THE POWER OF UNSTRUCTURED DATA: A STUDY OF THE IMPACT OF TACIT KNOWLEDGE ON BUSINESS PERFORMANCE

by

Armando E. Paladino

RICK DANIELS, PhD, Faculty Mentor and Chair

VIRGINIA HINRICHS, PhD, Committee Member

KATHLEEN HARGISS, PhD, Committee Member

Susan Tally, EdD, Dean, School of Technology

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

Capella University

December 2012

Abstract

This study examined the incorporation of tacit knowledge into corporate business intelligence and its impact on business performance, specifically analyzing individual productivity. Business productivity in relation to the use of knowledge has been investigated but using macro-dimensions not specifically oriented to individual workers' productivity. This study was based on externalization, one of the modes in the theory of organizational knowledge creation (that is, converting tacit knowledge into explicit knowledge). The findings on the literature stated that knowledge is the most important piece of business competitive advantage and that tacit knowledge is a key part of that knowledge. This research found that tacit knowledge did not influence individual engineers' productivity and as such did not affect business performance. Additionally, it found that tacit knowledge was not a factor that could be used to predict individual productivity. This research was the first attempt to investigate individual productivity in relation to tacit knowledge. The research discussion includes recommendations for future research and analyzes the possible causes for the obtained results.

Dedication

This dissertation is dedicated to several people who have provided inspiration, support, and guidance in my life through either words or example. This could not have been possible without the loving and firm care of God that was always behind them. To my wife, Leyla, who supported me even in my most difficult days; my father and mother, who even though are no longer with me will always be remembered with love and the assurance that they would be very proud of me today; my father- and mother-in-law, who have been an inspiration and, in the absence of my biological parents, have played an important role in my life; my stepdaughters Alexandra and Vanessa with whom I have a great friendship; my children, Isaac, Adita, Denise, and Jennifer, with whom I have always maintained a special relationship; and, finally, Ricardo, my nephew, whom I have always considered my second son, I dedicate this work.

Acknowledgments

First and most, a big thanks to almighty God, whose gave me the strength and clarity at all moments through the guidance of my Master Paramahansa Yogananda. To Dr. Lawrence Ness, my first instructor at Capella who surely prepared me for this long journey; Dr. Martha Hollis, from whom I learned quite a bit of statistics and best practices; Dr. Tsun Chow, who encouraged me to continue with my ideas related to my research and from whom I got the idea to register for my business intelligence Information Technology special class; the members of my committee, Virginia Hinrichs and Kathleen Hargiss, who were so kind in guiding me; and, finally Dr. Rick Daniels, my mentor whom I had the fortune to choose, I also say thank you.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

Introduction to the Problem

Based on the literature, it can be argued that knowledge is the most important piece of business competitive advantage, and that tacit knowledge is a key part of that knowledge. Additionally, corporations possess vast amounts of unstructured data waiting to be extracted and processed with the potential to be converted into knowledge. The knowledge obtained from unstructured data when incorporated in the BI process has the potential to support and enhance the current businesses (Chang Lee, Lee, & Kang, 2005; Goel, Rana, & Rastogi, 2010; Mezher, Abdul-Malak, Ghosn, & Ajam, 2005; Mundra, Gulati, & Vashisth, 2011).

Moreover, it can be argued that using externalization, one of the modes of the Nonaka (1994) theory (that is, converting tacit knowledge into explicit knowledge), can increase business performance. According to Cheung, Lee, and Wang (2005), 80% of organizations' data are stored in some form of unstructured data, and that these data hide an enormous potential in terms of supporting business performance and information in general. Transactions, databases, records, keys, and attributes typify the structured environment. E-mail, spreadsheets, medical records, documents, and reports typify the unstructured environment (W. H. Inmon & Nesavich, 2009). Unstructured data contain valuable corporate business information, and, according to Negash (2004), 60% of chief information officers and chief technology officers consider unstructured data very critical for the business. Why is it that unstructured data are not being processed and integrated into the business intelligence (BI) schema?

The answer to the previous question may be in the complexity to process the data (Abidin, Idris, & Husain, 2010; Rao, 2003; Srivastava & Cooley, 2003) or the failure to see its impact in business performance, but incorporating unstructured data into the business schema is not an impossible task. The knowledge obtained from unstructured data when incorporated in the BI process has the potential to support and enhance the current businesses (Chang Lee, Lee, & Kang, 2005; Goel, Rana, & Rastogi, 2010; Mezher, Abdul-Malak, Ghosn, & Ajam, 2005; Mundra, Gulati, & Vashisth, 2011). This research evaluated if incorporating tacit knowledge and unstructured data into the business intelligence schema would impact business performance.

Background of the Study

In his seminal work, Nonaka (1991) introduced the concept of perceiving the company as "a living organism. Much like an individual, it can have a collective sense of identity and fundamental purpose" (p. 8). Nonaka (1994) expanded his previous work and postulated the theory of organizational knowledge creation. He explained that knowledge possessed by individuals, organizations, and societies can be expanded through a spiral process in which tacit knowledge is converted into explicit knowledge, and then back into tacit. Tacit knowledge is hidden behind behaviors, skills competencies, and experiences (tacit actionable knowledge); articulated knowledge resides in individual thoughts and language use. Explicit knowledge resides inside computers in codified form, and by nature has a clear organization (Delen & Al-Hawamdeh, 2009).

Nonaka (1994) also provided an interpretation of tacit and explicit knowledge: "Polanyi classified human knowledge into two categories. 'Explicit' or codified knowledge refers to knowledge that is transmittable in formal, systematic language. On the other hand, 'tacit' knowledge has a personal quality, which makes it hard to formalize and communicate" (p. 16). Nonaka called the distinction between tacit and explicit knowledge the *epistemological dimension to organizational knowledge*. The exchange can take many forms and, based on these variations, different modes of knowledge conversion can be generated: (a) tacit to tacit, which is a shared experience and is called socialization, (b) explicit to explicit, in which modern computers play an important role, and it is called combination. The third and fourth modes are a combination of the first two, converting explicit into tacit, called internalization, and converting tacit into explicit, called externalization.

On the ontological dimension, the theory posits that individuals are the ones who create knowledge and that an organization should amplify this knowledge through the different levels of the firm. The key here is the constant dialogue in which the middle-up-down management leadership is the most suitable to crystallize the conversion and creation of knowledge. Nonaka (1994) used the metaphor of the orchestra, in which musicians play their part and the conductor coordinates the effort, producing a clean and coordinated melody. In this context, the middle manager is the bridge between employees and top management: "In sum, middle managers synthesize the tacit knowledge of both frontline employees and top management makes it explicit, and incorporates it into new technologies and products" (p. 32).

Zack (1999a) posited that the company's strategy is the most important context to guide knowledge management (KM). He had a clear vision when he wrote this more than 10 years ago as organizations are now more customer-oriented and KM is used to reach those customers. To succeed in today's business, corporations must be customer-focused, and one goal of KM is to provide a holistic view of the customer (Cader, 2007). In a sense, success depends greatly on knowledge after producing value from resources, and KM is the only promising medium to gain competitive advantage (Eftekharzadeh, 2008).

Zack (1999a) stated that the link between KM and a business strategy has been ignored and that companies must have a knowledge strategy. Zack stated that the best known corporate strategy approach uses the strengths, weaknesses, opportunities, and threats (SWOT) model, and that "application of the SWOT framework has been dominated over the last 20 years by Porter's 'five-force' model" (p. 127). The new perspective of strategic management deviates from the original model by Michael Porter as it focuses on internal resources and capabilities rather than the products produced by the resources. Knowledge, especially tacit knowledge, is difficult to acquire and cannot be purchased. The more intellectual resources a company has, the better equipped it is to compete; therefore, knowledge becomes the strategic resource for competitive advantage.

Zack (1999b) expanded his work and continued to popularize the concept of viewing organization knowledge as a strategic asset. The views of Zack differ from those of Nonaka (1994) because Zack considered explicit knowledge the most important asset and Nonaka put the relevance in tacit knowledge.

Nonaka and Toyama (2003) expanded on the concept of tacit and explicit knowledge. They described this process as occurring in a virtual environment that they named the *ba*: "Building on the concept that was originally proposed by the Japanese philosopher Kitaro Nishida (1921, 1970), we define ba as a shared context in motion, in which knowledge is shared, created, and utilized" (p. 6). Nonaka and Toyama reiterated that in today's world, knowledge is the most important source of a firm's sustainable competitive advantage, and that a new knowledge-based theory is needed that differs from the existing economic and organizational theory coinciding with Zack (1999a, 1999b). Socialization, externalization, combination, and internalization (SECI) is the model that describes the different phases of knowledge being converted through the spiral of knowledge (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Nonaka & von Krogh, 2009). The SECI model of creating knowledge and the ba environment are the dynamic creation of knowledge from tacit to explicit with the different variations.

Statement of the Problem

Business productivity in relation to the use of knowledge has been investigated, but using macro dimensions not specifically oriented to individual workers' productivity. For example, Chang Lee et al. (2005) designed and tested a new instrument that measures knowledge management performance using stock prices, price-earnings ratio (PER), and research and development (R&D) expenditure as the dependent variables. Using linear regression to measure the sample, Chang Lee et al. surveyed companies in the KOSDAQ (Korean Securities Dealers Automated Quotations) market in Korea, and the results showed strong significant numbers on the factors supporting the following hypothesis: When the Knowledge Management Performance Index is greater, stock prices, PER, and R&D are significantly better.

Singh (2008) conducted a survey on an Indian software company to investigate the impact of leadership styles on KM productivity. Even though Singh investigated explicit and tacit knowledge, those variables were used as the dependent variables being impacted by gender and leadership styles and not as predictors of productivity. Other authors have investigated business productivity in relation to KM, but using factors different from individual productivity. Whereas Goel et al. (2010) and Mezher et al. (2005) studied portal implementations, Mundra et al. (2011) examined competitive advantage as a variable for KM success.

Goel et al. (2010) argued that business gained competitive advantage after the use of KM, but they did not prove it statistically, and the results only showed possibilities at a higher level. Some of the findings from Mundra et al. (2011) revealed that companies are combining artificial intelligence to retrieve data from databases; that video conferences, e-mail, and chat groups are necessary tools to share knowledge, including tacit and explicit knowledge; and that one of the companies was able to reduce part of its training program from seven days to 4-5 hours with the help of knowledge, but this was a very broad finding that failed to point to the specifics of individual productivity. Mezher et al. (2005) conducted a case study on an engineering consulting company trying to demonstrate project management efficiency by the use of KM. They created a KM model and postulated that the use of knowledge would help to complete engineering projects in less time, but they did not test the model statistically.

Bosch-Sijtsema, Ruohamaki, and Vartianinen (2009) studied productivity on globally distributed teams as a whole, and even though individual productivity is mentioned in their framework, no provisions or details were shown on how to test the productivity. Britt (2009) attributed the increase of productivity in an insurance broker company to the use of knowledge obtained from a web-based model, a very vague and general assessment that failed to take into consideration other factors that may have contributed to the increase in insurance policies. Martins and Lopes dos Reis (2010) studied productivity under the human capital lens, more oriented to identify it as an asset that can be recorded along with the other assets that a company possesses. They used a very sophisticated software tool to measure their proposed framework but, again, this did not prove individual productivity as influenced by knowledge factors. Finally, Mahmood and Ali (2011) performed structural equation modeling after they operationalized several constructs to predict productivity. They concluded that knowledge sharing, organizational culture, and technology and rewards contributed to knowledge workers' productivity. However, they did not show how productivity was increased.

From August to October 2004, the University of New South Wales conducted research on KM studies. The research went back as far as 1892; 290 research papers were analyzed by two master of science students (The University of New South Wales, 2004). Only three research studies were identified to have investigated productivity and they did not address individual productivity:

- McCampbell, Clare, and Gitters (1999) conducted case studies on Teltech, Ernst & Young, Microsoft, and Hewlett-Packard to analyze the effect of KM in quality and productivity improvement, but the study was very general and did not include specifics.

- Zazzara (2001) mentioned that to elevate productivity and maintain clinical quality through the use of knowledge would be nirvana for the healthcare system but did not demonstrate specifics on how to achieve it.

- Filius, de Jong, and Roelofs (2000) prescribed three activities for organizations that are willing to improve productivity: activities that expand

> the individual or collective horizon, activities that consolidate knowledge, and informal or formal communication about the issue. However, again, this was a wide recommendation without any practical prescription.

All of the previous studies investigated the impact of KM in business performance, but they failed to show how KM impacted individual productivity.

Purpose of the Study

The purpose of this research was twofold. First, the intent of the study was to show that the incorporation of unstructured data into BI could increase business performance and motivates research around the designing and developing of new paradigms and ontologies to help with the complexity of inserting unstructured data into the data warehouses. Second, the study laid out the foundation for further research on KM factors that could contribute to business performance. The chosen research instrument was the Knowledge Management Assessment Tool (KMAT). The KMAT "was developed by Maier and Moseley (as cited in Singh, 2008) and consists of 30 statements to measure knowledge management practices of the organization" (Singh, 2008, p. 9). The tool measures five dimensions: knowledge identification and creation (KIC), knowledge collection and capture (KCC), knowledge storage and organization (KSO), knowledge sharing and dissemination (KSD), and knowledge application and use (KAU). Additionally, the tool has a mechanism to convert those five dimensions into only two dimensions: explicit knowledge and tacit knowledge.

Polanyi (2009) declared the importance of tacit knowledge and the difficulty to formalize and communicate it. In practical terms, he stated that tacit knowledge must not be excluded from the KM schema. This research was an attempt to demonstrate the importance of tacit knowledge, and its impact on increasing business performance.

By showing which of the dimensions measured by the research instrument affect business performance in a positive way, corporations can add more resources to that dimension to obtain performance gains. This can be done for each of the five dimensions of the KMAT as well as the two dimensions of explicit and tacit knowledge. So far, the literature shows studies on the impact of knowledge in business performance in general, but no study has been found that measures individual units of productivity. The research framework was a software and hardware engineering company that provides support to customers. Individual productivity was evaluated after incorporating and sharing tacit knowledge into the daily support operations. These findings could be used to predict productivity on engineering companies that provide customer support, similar to the one described in this research.

Rationale

W. H. Inmon and Nesavich (2009) commented about the way unstructured and structured data have developed: "It is amazing that at the same time that these worlds have grown up side by side, they have grown separately. It is as if these worlds exist in alternate universes" (p. xvii). Although it is true that structured and unstructured data have grown together in the different businesses, if the data are not processed, looked into, and intelligence extracted from it, the data are of no use. As W. H. Inmon and Nesavich noted, "Stated differently, organizations that look only at their structured data—usually transaction-based data—miss an entire class of information that waits to be used for the decision-making process" (p. 11). Analyzing feedback from customers and mining the web is the cornerstone of today's business (Kuechler, 2007); there is an enormous business potential when extracting business intelligence from unstructured data (B. Inmon, 2005).

As previously stated, there is a lot of potential in extracting knowledge from tacit knowledge and unstructured data. Several authors have investigated the impact of knowledge in business performance, but they did not study individual productivity. Chang Lee et al. (2005) proposed new metrics to measure business performance in relation to knowledge, but they concentrated on macro dimensions. Singh (2008) investigated explicit and tacit knowledge, but those variables were used as the dependent variables being impacted by gender and leadership styles and not as predictors to productivity. Goel et al. (2010) concluded,

> This paper has demonstrated that corporate sustainability is strongly linked to KM. Developing a KM strategy is the core to the concept of sustainability as improvement in the way knowledge assets are managed and reported, can lead to better corporate governance, facilitate continuous improvement, enhance stakeholder value and provide sustainable competitive advantage. (p. 114)

Bosch-Sijtsema et al. (2009) studied productivity on globally distributed teams as a whole, and although individual productivity is mentioned in their framework, no provisions or details were shown on how to test the productivity, leaving the model a mere concept, leaving open ends for future research. From August to October 2004, the University of New South Wales conducted research on KM studies. The research went back as far of 1892, and 290 research papers were analyzed by two master of science students (The University of New South Wales, 2004). Only three research studies were identified to have investigated productivity and they did not address individual productivity. This current research was the first attempt to address individual productivity to demonstrate the business value of sharing tacit knowledge; it laid out the foundation for more research and to continue developing ontologies to extract the same knowledge in an automated way other than the pure sharing in communities.

Research Questions

As previously stated, there is a lot of potential in extracting knowledge from tacit knowledge and unstructured data, and that lead to the following questions:

1. Can the incorporation of customer support unstructured data into the customer support schema increase business performance?

2. If unstructured customer support data are converted into explicit knowledge; can this converted data contribute to business performance by increasing engineers' productivity?

3. To what extent is there a significant decrease in the time to complete field engineers' tasks after unstructured data are incorporated into the BI framework?

4. Can the usage of any of the KMAT factors predict field engineers' time to complete tasks when unstructured data are incorporated into the BI framework?

The following hypotheses were developed from the research questions:

- $H1_0$: The inclusion of tacit knowledge in BI does not produce a significant difference in time to complete engineering tasks at customer sites.

- $H1_A$: The inclusion of tacit knowledge in BI produces a significant difference in time to complete engineering tasks at customer sites.

- $H2_0$: Tacit knowledge is not a factor that can be used to predict employees' productivity when included in BI.

- $H2_A$: Tacit knowledge is a factor that can be used to predict employees' productivity when included in BI.

The study tried to answer the questions and conducted analysis to accept or reject the hypotheses.

Significance of the Study

This research was an attempt to demonstrate the importance of tacit knowledge, and its impact on increasing business performance. According to Koskinen (2004), most of the attention in projects has been on codified explicit knowledge, neglecting the vast amount of information residing in tacit implicit knowledge. This study examined explicit (semi-structured) data, made explicit from tacit knowledge from individuals after they shared their experiences in online forums. The data were in a semi-structured format because they were still in free-form and had not been integrated into any relational or indexed database. Feghali and El-Den (2008) postulated that (a) ideas and opinions are the easiest form of tacit knowledge to share among virtual groups, and (b) these opinions and ideas can be progressively shared among virtual groups by the creation and sharing of documents. Tacit knowledge articulated in this way would supplement the face-to-face interaction that is missing in virtual environments. By a constant dialogue and refinement of the ideas expressed in the documents, the most hidden portions of the tacit knowledge can be discovered; "knowledge transformation among virtually dispersed group members is possible through the articulation of members' opinions and ideas into a shared document. This document provides the infrastructure for the interaction among the members by exposing them to each other's opinions" (Feghali & El-Den, 2008, p. 103).

The importance of knowledge residing in people's minds and the importance of sharing that knowledge was also emphasized by Brown and Duguid (2000): "Attending to knowledge, by contrast returns attention to people. . . . The importance of people as creators and carriers of knowledge is forcing organizations to realize that knowledge lies less in its databases than in its people" (p. 121). Furthermore, Brown and Duguid (2000) highlighted the value of communities of practice in the labs of particle physicists and biotechnologists..

Definition of Terms

Business intelligence (BI). "Emerged in early-1990s [,]. . . BI is a set of new technologies such as DW [Data Warehousing], OLAP [online analytical processing], and DM [Data Mart] which are used to handle with and analysis structured data in order to support decision making" (Zhou, Cheng, Chen, & Xiao, 2007, p. 5,468).

Business performance.

Business performance would include both operational performance (i.e. nonfinancial) as well as financial performance. In this framework several indicators can be considered, new product introduction, product quality, market share and effectiveness, and technological efficiency (Venkatraman & Ramanujam, 1986).

Data marts. "(Subsets of data warehouses) are conformed by following a standard set of attribute declarations called a data warehouse bus" (Arun & Atish, 2005, p. 79).

Data warehouse. "A subject-oriented, integrated, time-variant, and nonvolatile collection of data that supports managerial decision making" (Arun & Atish, 2005, p. 79).

Knowledge. "A fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers" (Davenport & Prusak, 1998, p. 5). Knowledge is a part of the traditional knowledge pyramid.

> Data are considered to be unprocessed raw representations of reality. Information is considered to be data that has been processed in some meaningful ways. Knowledge is considered to be information that has been processed in some

meaningful ways. Wisdom is considered to be knowledge that has been processed in some meaningful ways. (Faucher, Everett, & Lawson, 2008, p. 5)

This relation between data, information, knowledge, and wisdom is represented in Figure 1 and Table 1. This describes the graphical hierarchy and provides a brief description of each of the different phases, until the final one, wisdom, is reached.

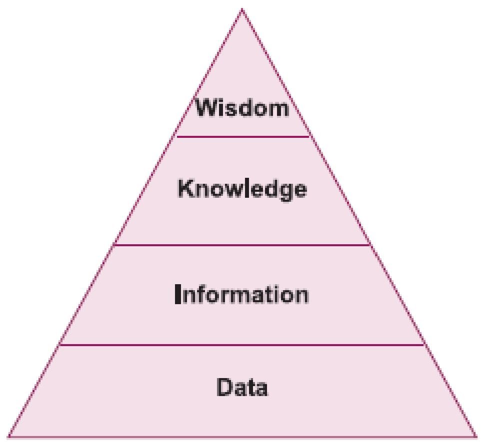

Figure 1. Traditional knowledge pyramid. From "Reconstituting Knowledge Management," by J.-B. P. L. Faucher, A. M. Everett, and R. Lawson, 2008, *Journal of Knowledge Management, 12*(3), 7. Copyright 2008 by Emerald Group Publishing Limited. Reprinted with permission.

Table 1. Distinctions Between Data, Information, Knowledge, and Wisdom

Level	Definition	Learning process	Outcome
Data	Raw facts	Accumulating truths	Memorization (data bank)
Information	Meaningful, useful data	Giving form and functionality	Comprehension (information bank)
Knowledge	Clear understanding of information	Analysis and synthesis	Understanding (knowledge bank)
Wisdom	Using knowledge to establish and achieve goals	Discerning judgments and taking appropriate actions	Better living/success (wisdom bank)

Note. From "Organizational Learning, Knowledge and Wisdom," by P. E. Bierly III, E. H. Kessler, and E. W. Christensen, 2000, *Journal of Organizational Change Management, 13*(6), p. 598. Copyright 2000 by Emerald Group Publishing Limited. Reprinted with permission.

Knowledge management (KM).

> Can be described as the management of the environment, making knowledge flow through the different phases of its life cycle. Thus, knowledge developed at one place in an organization can be made available to other units through an organizational knowledge repository. Companies survive with the continuous development of new knowledge based on creative ideas, daily experiences, and work in R&D departments. A company can only perform at its best if all available knowledge areas are combined. (Chang Lee et al., 2005, p. 472)

Online analytical process (OLAP).

> Takes the decision maker to new levels in data analysis. With OLAP, the decision maker's analysis interacts with the data contained within the system. It leverages the time-variant characteristics of the data warehouse to allow the strategist to look back in time as well as into the future. (Giovinazzo, 2003, p. 42)

Semantic web. "Is envisioned as an extension of the current web where, in addition to being human-readable using WWW browsers, documents are annotated with meta-information. This meta-information defines what the information (documents) is about in a machine-processable way" (Davies, Fensel, & van Harmelen, 2003, p. 4).

Structured data environment. "Is typified by transactions, databases, records, keys, and attributes" (W. H. Inmon & Nesavich, 2009, p. xvii). This information is always stored in the context of a data warehouse inside databases and stored with indexes for easy retrieval, with each record being uniquely identifiable.

Unstructured data environment. "Is typified by email, spreadsheets, medical records, documents, and reports. . . . The world of analytics and business intelligence has grown up around structured information" (W. H. Inmon & Nesavich, 2009, p. xvii). Even though some of this unstructured data have some form of structure, like the spreadsheets, the data as a whole are not stored in any form of relational databases, and there is no easy way to retrieve them by index, or to relate each file within each other.

> There is no format, structure, or repeatability to unstructured textual data. There is no one sitting on your shoulder telling you what to do when you write an email. . . . . In addition, there are other forms of text that occur well outside the email environs, such as contracts, warranties, spreadsheets, telephone books, advertisements, marketing materials, annual reports, and many more forms of textual information that are the fabric of the organization. (W. H. Inmon & Nesavich, 2009, p. 2)

These unstructured records are not related to each other and they are recorded in easy free form with no index or unique identifiers; processing these data under the current data-warehousing environment is a challenge due to no repeatability and no predictability of the data.

Assumptions and Limitations

Assumptions

The first assumption for this research was that tacit knowledge will produce an impact on business performance and that tacit knowledge is the most important piece in organizations (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Nonaka & von Krogh, 2009). The second assumption was that knowledge sharing and knowledge transfer also improve business performance (Antonova, Csepregi, & Marchev, 2011).

Limitations

The model for this research was particularly strong due to the nature of the company being surveyed. The company is spread around the world and each country resembles the corporate model; therefore, if a good random sample was collected and the statistical numbers were solid, the conclusions have the potential to be generalized but the generalization would be on similar industries only.

Because the study survey was mailed electronically to internal employees who were easily reached, the model had the potential to collect a good sample size. One of the limitations of this study was that the population sample was collected from a software/hardware engineering company and the results may not be generalized to other types of companies. A subsequent study can be conducted with the findings to reevaluate the hypotheses.

Conceptual Framework

The context for this research was an engineering company that sells hardware and software to customers and provides customer support to maintain its products. When a customer calls the support hotline, a preliminary analysis of the customer's issue is completed, and, if needed, an engineer is dispatched onsite to either troubleshoot the issue or comply with a specific task, for example, to replace a computer system board or processor.

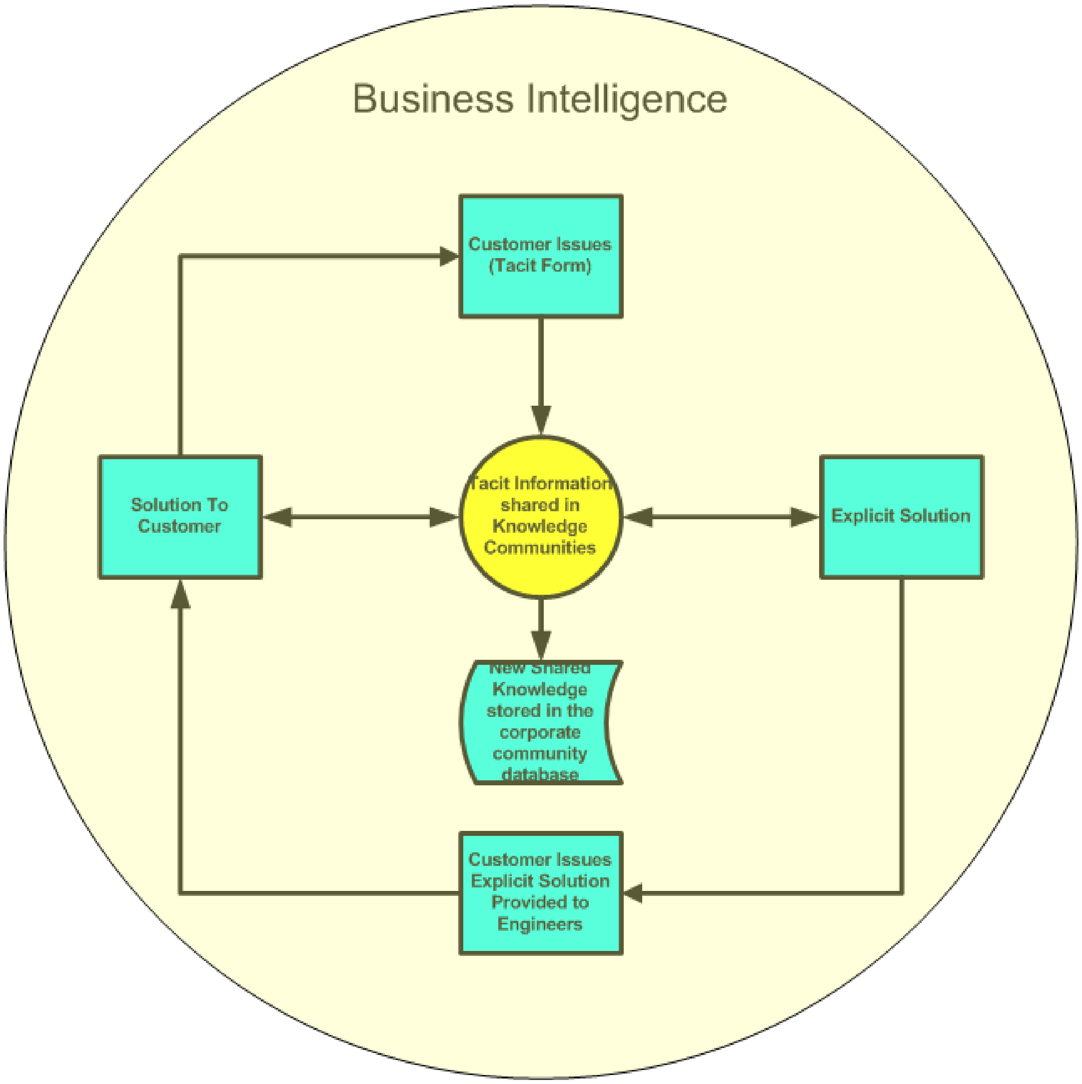

Figure 2. Research framework.

When engineers are dispatched, they are briefed with a problem description and possible solution to the problem. The preliminary solution is based on remote diagnosis, but more often, engineers encounter other issues, or the same one continues to manifest after the parts have been replaced or the suggested fix applied. Currently, the company measures the performance of these engineers in two ways: (a) the number of completed tasks and (b) the time taken to complete those tasks. For example, two engineers can complete the same numbers of tasks, but one will take more time to complete them.

Therefore, if possible it was important to consider both time to completion and number of tasks completed, and perhaps the best way to measure these is the ratio of time to completion over a number of tasks, but number of task completed in the month was chosen.

According to Koskinen (2004), most of the attention in projects has been on codified explicit knowledge, neglecting the vast amount of information residing in tacit implicit knowledge. This study looked into explicit (semi-structured) data, made explicit from tacit knowledge from individuals after sharing their experiences on online forums. The data were in semi-structured format because they were still in free-form and not integrated into any relational or indexed database. Feghali and El-Den (2008) postulated that (a) ideas and opinions are the easiest form of tacit knowledge to share among virtual groups, and (b) these opinions and ideas can be progressively shared among virtual groups by the creation and sharing of documents. Tacit knowledge articulated this way would supplement the face-to-face interaction that is missing in virtual environments. By a constant dialogue and refinement of the ideas expressed in the documents, the most hidden portions of the tacit knowledge can be discovered; "knowledge transformation among virtually dispersed group members is possible through the articulation of members' opinions and ideas into a shared document. This document provides the infrastructure for the interaction among the members by exposing them to each other's opinions" (Feghali & El-Den, 2008, p. 103).

The importance of knowledge residing in people's minds and the importance of sharing that knowledge was also emphasized by Brown and Duguid (2000): "Attending to knowledge, by contrast returns attention to people. . . . The importance of people as creators and carriers of knowledge is forcing organizations to realize that knowledge lies less in its databases that in its people" (p. 121). Furthermore, Brown and Duguid highlighted the value of communities of practice in the labs of particle physicists and biotechnologists.

Under the conducted research, customer support data that resided in people's minds in the form of tacit knowledge was shared through an online forum after engineers created and shared documents containing experiences from their tasks in the form of issues, recommendations, and advice in general. This forum was searched prior to the execution of new tasks. The forum was created on a corporate repository for engineers to share and discuss their experiences. These data continuously retro feed the community with more and refined knowledge about installations, tasks performed, and general support data. Knowledge shared and created from unstructured data is a circular process—it comes from customer issues and needs—engineers search possible solutions in the forum, and after reading the different comments, issues, past experiences, and suggestions, that information is used to find solutions. The new information was then used by field engineers to provide a solution to the customer, and this new information was fed back into the repository. As more knowledge is acquired, the customer gets better support, and more business is created; the result could be more productivity, and engineers could finish their tasks early, potentially completing more tasks (see framework in Figure 2).

The research is an experimental design.

> The simplest of all experimental designs is the two-group posttest-only randomized experiment. In design notation, it has two lines—one for each

> group—with an R at the beginning of each line to indicate that the groups were randomly assigned. (Trochim, 2006, para. 1)

As indicated by the *R*s in Figure 3, two groups of engineers were randomly assigned—one group of engineers did not belong to the new knowledge community (created for the experiment) and not exposed to the tacit knowledge and unstructured data variable, and one group (denoted by the *X*) was exposed to the tacit knowledge unstructured data variable. Because the researcher was interested in knowing if the experimental group's productivity was different after the exposure to the new knowledge, the group means were tested using an independent *t*-test analysis. After the control group's number of completed tasks was found to be significantly different from the experimental group's number, then regression analysis was conducted to determine which of the factors contributed the most to the results.

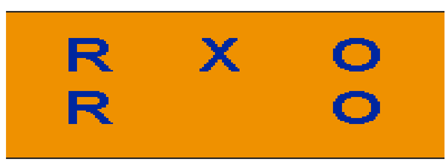

Figure 3. Two-group posttest-only randomized experiment. From "Two-Group Experimental Designs," by W. M. K. Trochim, 2006, retrieved from http://www.socialresearchmethods.net/kb/expsimp.php. Copyright 2006 by William M. K. Trochim. Reprinted with permission.

Multiple linear regression tests were conducted on the KMAT knowledge factors to predict productivity.

> Multiple regression analysis examines the relationship between a single dependent variable and two or more independent variables. It is a widely used analytic technique in organizational research and has been the most popular statistical technique for hypothesis testing for at least two decades (Weinzimmer, Mone, & Alwan, 1994). (Bates, 2005, p. 118)

The KMAT knowledge factors could not be confirmed using factor confirmatory analysis because the size of the sample. Additionally, structural equation modeling was performed on the results to support the regression tests. "Structural equation modeling (SEM) is a statistical methodology that takes a confirmatory (i.e., hypothesis testing) approach to the analysis of a structural theory bearing on some phenomenon" (Byrne, 2010, p. 3). The productivity was measured as a function of the different factors of knowledge, specifically tacit knowledge.

Organization of the Remainder of the Study

This dissertation comprises five chapters. Chapter 1 introduced the problem, provided with the background of the study, the problem statement, purpose and rationale as well as introduced the research questions and covered the assumptions and limitations. This chapter also presented the conceptual framework in a graphical representation. Chapter 2 provides the background for the theory of organizational knowledge creation and provided with an extensive literature review on the study of tacit and explicit knowledge covering the most important seminal authors. Chapter 3 provides a description of the sample, instrument used, methods of data collection, data analysis, validity and reliability, and ethical considerations. Chapter 4 provides the detailed analysis of the data collected and the conclusions that supported or rejected the hypotheses. Chapter 5 contains the final conclusions and recommendations for future research.

CHAPTER 2. LITERATURE REVIEW

Introduction

This research attempted to demonstrate the importance of tacit knowledge, and its impact in business performance. This research was aligned with Nonaka's (1994) view that tacit knowledge is the most important piece to increase business performance, and based on the literature argued that using externalization one of the modes of the Nonaka theory (converting tacit knowledge into explicit knowledge) could increase business performance. Since this research focused on individual productivity, a significant portion of the literature review concentrated on studies related to individual productivity in relation to tacit knowledge. Additionally, the literature review also analyzed several studies that investigated the impact of tacit and explicit knowledge in business performance.

Background

In the seminal piece, *The Knowledge-Creating Company*, Nonaka (1991) introduced the concept of perceiving a company as "not a machine but a living organism. Much like an individual, it can have a collective sense of identity and fundamental purpose" (p. 8). Nonaka (1994) expanded his previous work with the article, "A Dynamic Theory of Organizational Knowledge Creation," and postulated the theory of organizational knowledge creation. His writing explained that knowledge possessed by individuals, organizations, and societies can be expanded through a spiral process in which tacit knowledge is converted into explicit knowledge and then back into tacit knowledge. Tacit knowledge is hidden behind behaviors, skills, competencies, and experiences (tacit actionable knowledge) and articulated knowledge (implicit knowledge), which resides in individual thoughts and language use. Explicit knowledge resides inside computers in codified form and by nature has a clear organization (Delen & Al-Hawamdeh, 2009).

Nonaka (1994) also provided an interpretation of tacit and explicit knowledge: "Polanyi classified human knowledge into two categories. 'Explicit' or codified knowledge refers to knowledge that is transmittable in formal, systematic language. On the other hand, 'tacit' knowledge has a personal quality, which makes it hard to formalize and communicate" (p. 16). Nonaka called the distinction between tacit and explicit knowledge the *epistemological dimension to organizational knowledge.* The exchange can take many forms and, based on these variations, different modes of knowledge conversion can be generated. Tacit-to-tacit is a shared experience called *socialization.* Explicit-to-explicit, in which modern computers play an important role, is called *combination*. The third and four modes are a combination of the first two: converting explicit knowledge into tacit knowledge, called *internalization*; and converting tacit knowledge into explicit knowledge, called *externalization*.

On the ontological dimension, Nonaka's (1994) theory posits that individuals create knowledge and that organizations should amplify this knowledge through the different levels of the firm. The key here is the constant dialogue in which middle-up-down management leadership is the most suitable to crystallize the conversion and creation of knowledge. Nonaka used the metaphor of an orchestra in which musicians play their individual parts and the conductor coordinates their efforts, producing a clean and coordinated melody. In this context, the middle manager is the bridge between employees and top management: “In sum, middle managers synthesize the tacit knowledge of both frontline employees and top management makes it explicit, and incorporates it into new technologies and products” (p. 32). Zack (1999a) posited that a company’s strategy is the most important context to guide knowledge management (KM). He had a clear vision when he wrote this more than 10 years ago as organizations are now more customer-oriented and KM is used to reach to those customers. To succeed in today’s business environment, corporations must be customer-focused, and one goal of KM is to provide a holistic view of the customer (Cader, 2007). In a sense, success depends greatly on knowledge after producing value from resources, and KM is the only promising medium to gain a competitive advantage (Eftekharzadeh, 2008).

In his writing, Zack (1999a) stated that the link between KM and a business strategy has been ignored and that companies must have a knowledge strategy. Zack stated that the best-known corporate strategy approach uses the strengths, weaknesses, opportunities, and threats (SWOT) model, and that “application of the SWOT framework has been dominated over the last 20 years by Porter’s ‘five-force’ model” (p. 127). The new perspective of strategic management deviates from the original model by Michael Porter as it focuses on internal resources and capabilities rather than the products produced by the resources. Knowledge, especially tacit knowledge, is difficult to acquire and cannot be purchased. The more intellectual resources a company has, the better equipped it is to compete and, therefore, knowledge becomes the strategic resource for a competitive advantage.

Zack (1999b) expanded his work in his article, “Managing Codified Knowledge,” in which he continued popularizing the concept of viewing organization knowledge as a strategic asset. The views of Zack differ somewhat from those of Nonaka (1994) because Zack considered explicit knowledge the most important asset and Nonaka put the relevance in tacit knowledge.

Nonaka and Toyama (2003) expanded on the concept of tacit and explicit knowledge. They described this processing occurring in a virtual environment that they named the *ba*: “Building on the concept that was originally proposed by the Japanese philosopher Kitaro Nishida (1921, 1970), we define ba as a shared context in motion, in which knowledge is shared, created, and utilized” (p. 6). Nonaka and Toyama reiterated that in today’s world, knowledge is the most important source of a firm’s sustainable competitive advantage and that a new knowledge-based theory is needed that differs from the existing economic and organizational theory, coinciding with Zack (1999a, 1999b). Socialization, externalization, combination, and internalization (SECI) is a model that describes the different phases of knowledge being converted through the spiral of knowledge (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Nonaka & von Krogh, 2009). The SECI model of creating knowledge and the ba environment are the dynamic creation of knowledge from tacit knowledge to explicit knowledge with the different variations.

Regarding tacit knowledge, Polanyi (2009) stated, “I shall reconsider human knowledge by starting from the fact that we can know more than we can tell” (p. 4). He also declared, “We recognize the moods of the human face, without being able to tell, except quite vaguely, by what signs we know it” (p. 4) and classified this human characteristic as tacit knowledge, a knowledge that is hard to formalize and communicate. He further stated, “I think I can show that the process of formalizing all knowledge to the exclusion of any tacit knowing is self-defeating” (p. 20). As W. H. Inmon and Nesavich (2009) noted, “Stated differently, organizations that look only at their structured data—usually transaction-based data—miss an entire class of information that waits to be used for the decision-making process” (p. 11). Accordingly, this research attempted to demonstrate the importance of tacit knowledge and its impact in increasing business performance.

According to Koskinen (2004), most of the attention in projects has been on codified explicit knowledge, neglecting the vast amount of information residing in tacit implicit knowledge. This research examined explicit (semi-structured) data made explicit from tacit knowledge from individuals after sharing their experiences on online forums. The data were in a semi-structured format because they were still in free-form and had not been integrated into any relational and/or indexed database.

Feghali and El-Den (2008) postulated that (a) ideas and opinions are the easiest form of tacit knowledge to share among virtual groups and (b) that these opinions and ideas can be progressively shared among virtual groups by the creation and sharing of documents. Tacit knowledge articulated this way would supplement the face-to-face interaction that is missing in virtual environments. By a constant dialogue and refinement of the ideas expressed in documents, the most hidden portions of tacit knowledge can be discovered: “Knowledge transformation among virtually dispersed group members is possible through the articulation of members’ opinions and ideas into a shared document.

This document provides the infrastructure for the interaction among the members by exposing them to each other's opinions" (p. 103).

The importance of knowledge residing in people's minds and of sharing that knowledge is also emphasized by Brown and Duguid (2000): "Attending to knowledge, by contrast[,] returns attention to people. . . . The importance of people as creators and carriers of knowledge is forcing organizations to realize that knowledge lies less in its databases than in its people" (p. 121). Furthermore, Brown and Duguid highlighted the value of communities of practice in the labs of particle physicists and biotechnologists.

Unstructured Data and Tacit Knowledge

The importance of unstructured data as a source of knowledge has been highlighted by a series of authors. Kuechler (2007) stated that new laws mandate unstructured data to be monitored. He referred to the Sarbanes-Oxley Act of 2002, Securities and Exchange Commission Rule 17a-4, and the Health Insurance Portability and Accountability Act of 1996. These laws protect data so that private information does not become public. These mandates will facilitate the development of tools to process unstructured data. Kuechler focused his analysis on the unstructured text acquisition and analysis that can be useful in business intelligence (BI), customer relationship management, regulatory compliance, intellectual property management, call support, accounts payable/receivable analysis, and legal department support.

Abidin et al. (2010) reported that three-quarters of corporate data are in an unstructured format and that this represents an enormous opportunity for positive economic returns. They claimed that effectively managing this data will result in revenue and profitability. The dilemma with unstructured data resides in the tools to manage it because current tools are designed to extract data that are in a structured format. Seidler-de Alwis and Hartmann (2008) stated that compared to explicit knowledge research, research on tacit knowledge is relatively unexplored but that tacit knowledge can be a valuable source for a competitive advantage in business. Abidin et al. designed an automated tool to extract unstructured data from the web and convert it into XML format to be stored in Oracle databases. These data are structured, semi-structured, and unstructured. Unstructured data include multimedia files, documents, spreadsheets, news stories, e-mails, memorandums, reports, and web pages.

Abidin et al. (2010) pointed to existing algorithms for data classification that are based on different techniques such as k-nearest neighbor, naive Bayesian, and concept vector-based algorithm and presented a table to compare the different techniques. The prototype tool built by Abidin et al. is made of five layers: the user layer, the interface layer, the source layer, the XML layer, and the multimedia layer. Some of the techniques to find and store the data are to look at the HTML tags such as *src:* to determine if it is an image, video, or audio file.

Seidler-de Alwis and Hartmann (2008) examined tacit knowledge as a source of innovation. After describing the work of Nonaka (1991, 1994); Nonaka and Toyama (2003); and Nonaka and von Krogh (2009), Seidler-de Alwis and Hartmann stated that the first step in innovation is to make sure that relevant tacit knowledge is easily identified. They stated that this process implies a constant dialogue and brainstorming sessions because differences of opinion foster creativity. The current global competitive environment requires constant innovation to be able to compete; therefore, innovation is a must to gain a competitive advantage. This is achieved by facilitating the proper KM environment. "In recent literature, innovation is viewed in terms of the transfer of knowledge (e.g., Scarbrough, 2003)" (Seidler-de Alwis & Hartmann, 2008, p. 139). This need of innovation is extended beyond sharing between work groups and includes intra-groups and external knowledge sharing (KS). A great deal of innovation is implicit in tacit knowledge, which is difficult to transfer as individuals act based on intuition. Managers must facilitate the environment so that this knowledge, the source of competitive advantage, is shared and transmitted.

> Hall and Andriani (2002) argued that the major challenge of an organization should be the achievement of balance between the tacit knowledge developed by individuals and the explicit knowledge needed for effective communication and integration, which means to make the bulk of an organization's knowledge explicit and to render the company safe from employees walking away with their personal knowledge. (Seidler-de Alwis & Hartmann, 2008, p. 141)

Mingrui and Yongjian (2010) studied tacit knowledge under the lens of distributed cognition and stated that tacit knowledge explicating is a form of distributed cognition. The study examined the factors that comprise distributed cognition to make tacit knowledge explicating more efficient. Agreeing with multiple authors, Mingrui and Yongjian stated that 42% of organization knowledge resides in people's heads in tacit knowledge form, highlighting the importance of tacit knowledge utilization. Mingrui and Yongjian stated that tacit knowledge cannot be explained by language but only by demonstration. Similar to the Nonaka and Toyama (2003) spiral of knowledge theory, Mingrui and Yongjian situated tacit knowledge at one end, explicit knowledge (with all its structure) at the other end of the business context, and a kind of knowledge that cannot be structured or explained but that people can perceive somewhere in the middle. Mingrui and Yongjian stated, "It's particularly important that distributed cognition stresses interaction among individuals and technique tools in a specific cognitive activity. So distributed cognition is a system made up of cognitive subjects and environment" (p. 274). They stated that human factors, artifacts, sharing, and communication play an important role in distributed cognition.

Mingrui and Yongjian (2010) noted two kinds of tacit knowledge explicating: in the first kind, individuals use artifacts to promote their knowledge; in the second kind, individuals communicate to others their tacit knowledge and others can explain it by using others' explicit knowledge and artifacts. The last type of knowledge explained by Mingrui and Yongjian is what this research is interested in by discovering factors that promote knowledge. Mingrui and Yongjian described six variables that enhance the explicating of knowledge: (a) within individuals, (b) among individuals, (c) among artifacts, (d) in culture, (e) in environment, and (f) through time. They concluded by saying, "Communication is a necessary condition of distributed cognition, and shared information is pooled information, which can make someone who has the best resource apply the information for others' benefits" (p. 277).

Subashini (2010) argued that explicit and tacit knowledge are complementary because both types are essential, basically agreeing with both Nonaka (1994) and Zack (1999a), and pointed out that current technologies capture the data flow but not the knowledge flow. KM is the cycle from tacit knowledge to explicit knowledge and back to tacit knowledge. Proper use of tacit knowledge leads to innovation because it is a source of competitive advantage. One popular way to extract tacit knowledge is through brainstorming. Subashini cited some methods of extracting tacit knowledge such as Quick Think method, card technique, Metaplan technique, and morphological analysis without providing any specifics on the methods or proposing any framework for extraction. Subashini agreed with similar work on artificial intelligence (AI) by Grundspenkis (2007).

Smith (2001) provided a good explanation of the role of tacit and explicit knowledge in the workplace. More and more knowledge is replacing raw materials, labor, and capital as corporate assets. The demand for leaders that help workers to convert their intellect in productivity is increasing. Workers should be able to do more work in less time using collective shared knowledge. The importance of managing tacit knowledge is increased by the fact that outsourcing, downsizing, and termination result in lost knowledge that is not written anywhere in corporate database documents. "Reportedly, 99% of the work people do is knowledge-based (Wah, 199b)" (p. 312). Managers must devise ways to extract the tacit knowledge that otherwise will be lost. Some companies such as Xerox make KM 90% a social process and 10% infrastructure. Of the different areas to manage knowledge, working environments to share knowledge is one of the important ones. According to Smith, few companies handle tacit and explicit knowledge effectively. She recommended using rewards as one of the motivations for this important corporate activity. Smith stated, "The philosopher Polanyi (1967) described tacit knowledge as knowing more than we can tell, or knowing how to do something without thinking about it, like [riding] a bicycle" (p. 314). This highlights the importance of placing knowledge in KS communities. Smith concluded by presenting some case studies in which several companies implemented tacit knowledge sharing environments such as Andersen Consulting, Ernst & Young, and Canon, suggesting that more scholars support the idea that tacit knowledge can be converted into explicit knowledge as suggested by one of the basic patterns of knowledge conversion (Nonaka, 1991).

Wagner and Sternberg (1987) posited that common sense and practical intelligence are more important than IQ for managerial success. Wagner and Sternberg developed a model that identified tacit knowledge in managers as the most important factor for their success. They identified three kinds of tacit knowledge: content, which is concerned with the management of others or one's tasks; context, which is concerned with short-term or long-term accomplishments; and orientation, which is concerned with ideal quality or practicality of judgment. Using an inventory and a complex rating system, they interviewed several managers from the *Fortune* 500 list, prestigious universities, and experienced managers. Although the study did not reveal its exact quantitative method, the results showed that tacit knowledge and not IQ was the key factor for managers' success. If unstructured data are that important, how are the data to be processed?

Processing Unstructured Data

Zhou et al. (2007) declared that with the development of text-mining technologies, text-driven BI will be a remarkable characteristic of BI. Eighty percent of corporate data is in some form of unstructured data and text-driven BI is to apply text-mining technology to business. There are two methods to process text: one is to convert it to structured data and then process it with existing tools, and the other is to develop new ontologies to process the text directly. The latter method is the preferred one because it is less complicated. To provide an example of text mining, technology could convert the following sentence, *Analysis International Inc. (TAI) has bought the American Medical Records Processing (AMRP) for more than $130 million*. A software packager named Visual Text's Corporate Analyzer would create *Action: buy, Company 1: (name) TAI,* *Company 2: (name) AMRP, Amount: (>) 130000000*. There are four steps to text-driven BI: (a) text data collecting, (b) text preprocessing, (c) feature extraction, and (d) special information extraction and mining.

Plejie, Vujnovie, and Penco (2008) stated that one of the issues with KM is key employees quitting the company and taking corporate memory with them. Therefore, it is paramount that tacit unstructured knowledge be converted into repositories for later use. Unstructured data, which comprises almost 80% of corporate data, is difficult to process but it can enhance company knowledge by the same percent. "Some companies are enjoying especially high market caps because [of] their ability to be more effective in exploiting their most important asset: intellectual capital, which is to say, knowledge" (Plejie et al., 2008, p. 924). Plejie et al. also said that the value of business increasingly lies in intangibles: patents, software, research programs, ideas, expertise, and so on. If unstructured data can be converted into metadata, then this can be processed the same way as regular, structured data are.

Other authors are extracting data directly from the web. Abidin et al. (2010) designed an automated tool to extract unstructured data from the web and convert it into XML format to be stored in Oracle databases. These data are structured, semi-structured, and unstructured. Unstructured data includes multimedia files, documents, spreadsheets, news stories, e-mails, memorandums, reports, and web pages. Davies, Duke, Kings, and Mladenic (2005) authored a paper about a new trend in KM: semantic web-based knowledge. This technology, integrated with human language and knowledge discovery, provides a glimpse of what will be the next generation of KM. Semantic web and different projects around the world are helping to extract BI from otherwise unstructured data such as e-mails, memos, and so on. Davies et al. described different tools that are based on semantic web technology. Some of the trends are desktop search, categorization, integrated search, seamless search, personalized search, beyond search, visualization, and device independence.

Du, Li, and King (2009) proposed a method to semi-automate the extraction of ontologies from HTML. Their method, named OntoSpider, is aimed at facilitating the processing of the immense amount of data located on the web. "Ontology engineering involves various tasks, such as editing, evolving, and versioning, mapping, alignment, merging, and reusing, and extraction" (p. 320). The authors described various methods commonly used to accomplish the same tasks and then provided an overview of their method. Their method has the limitation that it cannot process multimedia and requires a website to be ontology-directed. For example, a university website is organized around the ontology of university, admissions, academics, and campus life. The expectation is that other like-business sites have a similar structure. The process consists of six phases: preparation, transformation, clustering, recognition, refinement, and revision. The entire process consists of a series of parsing procedures and is written in Java and Xerces2 Java parser with MySQL. The authors concluded by presenting various output screens from the tool.

Fan, Wallace, Rich, and Zhang (2006) presented a useful review of current text mining technologies. Fan et al. described the successful implementation of a text-mining project from the Dow Chemical Company; Dow added about 80 years of research information from Union Carbide after they merged with them. Fan et al. described the difficulty of processing unstructured data due to the current computers limitations when managing human linguistics. The current technologies to process unstructured data are information extraction, topic tracking, summarization, categorization, clustering, concept linkage, information visualization, and question answering. Three useful tables were presented in the study: a list of vendors with product names and web pages; a list of text-mining technology features by vendor name; and a list of practical applications in medicine, business, government, and education. They concluded with a list of steps to follow when implementing text mining: define goals, perform a return on investment study, talk to vendors and clients to inquire about experience, integrate text mining with existing information technology (IT) structure, and hire and train the right IT professionals. This was a short and concise paper but with valuable pointers for the inquirer.

Godbole and Roy (2008) developed a tool called Customer Satisfaction Analysis (I-TACS), a text-mining solution to help to improve customer satisfaction in service companies. They stated that service research is an emergent area in computer science and IT. Customer relation managing applications help with the context of KM and analytics, and with the continuous growing of unstructured text documents, extracting meaning information from text is becoming of paramount importance. Customer feedback is an important piece of service business, which depends on customer feedback for continuous improvement. Godbole and Roy cited Gartner, who estimated that 80% of all enterprise data is unstructured and that these data are growing. As stated in the voice of the customer in the form of e-mails, feedback, surveys, and text messages, service requests are an important part of improving customer satisfaction—therefore, the importance of extracting exact feedback for continuous business improvement. In today's business environment, the voice of the customer is analyzed manually and for a typical e-commerce client, the feedback amounts to 40,000 to 50,000 comments a month. A quality analyst examines a sample of customer satisfaction comments and assigns reason codes (reasons for dissatisfaction) to cases. This is very costly and inaccurate as the reports show the *what* but not the *why* and, additionally, only a small fraction of the total amount of feedback is analyzed. Automated computer text mining offers great accuracy and volume and is capable of producing graphics with drill-down capabilities. Godbole and Roy concluded by saying that I-TACS "can be used not only for verbatim analysis but [also] for analyzing any kind of textual data generated in contact centers and customer[-]facing service arms of companies" (p. 448).

Stating that knowledge work is difficult to measure, Bosch-Sijtsema et al. (2009) created a framework for knowledge work. Bosch-Sijtsema et al. added that previous research has failed to research knowledge work in distributed geographical areas. Bosch-Sijtsema et al. also reported that several authors acknowledged the difficulty of measuring productivity but that there is no single acknowledged way to measure or improve knowledge worker productivity. Knowledge workers are distinguished by the nature of their work, which is very unstructured and often deals with new technologies. After reviewing different authors in the literature, they presented the framework to measure knowledge worker productivity in distributed teams: "time spent by knowledge workers in different work modes and on different task; team structure and composition; team processes; physical, virtual and social workspaces; and organizational context" (Bosch-Sijtsema et al., 2009, p. 542). This was only a suggested framework and even though they mentioned individual productivity, they failed to point out the specifics of the measures.

Bose (2008) authored an article that described competitive intelligence (CI) and the common tools to process it. Bose differentiated CI from BI: "Sometimes CI is confused with business intelligence (BI). The difference between BI and CI is that BI is internal intelligence about and within one's own company, whereas CI is external intelligence about the firm's competitors" (p. 511). CI analyzes the capabilities of competitors to gain a competitive advantage. Bose continued by saying that a recent *Fortune* 500 survey showed that 55% of companies make use of competitive information in composing their business strategy. Bose stated that CI drives strategic decision making and market leadership. Bose then presented several tools to help with the collection and analysis phase of CI such as TexAnalysis, SharePoint, and Intelligence from Brimstone and Knowledge Works. The tools were also classified in (a) active collection tools, such as web search engines; (b) passive collection tools, such as software agents; and (c) analysis tools, such as clustering tools and concept linkage tools. Bose then provided an example of successful use of concept linkage tools in the form of a professor at the University of Chicago who was able to identify magnesium deficiency as a contributor to migraines by using these CI tools. To conclude his article Bose stated:

> The ability to remain cognizant of the competitors' likely strategies and moves, so as to prepare for countermoves to sustain or gain competitive advantage is what CI is to an organization. The ability to produce and use CI will become a necessity in the near future for most organizations. (p. 526)

Bara et al. (2009) proposed a model for BI development. Bara et al. started their article by stating that despite the substantial investment in BI, there is a complete absence of methods to measure its business value. A concise definition of BI is that it "utilizes a substantial amount of collected data during the daily operational processes, and transforms the data into information and knowledge to avoid the supposition and ignorance of the enterprises" (p. 100). BI is differentiated from enterprise resource planning systems in that enterprise resource planning are transaction-processing and weak on analytics and BI provides managers with the ability to extract information from different sources in a customized view summarized into meaningful business information.

Popovic, Turk, and Jaklic (2010) proposed a conceptual model of business value of BI systems. Popovic et al. argued that the measurement of the value of BI is not carried out due to the lack of measurement methods and that their model helps corporations evaluate the value of investing in BI systems. Further proof of the need to create standards and methods to measure the value of BI is the different definitions of BI. Vendors define the concept to suit their product. One of the many definitions is

> Applications, platforms, tools, and technologies that support the process of exploring business data, data relationships, and trends. BI provides an executive with timely and accurate information to better understand his or her business and to make more informed, real-time business decisions (Raisinghani, 2004). (Popovic et al., 2010, p. 7)

Ranjan (2008) authored a paper with the intent to find the justifications for corporations to include BI in their business. Ranjan emphasized that the quality and timelines of an organization's BI can mean not only the difference between profit and loss but also the difference between survival and bankruptcy, and that the purpose of BI is to transform a business from reactive to proactive. BI facilitates many activities, such as multidimensional analysis, data mining, forecasting business analysis, balanced scorecard preparation, and KM. Examples of BI decision-support databases include data warehouses, data marts, exploration warehouses, data mining databases, and web warehouses. Traditionally, businesses start collecting data not with BI in mind but because the technology allows it to be collected.

> According to Adelman et al. (2002), BI is a term that encompasses a broad range of analytical software and solutions for gathering, consolidating, analyzing and providing access to information in a way that is supposed to let an enterprise's users make better business decisions. (Ranjan, 2008, p. 463)

A successful BI ties IT with business to create a competitive advantage. Some of the components of BI are data warehouses, data sources, data marts, and query and reporting tools. Ranjan (2008) explained that the growing potential of BI is because systems that process BI are already in place. He added that corporations lack road maps and have no tools to implement BI, and concluded by providing guidelines to implement BI: (a) requirements from firm perspective such as reason for embracing a centralized, managed approach to BI; (b) details of users and corporate standards, such as a detail of a standard for BI tools: education and support; (c) details of databases, tools, and vendors such as details of BI database architectures, data warehouse, data marts, federated data access, online analytical processing (OLAP), and others; and (d) other requirements such as details of security measures, user authentication services, database maintenance, backup and recovery procedure, and disaster recovery procedure.

Sahay and Ranjan (2008) presented the concept of real-time BI in supply chain analytics. Today's business enterprises must analyze accurate and timely information, and supply chain corporations are no different from any other business. The nature of business is becoming more dynamic. Sahay and Ranjan argued that "in order to support firms that are service-oriented and desperately seeking customer loyalty and retentions, it is necessary to revisit BI concept that integrates and consolidates information in an organization" (p. 29). The goal of supply chain analytics is to extract meaningful information for decision makers. To configure supply chain functions, data across the supply chain is crunched, numbers are analyzed, and information is generated that may aid in important decisions to be made about prices and customer expectations. This is why real-time BI in supply chain analytics is needed.

The activities of extracting and processing data are not sufficient because the quality of the data may have been compromised and contaminated. The next section describes several techniques and attempts to process and clean data for more accurate analysis.

Data Mining and Consistency

Blake and Mangiameli (2009) conducted a case study to prove their hypothesis that data consistency is crucial in either form: structured or unstructured. Blake and Mangiameli started their article by defining the importance of data quality (DQ) and pointed out that the Association for Computing Machinery launched the *Journal of Data and Information Quality*, further qualifying this importance. The case study, which was not completed at the time of the publication of their article, tested two hypotheses: (a) The level of consistency found in structured data has a significant and negative impact on the outcomes of data mining, and (b) The level of consistency in unstructured data has a significant and negative impact on the outcomes of data mining. For the case study, they observed Do-Tel, a telecommunications company that—similar to other telecommunications company—values customer retention as one of the most important aspects of their business. Customers that leave the company are costly and difficult to replace. The relevance of this study is the metric that Blake and Mangiameli were developing to measure DQ. The research was still developing at the time of publication of their article but its preliminary results showed consistency with their theory.

Pintar, Vranic, and Skocir (2007) proposed an integrated data repository to improve the extracting transformation and loading (ETL) process in data mining. Currently, ETL is processed with historic data that was previously collected. Pintar et al. tried to switch the focus from analyzing past results to improved collection of higher quality data in the future. They acknowledged that the market offers all kind of tools to analyze data previously mined and that the problem is to have the right data with enough quality. The problem with ETL resides in transforming large distributed environments with different hardware and software and with no communication among members. This can occur in an academic environment or a business environment. Their proposal was to create an integrated data repository placed between the data warehouse and the distributed environment, and to collect data in real time as it is processed without interfering with normal business operations. Pintar et al. proposed an intelligent data source system that will know which information to collect; when, where, and how to collect it; and in what manner it should be sent to the appropriate storage. Their system uses XML with Extensible Stylesheet Language Transformations to store the logic and a Java loader module. Their proposal is scalable, secure, easy to use, and implementable in the current context of any business.

After providing a brief definition of data mining (DM) and BI, Wang and Wang (2008) stated that the key to succeed with DM implementation is to create collaboration and sharing between the front-end users and technology experts responsible for DM implementation. DM goes to data to find previously unknown relationships between the different pieces of data. BI is a group of elements and technologies designed to create effective business decisions. KM is the creation and application of knowledge to enhance performance. KM's main concern is with the human knowledge and not data, although knowledge can be extracted from data. What differentiates KM from BI is the explicit and tacit portion of knowledge that BI fails to address. Wang and Wang created a KS system for DM collaboration consisting of the following subjects: tasks, data, hypothesis, DM result, action, action outcome, internalization, and DM planning. Wang and Wang proposed a blog system to share between the technical phase and business phase of the model. The model was successfully tested in a very well-known supermarket study that showed that consumers that purchase beer are likely to purchase diapers at the same time. This test was conducted with master of business administration students who interacted in a BI, DM, and KM combination through collective sharing.

Kobielus (2008), in his article entitled, "Quality Really Is Job No. 1," pointed out the importance of DQ when combining customer relationship marketing applications with services-oriented architecture. Low-quality data, he said, undermines the key performance indicators crucial to performance management and this indeed compromises the BI corporate strategy. Some major vendors have been acquiring small companies that specialized in DQ capabilities. SAP acquired Business Objects and IBM has acquired Cognos. These companies compete with pure DQ vendors such as Microsoft, BEA Software, Progress Software, Software AG, and Tibco. Kobielus stated that prior to acquiring Business Objects, SAP was offering data-profiling and data-cleansing features but users were more interested in buying standalone DQ offerings from Business Objects. The literature suggests that companies are not yet ready to expend large amounts of money venturing in expensive propositions such as data profiling and only want to clean up data messes after they have been discovered. It seems that any initiative that can produce effective but inexpensive ways to clean and prevent dirty data is a good area for research.

Luebbers, Grimmer, and Jarke (2003) proposed a data audit tool and test generator that not only induce structure in large databases but also create simulated dirty data to test the algorithm. Data scrubbing tools fix issues such as assigning zip codes to cities, but data audit tools such as the one proposed by Luebbers et al. use data-mining algorithms that induce structure into databases. These tools can be tested if they can be measured. The tool proposed by Luebbers et al. generates data and then tests the efficiency of the tool. Luebbers et al. described DQ as having the following dimensions: accuracy or correctness, completeness, consistency, actuality, and relevance. They tested the tool and found that 30% of the errors could be identified. This low rate is explained by the fact that induced data corruption corrupts data regardless of whether the data are part of a generated rule. The assumption is that the data, although they can be dirty, are generated as part of a structure. Nevertheless, they concluded that data audit tools must be accompanied by data-scrubbing tools for holistic DQ management.

Reid and Catterall (2005) presented a comprehensive analysis of DQ issues after the implementation of a customer relationship marketing system. In essence, their strategy is to deal with DQ issues beforehand as otherwise they will only increase in cost to fix. Not fixing DQ issues prevents the benefits of good data in improving operational costs, customer satisfaction, and effective decision making. Reid and Catterall presented the results of a case study from a European telecommunications company and pointed out that they caught DQ issues too late—after the customer relationship marketing system had been implemented. Reid and Catterall cited Gartner, who pointed out that 75% of companies engaged in customer relationship marketing initiatives suffer from some kind of DQ issues and that information quality is getting worse when related to data warehousing. Some of the causes are poor data entry, missing data, lack of standards, multiple databases scattered throughout different departments, and older systems with poorly documented and obsolete data. Reid and Catterall stated that DQ is something one designs in and not something one discovers. Three main issues were discovered by the study: (a) Data entry people were not trained regarding the importance of entering good-quality data into the system and were more concerned about speed and quantities of calls than quality; (b) Bad data entered by customers on the web permitted customers to create new accounts as there was no check in place to see if accounts for them already existed; and (c) A poorly designed migration from the legacy system carried inconsistent data. Their recommendations were to (a) identify the main sources of data, (b) undertake an initial DQ assessment, (c) utilize the information to drive an immediate cleanup, and (d) create an ongoing DQ program including a data governor (or administrator) for ongoing consistency.

In an article published in the *Information Management Journal*, Swartz (2007) cited Gartner, which said that 25% of critical data in the *Fortune* 1000 companies is inconsistent. Companies overlook that they have dirty data and underestimate its size and impact on their business. Such data creates excessive expenses in the form of missed sales opportunities and mail-outs, and affects even some internal functions such as budgeting, manufacturing, and distribution. Some of the issues addressed by Swartz were existence of the data, validity, consistency, integrity, accuracy, and relevance.

Watson (2007) started his article by making an interesting distinction between data quality and data believability. Although they are related, and data must have quality, consistency, and completeness, data should also be suitable for business. Data believability relates to whether data are regarded as being true, real, and credible. According to Watson, perception is very important. Watson made several recommendations on how to improve data believability so that users feel comfortable that the data with which they are being provided is credible. Some of the factors that make data believable are the source of the data, the timeliness of the data, how well the data matches users' preconceptions, and users' understanding of and confidence in how the data are processed. Watson made three recommendations for making data more believable. At the application level, users must be allowed to see were the data is coming from if they so desire. At the metadata level, users may need to have more access to understand how the data is processed (this has traditionally been reserved for IT people only). At the governance level, user must be deeply involving in the operation side of the business, meet regularly, and allowed to make recommendations to solve common problems. Watson's recommendations are intrinsically related to DQ. Users must see the accuracy and consistency of data in order for them to accept it as credible. This assumes that the data are not only clean and credible but also processed in a way that makes sense for the business.

Future Developments and Knowledge Management

Artificial intelligence (AI) is rapidly spreading to areas other than the pure scientific field. Choy, Tan, and Chan (2007) developed a technology that utilizes OLAP, artificial neural networks (ANNs), and case-based reasoning (CBR) to predict future customer demands and allocate suitable suppliers during the order fulfillment process. Metaxiotis, Ergazakis, Samouilidis, and Psarras (2003) concurred that one of the key components of KM is AI and conducted a study to evaluate the potential and limitations of AI when used as a tool to help the KM process.

Choy et al.'s (2007) method, named *supplier knowledge management system*, was successfully tested in a case study of Farnell Newark-InOne Ltd. in Shanghai, China. The company is a market leader in provision and distribution of electronic, electrical, and mechanical components; health and safety products; and associated tools and services operating in over 22 countries and distributing to over 160 countries worldwide. Choy et al. provided explanations as to why the management of customer requirements on stock-keeping units (SKUs) has become a critical activity and that previous systems only took in consideration attributes such as price, quality, reliability, and delivery status when selecting suppliers but not customer demand. By creating this hybrid system of OLAP, ANNs, and CBR, Choy et al. created an integral system that produced excellent results when compared to traditional methods. OLAP typically consists of descriptive data and quantitative values to build a data cube in which applications can drill up and down for general or specific levels of detail. ANNs are information-processing paradigms designed in the way that the human brain processes information and can learn from its mistakes.

Finally, CBR is a problem-solving technique in which past mistakes are used to find solutions to particular problems.

The supplier knowledge management system leverages XML technologies in the front end to facilitate the integration of SKU raw data. The front-end system is linked with various databases to acquired SKU-related data such as past records of ordering points, quantities (size/volume), and frequency. The system is made of the customer demand advisory module (ONCAM) and the task allocation module (CTAM). The ONCAM module converts SKU raw data originally stored in the system into information that represents the future demand of the customers. Once the ordering point and quantity are determined, they are compared with the inventory to determine if ordering is necessary to move to the next module. The CTAM module relates the advisory parameters with the supplier capability and by using a series of algorithm that includes technical capabilities, quality level, and delivery status, a weighted score is created and the right supplier chosen. Choy et al.'s (2007) system produced impressive results after a 12-month implementation. This included a deviation in forecasting replenishment from 25% to 10%.

Metaxiotis et al. (2003) stated that the complexity of business globalization is making KM more necessary than ever. Some of the benefits of KM are the capability of employees to deliver work products for which they are responsible, the effectiveness of interpersonal work, and the degree to which innovations are captured, communicated, and applied. A definition of KM is that it "is concerned with the exploitation and development of the knowledge assets of an organization with view to furthering the organization's objectives. The knowledge to be managed includes both explicit, documented knowledge, and tacit, subjective knowledge" (p. 217). Most of the emphasis on KM has been on framework approaches and not on system and technologies. AI plays a major role in KM in three specific technologies: expert systems, ANNs, and intelligent agents.

Expert systems are computer systems with stored knowledge in the forms of facts, heuristics, and relationships gathered from the minds of experts. To operate, they use if-then rules and are able to achieve expert levels of problem solving. One of the drawbacks of these systems is that they cannot handle vague and unusual questions.

ANNs are based on how the biological nervous system works. In these systems, there is no explicit knowledge base; their function is to transform input data to outputs based on previous learning experiences. The advantage of this technology is that it can operate with incomplete data and demonstrate apparent intuition. One of the drawbacks is that its input must be presented in numeric form.

Intelligent agents are agents capable of autonomous action. They are owners of a great amount of knowledge, professional experience, and beliefs. They are like an object with a head and can reason by themselves. They can be used to assist in processing knowledge that at the end will create knowledge as well. The study of intelligent agents is the most important field of distributed artificial intelligence and this technology seems promising for KM.

Metaxiotis et al. (2003) concluded by saying that organizations must manage knowledge-related activities to become competitive, and that KM seems to be able to provide to decision-makers with an enhanced quality of support. According to Metaxiotis et al., companies such as General Electric, McKinsey & Company, Xerox, Microsoft, Ernst & Young, and Accenture have already identified the benefits of KM, and the World Bank spends 4% of its administrative budget on KM. The authors believed that the creation of hybrid systems (combining expert systems and ANN) can offer enhanced benefits.

Yoon, Broome, Singh, and Guimaraes (2005) compared expert systems against agent technologies. The article made an evaluation with the help of a case study in a reverse mortgage application. Yoon et al. started their article by emphasizing the importance of expert systems on AI and stated that many expert systems and agent techniques are applied to the new and emerging field of KM. A simple definition of KM is the process of capturing collective expertise and distributing it in a manner that produces a payoff. Expert systems are useful in accomplishing the knowledge-sharing task but have several shortcomings: (a) they are typically brittle, as they do not deal well with bending the rules; (b) they are typically isolated, as they do not interact with other systems; (c) as the system grows, they cannot avoid inconsistencies; and (d) they use humans as aids, and as users become expert in the system, there is no mechanism to completely automate knowledge-sharing tasks.

Intelligent agents, on the other hand, possess the following properties: (a) autonomy, as they can operate without direct intervention; (b) sociability, as they can cooperate with other agents or humans; and (c) adaptability, as they can modify their own behavior and adapt to changes. Intelligent agent technology is becoming very attractive in the new distributed environment, in which applications do not run solely on a central location and are very adaptable to dynamic environments. Expert systems rely on users to initiate reasoning but intelligent agents are autonomous. Expert systems have a fixed set of rules that defines their reasoning whereas intelligent agents interact with the environment to adapt to new conditions and are modular, allowing them to be reused in other environments. Some of the weaknesses of intelligent agent technology can be attributed to lack of maturity, because intelligent agents must have new software techniques specifically tailored to intelligent agent technology systems. Intelligent agents have difficulty decomposing goals and tasks, and new methods must be developed to lock intelligent agents out of environments in which they are not welcome.

Yoon et al. (2005) concluded their study by presenting the results of a case study in a reverse mortgage application and summarized the results in a table that shows intelligent agent technology as a winner in the new emergent world of versatility and dynamism.

To be competitive, companies are shifting to efficient KM initiatives. Global markets are characterized by a constant flow of information exchange between companies. To cope with the challenge, Dioşteanu and Cotfas (2009) proposed new ontologies after developing an agent-based KM framework that uses Java Agent Development Network (JADE) and Web Service Description Language (WSDL). This framework facilitates enterprise interoperability. The proposal was to use intelligent agents that are commonly used in manufacturing companies combined with semantic web technologies, in which enterprise web services are dynamically discovered. Along with JADE and WSDL, the framework would use a gateway that serves intelligent agents and services called *web service integration gateway*. This solution uses the following technologies: semantic web and ontology, multi-agent systems, and semantic web services. Dioşteanu and Cotfas concluded by saying that this framework can be applied to manage any kind of service and not just a supply chain management service, on which it was tested.

Hendler (2005) wrote a paper exploring the similarities between the semantic web and traditional AI knowledge representation. He referred to Tim Berners-Lee, the inventor of the World Wide Web, and Ora Lassila from Nokia, who posited that for knowledge representation to realize its full potential, it must be linked to a single global system. Hendler described three scenarios that no longer belong to the imagination but that the current state of the semantic web technologies made possible: (a) document metadata knowledge bases, with which tools can be created to tell who created documents, when, and with what application; (b) semantic annotation of non-text media, in which, for example, photos can be tagged with knowledge. By simply looking at a picture, one can find information related to a specific NASA mission, including commander names and more; and (c) large-scale knowledge infrastructure, in which the knowledge base has information about millions of people, places, things, transactions, and processes, and everything can be accessed by a computer anywhere and in which this data are indexed against thousands of ontologies. Hendler concluded by describing briefly the three technologies that will play an important role in extracting knowledge with the help of AI: resource description framework, which describes semantic web language; web ontology language, which describes ontologies; and resource description framework schema, which creates schemas from resource description frameworks.

Lau, Lee, and Ho (2005) demonstrated the usefulness of text mining in the hotel industry. Lau et al. stated that the large amount of free text-based information on the web can be used to extract information to help the hospitality industry get competitive, especially with variant information such as room rates and customer demographics and attitudes. One of the key issues regarding text mining is the construction of dictionaries that can be use to discover meaning from the large amounts of text that can be extracted with existing text-mining engines. Lau et al. provided a table with the most common commercially available text-mining tools, such as IBM Intelligent Miner for Text, SAS Text Mine, and SPSS LexiQuest Mine, among others. Text mining can be used in tourism, financial services, insurance, and manufacturing. Text-mining techniques comprise four steps: (a) definition of the mining concepts, (b) data collection, (c) dictionary construction, and (d) data analysis. To conclude, Lau et al. showed statistics of three case studies in which they used a commercially available text-mining tool to extract hotel profile information, room prices, and travel-related newsgroups.

Lausen et al. (2005) conducted a survey to evaluate web portals currently using semantic web technology. As stated in several research papers, semantic web allows information on the web to be understood by computers, helping humans with the process of extracting knowledge and meaning from the current web. Therefore, semantic web can improve information sharing at portals. Before starting the survey, Lausen et al. defined a *portal* as a site that (a) collects information for a group of users that have common interests, (b) is for a community to share and exchange information, and (c) is based on semantic web technologies. Four sites were surveyed as they complied with the criteria: two academic portals and two commercial portal technology infrastructures. The evaluation schema was as follows: Information access from the user's perspective, information processing features of the portal, and grounding technologies. Regarding grounding technologies, most of the portals used the traditional Tier 3 technology with a

Java Servlet user interface. Regarding information processing, all the portals were based on HTML forms. Regarding information access, the academic portals were used for document management and dissemination for research projects, and the commercial sites were oriented for development of web portals (e.g., a conference portal, or a portal for KM).

Lee, Upadhyaya, Rao, and Sharman (2005) addressed the security aspects of processing data with semantic web technologies. Lee et al. authored a paper describing future security protocols for the semantic web. They started their article by describing the importance of knowledge in decision making and cited Tim Berners-Lee's vision of the semantic web, in which flow is enhanced by machine-processable metadata. Semantic KM systems can capture a more articulated organizational knowledge that can be transferred more easily but must be kept secure. Traditional encryption and digital security is not enough and new technologies are in development, such as XLM encryption, SML digital signatures, XML key management specification, Extensible Access Control Markup Language, web services policy, Security Assertions Markup Language, and web services security.

Srivastava and Cooley (2003) proposed the concept of web business intelligence (WBI) as an emerging class of software with the main task of leveraging the enormous amount of data stored on static web pages. Srivastava and Cooley presented WBI architecture as well as a survey of the technologies behind it. WBI goes beyond the task of just extracting data from the web to include as its goal to manipulate and convert it into actionable knowledge. The architecture of WBI is made of (a) content acquisition, coming from communication, financial, and travel sites as some of its sources; (b) a profile database that is used for security and to store information from individual users; and (c) knowledge creation. Its information retrieval has many components: (a) a manual gather, in which individuals manually download and process data; (b) crawlers, which are automatic software agents that process and filter data for presentation; and (c) queries. One of the difficult parts of data extraction for WBI is the processing of unstructured and semi-structured data, which requires special techniques. Srivastava and Cooley concluded with a case study in which they showed a singular web page compiled from multiple sources, including separate accounts from different banks. Such pages are valuable for financial analysis but questions arise regarding privacy of data. This is assuredly a research area in progress.

Voth (2005) described several tools that use AI to extract intelligence from unstructured data, including e-mails, blogs, business records, manufacturer warranties, and other kind of text. Information is hiding in this data and with the help of AI, corporations can enhance their business, extracting information such as legal incompliance, cases and causes of problems, and market trends. These tools include (a) detecting workflows, which looks at e-mails generated by e-commerce transactions to discover workflows. This technology may help companies that facilitate shopping online by studying the patterns of online shopping; (b) OutBoxer, which investigates e-mails before they are released,“ which helps companies avoid circulation of inappropriate, offensive, and potentially illegal e-mails” (Voth, 2005, p. 5); (c) Attensity, which monitors mail to help with manufacturer warranties and helps companies understand how their products are behaving; (d) Brandpulse, which examines unstructured data for marketing purposes; and (e) Inxight, the "technology [of which] is used in counterterrorism intelligence efforts" (Voth, 2005, p. 6).

Conceptual Frameworks

Delen and Al-Hawamdeh (2009) presented a conceptual architecture to handle the complex issue of sharing and managing tacit knowledge. Managing explicit knowledge is easier than managing tacit knowledge, in which the complexity resides. Tacit knowledge is hidden behind behavior, skills, competencies, and experiences (tacit actionable knowledge) and articulated knowledge (implicit knowledge), which resides in individual thoughts and language use. Explicit knowledge resides inside computers in codified form and by nature has a clear organization. Tacit articulated knowledge can be transferred using e-mail, chat rooms, and discussion boards but actionable knowledge is more difficult, although not impossible, to transfer using video and multimedia technologies. The framework proposed to manage this combination was made of two modules: knowledge creation module, in which web crawlers, data/text mining tools, and manual entries feed the knowledge repository in the form of data nuggets; and the knowledge utilization module, which manages the knowledge manipulation using human experts and intelligence brokers. Intelligence brokers help user clients to make decisions and obtain the desired piece of knowledge. This subsystem makes the determination to contact a human expert after a series of decisions. Delen and Al-Hawamdeh succeeded in presenting this framework but did not show specific tools or methodologies to manipulate the unstructured data, which is the most difficult part of KM.

Arguing that organizations must have KM support to obtain optimal performance, Douglas (2009) examined systems of KM in large organizations and tested a prototype implementation of a new framework in which knowledge is packaged into objects. The prototype was tested using an action research approach and preliminary tests showed positive results. Douglas added that distributed cross-cultural corporations have an increasing need to communicate and share knowledge, and that there is a growing recognition of the knowledge transfer (KT) in social spaces. Douglas discussed that transmitting tacit knowledge is difficult to convey in structured documents and that the new trend of research is personal information management, in which employees collect and share their personal knowledge; therefore, a new design is needed to extract and communicate that knowledge.

Trying to bridge the gap between KM and AI, Grundspenkis (2007) proposed an agent-based framework, a rather conceptual model, for the purpose of building the structures for the new paradigm of KM and AI. The framework's two models, intelligent enterprise memory and intelligent organization KM system, are hoped to help researchers investigate directions in the development of intelligent systems for organizations. In his introduction, Grundspenkis argued that organizations possess large quantities of information but are poor in knowledge. Knowledge about an organization and business in general is a vital part of organizations' ability to react quickly to business demands and make decisions to accomplish strategic, tactical, and operational goals. There is currently a gap between the two schools of KM research—the school of the people track and the school of the IT track, in which researchers and practitioners are trying to construct KM-based systems with AI, groupware, and so on. Grundspenkis's idea was to treat knowledge as objects that can be managed as systems. Tacit knowledge can be elicited from humans and converted into explicit knowledge. The idea behind these conceptual models is to see organizations as systems with different objects and relationships between them. As KM continues its evolution, systems with different components and tools can evolve around those objects.

Knowledge that is considered organizational intellectual capital falls into three categories: individual, group, and enterprise. In KM, context models are used to provide knowledge services that capture knowledge and distribute it for general use. Grundspenkis (2007) presented his first model, named *intelligent enterprise memory*, which can support individual as well as enterprise-wide knowledge. The model has seven layers: knowledge source, knowledge acquisition, knowledge formalization, knowledge representation, knowledge processing, knowledge application, and knowledge user. Grundspenkis's second model expands on the concept of seeing organizational members such as decision-makers, researchers, secretaries, managers, advisers, and so on as a system supported by multi-agent subsystems for them to fulfill their activities. Grundspenkis posited that the conceptual model, named *organization knowledge management system*, functions like a human brain to fulfill the following functions: knowledge acquisition through sensors, knowledge formalization, representation and storage in the knowledge space, knowledge inference, knowledge sharing, and knowledge utilization. The model has three layers: an "engine room," (p. 455) a structural layer, and a "cooperation platform" (p. 455).

Mehta (2008) evaluated the framework of three software companies that implemented KM programs in a global context. The research provided evidence of strategic and cultural issues that influenced the successful implementation of KM. The implementation of KM is a complex one, and Mehta's research evaluated that complexity. For that purpose, a three-stage KM implementation was developed and three global companies with different goals were analyzed. These sequential capabilities were identified as articulating the KM strategic intent, facilitating knowledge flows to enable innovation, and assessing KM value. Articulating the KM defines the strategy for each company, facilitating knowledge creates the environment for employee and company participation, and assessing KM value evaluates both internal and external value. Each company had a different goal. Company 1 wanted to reduce software costs, Company 2 wanted to improve the quality of their software, and Company 3 wanted to enable virtual team works. The research resulted in four groups: (a) the content management group; (b) the evangelist group, which conducted events to promote the initiative; (c) the technology group, which developed the technologic structure; and (d) the process expert group. The important outcome of this research is that for each company studied, it was proven that KM successfully helped them achieve their goal, suggesting that valuable information is hidden behind the large amounts of companies' structured and unstructured data and that the concept can be applied to services in general.

Antonova et al. (2011) developed a framework that combined KT and KS and was tested among Hungarian and Bulgarian managers with 357 respondents. The authors argued that information technologies that facilitate knowledge are still moderately used to increase business performance. Some managers do not believe that information communication technology can be used to acquire knowledge more easily. Antonova et al. stated that the majority of organizations encourage employees to acquire and apply existing knowledge but not to share what they know. Antonova et al. stated that this may explain why the majority of organizations uses e-mail for communication, which is perceived as a tool for personal communication.

Knowledge Management and Productivity

"Insurance firms are using knowledge management solutions to increase productivity, save on expenses and to improve call center performance" (Britt, 2009, p. 10). Britt (2009) described the productivity increase of a London-based insurance broker after the adoption of web-enabled software that allows agents, underwriters, customers, and brokers to generate compliant insurance documents and capture data. Britt stated that according to Leslie Doel, one of the managers of Croton Stokes Wilson Ltd., the company increased their productivity to 710 insurance policies in 2008 from 642 the previous year while using one less agent. Britt did not address possible other reasons for the increase. Additionally, the study did not measure individual productivity in a rigorous scientific manner.

Chatzoglou, Vanezis, and Christoforidis (2005) provided an overview of KM including concepts such as OLAP, decision support systems, and data warehousing. Chatzoglou et al. presented a rather optimistic view of KM from survey results of the Delphi Consulting Group from 1997. Chatzoglou et al. showed that KM produced an increase of 89% in decision making, 84% in responsiveness to customers, 82% in efficiency of people and operations, 73% in innovation, and 73% in product/services. What is not clear is if these results were only for Greece, where the article was written. Chatzoglou et al. continued raising the bar of KM, describing the adoption of KM as a central strategy by World Bank, an organization owned by many governments. The purpose of Chatzoglou et al.'s article was to promote a KM tool designed to manage bank loans. Chatzoglou et al. stated that despite KM success, Greek banks have implemented very little KM. The authors cited several reasons why Greek banks have failed to implement KM: (a) There are no skills in KM techniques; (b) The benefits of KM are not understood; (c) There is lack of appropriate technology for KM; (d) There is no commitment to KM from senior management; (e) There is no funding for KM; and (f) The culture does not encourage KS.

Studies about KM and productivity started 14 years ago. Davenport and Prusak (1998) stated that if knowing how to do things is what creates value for companies, then knowledge is what defines a company. Their views in 1998 foresaw the future as now knowledge is considered the most important asset in companies' gaining a competitive advantage and—without the limitation of physical assets—is an unlimited resource. Davenport and Prusak were very explicit by saying, "The only sustainable advantage a firm has [come] from what it collectively knows, how efficiently it uses what it knows, and how readily it acquires and uses new knowledge" (p. xxiv). Davenport and Prusak made an important distinction between information or knowledge and the technology that delivers it. They provided an example that just because individuals have telephones does not mean that they will engage in brilliant phone conversations. Knowledge "is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers" (Davenport & Prusak, 1998, p. 5).

Davenport and Prusak (1998) stated that knowledge offers speed as it allows the possessor to act quickly with efficiency. They went on to say that studies show that two-thirds of the information and knowledge that managers possess come from face-to-face meetings or phone conversations. This statistics is aligned with other authors' giving more importance to tacit knowledge than explicit knowledge (Delen & Al-Hawamdeh, 2009; Koskinen, 2004; Mahmood & Ali, 2011; Nonaka, 1991; Plejie et al., 2008; Seidler-de Alwis & Hartmann, 2008; Subashini, 2010). Knowledge provides a competitive advantage because by the time other companies are able to produce what a specific company produced first, the company has already moved ahead to an innovation—therefore, the importance of managing a bank of knowledge that ultimately generates revenue and profits.

Hou and Chien (2010) created a knowledge construct based on different literature reviews and operationalized the variables that were examined using linear regression. Their survey was conducted with master of business administration students at National Chiayi University and senior executives from major companies in Taiwan. The results were subject to very rigorous statistical testing with highly reliable Cronbach's alpha values. The main conclusion was that increasing dynamic capabilities results in better KM competence environment. The hypothesis that dynamic capabilities support business performance was partially validated. Mahmood and Ali (2011) investigated knowledge worker productivity in Pakistan. They operationalized several constructs based on the literature to predict productivity using structural equation modeling and found that organizational culture, rewards, and technology impact productivity in a positive way. Mahmood and Ali highlighted the importance of knowledge in socioeconomic development and emphasized the difficulty in increasing knowledge worker productivity. According to Mahmood and Ali, knowledge workers have the potential to solve problems in less time if new knowledge is extracted and used properly. Mahmood and Ali recommended that organizations take care of their knowledge workers as they signify real assets for corporations. KS is an important piece of this equation "Knowledge sharing: Knowledge is described as a mixture of experiences, principles, and related knowledge. It begins and resides in the 'minds of knower,' described as tacit knowledge" (Mahmood & Ali, 2011, p. 28). Mahmood and Ali also emphasized the use of technology to help the sharing process in surmounting three types of barriers: temporal distance, physical distance, and social distance. Mahmood and Ali's literature review included rewards and incentive as an importance part of motivating employees. Mahmood and Ali's also studied culture that reveals the personality of organizations, and stated that organizations should promote social interaction networks to incentive KS. By using factor analysis, Cronbach's alpha tests, a Kaiser-Meyer-Olkin test, a Bartlett's test, and model fit indexes to support the rigor of the model, Mahmood and Ali supported three of their hypotheses: KS, organizational culture, and technology and rewards contribute to knowledge worker productivity. Mahmood and Ali did not provide details on how productivity was increased.

Singh (2008) conducted a survey of an Indian software company to investigate the impact of leadership styles on KM productivity. The goals of a company can be maximized by the right utilization and management of knowledge assets. It has been a common understanding that leadership plays an important role in firms' productivity, and this should be no different in managing knowledge. "Knowledge management is a formal, directed process of determining what information a company has that could benefit others and then devising ways to making it easily available to all concerned (Liss, 1999)" (p. 4). Singh's survey was directed to the two types of knowledge that firms possess: explicit, which is easy to communicate because it is in the form of hard data or codified procedures; and tacit, which applies to knowledge possessed by individuals and can only be communicated through conversation, storytelling, and so on. Singh further explained the different leadership styles as follows: (a) directive, which is high on regulating behavior but low on nurturing behavior; (b) supportive style, which is high on both regulating and nurturing behavior; (c) consulting, which is low on regulating behavior but high on nurturing behavior; and (d) delegating, which is low on both regulating and nurturing behavior. The study proposed three hypotheses: (a) Gender significantly differs in leadership styles and the art of practicing KM in the workplace; (b) Leadership styles are significantly related to an organization's KM practices; and (c) Leadership styles significantly predict the art of KM practices in an organization. Using rigorous statistical methods and variable validation, Singh concluded that directive and supportive styles of leadership are negatively associated with the art of KM. Consulting and delegating styles were found to be positively related with managing knowledge, and the delegating mode of leadership was found to be significant in predicting, creating, and managing knowledge. The first hypothesis was rejected as no evidence was found.

Martins and Lopes dos Reis (2010) proposed a model to measure the value of human capital so that it can be identified as another asset of companies. They pointed out that the previous literature has not been able to operationalize human capital and divided the different lines of research into three branches: (a) using human capital as a base, (b) measuring how executives can make financial decisions based on the human capital measurement, and (c) developing technical models to measure human capital. Using the software "Individual Differences Scaling or Perpetual Data (INDSCAL) developed by

Dr. J. D. Carrol and Jih Jie Chang" (p. 5), Martins and Lopes dos Reis concluded that knowledge has an economic value, has become one of the most important factors in the nation's economy, and is the main ingredient of what is bought and sold today. Knowledge is the raw material that is worked with in the new economic order. Intellectual capital—much more than natural resources, machinery, or even financial capital—seems to assume more and more of a major role in corporate assets.

Sharing Knowledge

Antonova et al. (2011) developed a model of KS and KT to examine their impact on organizations. To test the model, Antonova et al. conducted a survey of middle managers in Hungary and Bulgaria. Chowdhury (2005) investigated the role of affect- and cognition-based trust in complex KS. He successfully validated three hypotheses: (a) There is a positive relationship between the level of affect-based trust and the extent of complex KS between two individuals; (b) There is a positive relationship between the level of cognition-based trust and the extent of complex KS between two individuals; and (c) The impact of cognition-based trust on complex KS does not change if affect-based trust is present and vice versa. By using rigorous statistical validation, he demonstrated with individuals working in pairs that cognition-based trust is more significant that affect-based trust, although both of them showed significant beta weights in which the common denominator was trust. Douglas (2009) stated that distributed cross-cultural corporations are creating an increasing need to communicate and share knowledge and that there is growing recognition of KT in social spaces. Inthout, Vrancken, and Schrijnen (2010) used previous literature to design a system that combines tacit and explicit knowledge with emphasis on sharing.

Sharing knowledge had been described as an important component of tacit knowledge "Knowledge sharing: Knowledge is described as a mixture of experiences, principles, and related knowledge. It begins and resides in the 'minds of knower,' described as tacit knowledge" (Mahmood & Ali, 2011, p. 28). Mahmood and Ali (2011) also emphasized the use of technology to help the sharing process surmount three types of barriers: temporal distance, physical distance, and social distance. Delen and Al-Hawamdeh (2009) emphasized the difficulty of processing tacit knowledge and stated that managing explicit knowledge is easier than managing tacit knowledge, in which the complexity resides. Tacit knowledge is hidden behind behavior, skills, competencies, and experiences (tacit actionable knowledge) and articulated knowledge (implicit knowledge), which resides in individual thoughts and language use. Explicit knowledge resides inside computers in codified form and by nature has a clear organization. Tacit articulated knowledge can be transferred using e-mail, chat rooms, and discussion boards but actionable knowledge is more difficult, although not impossible, to transfer using video and multimedia technologies.

Reychav and Weisberg (2009) investigated the impact of tacit and explicit knowledge sharing on employees' rewards, performance, and intention to leave. The findings showed that sharing explicit knowledge created a positive impact on the receipt of monetary rewards, a positive indirect effect on employees' performance, and a negative indirect and a positive direct effect on employees' intention to leave. On the contrary, tacit knowledge sharing showed a positive indirect effect on employees' intention to leave, a positive and direct effect on performance, and a positive and direct effect on the receipt of nonmonetary rewards. The authors created a conceptual framework using the exchange theory as the foundation. The findings were tested using structural equation modeling, factor analysis, and Cronbach's tests. A year later, Reychav and Weisberg (2010) stated that to manage knowledge effectively, companies must implement methods to encourage KS behaviors in two main ways. The first involves explicit knowledge and is related to the capability to help create, store, and use explicitly documented knowledge mainly by using IT. The second step relates to tacit knowledge sharing through exchanges that can help to turn KS intention into actual KS behavior. Their findings suggest that companies should provide the proper environment for employees to share their knowledge, especially tacit knowledge, which is difficult to exchange, but at the same time give due importance to explicit knowledge in the corporate context.

Trust has also been studied in the context of sharing knowledge. Affect-based trust is grounded in relationships that are based on mutual care; cognition-based trust is based on competence. Similar to Antonova et al. (2011), Holste and Fields (2010) demonstrated that both are important when workers are sharing knowledge, although affect-based trust has a significant greater effect. The implication of these findings indicates that KM must "include a finer grained view of the nature of the social networks impacting the knowledge transfer and management process" (Holste & Fields, 2010, p. 128). A great amount of investment in knowledge is based on explicit knowledge, which is knowledge that is easy retrievable in the form of documents, reports, white papers, catalogs, presentation, and so on. The other form of knowledge is tacit knowledge, which is not easily retrievable and takes the form of stories, personal strategies, and metaphors. Trust is important when sharing tacit knowledge, which is intrinsically ingrained inside individuals' minds. So far, the industry has been so preoccupied in facilitating explicit knowledge that it has neglected the important and challenging task of facilitating tacit knowledge.

Holste and Fields's (2010) study supported the three following hypotheses: (a) Both affect-based and cognition-based trust in a coworker have positive relationships with the willingness to share and use tacit knowledge; (b) Affect-based trust of a coworker has a larger influence than cognition-based trust on an employee's willingness to share tacit knowledge; and (c) Cognition-based trust of a coworker has a larger influence than affect-based trust on an employee's willingness to use tacit knowledge. The findings suggest that companies should facilitate a sharing environment based on trust and knowledge because both are intimately related.

Kane, Robinson-Combre, and Berge (2010) emphasized the importance of social networks to enhance both e-learning and KM, and argued that both have a training and education application in the workplace. Kane et al. stated that there were 3.9 million e-learning students in 2007 and that the use of Web 2.0 is facilitating e-learning growth. Web 2.0 allows users to interact with an application as opposed to only obtaining a static web page, which is provided by Web 1.0. Social learning is learning within a group and social networks can facilitate this learning with different Web 2.0 tools. *Social networking* is defined as

> Web-based services that allow individuals to [a] construct a public or semi-public profile within a bounded system, [b] articulate a list of other users with whom they share a connection, and [c] view and traverse their list of connections and those made by others within the system. (p. 64)

Lopes (2008) tried to show evidence of an electronic knowledge management culture in which knowledge management systems store, organize, capture, disseminate, and create organizational knowledge. These systems are built on four repositories: people, systems, process, and culture. This study is aligned with Lopes regarding the view that sharing is an important piece of KM for support companies. By using nonparametric tests on 144 companies with a 0.05 significance level and an expected error of 3%, Lopes demonstrated that companies are moving into a KM culture. “Data represents isolated facts but when duly embodied and combined within a particular structure, information emerges. Once analyzed and used, this circulates as knowledge” (Lopes, 2008, p. 9). The success of this knowledge creation depends on individual sharing experiences. Lopes continued, explaining that intellectual capital is key for companies to succeed and that knowledge value is the real financial dimension of companies.

Summary

The roots of today’s tacit knowledge management research are in the theory of organizational knowledge creation (Nonaka, 1991) and the concept of the spiral process of knowledge, in which tacit knowledge is converted into explicit knowledge and then back into tacit knowledge. Processing tacit knowledge is complicated because “‘explicit’ or codified knowledge refers to knowledge that is transmittable in formal, systematic language. On the other hand, ‘tacit’ knowledge has a personal quality, which makes it hard to formalize and communicate” (Nonaka, 1994, p. 16). Nonaka (1994) called the distinction between tacit and explicit knowledge the *epistemological dimension to organizational knowledge*. The new perspective of strategic management deviates from the original model by Michael Porter as it focuses on internal resources and capabilities rather than the products produced by the resources. Knowledge, especially tacit knowledge, is difficult to acquire and cannot be purchased. The more intellectual resources a company has, the better equipped is to compete and, therefore, knowledge becomes a strategic resource for a competitive advantage. As W. H. Inmon and Nesavich (2009) noted, “Stated differently, organizations that look only at their structured data—usually transaction-based data—miss an entire class of information that waits to be used for the decision-making process” (p. 11). Accordingly, this research attempted to demonstrate the importance of tacit knowledge and its impact in increasing business performance.

New laws such as the Sarbanes-Oxley Act of 2002, Securities and Exchange Commission Rule 17a-4, and the Health Insurance Portability and Accountability Act of 1996 mandate unstructured data to be protected. The importance of unstructured data has been highlighted by several authors (Abidin et al., 2010; Kuechler, 2007; Seidler-de Alwis & Hartmann, 2008). Seidler-de Alwis and Hartmann (2008) examined tacit knowledge as a source of innovation. Mingrui and Yongjian (2010) studied tacit knowledge under the lens of distributed cognition. Subashini (2010) argued that explicit and tacit knowledge are complementary because both types are essential, basically agreeing with both Nonaka (1994) and Zack (1999a), and pointed that current technologies capture the data flow but not the knowledge flow. According to Koskinen (2004), most of the attention in projects has been on codified explicit knowledge, neglecting the vast amount of information residing in tacit implicit knowledge.

There is a new trend in investigation of processing unstructured data and using it as a source of innovation. Several authors have looked at different technologies to process this data (Bosch-Sijtsema et al., 2009; Du et al., 2009; Fan et al., 2006; Godbole

& Roy, 2008; Plejie et al., 2008; Zhou et al., 2007), others have linked it to BI (Bara et al., 2009; Popovic et al., 2010; Ranjan, 2008; Sahay & Ranjan, 2008), and a handful have worried about the data consistence (Blake & Mangiameli, 2009; Kobielus, 2008; Pintar et al., 2007; Wang & Wang, 2008).

Business productivity as related to the use of knowledge has been investigated using macro-dimensions but not related to individual workers' productivity. For example, Chang Lee et al. (2005) designed and tested a new instrument that measures KM performance using stocks price, price-earnings ratio (PER), and R&D expenditure as the dependent variables. Using linear regression to measure the sample, Chang Lee et al. surveyed companies in the Korean Securities Dealers Automated Quotations market and the results showed strong, significant numbers in the factors supporting the following hypothesis: When the Knowledge Management Performance Index is greater, stock prices, PER, and R&D are significantly better.

Singh (2008) conducted a survey of an Indian software company to investigate the impact of leadership styles on KM productivity. Even though Singh investigated explicit and tacit knowledge, those variables were used as the dependent variables being impacted by gender and leadership styles and not as predictors to productivity. Other authors investigated business productivity in relation to KM but using factors different from individual productivity. Whereas Goel et al. (2010) and Mezher et al. (2005) studied portal implementations, Mundra et al. (2011) examined competitive advantage as a variable for KM success.

Goel et al. (2010) argued that business gained a competitive advantage after the use of KM but they did not prove it statistically and the results only showed possibilities at a higher level. Some of the findings from Mundra et al. (2011) revealed that companies are combining AI to retrieve data from databases; that video conferences, e-mail, and chat groups are necessary tools to share knowledge, including tacit and explicit knowledge; and that one of the companies could reduce part of its training program from seven days to 4-5 hours with the help of knowledge, but this was a very broad finding that failed to point out the specifics of individual productivity. Mezher et al. (2005) conducted a case study of an engineering consulting company after it created a KM model and although he postulated that the use of knowledge would help to complete engineering projects in less time, the model was not tested statistically.

Bosch-Sijtsema et al. (2009) studied productivity on globally distributed teams as a whole and even though individual productivity was mentioned in their framework, no provisions or details were shown on how to test the productivity, leaving the model as a pure conceptual concept. Britt (2009) attributed the increase of productivity in an insurance broker company to the use of knowledge obtained from a web-based model, a very vague and general assessment that failed to take in consideration other factors that may have contributed to the increase in insurance policies. Martins and Lopes dos Reis (2010) studied productivity under the human capital lens oriented to identify it as an asset that can be recorded along with the other assets that a company possesses. They used a very sophisticated software tool to measure their proposed framework but, again, the study did not prove individual productivity as influenced by knowledge factors. Finally, Mahmood and Ali (2011) performed structural equation modeling after they operationalized several constructs to predict productivity. They concluded that KS, organizational culture, and technology and rewards contributed to knowledge worker productivity but failed to show how productivity was increased.

The University of New South Wales conducted research on KM studies from August to October 2004. The research went back as far as 1892 and 290 research papers were analyzed by two master of science students (The University of New South Wales, 2004). Only three research studies were identified to have investigated productivity and they did not address individual productivity. McCampbell et al. (1999) conducted case studies on Teltech, Ernst & Young, Microsoft, and Hewlett-Packard to analyze the effect of KM in quality and productivity improvement but the study was very general and did not include specifics. Zazzara (2001) mentioned that to elevate productivity and maintain clinical quality through the use of knowledge would be nirvana for the health care system but failed to demonstrate specifics on how to achieve it. Filius et al. (2000) prescribed three activities for organizations that are willing to improve productivity—activities that expand the individual or collective horizon, activities that consolidate knowledge, and informal or formal communication about the issue—but this, again, was just a wide recommendation without any practical prescription.

Sharing knowledge has been identified as a valid method of eliciting tacit knowledge to be converted into explicit knowledge "Knowledge sharing: Knowledge is described as a mixture of experiences, principles, and related knowledge. It begins and resides in the 'minds of knower,' described as tacit knowledge" (Mahmood & Ali, 2011, p. 28). Antonova et al. (2011) developed a model of KS and KT to examine their impact on organizations. Douglas (2009) stated that distributed cross-cultural corporations are creating an increasing need to communicate and to share knowledge, and that there is growing recognition of KT in social spaces. Inthout et al. (2010) used previous literature to design a system that would combine tacit and explicit knowledge with emphasis on sharing, and rewards and trust were investigated as important components of sharing knowledge (Antonova et al., 2011; Holste & Fields, 2010; Reychav & Weisberg, 2009).

Feghali and El-Den (2008) postulated that (a) ideas and opinions are the easiest form of tacit knowledge to share among virtual groups and (b) that these opinions and ideas can be progressively shared among virtual groups by the creation and sharing of documents. Tacit knowledge articulated this way supplements the face-to-face interaction that is missing in virtual environments. By constant dialogue and refinement of ideas expressed in documents, the most hidden portions of tacit knowledge can be discovered. "Knowledge transformation among virtually dispersed group members is possible through the articulation of members' opinions and ideas into a shared document. This document provides the infrastructure for the interaction among the members by exposing them to each other's opinions" (Feghali & El-Den, 2008, p. 103).

To conclude, since Nonaka (1991) developed the theory of organizational knowledge creation, unstructured data in the form of tacit knowledge has been identified as the most important component of business performance. Multiple investigations are being conducted to understand and process tacit knowledge, and authors are trying to understand its impact in business productivity, but so far, all these attempts have been concentrated in macro-dimensions and not on individual productivity. Sharing knowledge has been identified as a valid component to study the impact of tacit knowledge on business performance. This research examined explicit (semi-structured) data made explicit from tacit knowledge from individuals after sharing their experiences on online forums to investigate individual productivity as a source of business performance. As such, this investigation attempted to close a knowledge research gap by analyzing the impact of tacit knowledge on business performance under the lens of individual productivity.

CHAPTER 3. METHODOLOGY

Introduction

This chapter will describe in the detail the methodology used to conduct the research. It will describe the research design, sample, instrument, data collection, data analysis, validity and reliability and ethical considerations. The purpose of this research was twofold. First, the study intent was to show that the incorporation of unstructured data into business intelligence could increase business performance and could motivate research around the designing and developing of new paradigms and ontologies to help with the complexity of inserting unstructured data into data warehouses. Second, the study laid out the foundation for further research on knowledge management (KM) factors that could contribute to business performance. The chosen research instrument was the Knowledge Management Assessment Tool (KMAT). The KMAT "was developed by Maier and Moseley (2003) and consists of 30 statements to measure knowledge management practices of the organization" (Singh, 2008, p. 9). The tool measures five dimensions: knowledge identification and creation, knowledge collection and capture, knowledge storage and organization, knowledge sharing and dissemination, and knowledge application and use. Additionally, the tool has a mechanism to convert those five dimensions into only two dimensions: explicit knowledge and tacit knowledge.

The importance of tacit knowledge was highlighted several years ago by Polanyi (2009), who stated, "I shall reconsider human knowledge by starting from the fact that we can know more than we can tell" (p. 4). He also declared, "We recognize the moods of the human face, without being able to tell, except quite vaguely, by what signs we know it" (p. 5) and classified this human characteristic as tacit knowledge, a knowledge that is hard to formalize and communicate. He further stated, "I think I can show that the process of formalizing all knowledge to the exclusion of any tacit knowing is self-defeating" (p. 20).

As stated previously, this research attempted to demonstrate the importance of tacit knowledge and its impact on increasing business performance. By showing which of the dimensions measured by the research instrument affect business performance in a positive way, corporations could add more resources to that dimension to obtain performance gains. This can be done on each of the five dimensions of the KMAT as well as on the two dimensions of explicit and tacit knowledge. So far, the literature shows studies on the impact of knowledge in business performance in general but no study has been found that measures individual units of productivity. These findings can be used to predict productivity of engineering companies that provide customer support similar to the one described in this research. This research attempted to answer the following questions: (a) To what extent is there a significant decrease in the time to complete field engineers' tasks after unstructured data are incorporated into the business intelligence framework? and (b) Can the usage of any of the KMAT factors predict field engineers' time to complete tasks when unstructured data are incorporated into the business intelligence framework?

Research Design

The design research was an experimental design.

> The simplest of all experimental designs is the two-group posttest-only randomized experiment. In design notation, it has two lines—one for each group—with an R at the beginning of each line to indicate that the groups were randomly assigned. (Trochim, 2006, para. 1)

As indicated by the *R*s in Figure 3, two groups of engineers were randomly assigned. One group of engineers did not belong to the new knowledge community (created for the experiment) and not exposed to the tacit knowledge and unstructured data variable, and one group (denoted by the *X*) was exposed to the tacit knowledge unstructured data variable. To determine if the experimental group's productivity was different after exposure to the new knowledge, the group means were tested using an independent *t* test. When the control group's number of completed tasks was found to be significantly different from the treatment group's number, regression analysis was conducted to determine which of the factors contributed the most to the new results.

This study is based in the positivistic philosophy that try to obtain facts in term of relation among variables (Swanson, 2005). It used a quantitative methodological approach that measure frequency, means, variances and statistical analysis.

Multiple linear regression tests were conducted on the KMAT knowledge factors to predict productivity.

> Multiple regression analysis examines the relationship between a single dependent variable and two or more independent variables. It is a widely used analytic technique in organizational research and has been the most popular statistical technique for hypothesis testing for at least two decades (Weinzimmer, Mone, & Alwan, 1994). (Bates, 2005, p. 118)

The KMAT knowledge factors could not be confirmed using factor confirmatory analysis due to the size of the sample. Additionally, structural equation modeling was performed on the results to support the regression tests. “Structural equation modeling . . . is a statistical methodology that takes a confirmatory (i.e., hypothesis testing) approach to the analysis of a structural theory bearing on some phenomenon” (Byrne, 2010, p. 3). The productivity was measured as a function of the different factors of knowledge, specifically tacit knowledge.

The context for this research was an engineering company that sells hardware and software to customers and provides customer support to maintain their products. When a customer calls the support hotline, a preliminary analysis of the customer issue is completed and if needed, an engineer will be dispatched onsite to either troubleshoot the issue or comply with a specific task—for example, to replace a computer system board or a processor. When engineers are dispatched, they are briefed with a problem description and possible solution to the problem. The preliminary solution is based on remote diagnosis but more often, engineers encounter other issues, or the same issue continues to manifest after the parts have been replaced or the suggested fix has been applied. Currently, the company under study measures the performance of these engineers in two ways: (a) number of completed tasks and (b) time taken to complete those tasks. For example, two engineers can complete the same amount of tasks but one may take more time to complete them than the other one takes. Therefore, if possible, it was important to consider both time-to-completion and number of tasks completed, but for this research number of tasks completed in the month was chosen as the unit of measurement. The research model used on this research is shown in the conceptual framework in Figure 2.

According to Koskinen (2004), most of the attention in projects has been on codified explicit knowledge, neglecting the vast amount of information residing in tacit implicit knowledge. This study looked into explicit (semi-structured) data made explicit from tacit knowledge from individuals after sharing their experiences in online forums. The data were in semi-structured format because they were still in free-form and not integrated into any relational or indexed database. Feghali and El-Den (2008) postulated that (a) ideas and opinions are the easiest form of tacit knowledge to share among virtual groups and (b) these opinions and ideas can be progressively shared among virtual groups by the creation and sharing of documents. Tacit knowledge articulated this way supplements the face-to-face interaction that is missing in virtual environments. By a constant dialog and refinement of the ideas expressed in the documents, the most hidden portions of the tacit knowledge can be discovered: "Knowledge transformation among virtually dispersed group members is possible through the articulation of members' opinions and ideas into a shared document. This document provides the infrastructure for the interaction among the members by exposing them to each other's opinions" (p. 103).

The importance of knowledge residing in people's minds and the importance of sharing that knowledge is also emphasized by Brown and Duguid (2000): "Attending to knowledge, by contrast[,] returns attention to people. . . . The importance of people as creators and carriers of knowledge is forcing organizations to realize that knowledge lies less in its databases that in its people" (p. 121). Furthermore, Brown and Duguid highlighted the value of communities of practice in the labs of particle physicists and biotechnologists.

In the conducted research, customer support data that resided in people's minds in the form of tacit knowledge was shared through an online forum after engineers created and shared documents containing experiences of engineers going to support customers on-site. This sharing was in the form of issues, recommendations, and advice. This forum was then searched prior to the execution of new tasks. The forum was created on a corporate repository for engineers to share and discuss their experiences. This data continuously retro-fed the community with more and refined knowledge about installations, performed tasks, and general support data. Knowledge shared and created from unstructured data is a circular process. It comes from customer issues and needs. Engineers searched the forum and after reading the different comments, issues, past experiences, and suggestions, used the information to provide a possible solution. The new information was then used by field engineers to communicate a solution to the customer and this new information is fed back into the repository. As more knowledge is acquired, the customer gets better support and more business is created; the result could be more productivity and engineers could finish their task early, potentially completing more tasks.

Sample

The population for this research was a software and hardware support engineering companies that send engineers onsite to fulfill customer requests and troubleshoot and fix customer issues. The sample frame is a worldwide software and hardware computer company with more than 20,000 employees in the hardware service organization. The sample was selected randomly from its Latin America field service division of approximately 300 employees. The sample was selected from field engineers (male and female) specially trained to go onsite to service customers. The selection did not include remote engineers but only field engineers, although they could be in touch with remote engineers at the time of the service.

Recruiting was accomplished after obtaining permission from the Latin America field service director and then e-mailing the participants obtaining their consent. The selection was a random cluster sample obtained from a field engineer mailing list. The random sample selected two groups: control and experimental. There were two versions of the study survey: one for the group that completed the survey without participating in the sharing communities and one for the group that participated in the sharing communities.

> A cluster sample is a random sample in which members of the population sampled are embedded in a collection—that is, a cluster—of elements. For instance, instead of sampling employees, a researcher might sample work teams, which are composed of employees. (Passmore & Baker, 2005, p. 53)

The a original sample size was calculated using the recommendations of Barlett, Kotrlik, and Higgins (2001) based on "Cochran's (1977) sample size formula for both continuous and categorical data" (p. 44). With a table provided by Barlett et al., a minimum sample of 85 was calculated for a sample frame of 289 subjects. The suggested sample was also adequate to conduct the multiple regression and factor analysis because it suggested a ratio of 10:1 and the research used five factors for a minimum of 50. When the research started, the original estimation of the sample was reduced to 149 after managers and dispatch engineers were removed from the sample because they do not perform tasks at customer sites. The response rate was 26% with 39 responses; 25 from the control group and 14 from the experimental group. According to Cohen (1988), this represents an effect size of .40 with a 75% power

The hardware service division for this company is spread around the world and has similar characteristics in the way the customer is serviced. Customers call a toll-free number or use an automated web page to enter new requests and engineers are dispatched onsite. The Latin America service division is embedded in the support schema, representing a cluster sampling or a stratified sample (Passmore & Baker, 2005). The service division is also spread across different countries in Latin America and each country is a representation of the company as a whole. By choosing a sample from the Latin America division, the results have a high probability of being generalized, with representation across different cultures.

Instrumentation/Measures

The Knowledge Management Assessment Tool (KMAT) "was developed by Maier and Moseley (2003) and consists of 30 statements to measure knowledge management practices of the organization" (Singh, 2008, p. 9). The tool measures five dimensions: knowledge identification and creation, knowledge collection and capture, knowledge storage and organization, knowledge sharing and dissemination, and knowledge application and use. Additionally, the tool has a mechanism to convert the five dimensions into two dimensions: explicit knowledge and tacit knowledge.

This instrument was used and tested by Singh (2008), who conducted a survey of an Indian software company to investigate the impact of leadership styles on KM productivity. Singh's survey was directed to the two types of knowledge the firm possessed: explicit, which is easy to communicate because it is in the form of hard data or codified procedures; and tacit, which applies to knowledge possessed by individuals that can only be communicated through conversation, storytelling, and so on. For the survey, Singh used the KMAT survey tool, which revealed Cronbach's alpha coefficients higher than .90 for the explicit and tacit dimensions and 0.70 and higher for the five dimensions measured by the KMAT. Therefore, there was no requirement for a field test because the instrument has been previously validated, although a possible limitation of this research is the population sample because this research is smaller than the one conducted by Singh.

The constructs were based on the seminal work of Nonaka (1994). Nonaka, in his article, "A Dynamic Theory of Organizational Knowledge Creation," postulated the theory of organizational knowledge creation. His writing explained that knowledge possessed by individuals, organizations, and societies can be expanded through a spiral process in which tacit knowledge is converted into explicit knowledge and then back into tacit. Tacit knowledge is hidden behind behaviors, skills, competencies, and experiences (tacit actionable knowledge) and articulated knowledge resides in individual thoughts and language use. Explicit knowledge resides inside computers in codified form and by nature has a clear organization (Delen & Al-Hawamdeh, 2009).

This research also examined Polanyi (2009), who stated, "I shall reconsider human knowledge by starting from the fact that we can know more than we can tell" (p. 4). He also declared, "We recognize the moods of the human face, without being able to tell, except quite vaguely, by what signs we know it" (p. 5) and classified this human characteristic as tacit knowledge, a knowledge that is hard to formalize and communicate. He further stated, "I think I can show that the process of formalizing all knowledge to the exclusion of any tacit knowing is self-defeating" (Polanyi, 2009, p. 20).

Nonaka (1994) also provided an interpretation of Polanyi's (1966) tacit and explicit knowledge concept: "Polanyi classified human knowledge into two categories. 'Explicit' or codified knowledge refers to knowledge that is transmittable in formal, systematic language. On the other hand, 'tacit' knowledge has a personal quality, which makes it hard to formalize and communicate" (p. 16). Nonaka called the distinction between tacit and explicit knowledge the *epistemological dimension to organizational knowledge*. The exchange can take many forms and based on these variations, different modes of knowledge conversion can be generated. Tacit-to-tacit, a shared experience, is called *socialization*. Explicit-to-explicit, in which modern computers play an important role, is called *combination*. The third and four modes are a combination of the first two: converting explicit into tacit, called *internalization*; and converting tacit into explicit, called *externalization*. On the ontological dimension, the theory posits that individuals are the ones that create knowledge and that an organization should amplify this knowledge through different levels of the firm. The key here is a constant dialog, in which middle-up-down management leadership is the most suitable to crystallize the conversion and creation of knowledge.

Nonaka and Toyama (2003) expanded on the concept of tacit and explicit knowledge. They described a process occurring in a virtual environment that they named the *ba*: "Building on the concept that was originally proposed by the Japanese philosopher Kitaro Nishida (1921, 1970), we define ba as a shared context in motion, in which knowledge is shared, created, and utilized" (Nonaka & Toyama, 2003, p. 6). Nonaka and Toyama reiterated that in today's world, knowledge is the most important source of a firm's sustainable competitive advantage and that a new knowledge-based theory is needed that differs from the existing economic and organizational theory coinciding with Zack (1999a, 1999b).

The performance construct is based on Zack's (1999a) study, who stated that the link between KM and business strategy has been ignored and that companies must have a knowledge strategy. Zack stated that the best-known corporate strategy approach uses the strengths, weaknesses, opportunities, and threats (SWOT) model, and that "application of the SWOT framework has been dominated over the last 20 years by Porter's 'five-force' model" (p. 127). The new perspective of strategic management deviates from the original model by Michael Porter as it focuses on internal resources and capabilities rather than the products created by the resources. Knowledge, especially tacit knowledge, is difficult to acquire and cannot be purchased. The more intellectual resources a company has, the better equipped it is to compete; therefore, knowledge becomes the strategic resource for competitive advantage. Zack (1999b) expanded his work in his article, "Managing Codified Knowledge," in which he continued popularizing the concept of viewing organization knowledge as a strategic asset. The views of Zack differ somewhat from that of Nonaka (1994) because Zack considered explicit knowledge the most important asset and Nonaka put the relevance in tacit knowledge.

This research attempted to demonstrate the importance of tacit knowledge and its impact in increasing business performance. It is aligned with Nonaka's (1994) view that tacit knowledge is the most important part of increasing business performance, and argues that using externalization as one of the modes of the Nonaka theory (that is, converting tacit knowledge into explicit knowledge) can increase business performance. This research investigated the impact of tacit knowledge on business performance after evaluating individual engineers' productivity and it was based on the previously discussed theory of organizational knowledge creation from tacit knowledge (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Nonaka & von Krogh, 2009; Polanyi, 2009). Tacit and explicit knowledge were the independent variables and individual workers' productivity was the dependent variable.

As stated previously, the research contained three main constructs and two levels of participants. The constructs (tacit knowledge, explicit knowledge, and business performance) were operationalized with the help of the KMAT, which takes abstractions and converts them into independent and dependent variables. The two levels of participants were engineers, in which one group was exposed to the tacit knowledge exchanged variable and one group was not exposed to it. The model investigated the following research questions: (a) Can the incorporation of customer support unstructured data into the customer support schema increase business performance? (b) If unstructured customer support data are converted into explicit knowledge, can this converted data contribute to business performance by increasing engineers' productivity? (c) To what extent is there a significant decrease in the time to complete field engineers' tasks after unstructured data are incorporated into the BI framework? and (d) Can the usage of any of the KMAT factors predict field engineers' time to complete tasks when unstructured data are incorporated into the BI framework?

The following hypotheses were developed from the research questions.

- $H1_0$: The inclusion of tacit knowledge in BI does not produce a significant difference in time to complete engineering tasks at customer sites.

- $H1_A$: The inclusion of tacit knowledge in BI produces a significant difference in time to complete engineering tasks at customer sites.

- $H2_0$: Tacit knowledge is not a factor that can be used to predict employees' productivity when included into BI.

- $H2_A$: Tacit knowledge is a factor that can be used to predict employees' productivity when included into BI.

The independent variables were tacit knowledge and explicit knowledge, and the dependent variable was the number of tasks performed by engineers using and not using tacit knowledge to complete their tasks. Tacit knowledge was incorporated into the experiments after field support engineers shared their knowledge and experience in a knowledge community.

The two independent variables were constructed from a set of 30 questions obtained from a survey. The answers were evaluated on a Likert scale using a range of values from 1 (*Strongly disagree*) to 6 (*Strongly agree*). These 30 questions were converted into five factors and the five factors into two factors: tacit and explicit knowledge. The dependent variable was obtained from the same survey after the engineers indicated how many tasks they have completed for a particular month. The constructs were based on the theory of organizational knowledge creation from tacit knowledge (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Nonaka & von Krogh, 2009; Polanyi, 2009). After each dimension was measured, the scores were grouped into factors and calculations performed for each factor.

The five factors of knowledge were obtained by grouping the 30 questions into five dimensions and the scores are interpreted as shown in Table 2. The two factors of knowledge were obtained by grouping the 30 questions into two dimensions (tacit and explicit knowledge). The scores for each dimension are interpreted as shown in Table 3.

Table 2. Five Knowledge Management Dimensions Scores

Score	Explanation
31–36	The organization/department exhibits highly effective KM practices on this dimension.
26–30	The organization/department exhibits very effective KM practices on this dimension.
21–25	The organization (or department) exhibits moderately effective KM practices on this dimension.
16–20	The organization/department exhibits moderately ineffective KM practices on this dimension.
11–15	The organization/department exhibits very ineffective KM practices on this dimension.
6–10	The organization/department exhibits extremely ineffective KM practices on this dimension

Table 3. Two Knowledge Management Dimensions Scores

Score	Explanation
79–90	The organization/department exhibits highly effective KM practices on this dimension.
66–78	The organization/department exhibits very effective KM practices on this dimension.
53–65	The organization/department exhibits moderately effective KM practices on this dimension.
40–52	The organization/department exhibits moderately ineffective KM practices on this dimension.
27–39	The organization/department exhibits very ineffective KM practices on this dimension.
15–26	The organization/department exhibits extremely ineffective KM practices on this dimension

After the scores were calculated, the highest scores on tacit or explicit knowledge could determine the higher usage of that dimension. The same calculations were made for the five dimensions that showed in which dimension the company was stronger in terms of knowledge. Additionally, the study survey inquired about number of engineering tasks completed as the research attempted to evaluate a relationship between dimensions of knowledge and number of engineering tasks completed.

Data Collection

The method of data collection was an electronic intranet survey using the corporate survey tool. According to Puleston (2011), there are seven best practices to consider to succeed in administering online surveys: (a) look at the online survey as a form of communication and use all the graphic tools available to make it more attractive, (b) engage participants from the beginning and do not start with the survey directly but "[break] the information into sound bites, telling a story, adding some imagery and humor, [resulting] in respondents investing more time in the survey and giving more thoughtful feedback" (p. 558), (c) adopt more creative questioning methods recently developed by the industry and technology, (d) understand the critical use of imagery and use it effectively, (e) learn from social psychology techniques, (f) learn from quality researchers, and (g) use a pilot if possible. After the sample was randomly chosen, the e-mails to the subjects were separated by the control group and the experimental group and an introductory mail was prepared explaining the purpose of the survey. For the experimental group, the introductory mail clearly explained that they had to wait at least two months before they answered the questions of the survey and that they should engage immediately in the creation, writing, and sharing of the information on the engineering community forum. Their names were entered into the forum based on the random sample. After two months of sharing, they completed the survey that included answering how many tasks they had completed for the last month of sharing. For the control group, they were asked to complete the survey at their convenience after that particular month ended because they would be entering the number of tasks completed for that month. Because the survey was answered inside the intranet, the delivery of the answers happened automatically by sending the survey to an e-mail address.

Data Analysis

The survey collected descriptive statistics on age and gender. The data for the inferential statistics was compiled based on the answers of the KMAT survey. The KMAT "was developed by Maier and Moseley (2003) and consists of 30 statements to measure knowledge management practices of the organization" (Singh, 2008, p. 9). The tool measures five dimensions: knowledge identification and creation, knowledge collection and capture, knowledge storage and organization, knowledge sharing and dissemination, and knowledge application and use. Additionally, the tool has a mechanism to convert the five dimensions into two dimensions: explicit knowledge and tacit knowledge. The tool provides a formula to calculate the five dimensions and two dimensions based on the question number and the answers provided by the responders.

Even though the tool has been previously validated for consistence, the research calculated the Cronbach's alpha coefficients of all the dimensions of the KMAT. The Cronbach's alpha measures the reliability of the scales: "[the] degree to which instrument items are homogeneous and reflect the same underlying constructs" (Cooper & Schindler, 2008, p. 293). Cronbach's values higher than .70 are considered good coefficient values

(Ab Hamid, Mustafa, Idris, Abdullah, & Suradi, 2011; Newbert, 2008; Xiao & Kim, 2009). Conducting confirmatory factor analysis with the 30 questions was not possible due to the sample size. According to Field (2009), one of the uses of factor analysis is "to reduce a data set to a more manageable size while retaining as much of the original information as possible" (p. 628). Tarafdar, Tu, Ragu-Nathan, and Ragu-Nathan (2007) performed factor analysis to reduce 39 items that originally were identified as conditions that produce stress to 36. Factor analysis grouped them into five factors that describe reasons why information computer technology creates stress and three factors that corresponded to role conflict, role overload, and productivity.

Additionally, testing hypotheses about the structures of latent variables could not be done with the Analysis of Moment Structures (AMOS) software using confirmatory factor analysis for the reasons previously described.

In regards of using AMOS Ab Hamid et al. (2011) stated:

> In short, [the questionnaire] proved to be reliable. In order to validate the instrument, this study also considered construct validation using [AMOS] with maximum likelihood . . . to analyze the data. This approach is called . . . confirmatory factor analysis[,] which is more advanced as the hypothesized are based on the underpinning theory. (p. 88)

Similarly, Hui-Ling and Yu-Hsuan (2011) used confirmatory factor analysis to test for leadership performance on management teams.

Once the validity of the constructs was demonstrated, the significant differences were tested using independent *t* tests to test the statistical impact of using tacit knowledge in one group of engineers and then linear regression to discover predictability factors from tacit and explicit knowledge on number of tasks performed.

Validity and Reliability

The design was an experimental design. The sample, represented by a group of engineers, was randomly selected and half of the sample was randomly assigned to have access to an online community in which they could share knowledge related to engineering tasks. The data were collected via an online survey in the following manner: the first group (the control group) returned the survey indicating the number of tasks performed for the last month among the answers provided by the KMAT instrument. The second group waited to respond at least two months after they have been sharing their knowledge and experience using the online community of knowledge. The choosing of both random selection and random assignment complied with internal and external validity of the sample. Similar to Singh's (2008) study, this research tested the KMAT Cronbach's alpha coefficients. Values higher than 0.70 are considered to have high reliability (Ab Hamid et al., 2011; Newbert, 2008; Xiao & Kim, 2009).

Ethical Considerations

The survey was voluntary and no coercion of any kind was implied. Participants were briefed with the purpose of the research and notified that they could decline to participate with no consequences of any kind. The research presented no greater than minimal risk and the participants contributed to the benefit of the whole population by providing statistical evidence that engaging in the kind of activities intended for the experiment could benefit the hardware engineering community. Due to the random nature of the sample, there was no privilege of any kind in choosing the participants and the ones randomly chosen by the algorithm had the ability to decline participation.

CHAPTER 4. RESULTS

This chapter presents the result of a quantitative study that investigated the usage of tacit knowledge into business performance. As explained in the previous chapters, the research design was an experimental design that evaluated the impact of field support engineers using tacit knowledge to complete customer tasks. The study attempted to answer the following questions:

1. Can the incorporation of customer support unstructured data into the customer support schema increase business performance?
2. If unstructured customer support data are converted into explicit knowledge, can this converted data contribute to business performance by increasing engineers' productivity?
3. To what extent is there a significant decrease in the time to complete field engineers' tasks after unstructured data are incorporated into the BI framework?
4. Can the usage of any of the KMAT factors predict field engineers' time to complete tasks when unstructured data are incorporated into the BI framework?

The following hypotheses were developed from the research questions:

- $H1_0$: The inclusion of tacit knowledge in BI does not produce a significant difference in time to complete engineering tasks at customer sites.
- $H1_A$: The inclusion of tacit knowledge in BI produces a significant difference in time to complete engineering tasks at customer sites.
- $H2_0$: Tacit knowledge is not a factor that can be used to predict employees' productivity when included into BI.
- $H2_A$: Tacit knowledge is a factor that can be used to predict employees' productivity when included into BI.

The context for this research was an engineering company that sells hardware and software to customers and provides customer support to maintain its products. When a customer calls the support hotline, a preliminary analysis of the customer's issue is completed, and, if needed, an engineer is dispatched onsite to either troubleshoot the issue or to comply with a specific task—for example, to replace a computer system board or processor. When engineers are dispatched, they are briefed with a problem description and a possible solution to the problem. The preliminary solution is based on remote diagnosis, but more often, engineers encounter other issues, or the same one continues to manifest after the parts have been replaced or the suggested fix applied.

This research attempted to demonstrate the importance of tacit knowledge and its impact in increasing business performance. It is aligned with Nonaka's (1994) view that tacit knowledge is the most important part of increasing business performance, and argued that using externalization, one of the modes of the Nonaka theory (that is, converting tacit knowledge into explicit knowledge), can increase business performance. In this research context, the field support engineers converted their tacit knowledge into explicit knowledge using a knowledge community that was created for the experiment. This knowledge is based on their past and current experiences on solving customer problems. Once the tacit knowledge is converted into explicit knowledge, engineers can search for that knowledge in the tacit knowledge group that was created for the experiment. The research framework is shown in Figure 2.

The hypothesis about the time to complete tasks could be have been investigated by calculating the ratio of completed tasks to days in the month of completion. The other option was simply to analyze the numbers of tasks completed in the month and detects any significant difference between the group that was not exposed to tacit knowledge and the group that was exposed to tacit knowledge; the latter was chosen.

The remaining sections contain the description of the sample, the summary of the results a detailed analysis of the results and the conclusion.

Description of the Sample

The sample was chosen from a division of field engineers spread across multiple geographical areas. A random process was used to divide the invitations to the control and the experimental group. The sample was collected from one of the regional field support engineer division made of 149 engineers with a response rate of 26% with 39 responses; 25 from the control group and 14 from the experimental group. According to Cohen (1988), this represents an effect size of .40 with a 75% power. The original estimation of the sample was reduced after managers and dispatch engineers were removed from the sample because they do not perform tasks at customer sites.

The demographic of the sample is described as follows. There were a total of 39 engineers 25 for the control group and 14 for the experimental group. The age distribution is depicted in Table 4. There were four engineers between 24 and 29 years old; 16 between 30 and 35; eight between 36 and 41; seven between 42 and 47; and two each between 48 and 53, and between 54 and 60 years old. The number of tasks executed for each group can be seen in Tables 5 and 6. The mean numbers of tasks was 4.64, with a standard deviation of 1.87, a median of 5, and a mode of 4. The execution of tasks is very well distributed among engineers as can be seen by the closeness of the mean and median. (A perfect normal distributed means and median will be the same.)

Table 4. Sample Age

Age group	Frequency	Percent	Valid percent	Cumulative percent
2 (24–29 years)	4	10.3%	10.3%	10.3%
3 (30–35 years)	16	41.0%	41.0%	51.3%
4 (36–41 years)	8	20.2%	20.5%	71.8%
5 (42–47 years)	7	17.9%	17.9%	89.7%
6 (48–53 years)	2	5.1%	5.1%	94.9%
7 (54–60 years)	2	5.1%	5.1%	100.0%
Total	39	100.0%	100.0%	

Table 5. Sample Number of Tasks

Number of engineers	39
Mean	4.64
Std error of mean	.30
Median	5
Mode	4
Std deviation	1.87

Table 6. Sample Distribution of Number of Tasks

Distribution	Frequency	Percent	Valid percent	Cumulative percent
0.60	1	2.6%	2.6%	2.6%
1.20	1	2.6%	2.6%	5.1%
1.40	1	2.6%	2.6%	7.7%
1.80	2	5.1%	5.1%	12.8%
2.00	2	5.1%	5.1%	17.9%
3.20	1	2.6%	2.6%	20.5%
3.80	1	2.6%	2.6%	23.1%
4.00	6	15.4%	15.4%	38.5%
4.20	1	2.6%	2.6%	41.0%
4.80	2	5.1%	5.1%	46.2%
5.00	5	12.8%	12.8%	59.0%
5.20	1	2.6%	2.6%	61.5%
5.40	3	7.7%	7.7%	69.2%
5.60	1	2.6%	2.6%	71.8%
5.80	1	2.6%	2.6%	74.4%
6.00	3	7.7%	7.7%	82.1%
6.20	3	7.7%	7.7%	89.7%
6.80	1	2.6%	2.6%	92.3%
7.00	1	2.6%	2.6%	94.9%
8.00	1	2.6%	2.6%	97.4%
9.00	1	2.6%	2.6%	100.0%

The gender distribution as seen in Table 7 shows a predominantly male group and this is related to a natural phenomenon in the researched company because this was a

100%-random sample. In looking at the descriptive of the survey answers in Table 8, it is noted that the mean for the five factors range from 27.40 to 28.24 on the control group and between 28.36 and 29.50 on the experimental group, showing a slightly higher mean on the experimental group. On the two factors, tacit and explicit knowledge, the mean is 69.84 and 70.36 for the control group and 71.64 and 72.86 for the experimental group, also reflecting a slight higher number on the experimental group; both groups showed very effective KM practices on all the dimensions of the KMAT as shown on Table 2 making the factors suitable to conduct the linear regression. Finally, on the number of tasks the means are 1.64 on the control group and 1.65 on the experimental group. It is noted that the number of tasks were converted using a valid statistical undisclosed conversion to preserve the anonymity of the researched company.

Table 7. Sample Gender

Gender	Frequency	Percent	Valid percent	Cumulative percent
Male	38	97.4%	97.4%	97.4%
Female	1	2.6%	2.6%	100.0%
Total	39	100.0%	100.0%	

Table 8. Survey Answers by Group

Group	Survey answer	*N*	Min	Max	Mean	*SE*	*SD*
Control	Knowledge identification and creation	25	19	36	28.12	0.819	4.096
	Knowledge collection and capture	25	20	36	27.40	0.735	3.674
	Knowledge storage and organization	25	22	36	28.24	0.784	3.919
	Knowledge sharing and dissemination	25	22	36	28.20	0.742	3.708
	Knowledge application and use	25	20	36	28.24	0.863	4.314
	Explicit knowledge management practices	25	55	90	70.36	1.837	9.187
	Tacit knowledge management practices	25	49	90	69.84	1.901	9.507
	Converted tasks	25	1.80	9.00	5.32	0.327	1.639
Experimental	Knowledge identification and creation	14	18	36	29.50	1.199	4.485
	Knowledge collection and capture	14	15	36	29.07	1.458	5.456
	Knowledge storage and organization	14	15	36	28.71	1.377	5.150
	Knowledge sharing and dissemination	14	16	36	28.86	1.275	4.769
	Knowledge application and use	14	18	36	28.36	1.077	4.031
	Explicit knowledge management practices	14	41	90	71.64	3.249	12.157
	Tacit knowledge management practices	14	41	90	72.86	3.063	11.461
	Converted tasks	14	0.60	5.40	3.41	0.441	1.651

Summary of Results

This section contains a brief description of the results in relation to each hypothesis followed by a detailed analysis.

Alternative Hypothesis 1 was rejected and Null Hypothesis 2 was accepted. On average, the engineers in the control group created larger numbers of tasks (M = 5.32, SE = 3.27) than the experimental group (M = 3.41, SE = 4.41). This difference was significant (t[37] = 3.47, p < .005) and represented a medium-sized effect (r = .30).

Alternative Hypothesis 2 was rejected and Null Hypothesis 2 was accepted. Neither explicit knowledge nor tacit knowledge were valid predictors for the number of tasks as seen on *explicit knowledge* (t[36] = 1.83, $p > .05$) and *tacit knowledge* (t[36] = -1.63, $p > .05$). Even though neither of the two factors showed a causal relation to the number of tasks completed, explicit knowledge showed a number closer to the statistical significance.

Details of Analysis and Results

An exploratory analysis was conducted to determine if the data were fit to conduct parametric tests. According to Field (2009), assumptions of parametric tests are normally distributed data, homogeneity of variances, interval data, and independence. The interval compliance of the data was obtained due to the nature of the number of tasks completed at month-end "continuous variable is one that gives us a score for each person and can take on any value on the measurement scale we are using. The first type of continuous variable that you might encounter is an internal variable" (Field, 2009, p. 9). To preserve the anonymity of the researched company, a multiple of the number of tasks were calculated to masquerade the data. The independency was easily accomplished because the control group was separated from the experimental group, including all the communications.

Normal Distribution

To test if the sample was normally distributed, the following tests were conducted. First, a visual inspection on the histogram, descriptive and P-P was performed. A P-P plot, or probability-probability plot, is a useful graph that can be used to test normal distribution. The P-P plot and histogram for number of tasks as shown in Figures 4 and 5 represent what appears to be a normally distributed data with small values of negative skewness and positive kurtosis as seen in Table 9.

Although visual inspection and the P-P plot can be sufficient, the Kolmogorov-Smirnov (K-S) and Shapiro-Wilk tests offers a method to test whether the distribution as a whole deviates from a comparable normal distribution (Field, 2009). By comparing the distribution of the data to a different normally distributed data set with the same mean and standard deviation, the data were found to be coming from a normal distribution and was supported by the visual of the Q-Q plots. Field (2009) stated, "A Q-Q plot is very similar to the P-P plot . . . except that it plots the quintiles of the data set instead of every individual score in the data" (p. 145). The output, displayed in Figure 6, showed little differences between the Q-Q plots and the P-P plots, and the K-S test shown in Table 10 revealed a non-significant difference, with scores of .136 for number of tasks completed. The percentage of number of tasks ($D[39] = .136$, $p > .05$) was normal.

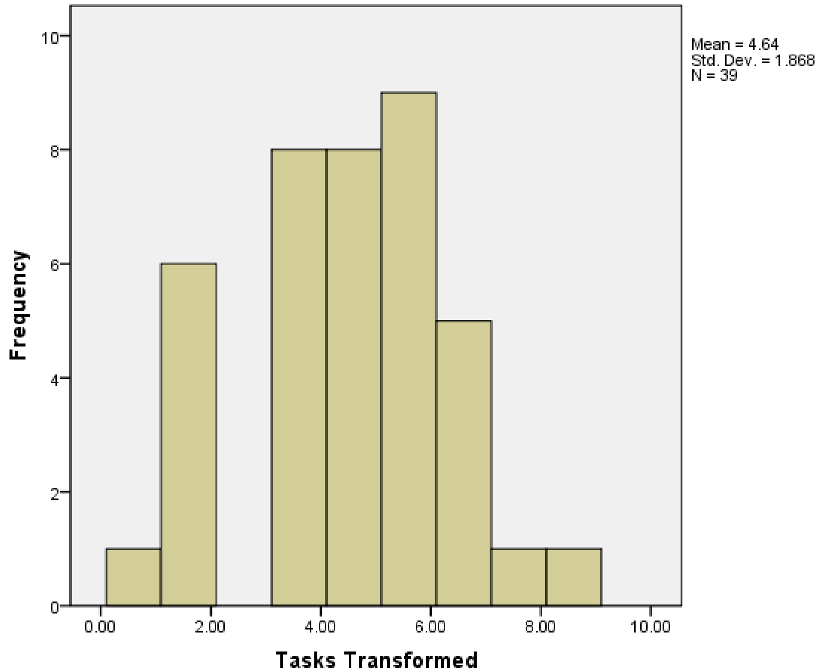

Figure 4. Histogram of tasks.

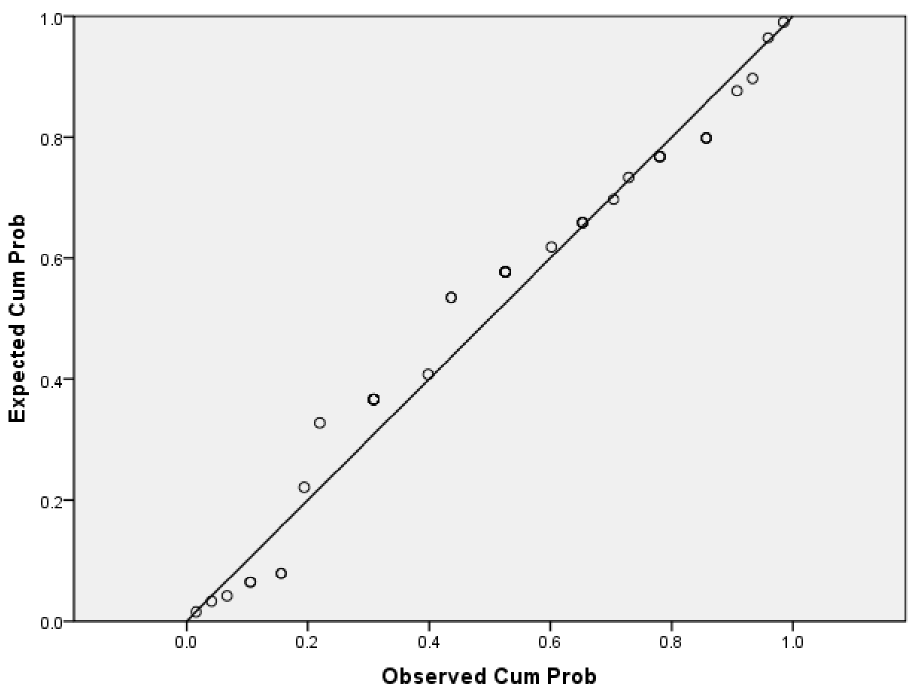

Figure 5. P-P plot of tasks.

Table 9. Task Descriptive

Mean	4.63
Standard deviation	1.87
Skewness	-.254
Kurtosis	.097

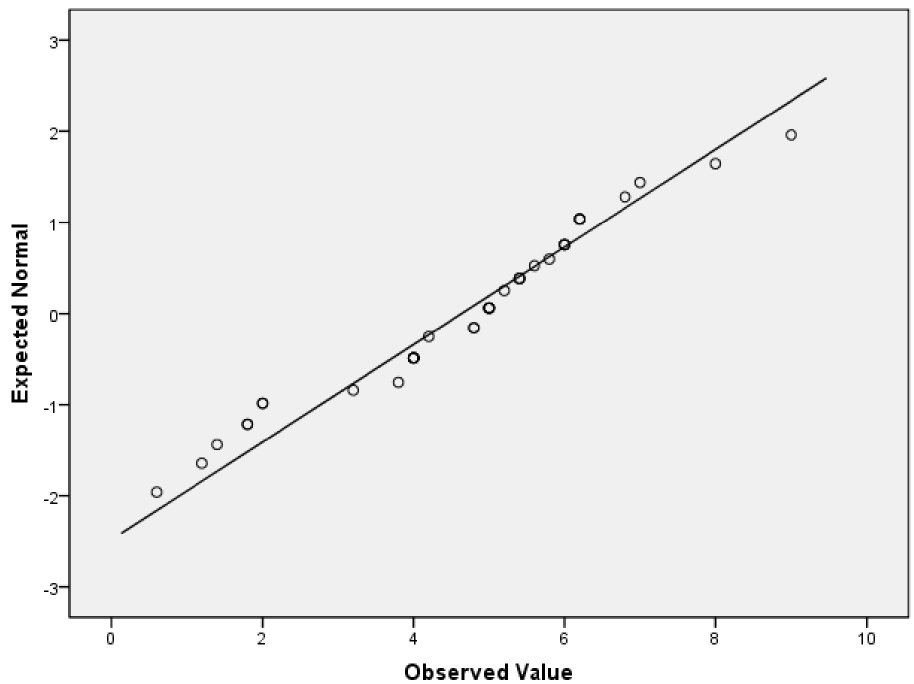

Figure 6. Q-Q plot of tasks.

Table 10. Tests of Normality

Kolmogorov-Smirnov			Shapiro-Wilk		
Statistic	*df*	Sig.	Statistic	*df*	Sig.
.136	39	.067	.961	39	.195

Homogeneity of Variances

The next step was to test for homogeneity of variances. Using the Levene's test, the null hypothesis that the variances in different groups were equal was tested. The test output is shown in Table 11 and indicated that for the number tasks, the variances were equal for both groups (F[1, 37], non significant).

Table 11. Test of Homogeneity of Variance

	Levene statistic	*df*1	*df*2	Sig.
Based on mean	.507	1	37.000	.481
Based on median	.073	1	37.000	.788
Based on median and with adjusted *df*	.073	1	36.959	.788
Based on trimmed mean	.443	1	37.000	.510

Factor Analysis

The size of the field engineer division was not large enough to conduct factor analysis, the SSPS test produced a zero determinant, and the matrix was not positive-definite, so the tests concentrated on the reliability analysis.

Reliability Analysis

The KMAT instrument used in this dissertation has been previously validated by Singh (2008), who obtained very reliable Cronbach's numbers. This research sample is considerable smaller than the one collected by Singh and that could be a limitation so, to confirm the internal consistency of the instrument, Cronbach's alpha coefficients were calculated on all the dimensions of the KMAT tool for this research. The calculated numbers are shown in Table 12. The five factors have Cronbach's α greater than .718 all the way to .846, and the two factors, explicit and tacit knowledge, have Cronbach's α greater than .914. Cronbach's values higher than .70 are considered good coefficient values (Ab Hamid et al., 2011; Newbert, 2008; Xiao & Kim, 2009).

Table 12. Knowledge Factors

Knowledge factor	Cronbach's α
Knowledge identification and creation	.821
Knowledge collection and capture	.794
Knowledge storage and organization	.811
Knowledge sharing and dissemination	.718
Knowledge application and use	.846
Explicit knowledge management practices	.914
Tacit knowledge management practices	.923

Having established that the sample complied with the assumptions to run parametric tests, an independent *t* test was performed to compare both groups. "The independent *t* test is used in situations in which there are two experimental conditions and different participants have been used in each condition" (Field, 2009, p. 334). The statistical tool (SSPS) provides for the proper calculation when the two group sizes are not equal, as in this case. The first output from the *t* test is shown in Table 13. This output was used for further analysis.

Table 13. Group Statistics

Group	N	Mean	Std deviation	Std error mean
Control	25	5.3200	1.63911	.32782
Experimental	14	3.4143	1.65197	.44151

The numbers of participants were 25 for the control group with mean of 5.32 and a standard deviation of 1.63 and a standard error or .32. The experimental group participants were 14 with a mean of 3.41, a standard deviation of 1.65 and a standard error of .44. Previous tests analyzed the homogeneity of the variance confirming that the variances in the different groups were equal. The output of Table 11 indicated that for the number tasks, the variances were equal for both groups (F[1, 37], non-significant). The output of the *t* test as shown in Table 14 confirmed that the two groups have equal variances because the Levene's test is non-significant ($p > .05$) and the row values for equal variances assumed were used for the analysis.

Table 14. Independent Samples Test

	Levene's test for equality of variances		*T* test for equality of means						
								95% confidence interval of the difference	
	F	Sig.	t	df	Sig. (two-tailed)	Mean diff	Std error diff	Lower	Upper
Equal variances assumed	.507	.481	3.473	37.000	.001	1.90571	.54866	.79402	3.01741
Equal variances not assumed			3.466	26.863	.002	1.90571	.54991	.77713	3.03430

In looking at the two-tailed value, it can be concluded that there was a significant difference between the two samples because $p < .005$. To calculate the effect size the equation shown in Figure 7 was used as suggested by field "calculating the effect size" (Field, 2009, p. 332).

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Figure 7. *T*-test effect size.

The calculated effect size was .30 and this represents a medium effect size. The analysis showed the value of the mean differences to be a positive 1.90 and after looking at Table 13 that showed the value of the control group to be greater than that of the experimental group (5.32 > 3.41) can be concluded that the control group created a higher number of tasks than the experimental group, therefore supporting Null Hypothesis 1, *The inclusion of tacit knowledge in BI does not produce a significant difference in time to complete engineering tasks at customer sites*. On average, the engineers in the control group created larger numbers of tasks (M = 5.32, SE = 3.27) than the experimental group (M = 3.41, SE = 4.41). This difference was significant (t[37] = 3.47, $p < .005$). This represented a medium-sized effect (r = .30). The final analysis expands into the possible reasons for this output.

Even though Table 10 indicated that this was a normally distributed sample, the values were small and to reinforce the parametric tests, this research conducted an additional nonparametric test "Nonparametric tests are sometimes known as assumption-free tests because they make fewer assumptions about the type of data which they can be used" (Field, 2009, p. 540). Furthermore, "when you want to test difference between two conditions and different participants have been used in each condition they you have two choices: the Mann-Whitney test . . . and the Wilcoxon rank-sum test" (Field, 2009, p. 540). This research used the Mann-Whitney test; the output is shown in Table 15.

The effect size was calculated using the equation shown in Figure 8. The result was an effect size of -.50, representing a large effect size. To conclude, the number of tasks were significant different between the control group and the experimental group (U = 61.5, r = 0.50), again supporting Null Hypothesis 1 and rejecting Alternative Hypothesis 1.

Table 15. Test Statistics

Mann-Whitney U	61.500
Wilcoxon W	166.500
Z	-3.335
Asymp. sig. (two-tailed)	.001

$$r = \frac{Z}{\sqrt{N}}$$

Figure 8. Mann-Whitney effect size.

This analysis continued to test Null Hypothesis 2, *tacit knowledge is not a factor that can be used to predict employees' productivity when included into BI.* The analysis was conducted on the group as a whole, using a linear regression between the two factors:

explicit knowledge and tacit knowledge. After running the regression, the overall model summary and ANOVA table were produced (see Tables 16 and 17).

Table 16. Model Summary

				Change statistics					
R	R^2	Adjusted R^2	Std error of estimate	R^2 change	*F* change	*df*1	*df*2	Sig. *F* change	Durbin-Watson
.293[a]	.086	.035	1.83457	.086	1.692	2	36	.199	1.514

Note. Dependent variable: Tasks transformed.

[a]Predictors: (constant), tacit knowledge management practices, explicit knowledge management practices.

Table 17. ANOVAs

	Sum of squares	*df*	Mean square	*F*	Sig.
Regression	11.387	2	5.693	1.692	.199[a]
Residual	121.163	36	3.366		
Total	132.550	38			

Note. Dependent variable: Tasks transformed.

[a]Predictors: (constant), tacit knowledge management practices, explicit knowledge management practices.

Initially, interpreting the R^2 value in the model summary tells that this model explains 8.6% of the number of tasks completed overall. When looking at the *F*-ratio or *F*-statistic (*F*) value on the ANOVA table, a non-significant fit is found. The $F(2, 36) = 1.69$, $p > .05$, and signifies that this model is not a better predictor of tasks than if the mean of number of tasks completed had been used. When looking at the adjusted square on the model summary the adjusted square shrank from 0.08 to 0.03. This adjusted R^2

gave some indication of how well the model generalized. Ideally, the value should be the same or very close to the value of R^2 (Field, 2009, p. 235). The data indicated that the model does not generalize well.

Next, the analysis concentrated on the individual predictors because the model summary included all the predictors at once. For sake of space, some of the output from SPSS was removed and only the important values for the analysis were kept (see Table 18).

Table 18. Coefficients

	Unstandardized coefficients		Standardized coefficients		
Model	*B*	Std error	Beta	*t*	Sig.
Constant	3.623	2.130		1.701	.098
Explicit knowledge	.135	.073	.737	1.836	.075
Tacit knowledge	-.120	.074	-.658	-1.638	.110

Neither explicit knowledge nor tacit knowledge were valid predictors for the number of tasks as seen on *explicit knowledge* (t[36] = 1.83, $p > .05$) and *tacit knowledge* (t[36] = -1.63, $p > .05$). The negative number on the tacit knowledge predictor may be an indication of multicollinearity "multicollinearity: a situation in which two or more variables are very closely linearly related" (Field, 2009, p. 790) this condition was analyzed further in the example. Even though neither of the two factors showed a causal relation to the number of tasks completed, explicit knowledge showed a slightly closer number to the statistical significance. This second test also support Null Hypothesis 2, *tacit knowledge is not a factor that can be used to predict employees' productivity when included into BI,* and rejects Alternate Hypothesis 2. The AMOS model is shown in Figure 9.

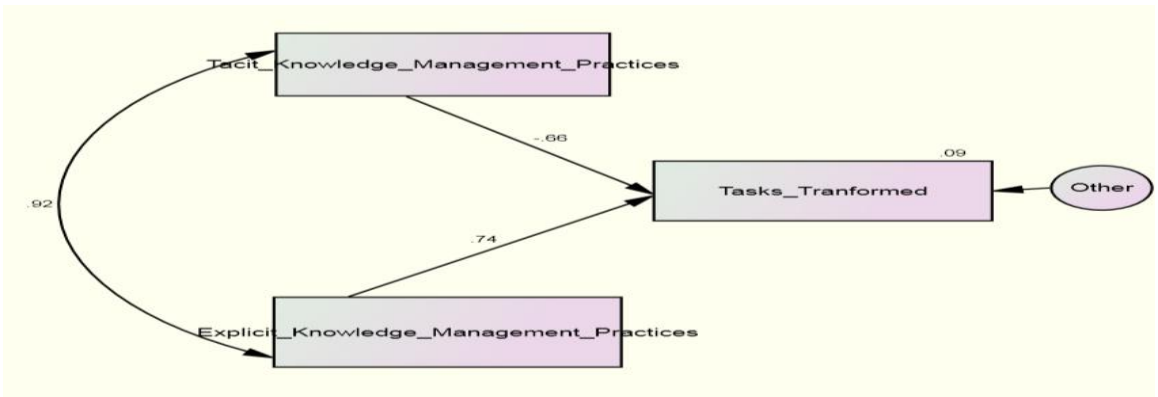

Figure 9. AMOS regression model.

To complement the analysis, several assumptions for linear regression were checked including homoscedasticity, independence of errors, normality of errors, and multicollinearity.

Residuals

According to Field (2009), in an ordinary sample, it would be expected that 95% of cases would have standardized residuals within about ± 2. In addition, 99% of cases should fall within ± 2.5 (p. 244). There were no cases with standardized residuals greater than 3. Table 19 shows that two out of 39 cases were above two, which is within the expected 5%.

Table 19. Casewise Diagnostics

Case number	Std residual	Tasks transformed	Predicted value	Residual
18	2.494	9.00	4.4245	4.57550
28	-2.016	1.40	5.0990	-3.69896

Note. Dependent variable: Tasks transformed.

Normality of Errors

The histogram of the standardized residuals showed a small kurtosis in distribution, also confirmed by the P-P plot of standardized residuals (see Figures 10 and 11). This indicated that the normality assumption was fulfilled because the small value of kurtosis is not much of a concern.

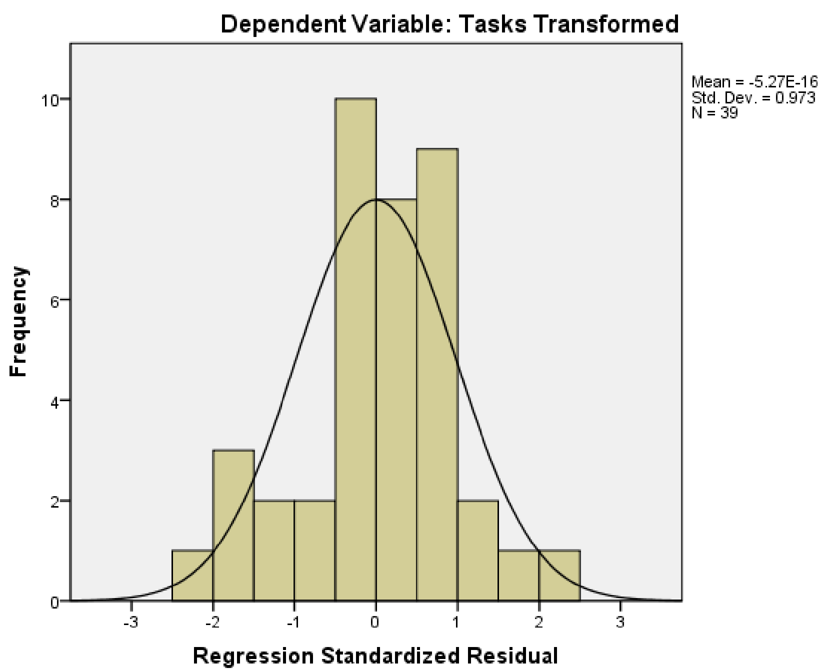

Figure 10. Histogram of standardized residuals.

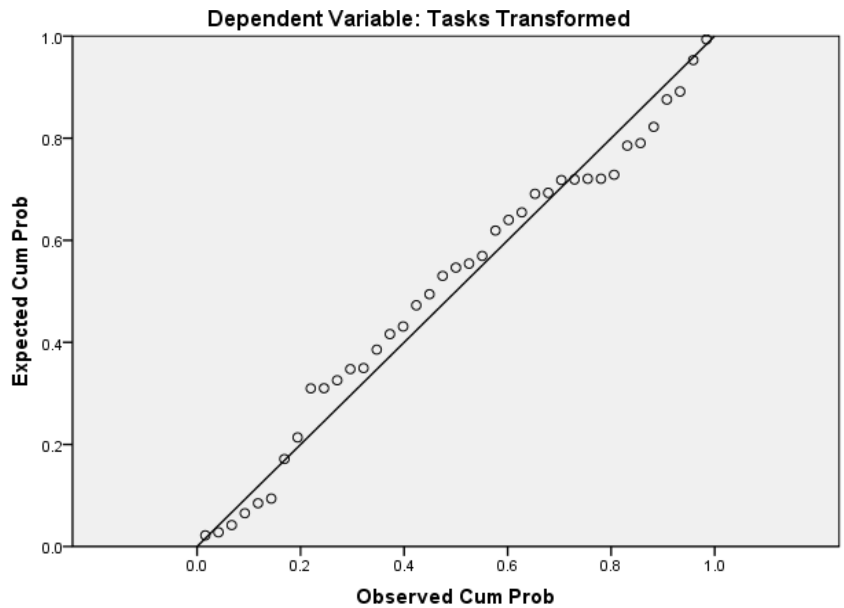

Figure 11. P-P plot of standardized residuals.

Homoscedasticity and Independence of Errors

When looking at the scatterplot of the regression standardize predicted value, it was clear that it did show a random pattern, indicating that the homoscedasticity had been fulfilled (see Figures 12, 13, and 14). In looking at Table 16, the values of Durbin-Watson statistic fall within the recommended values of 1–3, suggesting an independence of errors. "This statistic informs us about whether the assumption of independent errors is tenable. . . . As a conservative rule I suggested that values less than 1 or greater than 3 should definitely raise alarms bells" (Field, 2009, p. 236).

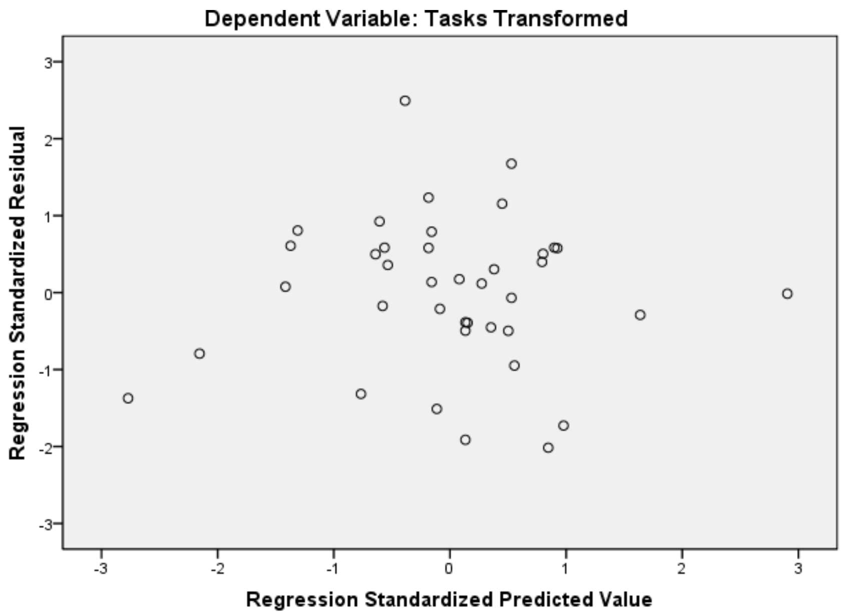

Figure 12. Scatterplot of tasks.

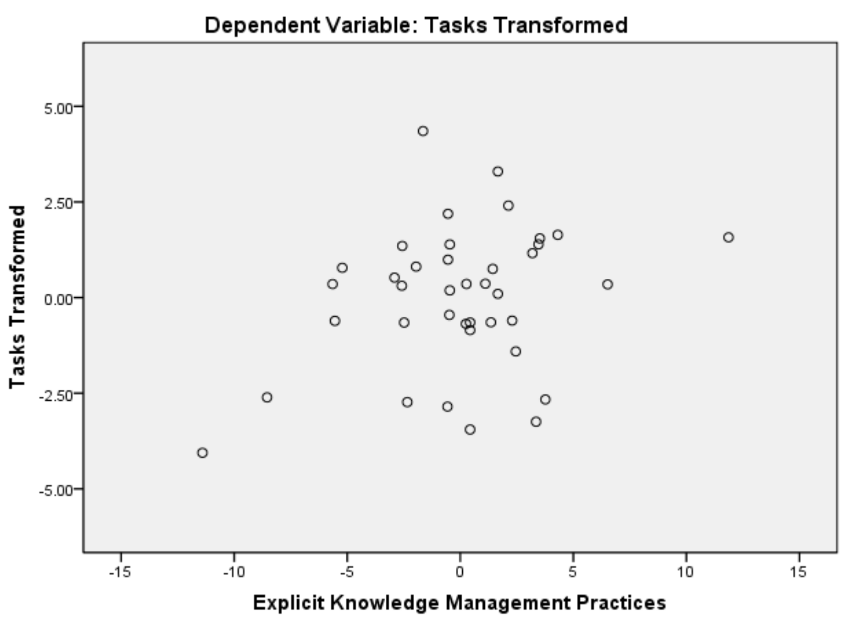

Figure 13. Scatterplot of explicit knowledge partial regression.

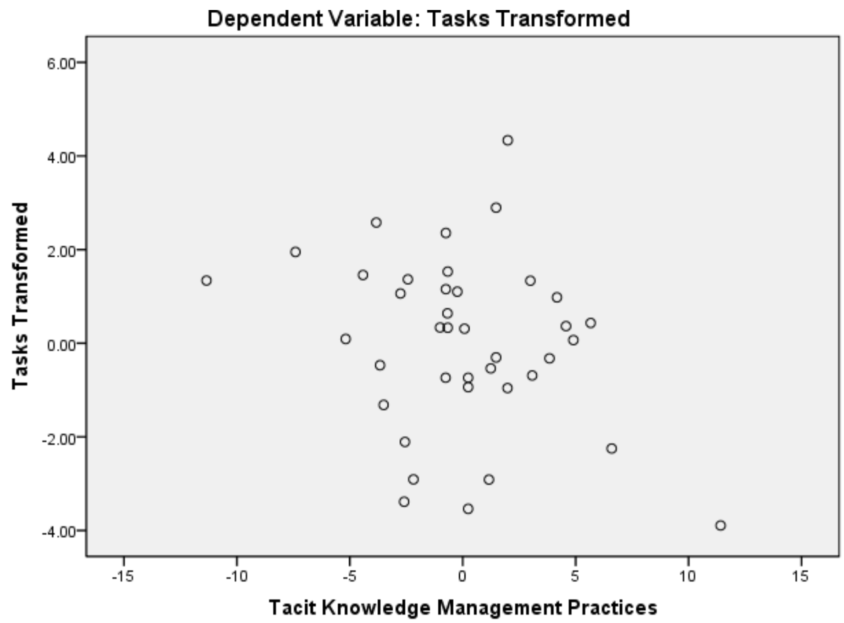

Figure 14. Scatterplot of tacit knowledge partial regression.

Multicollinearity

If the largest Variance Inflation Factor (VIF) is greater than 10, there is a cause for concern. If the average VIF is substantially greater than one, the regression may be biased. Tolerance below 0.1 indicates a serious problem and tolerance below 0.2 indicates a potential problem (Field, 2009). For purposes of space, some columns were removed from Table 20, but the VIF values in it reveal the following: explicit knowledge and tacit knowledge indicates multicollinearity in the data. This can be seen with the tolerance values below 0.2, and the average VIF substantially greater than 1. Additionally, Table 21 shows a correlation between tacit knowledge and explicit knowledge. This may explains the reason why tacit knowledge showed a negative value in the linear regression and provides some indication that both variables in the regression could be reporting the same thing; this is related to multicollinearity and it is one of the outcomes of multicollinearity bias.

Table 20. Collinearity Statistics

Model	Collinearity statistics	
	Tolerance	VIF
Explicit knowledge	.157	6.357
Tacit knowledge	.157	6.357

Table 21. Collinearity Diagnostics

	Eigenvalue Value	Condition Index	(Constant)	Explicit knowledge management practices	Tacit knowledge management practices
1	2.986	1.000	.00	.00	.00
2	.013	15.315	1.00	.04	.04
3	.002	42.862	.00	.96	.96

Additional Regression

To complement the analysis, a regression was performed on the five factors against the numbers of tasks but neither of the factors showed any significance as all the predictors were $p > .05$ (see Table 22).

Table 22. Five Factors Coefficients

Model	Unstandardized coefficients		Standardized coefficients		
	B	Std error	Beta	*t*	Sig.
(Constant)	3.362	2.176		1.545	.132
Knowledge identification and creation	-0.391	0.206	-.887	-1.898	.066
Knowledge collection and capture	0.017	0.214	.041	0.081	.936
Knowledge storage and organization	-0.010	0.236	-.023	-.041	.967
Knowledge sharing and dissemination	0.139	0.164	.302	0.848	.403
Knowledge application and use	0.294	0.197	.655	1.497	.144

Conclusion

This study attempted to answer the following questions:

1. Can the incorporation of customer support unstructured data into the customer support schema increase business performance?

The incorporation of customer support unstructured data into the support schema did not increase business performance; the tacit knowledge sharing from experimental group did not produce the expected results. The expectations were that the experimental group was going to create more tasks than the control group.

2. If unstructured customer support data are converted into explicit knowledge, can this converted data contribute to business performance by increasing engineers' productivity?

The knowledge documents created by the experimental group failed to increase the number of tasks when compared against the control group.

3. To what extent is there a significant decrease in the time to complete field engineers' tasks after unstructured data are incorporated into the BI framework?

There was not a significant decrease in the time to complete engineers' tasks because the experimental group ended creating less number of tasks than the control group.

4. Can the usage of any of the KMAT factors predict field engineers' time to complete tasks when unstructured data are incorporated into the BI framework?

As demonstrated on the last linear regression on Table 22, none of the KMAT factors was good predictor of engineers' productivity

The following hypotheses were developed from the research questions:

- $H1_0$: The inclusion of tacit knowledge in BI does not produce a significant difference in time to complete engineering tasks at customer sites.

- $H1_A$: The inclusion of tacit knowledge in BI produces a significant difference in time to complete engineering tasks at customer sites.

- $H2_0$: Tacit knowledge is not a factor that can be used to predict employees' productivity when included into BI.

- $H2_A$: Tacit knowledge is a factor that can be used to predict employees' productivity when included into BI.

Neither of the alternative hypotheses was accepted concluding that tacit knowledge did not make any difference in the engineers' productivity, contradicting the finding of the literature that this brand of knowledge should have made an impact in business performance.

CHAPTER 5. DISCUSSION AND RECOMMENDATIONS

This chapter will discuss the results from the statistical analysis; it also contains discussion from the conclusions in relation to the literature and the field, discuss limitations and recommendations for future study, and finally, the dissertation conclusion.

In the data analysis, Alternative Hypothesis 1 was rejected and Null Hypothesis 1 accepted. On average, the engineers in the control group created larger numbers of tasks ($M = 5.32$, $SE = 3.27$) than the experimental group ($M = 3.41$, $SE = 4.41$). This difference was significant ($t[37] = 3.47$, $p < .005$) and represented a medium-sized effect ($r = .30$).

Alternative Hypothesis 2 was rejected and Null Hypothesis 2 accepted. Neither explicit knowledge nor tacit knowledge were valid predictors for the number of tasks as seen on *explicit knowledge* ($t[36] = 1.83$, $p > .05$) and *tacit knowledge* ($t[36] = -1.63$, $p > .05$). Even though neither of the two factors showed a causal relation to the number of tasks completed, explicit knowledge showed a closer number to the statistical significance.

This research attempted to demonstrate the importance of tacit knowledge, and its impact in business performance. This research was aligned with Nonaka's (1994) view that tacit knowledge is the most important piece to increase business performance, and based on the literature argued that using externalization one of the modes of the Nonaka theory (converting tacit knowledge into explicit knowledge) could increase business performance. This research investigated the impact of tacit knowledge into business performance after evaluating individual engineers' productivity and was based on the theory of organizational knowledge creation originated from tacit knowledge (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Polanyi, 2009). Tacit and explicit knowledge became the independent variables and individual workers productivity the dependent variable.

The research contained three main constructs and two-level participants. The constructs (tacit knowledge, explicit knowledge and business performance), were operationalized with the help of the KMAT instrument that takes the abstractions and converted them into independent and dependent variables. The two-level participants were two groups of engineers, of which only one group was exposed to the tacit knowledge exchange variable. The model investigated the following research questions: (a) Can the incorporation of customer support unstructured data into the customer support schema increase business performance? and (b) If unstructured customer support data are converted into explicit knowledge; can this converted data contribute to business performance by increasing engineers' productivity? The independent variables were tacit knowledge and explicit knowledge, and the dependent variable the number of tasks performed by engineers using and not using tacit knowledge to complete their tasks. Tacit knowledge was incorporated into the experiments by the sharing of experiences in a knowledge community.

The constructs were based on the theory of organizational knowledge creation that is originated from tacit knowledge (Nonaka, 1991, 1994; Nonaka & Toyama, 2003; Polanyi, 2009). Nonaka (1994), in his article "A Dynamic Theory of Organizational

Knowledge Creation," postulated the theory of organizational knowledge creation. His writing explained that knowledge possessed by individuals, organizations, and societies can be expanded through a spiral process in which tacit knowledge is converted into explicit knowledge, and then back into tacit. Tacit knowledge is hidden behind behaviors, skills competencies and experiences (tacit actionable knowledge) and articulated knowledge resides on individual thoughts and language use. Explicit knowledge resides inside computers in codified form, and by nature has a clear organization (Delen & Al-Hawamdeh, 2009). Regarding tacit knowledge, Polanyi (2009) stated, "I shall reconsider human knowledge by starting the fact that we can know more than we can tell" (p. 4). He also declared, "We recognize the moods of the human face, without being able to tell, except quite vaguely, by what signs we know it" (p. 4), and he classified this human characteristic as *tacit knowledge*, a knowledge that is hard to formalize and communicate. He further stated, "I think I can show that the process of formalizing all knowledge to the exclusion of any tacit knowing is self-defeating" (p. 20)

Nonaka (1994) called the distinction between tacit and explicit knowledge the *epistemological dimension to organizational knowledge*. The exchange can take many forms and based on these variations, different modes of knowledge conversion can be generated. Tacit-to-tacit is a shared experience and is called *socialization*. Explicit-to-explicit is a mode in which modern computers play an important role and is called *combination*. The third and four modes are a combination of the first two: converting explicit into tacit, called *internalization*; and converting tacit into explicit, called *externalization*. On the ontological dimension, the theory posits that individuals are the ones that create knowledge and that organization should amplify this knowledge through the different levels of the firm. The key here is the constant dialog in which the middle-up-down management leadership is the most suitable to crystallize the conversion and creation of knowledge.

This chapter continues the results discussion and attempts to provide answers for the obtained results. It also provides suggestions on how to improve the research with subsequent investigations around the same topic.

Discussion of the Results

As seen in the introduction to this chapter, both alternative hypotheses were rejected, accepting the null hypotheses and clearly contradicting the expectations that tacit knowledge was going to produce an impact on business performance.

Alternative Hypothesis 1 was rejected and Null Hypothesis 1 accepted. On average, the engineers in the control group created larger numbers of tasks ($M = 5.32$, $SE = 3.27$) than the experimental group ($M = 3.41$, $SE = 4.41$). This difference was significant ($t[37] = 3.47$, $p < .005$) and represented a medium-sized effect ($r = .30$).

Alternative Hypothesis 2 was rejected and Null Hypothesis 2 accepted. Neither explicit knowledge nor tacit knowledge were valid predictors for the number of tasks as seen on *explicit knowledge* ($t[36] = 1.83$, $p > .05$) and *tacit knowledge* ($t[36] = -1.63$, $p > .05$). Even though neither of the two factors showed a causal relation to the number of tasks completed, explicit knowledge showed a number slightly closer to the statistical significance. There are many reasons why the analysis produced such results and each one is discussed in detail as follows.

Amount and Quality of Shared Information

The knowledge shared (via documents created from tacit knowledge) was of high quality and ranged from complex tape drive replacements, updating firmware, booting problems and improvements on existing documents between others. Although this is a one-to-many relation, one document can help multiple engineers who are executing the same task. The community created 15 documents altogether and this may be a limitation, and perhaps more participation could have made a difference. The knowledge management community did not have any tool to evaluate how many times the documents were consulted nor it was possible to observe the amount of personal e-mails that engineers exchanged, but the same analysis apply: increasing the amount of participation could have increased the number of tasks completed.

Time Impact of Consulting KM Community

The assumption of this research was that consulting the KM community prior to going onsite, was going to produce a minimal impact, but there is always the possibility that it added extra time to the engineers, therefore impacting the time to complete the number of tasks in a day. The researcher thinks that this was not a factor on this research.

Seasonal Factors

The researched company was a hardware and software customer support company with the same limitations and fluctuations of any economic market and as such, it is not exempted from the same external influences. One influence could have been a seasonal factor related to a high number of support calls or to new equipment bought in the particular period when the control group completed the survey, therefore producing larger number of tasks than the experimental group. The higher the number of the support calls, the more possibilities for engineers to complete tasks at customer sites. A better model could have been for the control group to wait for two months while the experimental group was being exposed to tacit knowledge, and then to have both groups completing the survey. The researcher did not want the control group to wait for the experimental group to finish the experiment, to avoid any contamination or a casual exposure to any of the experimental members and subsequently to be exposed to the tacit knowledge independent variable. The researcher is suggesting a new model that is discussed in the research design improvement section.

Cultural Factors

The research company is extended across Latin American countries and the random process chose engineers across different areas. There is a possibility that the acceptance process could have been naturally biased by some regions creating more participants than others regions. This phenomenon could have created the chance of some regions producing more support business than others do, due to a natural economic grow.

Sample Size

Another possibility to obtain the non-expected results could have been the size of the sample combined with the cultural factors. Perhaps a larger sample that combines more regions and cultures can produce different results.

Research Design

This research used the two-group posttest-only randomized experiment as seen in Figure 3. Another possible design is to use a pretest-posttest control group as seen in Figure 15. A pretest-posttest control group design is “to determine the effects of a treatment by comparing a treatment group with a controlled group sample" (Russ-Eft & Hoover, 2005, p. 86).

R------------O-------X--------O

R------------O-----------------O

Figure 15. Pretest-posttest control group experiment.

The difference of this proposed new design against the one used is that the control group and the experimental group are tested twice, once before the experiment and once after the experiment, and this would give a more clear picture of the influence of the independent variable into the experiment.

Research Framework

This research argued that using *externalization* one of the modes of the Nonaka (1994) theory (that is, converting tacit knowledge into explicit knowledge) could have increased business performance. This is one of the modes of knowledge exchange as postulated by Nonaka. The exchange can take many forms and, based on these variations, different modes of knowledge conversion can be generated. Tacit-to-tacit is a shared experience called *socialization*. Explicit-to-explicit, in which modern computers play an important role, is called *combination*. The third and four modes are a combination of the first two: converting explicit knowledge into tacit knowledge, called *internalization*; and converting tacit knowledge into explicit knowledge, called *externalization*.

Perhaps an improved research framework could expand the research framework used in this experiment that was described in Figure 2 and could include the *socialization* dimension into the framework. This inclusion may be shaped in the form of a social network channel to complement the *tacit to explicit* with a *tacit to tacit* knowledge exchange. This social network tool must provide the ability to record all the exchanges to help with the final evaluation and could include a qualitative analysis of the exchanges. The new framework is displayed in Figure 16, and it is depicted by a circle around the tacit to explicit dimension and represents the *socialization* dimension.

Discussion of Conclusions in Relation to Literature and Field

The constructs of this research were based on the seminal work of Nonaka (1994) Nonaka, in his article, “A Dynamic Theory of Organizational Knowledge Creation,” postulated the theory of organizational knowledge creation. His writing explained that knowledge possessed by individuals, organizations, and societies can be expanded through a spiral process in which tacit knowledge is converted into explicit knowledge and then back into tacit.

This was the first attempt to evaluate the impact of tacit knowledge into individual productivity and it cannot be concluded that the results as such contradicted the previous findings about the importance of tacit knowledge into business performance, but indicated that more tuning is needed on the different factors that may have contributed to the obtained results.

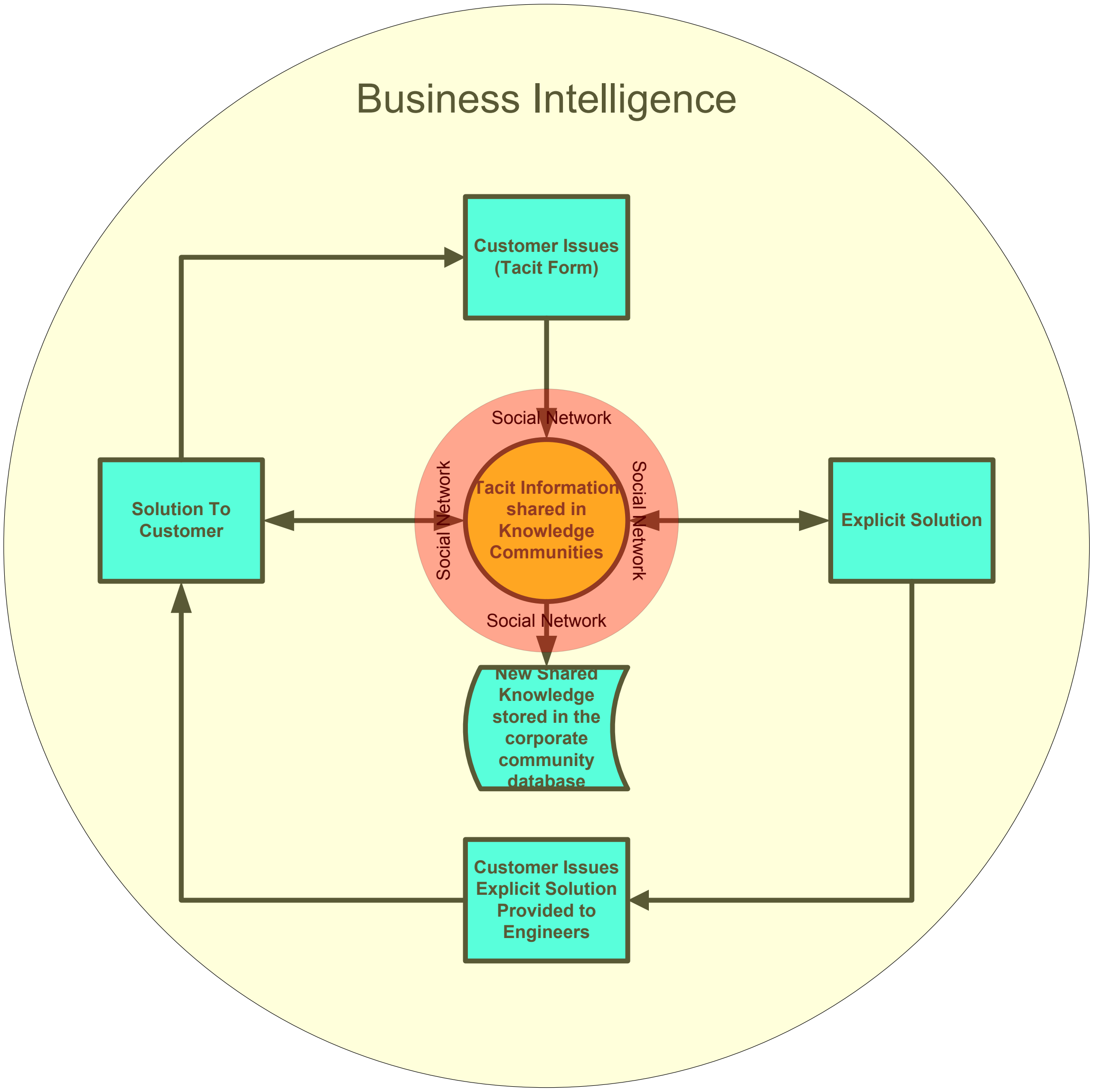

Figure 16. Possible new research framework.

When the researcher was compiling the results and had sent an e-mail to the participants indicating the end of the experiment, one of the participants sent an e-mail expressing his satisfaction with the experiment, and stated that it was a positive initiative. He also said that the synergy created due to the experiment had motivated engineers to share their tacit knowledge a phenomenon that was not occurring before. Perhaps extending the length of the experiment, changing the design, or the design framework and introducing more variables into the research, can provide more indications of what else is influencing the engineers' productivity and obtain different results.

Limitations

One of the limitations of this research was the size of the sample. Originally the field engineer division was reported to be 300 employees, but when creating the invitations, the researcher found that a good quantity of those 300 were managers, dispatch engineers and onsite engineers that had to be removed from the sample because they do not execute tasks at customers, so the sample frame was reduced. The response rate was what was expected from an Internet survey, but a larger response could have produced different results. Another limitation of this study is that the population sample was collected from a software/hardware engineering company and the results may not be generalized to other types of companies.

Recommendations for Further Research or Intervention

The theory of tacit knowledge is solid and contains many factors of influence. The following recommendations may help to shed some light to get to conclusions on the impact of tacit knowledge into individual productivity. The investigation found that the control group executed a larger amount of tasks than the experimental group. The two group results were spread two months apart to allow the experimental group to receive the influence of the dependent variable tacit knowledge. One recommendation is to change the research design to a pretest-posttest design as seen in Figure 15, extend the time of the tacit knowledge influence, and expand the research to more geographical areas. Additionally, culture, age, and other factors could be included in the new design as indirect variables of influence as well as including socialization into the research framework.

Conclusion

This research found that tacit knowledge did not influence individual engineers' productivity and as such affecting business performance. Additionally, it found that tacit knowledge was not a factor that could be used to predict individual productivity. This research was the first attempt to investigate individual productivity in relation to tacit knowledge and created more questions than answers by providing preliminary results that can be used to expand the research that is based on Nonaka (1991) and Polanyi (2009). Tacit knowledge is hidden behind behaviors, skills, competencies, and experiences (tacit actionable knowledge) and articulated knowledge (implicit knowledge), which resides in individual thoughts and language use. Explicit knowledge resides inside computers in codified form and by nature has a clear organization (Delen & Al-Hawamdeh, 2009). The very nature and difficulty to measure tacit knowledge is what makes this investigation very valuable to pursue. This research laid out a very important foundation for the investigation of the impact of tacit knowledge on business performance.

REFERENCES

Ab Hamid, M. R., Mustafa, Z., Idris, F., Abdullah, M., & Suradi, N. R. M. (2011). Measuring value-based productivity: A confirmatory factor analytic (CFA) approach. *International Journal of Business & Social Science, 2*(6), 85–93.

Abidin, S. Z. Z., Idris, N. M., & Husain, A. H. (2010, March). *Extraction and classification of unstructured data in WebPages for structured multimedia database via XML.* Paper presented at the International Conference on Information Retrieval & Knowledge Management, Shah Alam, Selangor, Malaysia.

Antonova, A., Csepregi, A., & Marchev, A., Jr. (2011). How to extend the ICT used at organizations for transferring and sharing knowledge. *IUP Journal of Knowledge Management, 9*(1), 37–56. Retrieved from Business Source Complete database.

Arun, S., & Atish, P. S. (2005). A comparison of data warehousing methodologies. *Communications of the [Association for Computing Machinery] ACM, 48*(3), 79–84. doi:10.1145/1047671.1047673

Bara, A., Botha, I., Diaconiţa, V., Lungu, I., Velicanu, A., & Velicanu, M. (2009). A model for business intelligence systems' development. *Informatica Economica, 13*(4), 99–108.

Barlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal, 19*(1), 43–50.

Bates, R. A. (2005). Multivariate research methods. In R. A. Swanson & E. F. Holton (Eds.), *Research in organizations: Foundations and methods of inquiry* (pp. 115–141). San Francisco, CA: Berrett-Koehler.

Bierly, P. E., III, Kessler, E. H., & Christensen, E. W. (2000). Organizational learning, knowledge and wisdom. *Journal of Organizational Change Management, 13*(6), 595–618.

Blake, R. H., & Mangiameli, P. (2009). Structured or unstructured: In either form, consistency is critical. In *Proceedings for the [Northeast Region Decision Sciences Institute] NEDSI* (pp. 179–184).

Bosch-Sijtsema, P. M., Ruohomaki, V., & Vartiainen, M. (2009). Knowledge work productivity in distributed teams. *Journal of Knowledge Management, 13*(6), 533–546. doi:10.1108/13673270910997178

Bose, R. (2008). Competitive intelligence process and tools for intelligence analysis. *Industrial Management + Data Systems, 108*(4), 510-528.

Britt, P. (2009). Insurers improve productivity with KM. *[Knowledge Management] KM World, 18*(9), 10–29.

Brown, J. S., & Duguid, P. (2000). *The social life of information*. Boston, MA: Harvard Business School Press.

Byrne, B. M. (2010). *Strucural equation modeling with AMOS* (2nd ed.). New York, NY: Routledge, Taylor & Francis Group.

Cader, Y. (2007). Knowledge management and knowledge-based marketing. *Journal of Business Chemistry, 4*(2), 46–58. Retrieved from Business Source Complete database.

Chang Lee, K., Lee, S., & Kang, I. W. (2005). KMPI: Measuring knowledge management performance. *Information & Management, 42*(3), 469–482. doi:10.1016/j.im.2004.02.003

Chatzoglou, P. D., Vanezis, P., & Christoforidis, S. (2005, September). *Knowledge management practices in the financial sector.* Paper presented at the Knowledge-Based Economy: Management of Creation & Development.

Cheung, C. F., Lee, W. B., & Wang, Y. (2005). A multi-facet taxonomy system with applications in unstructured knowledge management. *Journal of Knowledge Management, 9*(6), 76–91.

Chowdhury, S. (2005). The role of affect- and cognition-based trust in complex knowledge sharing. *Journal of Managerial Issues, 17*(3), 310–326.

Choy, K. L., Tan, K. H., & Chan, F. T. S. (2007). Design of an intelligent supplier knowledge management system—An integrative approach. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 221*(2), 195–211. doi:10.1243/09544054jem627

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Erlbaum.

Cooper, D., & Schindler, P. (2008). *Business research methods* (10th ed.). Boston, MA: McGraw-Hill.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Boston, MA: Harvard Business School Press.

Davies, J., Duke, A., Kings, N., & Mladenic, D. (2005). Next generation knowledge access. *IUP Journal of Knowledge Management, 9*(5), 64-84.

Davies, J., Fensel, D., & van Harmelen, F. (2003). *Toward the semantic web: Ontology-driven knowledge management*. West Sussex, England: Wiley.

Delen, D., & Al-Hawamdeh, S. (2009). A holistic framework for knowledge discovery and management. *Communications of the [Association for Computing Machinery] ACM, 52*(6), 141–145. Retrieved from Business Source Complete database.

Dioşteanu, A., & Cotfas, L. (2009). Agent based knowledge management solution using ontology, semantic web services and GIS. *Informatica Economica, 13*(4), 90–98. Retrieved from Business Source Complete database.

Douglas, I. (2009). An object and performance framework for implementation of web-based knowledge sharing technology. *Journal of Theoretical & Applied Electronic Commerce Research, 4*(1), 57–71.

Du, T. C., Li, F., & King, I. (2009). Managing knowledge on the web—Extracting ontology from HTML web. *Decision Support Systems, 47*(4), 319–331. doi:10.1016/j.dss.2009.02.011

Eftekharzadeh, R. (2008). Knowledge management implementation in developing countries: An experimental study. *Review of Business, 28*(3), 44–58. Retrieved from Business Source Complete database.

Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the [Association for Computing Machinery] ACM, 49*(9), 77–82.

Faucher, J.-B. P. L., Everett, A. M., & Lawson, R. (2008). Reconstituting knowledge management. *Journal of Knowledge Management, 12*(3), 3–16. doi:10.1108/13673270810875822

Feghali, T., & El-Den, J. (2008). Knowledge transformation among virtually-cooperating group members. *Journal of Knowledge Management, 12*(1), 92–105. doi:10.1108/13673270810852412

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.

Filius, R. e., de Jong, J. A., & Roelofs, E. C. (2000). Knowledge management in the HRD office: A comparison of three cases. *Journal of Workplace Learning, 12*(7), 286–295.

Giovinazzo, W. A. (2003). *Internet-enabled business intelligence*. Upper Saddle River, NJ: Prentice Hall PTR.

Godbole, S., & Roy, S. (2008, July). *Text to intelligence: Building and deploying a text mining solution in the services industry for customer satisfaction analysis.* Paper presented at the IEEE International Conference on Services Computing, Honolulu, Hawaii.

Goel, A., Rana, G., & Rastogi, R. (2010). Knowledge management as a process to develop sustainable competitive advantage. *South Asian Journal of Management, 17*(3), 104–116. Retrieved from Business Source Complete database.

Grundspenkis, J. (2007). Agent based approach for organization and personal knowledge modelling: Knowledge management perspective. *Journal of Intelligent Manufacturing, 18*(4), 451–457.

Hendler, J. (2005). Knowledge is power: A view from the semantic web. *[Artificial Intelligence] AI Magazine, 26*(4), 76–84.

Holste, J. S., & Fields, D. (2010). Trust and tacit knowledge sharing and use. *Journal of Knowledge Management, 14*(1), 128–140. doi:10.1108/13673271011015615

Hou, J.-J., & Chien, Y.-T. (2010). The effect of market knowledge management competence on business performance: A dynamic capabilities perspective. *International Journal of Electronic Business Management, 8*(2), 96–109. Retrieved from Business Source Complete database.

Hui-Ling, T., & Yu-Hsuan, C. (2011). Effects of empowering leadership on performance in management team. *Journal of Chinese Human Resources Management, 2*(1), 43–60. doi:10.1108/20408001111148720

Inmon, B. (2005). Looking ahead: Unstructured data. *DM Review, 15*(12), 81–87.

Inmon, W. H., & Nesavich, A. (2009). *Tapping into unstructured data*. New York, NY: Pearson Education.

Inthout, R., Vrancken, J., & Schrijnen, P. (2010). Wiki-based knowledge management in a transport consultancy, a case study. *Electronic Journal of Information Systems Evaluation, 13*(2), 133–142. Retrieved from Business Source Complete database.

Kane, K., Robinson-Combre, J., & Berge, Z. (2010). Tapping into social networking: Collaborating enhances both knowledge management and e-learning. *VINE, 40*(1), 62–70. doi:10.1108/03055721011024928

Kobielus, J. (2008). Quality really is job no. 1. *[Customer Relationship Management] CRM Magazine, 12*(1), 48.

Koskinen, K. U. (2004). Knowledge management to improve project communication and implementation. *Project Management Journal, 35*(2), 13–19. Retrieved from Business Source Complete database.

Kuechler, W. L. (2007). Business applications of unstructured. *Communications of the [Association for Computing Machinery] ACM, 50*(10), 86–93.

Lau, K.-N., Lee, K.-H., & Ho, Y. (2005, August). Text mining for the hotel industry. *Cornell Hotel and Restaurant Administration Quarterly, 46*(3), 344–362.

Lausen, H., Ding, Y., Stollberg, M., Fensel, D., Hernandez, R. L., & Han, S.-K. (2005). Semantic web portals: State-of-the-art survey. *Journal of Knowledge Management, 9*(5), 40–49.

Lee, J., Upadhyaya, S. J., Rao, H. R., & Sharman, R. (2005). Secure knowledge management and the semantic web. *Communications of the [Association for Computing Machinery] ACM, 48*(12), 48–54. Retrieved from Business Source Complete database.

Lopes, I. T. (2008). Towards an electronic knowledge management culture. *[Institute of Chartered Financial Analysts of India] ICFAI Journal of Knowledge Management, 6*(4), 7–24. Retrieved from Business Source Complete database.

Luebbers, D., Grimmer, U., & Jarke, M. (2003, September). *Systematic development of data mining-based data quality tools*. Paper presented at the 29th International Conference on Very Large Databases, Berlin, Germany.

Mahmood, S., & Ali, B. (2011). An empirical investigation on knowledge workers productivity in telecom sector of Pakistan. *Information Management and Business Review, 3*(1), 27–38.

Martins, A., & Lopes dos Reis, F. (2010). The model of intellectual capital approach on the human capital vision. *Journal of Management Research, 2*(1), 1–20.

McCampbell, A. S., Clare, L. M., & Gitters, S. H. (1999). Knowledge management: The new challenge for the 21st century. *Journal of Knowledge Management, 3*(3), 172–179.

Mehta, N. (2008). Successful knowledge management implementation in global software companies. *Journal of Knowledge Management, 12*(2), 42–56. doi:10.1108/13673270810859505

Metaxiotis, K., Ergazakis, K., Samouilidis, E., & Psarras, J. (2003). Decision support through knowledge management: The role of the artificial intelligence. *Information Management & Computer Security, 11*(5), 216–221.

Mezher, T., Abdul-Malak, M. A., Ghosn, I., & Ajam, M. (2005). Knowledge management in mechanical and industrial engineering consulting: A case study. *Journal of Management in Engineering, 21*(3), 138–147. doi:10.1061/(asce)0742-597x(2005)21:3(138)

Mingrui, H., & Yongjian, L. (2010). Exploiting distributed cognition to make tacit knowledge explicating. *Journal of Software Engineering & Applications, 3*(3), 273–279. doi:10.4236/jsea.2010.33033

Mundra, N., Gulati, K., & Vashisth, R. (2011). Achieving competitive advantage through knowledge management and innovation: Empirical evidences from the Indian IT sector. *IUP Journal of Knowledge Management, 9*(2), 7–25. Retrieved from Business Source Complete database.

Negash, S. (2004). Business intelligence. *Communications of [Association for Information Systems] AIS,* 13, 177–195.

Newbert, S. L. (2008). Value, rareness, competitive advantage, and performance: A conceptual-level empirical investigation of the resource-based view of the firm. *Strategic Management Journal, 29*(7), 745–768.

Nonaka, I. (1991). *The knowledge-creating company*. Boston, MA: Harvard Business Press.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14–37. Retrieved from Business Source Complete database.

Nonaka, I., & Toyama, R. (2003). The knowledge-creating theory revisited: Knowledge creation as a synthesizing process. *Knowledge Management Research & Practice, 1*(1), 2–10.

Nonaka, I., & von Krogh, G. (2009). Tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory. *Organization Science, 20*(3), 635–652.

Passmore, D. L., & Baker, R. M. (2005). Sampling strategies and power analysis. In R. A. Swanson & E. F. Holton (Eds.), *Research in organizations: Foundations and methods of inquiry* (pp. 45–55). San Francisco, CA: Berrett-Koehler.

Pintar, D., Vranic, M., & Skocir, Z. (2007, September). *Intelligent data sources and integrated data repository as a foundation for business intelligence analysis.* Paper presented at the Software, Telecommunications and Computer Networks (SoftCOM) 15th International Conference, Split, Croatia.

Plejie, B., Vujnovie, B., & Penco, R. (2008, December). *Transforming unstructured data from scattered sources into knowledge.* Paper presented at the IEEE International Symposium on Knowledge Acquisition and Modeling Workshop (KAM), Wuhan, China.

Polanyi, M. (2009). *The tacit dimension*. Chicago, IL: The University of Chicago Press.

Popovic, A., Turk, T., & Jaklic, J. (2010). Conceptual model of business value of business intelligence systems. *Management: Journal of Contemporary Management Issues, 15*(1), 5–30.

Puleston, J. (2011). Improving online surveys. *International Journal of Market Research, 53*(4), 557–560. doi:10.2501/ijmr-53-4-557-562

Ranjan, J. (2008). Business justification with business intelligence. *VINE, 38*(4), 461–475.

Rao, R. (2003). From unstructured data to actionable intelligence. *IT Professional, 5*(6), 29–35.

Reid, A., & Catterall, M. (2005). Invisible data quality issues in a CRM implementation. *Journal of Database Marketing & Customer Strategy Management, 12*(4), 305–314.

Reychav, I., & Weisberg, J. (2009). Good for workers, good for companies: How knowledge sharing benefits individual employees. *Knowledge and Process Management, 16*(4), 186–197. doi:10.1002/kpm.335

Reychav, I., & Weisberg, J. (2010). Bridging intention and behavior of knowledge sharing. *Journal of Knowledge Management, 14*(2), 285–300. doi:10.1108/13673271011032418

Russ-Eft, D., & Hoover, A. L. (2005). Experimental and quasi-experimental designs. In R. A. Swanson & E. F. Holton (Eds.), *Research in organizations: Foundations and methods of inquiry* (pp. 75–95). San Francisco, CA: Berrett-Koehler.

Sahay, B. S., & Ranjan, J. (2008). Real time business intelligence in supply chain analytics. *Information Management & Computer Security, 16*(1), 28–48.

Seidler-de Alwis, R., & Hartmann, E. (2008). The use of tacit knowledge within innovative companies: Knowledge management in innovative enterprises. *Journal of Knowledge Management, 12*(1), 133–147. doi:10.1108/13673270810852449

Singh, S. K. (2008). Role of leadership in knowledge management: A study. *Journal of Knowledge Management, 12*(4), 3–15.

Smith, E. A. (2001). The role of tacit and explicit knowledge in the workplace. *Journal of Knowledge Management, 5*(4), 311–321.

Srivastava, J., & Cooley, R. (2003). Web business intelligence: Mining the web for actionable knowledge. *[Institute for Operations Research and the Management Sciences] INFORMS Journal on Computing, 15*(2), 191–207.

Subashini, R. (2010). Tacit knowledge—The ultimate essence of an organization. *Advances in Management, 3*(8), 36–39. Retrieved from Business Source Complete database.

Swanson, R. A. (2005). The Challenge of Research in Organizations. In R. A. Swanson & E. F. Holton (Eds.), *Research in organizations: Foundations and Methods of Inquiry* (pp. 3-26). San Francisco: Berrett-Koehler.

Swartz, N. (2007). Gartner warns firms of 'dirty data.' *Information Management Journal, 41*(3), 6.

Tarafdar, M., TU, Q., Ragu-Nathan, B. S., & Ragu-Nathan, T. S. (2007). The impact of technostress on role stress and productivity. *Journal of Management Information Systems, 24*(1), 301–328.

The University of New South Wales. (2004). *Knowledge management selected abstracts and citations*. Retrieved from http://med.unsw.edu.au/medweb.nsf/resources/reports1/$file/Knowledge+Management+categories+061004+final.pdf

Trochim, W. M. K. (2006). *Two-group experimental designs*. Retrieved from http://www.socialresearchmethods.net/kb/expsimp.php

Venkatraman, N., & Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *Academy of Management Review, 11*(4), 801–814.

Voth, D. (2005). Minding your business with AI. *IEEE Intelligent Systems, 20*(4), 4–7.

Wagner, R. K., & Sternberg, R. J. (1987). Tacit knowledge in managerial success. *Journal of Business & Psychology, 1*(4), 301–312.

Wang, H., & Wang, S. (2008). A knowledge management approach to data mining process for business intelligence. *Industrial Management + Data Systems, 108*(5), 622–634. doi:10.1108/02635570810876750

Watson, H. (2007). The believability of data. *Business Intelligence Journal, 12*(4), 4-6.

Xiao, G., & Kim, J.-O. (2009). The investigation of Chinese consumer values, consumption values, life satisfaction, and consumption behaviors. *Psychology & Marketing, 26*(7), 610–624.

Yoon, V., Broome, B., Singh, R., & Guimaraes, T. (2005). Using agent technology for company knowledge management. *Information Resources Management Journal, 18*(2), 94–113.

Zack, M. H. (1999a). Developing a knowledge strategy. *California Management Review, 41*(3), 125–145. Retrieved from Business Source Complete database.

Zack, M. H. (1999b). Managing codified knowledge. *Sloan Management Review, 40*(4), 45–58. Retrieved from Business Source Complete database.

Zazzara, P. (2001). Operationalizing knowledge management in health care. *Topics in Health Information Management, 21*(3), 1–7.

Zhou, N., Cheng, H., Chen, H., & Xiao, S. (2007, September). *The framework of text-driven business intellappendigence.* Paper presented at the Wireless Communications, Networking and Mobile Computing (WiCom) International Conference, Shanghai, China.