effect of formal and informal socialization mechanisms. *Journal of Product Innovation Management*, 26, 156–172.

Leana, C.R. and Van Buren, H.J., III. (1999). Organizational social capital and employment practices. *Academy of Management Review*, 24, 538-555.

Lechner, C., Frankenberger, K. and Floyd, S.W. (2010). Task contingencies in the curvilinear relationship between inter-group networks and performance. *The Academy of Management Journal,* (53)4, 865-889.

Lee, C., Lee, K. and Pennings, J.M. (2001). Internal capabilities, external networks and performance: a study on technology-based ventures. *Strategic Management Journal*, 22, 615–640.

Lenoir, T., ed. (1998). *Inscribing science: scientific texts and the materiality of communication*. Stanford: Stanford University Press.

Levin, D.Z. and Cross, R. (2004). The strength of weak ties you can trust: the mediating role of trust in effective knowledge transfer. *Management Science*, 50(11), 1477-1490.

Li, Z. and Zhu, T. (2009). Study on the influence of social capital to informal knowledge transfer among individuals. Proceedings of the 2009 International Symposium on WISA, pp. 355-358.

Liebowitz, J. (2004). Conceptualizing and implementing knowledge management. In P.E. Love, P.S. Fong, and Z. Iraní, eds. *Management of knowledge in project environments*. Oxford: Elsevier, pp. 1-19.

Lin, C., Wu, J.C. and Yen, D.C. (2011). Exploring barriers to knowledge flow at different knowledge management maturity stages. *Information & Management,* 49(1), 10-23.

Lin, M-J.J., Hung, S-W. and Chen, C-J. (2009). Fostering The Determinants of Knowledge Sharing in Professional Virtual Communities. *Computers in Human Behavior*, 25(4), 929-939.

Liu, J. and Belkin, N. J. (2011). Knowledge examination in multi-session tasks. Proceedings of the 5th Workshop on Human-Computer Interaction and Information Retrieval (HCIR 2011).

Losada, M. and Heaphy, E. (2004). The role of positivity and connectivity in the performance of business teams: a nonlinear dynamics model. *American Behavioral Scientist*, 47(6), 740-765.

Louch, H. (2000). Personal network integration: transitivity and homophily in strong-tie relations. *Social Networks*, 22, 45–64.

Lusch, R.F. (2006). Service-dominant logic: reactions, reflections and refinements. *Marketing Theory*, 6(3), 281-288.

Lusch, R.F., Vargo, S.L. and Tanniru, M. (2009). Service, value networks and learning. *Journal of the Academy of Marketing Science*. 38(1),19-31.

Lyles, M.A. and Salk, J.E. (1996). Knowledge acquisition from foreign parents in international joint ventures: an empirical examination in the Hungarian context. *Journal of International Business Studies*, 38(1), 3-18.

Lyles, M.A., van Wijk, R., and Jansen, J.J. (2008). Inter- and intra-organizational knowledge transfer: a meta-analytic review and assessment of its antecedents and consequences. *Journal of Management Studies*, 45(4), 830-853.

Makino, S. and Delios, A. (1996). Local knowledge transfer and performance: implications for alliance formation in Asia. *Journal of International Business Studies*, 27, 905–27.

Malone, T.W., Laubacher, R. and Dellarocas, C. (2009). Harnessing Crowds: mapping the genome of collective Intelligence. Working paper No. 2009-001. Cambridge, MA: MIT Center for Collective Intelligence, Massachusetts Institute of Technology.

March, J. C., and Cyert, R.M. (1963). *A behavioral theory of the firm*. Englewood Cliffs, N.J: Prentice-Hall.

Marin, A. and Wellman, B. (2011). Social network analysis: an introduction. In J. Scott and P. J. Carrington, eds. *The SAGE handbook of social network analysis*. London and Thousand Oaks, CA: Sage Publication, pp. 11-25.

Maurer, I. Bartsch, V. and Ebers, M. (2011). The value of intra-organizational social capital: how it fosters knowledge transfer, innovation performance, and growth. *Organization Studies*, 32(2), 157-185.

McAllister, D.J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38(1), 24-59.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. London: Chapman and Hall.

McLaughlin, S., Paton, R.A. and Macbeth, D.K. (2008). Barrier impact on organizational learning within complex organizations. *Journal of Knowledge Management*, 12(2), 107-123.

McPherson, J.M., Smith-Lovin, L. and Cook, J.M. 2001. Birds of a feather: homophily in social networks. *Annual Review of Sociology,* 27, 415–444.

Mead, S. (2001). Using social network analysis to visualize project teams. *Project Management Journal*, 32(4), 32-38.

Milgrom, P. and Roberts, J. (1990). The economics of modem manufacturing: technology, strategy, and organization. *American Economic Review,* 80(3), 511-528.

Mingers, J. and Willcocks, L. (2004). *Social theory and philosophy for information systems.* Chichester: John Wiley & Sons.

Mintzberg, H. (1980). Structure in 5's: a synthesis of the research on organization design. *Management Science*, 26(3), 322-341.

Moorthy, S. and Polley, D.E. (2010). Technological knowledge breadth and depth: performance impacts. *Journal of Knowledge Management*. 14(3),359-377.

Moran, P. and Ghoshal, S. (1996). Theories of economic organization: the case for realism and balance. *The Academy of Management Review*, 21(1), 58.

Morgan. G. (1997). *Images of* organization, 2nd edn. Newbury Park, CA: Sage Publications.

Moutinho, L. and Hutcheson, G. (2011). *The SAGE dictionary of quantitative management research*, London: Sage Publications.

Mowery, D. C., Oxley, J. E. and Silverman, B. S. (1996). Strategic alliances and interfirm knowledge transfer. *Strategic Management Journal*, 17, 77-91.

Mu, J., Peng, G. and Love, E. (2008). Interfirm networks, social capital, and knowledge flow*, Journal of Knowledge Management*, 12(4), 86-100.

Nachmias, D. and Nachmias, C. (1981). *Research methods in the social sciences*. New York: St. Martin's Press.

Nahapiet, J. and Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review*, 23, 242-66.

Nardi, B., Whittaker, S., Isaacs, E., Creech, M., Johnson, J. and Hainsworth, J. (2002). Integrating communication and information through ContactMap. *Comm. of the ACM*, 45, 89-95.

Neergaard, H. (2005). Networking activities in technology-based entrepreneurial teams. *International Small Business Journal* 23(3), 257–78.

Nelder, J. and Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society,* 135, 370-384.

Nelson, R. and Winter, S. (1982). *An evolutionary theory of economic change.* Cambridge, MA: Belknap Press of Harvard University Press.

Neuman, W. (2000). *Social research methods: qualitative and quantitative approaches.* 4th edn. Boston: Allyn and Bacon.

Newman, M.E. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E, Statistical, nonlinear, and soft matter physics*, 64(1), Pt 2, 016132.

Ni, C., Sugimoto, C. and Jiang, J. (2011). Degree, closeness, and betweenness: application of group centrality measurements to explore macro-disciplinary evolution diachronically. Proceedings of ISSI.

Nie, K., Lin, S., Ma, T. and Nakamori, Y. (2010). Connecting informal networks to management of tacit knowledge. *Journal of Systems Science and Systems Engineering*, 19(2), 237-253.

Nissen, M.E. (2005). *Harnessing knowledge dynamics: Principled organizational knowing and learning*. Hershey, PA: Idea Group Publishing.

Nissen, M.E. and Levitt, R.E. (2004). Agent based modeling of knowledge flows: illustration from the domain of information systems design. Hawaii International Conference of Systems Sciences. Waikoloa, HI.

Noethen, D. and Voelpel, S. (2010). A multilevel investigation of predictors of knowledge seeking and sharing in teams. Academy of Management Annual Meeting Proceedings. Montreal.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, *5*(1), 14-37.

Nonaka, I. and von Krogh, G. (2009). Perspective--tacit knowledge and knowledge conversion: controversy and advancement in organizational knowledge creation theory. *Organization Science*, 20(3), 635-652.

Nonaka, I., Toyama, R. and Konno, N. (2000). SECI, BA and leadership: a unified model of dynamic knowledge creation. *Long Range Planning*, 33(1), 5-34.

Nooteboom, B.B., Vanhaverbeke, W., Duysters, G., Gilsing, V., and Van Den Oord, A. (2006). Optimal cognitive distance and absorptive capacity. *Research Policy*, 33.

Novak, T.P., Hoffman, D.L. and Yung, Y.F. (2000). Measuring the customer experience in online environments: a structural modeling approach. *Marketing Science* 19(1), 22-42.

Oh, H., Labianca, G. and Chung, M. (2006). A multi-level model of group social capital. *Academy of Management Review*, 31, 569–582.

Opsahl, T., Agneessens, F. and Skvoretz, J. (2010). Node centrality in weighted networks: generalizing degree and shortest paths. *Social Networks*, 32(3), 245-251

O'Reilly, C.A., Caldwell, D.F. and Barnett, W.P. (1989). Work group demography, social integration and turnover. *Administrative Science Quarterly,* 34, 21-37.

Palich, L. and Gomez-Mejia, L. (1999). A theory of global strategy and firm efficiencies: considering the effects of cultural diversity. *Journal of management*, 25(4), 587-606.

Pearce, J.A. and David, F.R. (1983). A social network approach to organizational design-performance. Academy of Management. *The Academy of Management Review*, 8(3), 436-444.

Pearson, A., Carr, J. C. and Shaw, J. (2008). Clarifying the familiness construct: a social capital perspective. *Entrepreneurship Theory and Practice*, 32, 949–69.

Pelled, L.H., Eisenhardt, K.M. and Xin, K.R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, 44, 1–28.

Peng, M.W. (2009). *Global business*. Mason, OH: South-Western Cengage Learning.

Penrose, E.T. (1995). *The theory of growth of the firm*. 3rd edn. New York: Oxford University Press.

Pepe, A. (2011). The relationship between acquaintanceship and coauthorship in scientific collaboration networks. *Journal of the American Society for Information Science and Technology*, 62(11), 2121–2132. doi: 10.1002/ asi.21629.

Perer, A. and Shneiderman, B. (2006). Balancing systematic and flexible exploration of social networks. *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 693-700.

Pfeffer, J. (1983). Organizational Demography. In L.L. Cummings and B. Staw, eds. *Research in Organizational Behavior, Volume 5*. Greenwich, CT: JAI Press, pp. 299-357.

Podolny, J.M. and Page, K. L. (1998). Network forms of organization. *Annual Review of Sociology*, 24(1), 57-76.

Polanyi, M. (1966). *The tacit dimension*. London: Routledge &Kegan Paul.

Porter, M.E. (1985). *Competitive advantage: creating and sustaining superior performance.* New York: Free Press.

Powell, W.W. (1990). Neither market nor hierarchy: network forms of organization. *Research on Organizational Behavior,* 12, 295–336.

Powell, W.W., Koput, K.W. and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology*. Administrative Science Quarterly*, 41.

Pratt, M.G. (2001). Social identity dynamics in modern organizations: An organizational psychology/organizational behavior perspective. In M. Hogg and D.J. Terry, eds. *Social identity processes in organizational contexts*. Philadelphia, PA: Psychology Press, pp. 13–30.

Priem, R.L. and Butler, J.E. (2001). Is the resource-based 'view' a useful perspective for strategic management research? *Academy of Management Review*, 26(1), 22-40.

Proctor, C. (1953). Informal social systems. In C.P. Loomis, J.O. Moralis, R.A. Clifford and O.E. Leonard, eds. *Turrialba*. Glencoe, IL: Free Press, pp. 73-88.

Proulx, S.R.; Promislow, D.E.L. and Phillips, P.C. (2005). Network thinking in ecology and evolution. *Trends in Ecology and Evolution* 20(6), 345–353.

Punch, K.F. (1998). *Introduction to social research: quantitative and qualitative approaches.* Thousand Oaks, CA: Sage.

Putnam, R.D. (2000). *Bowling alone: the collapse and revival of American community*. New York: Simon & Schuster.

Putnam, R.T. and Borko, H. (1997). Teacher learning: Implications of new views of cognition. In Biddle, B.J., Good, T.L. and Goodson, I.F., eds. *International handbook of teachers and teaching, Volume II*. Dordrecht and Boston: Kluwer, pp. 1223-1296.

Putterman, L. and Krsozner, R. S., eds. (1996). *The economic nature of the firm: a reader*. Cambridge: Cambridge University Press.

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

Reed, R. and Defillippi, R.J. (1990). Causal ambiguity, barriers to imitation, and sustainable competitive advantage. *The Academy of Management Review*, 15(1), 88-102.

Reed, M.S., Fraser, E.D.G. and Dougill, A.J. (2006). An adaptive learning process for developing and applying sustainability indicators with local communities. *Ecological Economics*, 59, 406-418.

Rice, R. and Steinfeld, C. (1994). Experiences with new forms of organizational communication via electronic mail and voice messaging. In J.E. Andriessen and R. Roe, eds. *Telematics and Work*. E. Sussex, UK: Lea, pp. 109-137.

Rodan, S. (2010). Structural holes and managerial performance: Identifying the underlying mechanism. *Social Networks*, 32(3), 168-179.

Rogers, E. (1962). *Diffusion of innovations*. New York: Free Press.

Rogers, E. ([1962] 1983). *The Diffusion of Innovation*. Reprint. New York: Free Press.

Rollag, K., Parise, S. and Cross, R. (2005). Getting new hires up to speed quickly. *MIT Sloan Management Review,* 46(2), 35-35.

Rossignoli, C. (2009). The contribution of transaction cost theory and other network-oriented techniques to digital markets. *Information Systems and E-Business Management,* 7(1), 57-79.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika,* 31, 581-603.

Sahay, Y.P. and Gupta, M. (2011). Role of organization structure in innovation in the bulk-drug industry. *Indian Journal of Industrial Relations*, 46(3), 450-464.

Sakia, R.M. (1992). The Box-Cox transformation technique: A review. *The Statistician*, 41, 169-178.

Sarantakos, S. (1998). *Social research*. South Yarra, VIC: Penguin (Australia).

Saunders, M., Thornhill, A. and Lewis, P. (2006). *Research methods for business students*. Essex, England: Prentice Hall.

Scandura, T.A., and Williams, E.A. (2000). Research methodology in management: current practices, trends, and implications for future research. *The Academy of Management Journal*, 43(6), 1248-1264.

Schraw, G. (2006). Knowledge: structure and processes. In P. Alexander and P. Winne, eds. *Handbook of educational psychology*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 245-264.

Schreyogg, G. and Kliesch-Eberl, M. (2007). How dynamic can organizational capabilities be? Towards a dual-process model of capability dynamization. *Strategic Management Journal*, 28, 913-933.

Schulz, M. (2001). The uncertain relevance of newness: organizational learning and knowledge flows. *The Academy of Management Journal*, 44(4), 661.

Schumpeter, J. (1950). *Capitalism, socialism and democracy.* London and New York: Routledge.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., et al. (2009). Economic networks: the new challenges. *Science*, 325(5939), 422-425.

Scott, W.R. (1981). *Organizations: rational, natural, and open systems*. Englewood Cliffs, NJ: Prentice Hall.

Seidman, S. (2004). *Contested knowledge: social theory today*. Malden, MA: Blackwell Publishing.

Selnes, F. (1998). Antecedents and consequences of trust and satisfaction in buyer–seller relationships. *European Journal of Marketing* 32(3/4), 305–322.

Shen, Z. and Ma, K. (2008). Mobivis: a visualization system for exploring mobile data. In Proceedings of IEEE Pacific Visualization Symposium, pp. 175–182.

Sheriff, R.E. (2012). An evaluation of students' and lecturers' use of technologies: an engineering case study. *Engineering Education*, 7(1), 33-46.

Simon, F. and Tellier, A. (2011). How do actors shape social networks during the process of new product development? *European Management Journal*, 29(5), 414–430.

Simonin, B.L. (1999). Transfer of marketing know-how in international strategic alliances: an empirical investigation of the role and antecedents of knowledge ambiguity. *Journal of International Business Studies*, 30(3), 463-490.

Singh, J. (2005). Collaboration networks as determinants of knowledge diffusion processes. *Management Science*, 51, 756–770.

Skyrms, B. and Pemantle, R. (2001). A dynamic model of social network formation. *Proceedings of the National Academy of Sciences of the United States of America*, 97(16), 9340-9346.

Snell, A. (2006). Researching onboarding best practice: using research to connect onboarding processes with employee satisfaction, *Strategic HR Review*, 5(6), 32-35.

Squire, B.C. (2005). Knowledge transfer within vertical supplier relations: moderators and performance outcomes. PhD thesis, University of Bath.

Stecher, K., Counts, S., Cheng, L., Williams, S. and Turski, A. (2000). Salsa: leveraging email to create a social network for the enterprise. Proceedings of the Third International ICWSM Conference.

Stewart, A. (2003). Help one another, use one another: Toward an anthropology of family business. *Entrepreneurship Theory & Practice*, 27(4), 383–396.

Sturdy, A., Clark, T., Fincham, R. and Handley, K. (2009) Between innovation and legitimation: boundaries and knowledge flow in management consultancy. *Organization*, 16(5), 627-653.

Stutzman, F. (2006). An Evaluation of Identity-Sharing Behavior in Social Network Communities. *International Digital and Media Arts Journal*, 3(1), 10-18.

Subramaniam, M. and Youndt, M.A. (2005). The influence of intellectual capital on the types of innovative capabilities. *Academy of Management Journal,* 48, 450–463.

Szulanski, G. (1996). Exploring internal stickiness: impediments to the transfer of best practice within the firm. *Strategic Management Journal*, 17 (Winter Special Issue), 27-43.

Szulanski, G. (2000). The process of knowledge transfer: a diachronic analysis of stickiness. *Organizational Behavior and Human Decision Processes*, 82(1), 9-27.

Szulanski, G. and Jensen, R.J. (2006). Presumptive adaptation and the effectiveness of knowledge transfer. *Strategic Management Journal*, 27(10), 937-957.

Taylor, A.M. (1971). *Imagination and the growth of science*. New York: Shocken.

Teece D.J. (2009). *Dynamic capabilities and strategic management: organizing for innovation and growth*. Oxford and New York: Oxford University Press.

Teece, D.J., Pisano, G. and Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509-533.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: effects of network position and absorptive capacity on business unit innovation and performance. *The Academy of Management Journal,* 44(5), 996-1004.

Tsai, W. (2002). Social structure of 'coopetition' within a multiunit organization: coordination, competition, and intraorganizational knowledge sharing. *Organization Science*, 13(2), 179-190.

Tsai, W. and Ghoshal, S. (1998). Social capital and value creation: the role of intrafirm networks. *Academy of Management Journal*, 41(4), 464-476.

Tsang, E.W. (2002). Acquiring knowledge by foreign partners from international joint ventures in a transition economy: learning-by-doing and learning myopia. *Strategic Management Journal*, 23(9), 835-854.

Tufte, E.R. (1983). The visual display of quantitative information. Cheshire, CT: Graphics Press.

Tukey, J.W. (1972). Some graphic and semigraphic displays. In T.A. Bancroft, ed. *Statistical papers in honor of George W. Snedecor*. Ames: Iowa State University Press, pp. 293-316.

Tyler, J.R., Wilkinson, D.M. and Huberman, B.A. (2005). E-Mail as spectroscopy: automated discovery of community structure within organizations. *The Information Society*, 21(2), 143-153.

Unterseher, M., Jumpponen, A., Opik, M., Tedersoo, L., Moora, M., Dormann, C.F. and Schnittler, M. (2011). Species abundance distributions and richness estimations in fungal metagenomics: lessons learned from community ecology. *Molecular Ecology,* 20, 273-285.

Uzzi, B. (1997). Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly*, 42, 35–67.

Vaara, E., Sarala, R., Stahl, G.K. and Björkman, I. (2012). The impact of organizational and national cultural differences on social conflict and knowledge transfer in international acquisitions. *Journal of Management Studies*, 49(1), 1-27.

Van de Ven, A. H. (1986). Central problems in the management of innovation. *Management Science*, 32(5), 590-607.

Van de Ven, A.H., Delbecq, A. L. and Koenig, R. (1976). Determinants of coordination modes within organizations. *American Sociological Review*, 41(2), 322-338.

Van Wijk, R., Jansen, J.J.P. and Lyles, M.A. (2008). Inter- and intra-organizational knowledge transfer: a meta-analytic review and assessment of its antecedents and consequences. *Journal of Management Studies*, 45(4), 830-853.

Vargo, S.L. and Lusch, R.F. (2008a). Service-dominant logic: continuing the evolution. *Journal of the Academy of Marketing Science*, 36(1), 1-10.

Vargo, S.L. and Lusch, R.F. (2008b). From goods to service(s): divergences and convergences of logics. *Industrial Marketing Management*, 37(3), 254-259.

Venkatraman, N. and Ramanujam, V. (1986). Measurement of business performance in strategy research. *Academy of Management Review*, 11(4), 801-814.

Vera, D. and Crossan, M. (2006). Organizational learning and knowledge management: toward an integrative framework. In Easterby-Smith, M. and Lyles, M., eds. *Handbook of Organizational Learning and Knowledge Management*. Malden, MA: Blackwell, pp. 122-141.

Villena, V.H., Revilla, E. and Choi, T.Y. (2011). The dark side of buyer-supplier relationships: a social capital perspective. *Journal of Operations Management,* 29(6), 561-576.

Von Hippel, E. (1988). *Sources of innovation*. New York: Oxford University Press.

Von Hippel, E. (1994). 'Sticky information' and the locus of problem solving: implications for innovation. *Management Science*, 40(4), 429-439.

Walsham, G. (1995). Interpretive case studies in IS research: nature and method. *European Journal of Information Systems* 4(2), 74–81.

Walter, J., Lechner, C. and Kellermanns, F.W. (2007). Knowledge transfer between and within alliance partners: private versus collective benefits of social capital. *Journal of Business Research*, 60(7), 698–710.

Wasko, M.M. and Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 29(1), 35–57.

Wasserman, S. and Faust, K. (1994). *Social network analysis: methods and applications*. New York: Cambridge University Press.

Watts, A. (2001). A dynamic model of network formation. *Games and Economic Behavior*, 34(2), 331-341.

Wellman, B. (1983). Network analysis: some basic principles. *Sociological Theory,* 1, 155-200.

Wellman, B. (2002). Little boxes, glocalization, and networked individualism from little boxes to social networks. In M. Tanabe, P. van den Besselaar and T. Ishida, eds. *Digital Cities II: Computational and Sociological Approaches.* Berlin: Springer-Verlag, pp. 11-25.

Wernerfelt, B. (1984). A Resource-based View of the Firm. *Strategic Management Journal*, 5(2), 171-180.

Wernerfelt, B. (1995). The resource-based view of the firm: ten years after. *Strategic Management Journal*, 16, 171–174.

Whitford, J. (2005). The new old economy: networks, institutions, and the organizational transformation of American manufacturing. Oxford: Oxford University Press.

Whittaker, S. and Sidner, C. (1996). Email overload: exploring personal information management of email. In Proceedings of the SIGCHI conference on human factors in computing systems: common ground table of contents. Vancouver, Canada, pp. 276 - 283.

Widén-Wulff, G. and Ginman, M. (2004). Explaining knowledge sharing in organizations through the dimensions of social capital. *Journal of Information Science*, 30(5), 448-458.

Wu, L., Lin, C.-yung, Stern, N.Y.U. and Brynjolfsson, E. (2009). Value of social network – a large-scale analysis on network structure impact to financial revenue of information technology consultants working paper abstract. Winter Information Systems Conference, Salt Lake City, UT.

Wu, L., Waber, B., Aral, S., Brynjolfsson, E. and Pentland, A. (2008). Mining face-to-face interaction networks using sociometric badges: predicting productivity in an IT configuration task. International Conference on Information Systems. Paris, France.

Wu, W.L., Hsu, B.F. and Yeh, R.S. (2007). Fostering the determinants of knowledge transfer: a team-level analysis. *Journal of Information Science,* 7(33), 326-339.

Yang, J., Brashear Alejandro, T.G. and Boles, J.S. (2011). The role of social capital and knowledge transfer in selling center performance. *The Journal of Business & Industrial Marketing,* 26(3), 152-161.

Yli-Renko, H., Autio, E. and Sapienza, H.J. (2001). Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, 22(6-7), 587-613.

Zack, M.H. (1999). Managing codified knowledge. *Sloan Management Review,* 40(4), 45-58.

Zaheer, A. and Bell, G.G. (2005). Benefiting from network position: firm capabilities, structural holes, and performance. *Strategic Management Journal*, 26(9), 809-825.

Zaheer, A. and Soda, G. (2009). Network evolution: The origins of structural holes. *Administrative Science Quarterly*, 54, 1-31.

Zahra, S. and George, G. (2002). Absorptive capability: a review, reconceptualization, and extension, *Academy of Management Review*, 27(2), 185–203.

Zander, U. and Kogut, B. (1995). Knowledge and the speed of the transfer and imitation of organizational capabilities: an empirical test. *Organization Science*, 6(1), 76-92.