Volume 9, Number 1 **ISSN 1524-7252**

ACADEMY OF INFORMATION AND MANAGEMENT SCIENCES JOURNAL

An official Journal of the

Academy of Information and Management Sciences

James W. Carland, Jr.
Editor
Western Carolina University

Editorial and Academy Information
may be found on the Allied Academies web page
www.alliedacademies.org

The Academy of Information and Management Sciences is a subsidiary of the Allied Academies, Inc., a non-profit association of scholars, whose purpose is to support and encourage research and the sharing and exchange of ideas and insights throughout the world.

Whitney Press, Inc.

Printed by Whitney Press, Inc.
PO Box 1064, Cullowhee, NC 28723
www.whitneypress.com

Allied Academies is not responsible for the content of the individual manuscripts. Any omissions or errors are the sole responsibility of the individual authors. The Editorial Board is responsible for the selection of manuscripts for publication from among those submitted for consideration. The Publishers accept final manuscripts in digital form and make adjustments solely for the purposes of pagination and organization.

The *Academy of Information and Management Sciences Journal* is published by the Allied Academies, Inc., PO Box 2689, 145 Travis Road, Cullowhee, NC 28723, USA, (828) 293-9151, FAX (828) 293-9407. Those interested in subscribing to the *Journal*, advertising in the *Journal*, submitting manuscripts to the *Journal,* or otherwise communicating with the *Journal*, should contact the Executive Director at info@.alliedacademies.org.

ACADEMY OF INFORMATION AND MANAGEMENT SCIENCES JOURNAL EDITORIAL BOARD MEMBERS

ACADEMY OF INFORMATION AND MANAGEMENT SCIENCES JOURNAL

CONTENTS

LETTER FROM THE EDITOR

Welcome to the *Academy of Information and Management Sciences Journal,* the official journal of the Academy of Information and Management Sciences. The Academy is one of several academies which collectively comprise the Allied Academies. Allied Academies, Incorporated is a non-profit association of scholars whose purpose is to encourage and support the advancement and exchange of knowledge.

The editorial mission of the *AIMSJ* is to publish empirical and theoretical manuscripts which advance the disciplines of Information Systems and Management Science. All manuscripts are double blind refereed with an acceptance rate of approximately 25%. Manuscripts which address the academic, the practitioner, or the educator within our disciplines are welcome. And, diversity of thought will always be welcome.

Please visit the Allied Academies website at www.alliedacademies.org for continuing information about the Academy and the *Journal*, and about all the Allied Academies and our collective conferences. Submission instructions and guidelines for publication are also provided on the website.

Jim Carland
Western Carolina University
www.alliedacademies.org

2+2 TIER BANDED FRAMEWORKS OF INTERCONNECTEDNESS: INDUSTRY STRUCTURE DETERMINANTS

SeungJae Shin, Mississippi State University - Meridian
Jack Tucci, Mississippi State University - Meridian
Martin B. H. Weiss, University of Pittsburgh
Hector Correa, University of Pittsburgh*

ABSTRACT

The Internet industry is generally considered to be vertically structured with the Internet Backbone Provider (IBP- long distance service carrier) in the upstream and Internet Service Providers (ISP) in the downstream. Although there are many ISPs and IBPs in each stream, both markets are considered independent oligopolies in that there are a few dominant providers for both ISPs and IBPs. The market leaders in each market create their own hierarchical tier and it is generally accepted that the Internet industry structure has evolved into a four-tier hierarchical structure. To understand the Internet industry, it is necessary to understand interconnection between ISPs and IBPs. The key element as an industry structural determinant is peering interconnection and the relationship created by that interconnectedness. Peering interconnection occurs within the same tier and the transit interconnection between the different tiers. This paper examines the internet industry structure using market share and interconnection strategies.

INTRODUCTION

The Internet industry is dynamic. It consists of millions of computers and switching devices. The number of Internet Service Providers (ISPs) increased rapidly from the mid 90' and the structure of the industry continues to change. It is widely accepted that today's Internet industry has vertical structure: over 40 Internet Backbone Providers (IBPs) including 5 top-tier backbones constitute the upstream industry (Kende, 2000) and over 10,000 ISPs for accessing the Internet make up the downstream industry (Weinberg, 2000). A backbone provider service is critical for any ISPs desiring to connect to the Internet. There are manifold interconnections between ISPs and IBPs. Moving data from one interconnection (tier-to-tier (IBP-to-ISP), or IBP-to-IBP, ISP-to-ISP) to another is the catalyst changing the Internet market structure. Compounding this is when users change mode of access from narrowband (dial-up) to broadband (DSL or Cable Modem). Narrowband dial-up access has been a major way to connect the internet, but in the summer of 2004, the number of broadband internet users surpassed the number of narrowband dial-up users. In this

paper, we analyze the dynamic internet industry (both IBP and ISP) using market share and Internet interconnection strategies and dissect its complicated industry structure.

VERTICAL INDUSTRY STRUCTURE

Overview of Internet Industry

The Internet industry integrates the equipment, software, and organizational infrastructure required for Internet communications. As a rough approximation it can be said it is divided into two components: IBPs that transfer communications in bulk among network exchange points, and ISPs that (1) receive communications from individuals or institutions and transfer them to an IBP's network, and (2) receive communications from a IBP and transfer them to their destination. ISPs are used to refer to any company who can offer Internet connectivity. Some people use ISPs as a general term including IBPs. Some people argue that ISPs can be differentiated from other types of online information services, such as CompuServe or American On Line, because ISPs do not provide content but they focus only on providing Internet connectivity.

Generally speaking, the Internet industry has a vertical structure: Upstream IBPs provide an intermediate good and downstream ISPs using this input sell connectivity to their customers. Simplified, we suggest the analogy that the relationship between IBPs and ISPs is very much like that of "wholesalers and retailers."

In reality the internet is much more complex. The ISPs themselves are networks of users that may directly exchange information among each other. In addition, the IBPs may provide services directly to users and also may interconnect with other IBPs. A gestalt perspective gives the understanding that the internet is a network of networks that is accessible in many parts of the world. Since the telephone industry is tightly intertwined with the Internet industry, we begin by with its examination.

Telephone Industry and its Relationship with Internet Industry

Public Switched Telephone Network (PSTN) was designed and optimized for the transmission of the human voice. In the United States, telephone service is divided into two industries: (1) local telephone service provided by Local Exchange Carriers (LECs) and (2) long distance telephone service provided by Interexchange Carriers (IXCs). This structure creates a vertical hierarchy: Upstream IXCs provide the connection between LECs, and the downstream LECs have direct access to telephone users. The hierarchical structure of telephone industry has a strong impact on the Internet industry. This is due to the fact that many of these companies provide some type of service to either or both the ISP and IBP and/or the end user. End users in this case can be either public, private, or governmental users.

Traditionally, a LEC was a monopoly that served a specific geographic region without competition. Even after deregulation, LECs are still considered by many to be a local monopoly, especially for residential customers. In the U.S., the local telephone services provided flat-rate billing, that is, a telephone user can originate local calls as many times and as long as he wishes with only monthly flat rate. This type of billing system has been a great influence on the growth of Internet access market. This is not true for foreign countries where consumers often pay for time and distance for each call, whether it is local or long distance. Although not the topic of this paper, it is a good example why the internet is so expensive in countries where time and distance charges are levied against each call as compared to local telephone rates in the U.S.

The long distance market is now generally considered to be a very competitive market, though it too was a monopoly less than thirty years ago. Users can make a long distance call with pre-selected IXC through their LEC. Any IXC that wishes to handle calls originating in a local service area can build a switching office, called a Point of Presence (POP). The function of the POP is to interconnect networks so that any site where networks interconnect may be referred to as a POP.

Dial-up access using PSTN is the most universal form of Internet access. In the U.S., a modem call is typically a local call without a per-minute charge. ISP's lines are treated as a business telephone user not as a carrier, so they are not required to pay the measured Common Carrier Line Charge (CCLC) for originating and terminating calls, which recovers part of the cost of the local loop. The switching system in LEC's POP connects calls between Internet users and ISP's modem pool so the LECs' facilities support dial-up Internet communications. In addition, IBPs and large ISPs often construct their backbone networks by leasing lines from IXCs and LECs. As a result, we can say that telephone industry provides the basic infrastructure for the Internet industry.

INTERNET INTERCONNECTION STRATEGY

Background of Internet Interconnection

There are two types of Internet interconnection among ISPs and IBPs: peering and transit. The only difference among theses types is in the financial rights and obligation that they generate to their customers (Weiss & Shin, 2004). To understand the relationship between peering and transit, it is necessary to recall the non-commercial origin of the Internet. During the Internet's early development, there was only one backbone and only one customer, the military, so interconnection was not an issue. In the 1980s, as the Internet was opened to academic and research institutions, the National Science Foundation (NSF) funded the NSFNET as an Internet backbone. Around that time, the Federal Internet Exchange (FIX) served as a first point of interconnection between federal and academic networks. At the time that commercial networks began appearing, general commercial activity on the Internet was restricted by Acceptable Use Policy (AUP), which prevented the commercial networks from exchanging traffic with one another using the NSFNET as the backbone

(Kende, 2000). In the early 1990s, a number of commercial backbone operators including PSINet, UUNET, and CerfNET established the Commercial Internet Exchange (CIX) for the purpose of interconnecting theses backbones and exchanging their end users' traffic. The NSF decided to stop operating the NSF backbone which was replaced by four Network Access Points (NAPs) (Minoli and Schmidt, 1998). NAPs are public interconnection points where major providers interconnect their network and these connections consist of high-speed switchs or a network of switches to which a number of routers can be connected for the purpose of traffic exchange. The function of NAPs is similar to major airport hubs; all ISPs and IBPs are gathered at the NAPs to connect each other. The NSF required that any ISP receiving government contracts or receiving money from public universities must connect to all of the NAPs. After the advent of CIX and NAPs, commercial backbones developed and a system of interconnection known as peering quickly evolved.

Peering Interconnection Strategy

The term "peering" is sometimes used generically to refer to Internet interconnection with no financial settlement, known as a "Sender Keeps All (SKA)" or "Bill and Keep," which can be thought of as payments or financial transfers between ISPs in return for interconnection and interoperability (Cawley, 1997). Peering can be divided into several categories: (1) according to its openness, it can be private peering or public peering, (2) according to the numbers of peering partners it can be Bilateral Peering Arrangement (BLPA) or Multilateral Peering Arrangement (MLPA), and (3) according to the market in which it occurs, it can be primary peering in the backbone market or secondary peering in the downstream market (Weiss and Shin, 2004). A peering arrangement is based on equality, that is, ISPs of equal size would peer. The measures of size could be (i) geographic coverage, (ii) network capacity, (iii) traffic volume, (iv) size of customer base, or (v) a position in the market. The ISPs would peer when they perceive equal benefit from peering based on their own subjective terms (Kende, 2000).

The original four NAPs were points for public peering. As the Internet traffic grew, the NAPs suffered from congestion. Therefore, direct circuit interconnection between two large IBPs was introduced (called bilateral private peering) which takes place at a mutually agreed place of interconnection. This private peering is opposed to public peering that takes place at the NAPs. It is estimated that 80 percent of Internet traffic is exchanged via private peering (Kende, 2000).

Before the Internet privatization, the NSF was responsible for the operation of the Internet. There are probably around 50 major NAPs world-wide in the internet, most of which are located in the U.S. (Moulton, 2001). As the Internet continues to grow, NAPs suffer congestion because of the enormous traffic loads. Because of the resulting poor performance, private direct interconnections between big IBPs were introduced, called peering points.

From the interconnection perspective, NAPs are the place for public peering. Anyone who is a member of NAP can exchange traffic based on the equal cost sharing. Members pay for their own router to connect to the NAP plus the connectivity fee charged by the NAP. Historically, in

public peering, there was no discrimination for interconnection among the service providers (no priority given or taken based on usage). On the other hand, the direct interconnection between two equal sized IBPs is bilateral private peering, which takes place at the mutually agreed place of interconnection.

According to Block (Cukier, 1999), there are two conditions necessary for the SKA peering, that is, peering with no settlement, to be viable: (1) The traffic flows should be roughly balanced between interconnecting networks; and (2) the cost of terminating traffic should be low in relation to the cost of measuring and billing for traffic. The conclusion drawn from the above observations is that peering is sustainable under the assumption of mutual benefits and avoidance of costly, unnecessary traffic measuring. Nevertheless, peering partners would make a peering arrangement if they each perceive that they have more benefits than costs from the peering arrangement. Most ISPs in the U.S. historically have not metered traffic flows and accordingly have not erected a pricing mechanism based on usage. Unlimited access with a flat rate is a general form of pricing structure in the Internet industry. Finally, peering makes billing simple: no metering and no financial settlement.

Peering benefits come mainly from the network externality. Network externalities arise when the value or utility that a customer derives from a product or service increases as a function of other customers of the same or compatible products or services; that is, the more users there are, the more valuable the network is (Kende, 2000). There are two kinds of network externalities in the internet. One is direct network externality: the more E-mail users, the more valuable the internet. The other is indirect network externality: the more internet users there are, more web content will be developed which makes the internet even more valuable for its users. The ability to provide direct and indirect network externalities to customers provides an almost overpowering incentive for ISPs to cooperate with one another by interconnecting their networks (Kende, 2000). Contributing to the motivation for peering is lower latency because cooperation makes it necessary for only one hop to exchange traffic between peering partners.

Transit Interconnection Strategy

Transit is an alternative arrangement between ISPs, in which one pays another to deliver traffic between its customers and the customers of other provider. The relationship of transit arrangement is hierarchical: a provider-customer relationship. Unlike a peering relationship, a transit provider will route traffic from the transit customer to its peering partners. With transit agreements, usually small IBPs are able to receive and send communications using the facilities of large IBPs, and must pay a fee for these services. A concern related to transit is that while small IBPs do not have to pay in the case of peering through NAPs, they must pay a transit fee if they directly connect to one of the large IBPs. Before the commercialization of the Internet, carriers interconnected without a settlement fee, regardless of their size. However, after the Internet's commercialization, the large IBPs announced their requirements for a peering arrangement, and any carrier who could

not meet those terms would be required to pay a transit fee in addition to the interconnection costs. An IBP with many transit customers has a better position when negotiating a peering arrangement with other IBPs. Another difference between peering and transit is existence of a Service Level Agreement (SLA), which describes outage and service objectives, and the financial repercussion for failure to perform. In a peering arrangement, there is no SLA to guarantee rapid resolution of problems. In case of an outage, both peering partners may try to resolve the problem, but it is not mandatory. This is one of the reasons peering agreements with a company short on competent technical staff are broken. In a transit arrangement it is a contract and customers could ask the transit provider to meet the SLA. Many e-commerce companies prefer transit to peering for this reason. A one minute outage cause the IBP, ISP, and the customer losses, hence, rapid recovery is critical to their business. Furthermore, in the case of transit, there is no threat to quit the relationship while in the case of peering a non-renewal of the peering agreement is a threat. ISPs are not permitted to form transit relationship over public NAPs because these are designed as a neutral meeting place for peering. When purchasing transit service, ISPs will consider other factors beside low cost: performance of the transit provider's backbone, location of access nodes, number of directly connected customers, and a market position.

ANALYSIS OF BACKBONE AND ACCESS MARKETS

Internet Backbone Market

With some simplification, it can be said that the IBPs receive communications in bulk from POPs or NAPs and distribute them to other POPs or NAPs close to the destination. To make the Internet a seamless network, the IBPs have multiple POPs distributed over the whole world. Most frequently they are located in large urban centers. These POPs are connected to each other with owned or leased optical carrier lines. Typically, these lines are 622 Mbps (OC-12) or 2.488 Gbps (OC-48) circuits or more, as defined by the SONET hierarchy, a standard for connecting fiber-optic transmission systems. These POPs and optical carrier lines make up the IBP backbone network. The IBP's POPs, are also connected with the POPs of many ISPs. The relationship between an ISP's POP and IBP's is the same as that of ISPs and IBPs.

According to Erickson (2001), the North American backbone market had around 36 IBPs in the first quarter of 2001. However, these numbers misinterpret the Internet backbone market structure because this market is highly concentrated. There were 11,888 transit interconnections between backbone and access markets in April 2000 (McCarthy, 2000). Counting by the number of connections to downstream market, MCI/Worldcom is a dominant player in the backbone market with 3,145 connections and Sprint is the second largest backbone provider with 1,690 connections and AT&T (934 connections) and C&W (851 connections) are the third and the fourth.

Several of the large IBPs are subsidiaries of large telephone companies such as AT&T, MCI/WorldCom, Sprint, etc. Since these companies own the infrastructure needed for telephone

services, they are very favorably positioned to provide the facilities and equipment required by the IBPs. In addition, due to their economies scale, they are able to offer large volume discount rates or bundling agreements of both telephone and Internet lines for the services they provide. This is possible because the Internet industry is lightly, if at all, regulated. In particular, there are no regulations with respect to the tariffs that can be charged for the services provided. From these observations it follows that the large IBPs, supported by the large telephone companies, are in a position to capture large shares of the upstream market.

According to the Carlton and Perloff (1999), the most common measure of concentration in an industry is the share of sales by the four largest firms, called the 'C4' ratio. Generally speaking, if the C4 ratio is over 60, the market is considered a tight oligopoly. For the upstream backbone market this ratio based on the 1999 U.S. backbone revenue (Worldcom 38%, Genuity 15%, AT&T 11%, Sprint 9%) is 73, which shows high concentration in the market. The entry barrier is also high because there is a large sunk cost for nationwide backbone lines and switching equipment. The number of IBPs for the past three years shows just how high the entry barrier in the backbone market is: 43 (1999), 41(2000), and 36 (2001) (Erickson, 2001). The slight reduction for last three years is caused by mergers and acquisitions and reclassification. According to the number of players, we conclude that the overall backbone market is stable although oligarchic. In addition there are significant economies of scale and the rapid pace of technological change generates a large amount of uncertainty about the future return on investments. It is not easy to enter this market without large investments and advanced technology.

The interconnection price is usually determined by the provider's relative strength and level of investment in a particular area (Halabi, 2000). It is certain T-1 transit price has been decreasing continuously. In 1996, the internet connectivity for T-1 was $3,000 per month with $1000 setup fee (Halabi, 1997). According to Martin (2001), the average price of T-1 connection in 1999 was $1,729. In 2000, it was $1,348. In 2001, it was $1,228. One of reasons for decreasing T-1 interconnection price is advent of substitute services for T-1 line, such as wireless Internet access technology (LMDS, Satellite), digital subscriber line (DSL) technology, and cable-modem technology, which exerts a downward pressure on T-1 prices. As technology continues to improve and data transmission rates increase, pressure will continue to maintain a cap on service.

Internet Access Market

An ISP's product is public access to the Internet, which includes login authorization, e-mail services, some storage space, and possibly personal web pages. The ISP's coverage area is usually determined by the existence of an ISP's POP within the local telephone area. ISPs are classified as local, regional, and national according to the scope of their service coverage. The distribution of ISPs is presented in the Table 2.

Among 307 telephone area codes in U.S., the largest ISP covers 282 area codes and the smallest covers only 1 area code. The ISPs with 1 to 10 area codes constitute 79.81% of the total

number of ISPs. This explains that most of ISPs in the downstream market are small, local companies. Some of these small ISPs are subsidiaries or affiliates of CLECs (Competitive Local Exchange Carriers), which are small telephone companies established in the 1990s as a result of telephone industry deregulation.

Table 1: Distribution of ISPs by their coverage
(Erickson, 2001)

Telephone area codes covered by ISP	Percentage	Type
1	35.14%	Local
2-10	44.67%	Local / Regional
11-24	4.11%	Regional
25-282	16.08%	National

AOL-Time Warner is a dominant player in the dial-up access market. According to Goldman (2004), AOL-Time Warner had 22.8 million subscribers in the 3rd Quarter of 2004. AOL-Time Warner's subscribers are 22.8 million (24%) out of 81.1 million U.S. subscribers (Goldman, 2004). The Table 2 shows top 10 dial-up ISPs ranked by the number of subscribers.

Table 2: Top 10 U.S. ISPs (Q3 2004)
(Goldman, 2004)

Rank & ISP	Subscribers	Market Share	Rank & ISP	Subscribers	Market Share
(1) AOL	22.8M	24.0%	(6) Road Runner	3.9M	4.1%
(2) United Online	6.6M	6.9%	(7) Verizon	3.3M	3.5%
(3) Comcast	6.5M	6.8%	(8) Coax	2.4M	2.5%
(4) EarthLink	5.2M	5.7%	(9) BellSouth	1.9M	2.0%
(5) SBC	4.7M	4.9%	(10) Charter	1.8M	1.9%

In the downstream access market the C4 ratio is 43 and would be defined as relatively concentrated. However, the entry barrier in the downstream market is much lower than in the backbone market. Since subscribers can utilize the PSTN line to connect ISPs' modems and ISPs purchase business telephone lines from a LEC, ISPs for dial-up service do not have to invest in access lines to individual subscribers. They can build POPs to link to the PSTN and other ISPs. Since a T-1 lines prices and telecom equipment prices are currently dropping quickly, a large number of small ISPs are possible, especially in the less densely populated areas. The number of

North American ISPs for the past several years is an evidence of low entry barrier in the downstream market: 1447 (February 1996), 3640 (February 1997), 4470 (February 1998), 5078 (March 1999), 7463 (April 2000), and 7288 (March 2001) (Erickson, 2001).

Most ISPs provide unlimited Internet access with a monthly flat rate. For major national ISPs, the price ranges generally from $0 to $25 per month dependent on the level of service. Some ISPs provide Internet access service with zero monthly subscription fees to their customers; their revenues depend completely on Internet advertising income. Some base their service on speed, while others on memory usage as an upgrade to their standard service.

ISPs are free to make local peering arrangements with other ISPs. Cremer and Tirole (2000, p445) call this local secondary peering. The Pittsburgh Internet Exchange (PITX) is an example of local peering arrangement. Without this local peering, all network traffic passing from one Pittsburgh network to another had to be sent through Washington, D.C., Chicago, or New York City. The sending and receiving networks pay an unnecessary cost for this inefficient handling of data that should have remained local. Participants in this local exchange point reduce their costs and improve performance and reliability for their local Internet traffic with the equal basis of cost recovery. However, this kind of peering is confined to local internet traffic. Outbound traffic (connected to other networks through an IBP) to other areas still has to depend on the IBP's transit service.

Broadband Internet Market

The Internet access technologies are roughly divided into two categories: narrow band access using dial-up modem technology and broadband access such as Cable-Modem, DSL, and wireless broadband access technology. Among the above broadband access technologies, DSL and Cable-Modem are the two dominant broadband access methods. According to Vara (2004), in July of 2004, more than half of U.S. internet users connected to the internet using a broadband service. It was the first time high-speed broadband internet connection had more market share than dial-up connection. The broadband service providers usually confine their business to high density regions because broadband service requires large investments in "advanced" (read expensive) technologies. Internet users in the rural area rarely have a chance to enjoy the benefit of the higher speed access that broadband services offer.

According to Leichtman Research Group, the twenty largest cable and DSL providers in the US account for about 95% of the market in high speed internet access. The top broadband providers now account for over 30.9 million high-speed Internet subscribers, with cable having nearly 18.8 million broadband subscribers, and DSL trailing behind at 12.2 million subscribers. If we confine the access market into the broadband technology, the C4 ratio in this market is 55% which is considered oligarchic. The following table shows the top 10 broadband access providers in the U.S.

Table 3: Top 10 U.S. Broadband ISPs (Q3 2004)
(Lehichman Research Group, http://www.itfacts.biz)

Rank & ISP	Subscriber	Market Share	Rank & ISP	Subscriber	Market Share
(1) Comcast (Cable)	6.5M	20%	(6) Bell South (DSL)	1.9M	6%
(2) SBC (DSL)	4.7M	14%	(7) Charter (Cable)	1.8M	6%
(3) Time Warner (Cable)	3.7M	11%	(8) Adelphia (Cable)	1.3M	4%
(4) Verizon (DSL)	3.3M	10%	(9) Cablevision (Cable)	1.3M	4%
(5) Cox (Cable)	2.4M	7%	(10) Qwest (DSL)	1.0M	3%

DE-PEERING AND FOUR-TIER HIERARCHICAL STRUCTURE

In 1996, AGIS was the first IBP to unilaterally terminate peering arrangements. After that, a series of IBPs announced that they were ending peering with many of their previous peering partners and were no longer accepting peering arrangements from other networks whose infrastructure would not allow the exchange of similar levels of traffic access. Instead of peering, they would charge those smaller ISPs for transit. Finally, the large IBPs moved away from public NAPs to private peering or maintained relatively small capacities like T3 in the NAPs and then placed themselves in a new hierarchy, so called top-tier IBPs (Jew and Nicolls, 1999). Most top-tier IBPs are subsidiaries or affiliates of the major facilities-based telecommunication carriers. They are UUNET (Worldcom), C&W, Genuity, AT&T WorldNet, and Sprint, the 'so called' Big 5. They don't need transit service from others and they make peering arrangement with each other. Over 80% of the U.S. backbone traffic is estimated to pass through their systems and switches (Weinberg, 2000). Other non Big 5 IBPs make peering arrangements among themselves and simultaneously purchase transit services from the Big 5.

There are two cases for which peering is generally refused: (1) Regional IBPs which do not have a national backbone network and (2) content providers or web hosting companies, so called web farms. The main reason for this refusal is a free-rider issue. Peering partners generally meet in a number of geographically dispersed locations. In order to decide where to pass traffic to another, they have adopted what is known as "hot-potato routing," where an ISP will pass traffic to another backbone at the earliest point of exchange. Under the hot-potato routing rule, someone who does not have a national backbone network must transport its traffic on the others' backbone networks. In addition to that, asymmetric traffic patterns, which occur in file transfer or web surfing, result in increased capacity costs without commensurate revenues.

Some of the Big 5 recently disclosed their policy for peering but some of them still do not. There is an unwritten rule shared by the Big 5 about their peering standard: (i) A coast to coast national backbone with a certain level of bandwidth requirement, (ii) a number of presences in the major exchange points, (iii) 7 days by 24 hours Network Operation Center (NOC) and highly

experienced technical staffs, and (iv) a certain level of traffic ratio between inbound and outbound, usually 1:4. It is indeterminate what the exact requirements for the private peering are since peering agreements are under non-disclosure. Without a doubt, these requirements could be a significant entry barrier for any new entrant.

PSINet, which was one of the large IBPs, used a peering standard called "open peering policy", that was different from the Big 5. It would peer with any ISP including local, regional, and national except for companies whose primary business was web hosting or content collection. Some of the Big 5 did not want to peer with PSINet, because some of PSINet's private peering partners are transit customers of the Big 5. Whenever the Big 5 upgrade their networks, they upgrade their peering policy. From the tier-2 IBP's point of view, peering requirements are getting tougher and tougher. Nobody can enter into the top tier group without their approval. This cartel-like behavior has been an important issue in the Internet industry for several years and will eventually be a sticking point with the Federal Trade Commission in the future.

After being refused peering in the backbone market, ISPs in the downstream access market, usually operating in a limited geographic region, tried to peer among themselves. Cremer and Tirole (2000) in their paper call this kind of peering "local secondary peering." This is a major factor in proliferation of local and regional Internet exchange points. These smaller exchange points (compared to NAPs) referred to as Metropolitan Exchange Points (MXPs) (OECD, 1998).

Most of Internet exchange points, or POPs of major IBPs are located near the metropolitan areas, which are far from the rural areas. The local ISPs in a rural area have to lease long expensive lines to reach an interconnection point. The long distance from private or public peering points is an additional obstacle to overcome for the rural ISPs. Through de-peering in one tier and peering in lower tier in both markets, the four-tier hierarchical structure has emerged; in the backbone market, tier-1 IBPs with their nationwide backbones interconnect each other and make a core network in the Internet and tier-2 IBPs with their regional backbones interconnect each other and pay tier-1 IBPs for connectivity to rest of internet, which mean they are customers of tier-1 providers. A few of nationwide big ISPs are also a member of tier-2 group. In the access market, tier-3 regional ISPs are customers of tier-1 or tier-2 connecting them to the rest of internet. Local, small ISPs are tier-4 providers and they are customers of higher tier providers. However, the demarcation between the tiers is not clear. In the following section we explain how peering decisions are made.

PEERING DECISION-MAKING PROCESS

An interconnection strategy may be different according to its priority. If expense of interconnection is the number one issue, ISPs will try to find as many peering partners as they can and try to choose minimum combination costs among them. Or if performance is the top priority, they may prefer private peering or transit to public peering. All interconnection decisions start from the analysis of their own traffic. An ISP should try to find the available options and negotiate with

their interconnection partners for interconnection methodology, interconnection line capacity, interconnection settlement, etc. This process will be explained below in detail. (Norton, 1999)

Phase I: Identification of ISP's Traffic

The costs of peering and transit vary according to the distance of the ISPs' POP and interconnection point. Generally, the cost of transit is more expensive than that of peering. Before deciding on a transit or peering arrangement, the ISP may systematically sample inbound and outbound traffic flows and then map these flows to the originating Autonomous System (AS), which is defined as a collection of networks that are under the administrative control of a single organization and that share a common routing strategy. Calculations are made to determine where to reduce the load on the expensive transit paths.

Phase II: Finding Potential Interconnection Partners

Based on the traffic map and the aforementioned analysis, ISPs try to find interconnection partners. Because peering policies are often exposed only under Non-Disclosure Agreements (NDA), it is not easy to know them in advance of negotiations. It is reasonable for an ISP to find its peering partners in its own level of Internet industry hierarchy. If a top-tier IBPs makes a peering arrangement with a second tier IBP, then the latter could be the formers' competitor. Therefore, a higher tier ISP would prefer selling transit service to lower tier ISPs and have an incentive to reduce the number of their own competitors. Many ISPs, except for top tier IBPs, have adopted a hybrid approach to interconnection, peering with a number of ISPs and paying for transit from one or more IBPs in order to have access to the transit provider as well as the peering partners of the transit provider.

Phase III: Implementing Interconnection Methodology

Since peering is seen as being of mutual benefit, both parties explore the interconnection methods that will most effectively exchange traffic. Both parties decide (1) how many interconnection points they have, (2) where to locate the interconnection points, (3) how they interconnect, private peering or public peering, (4) what line capacity they will use, (5) settlement free or settlement involved, etc.

Table 6 illustrates comparison of per Mega-bit cost of transit, private peering, and public peering. If we compare cost per Mbps shipped (CPMS) per month of OC3 capacity, the order from the cheapest is public peering ($30), private peering ($64~$129), and transit ($464).

Table 4: Per M-bit Cost Comparison
(AT&T (Transit), Norton (2000, Private Peering), and Chicago NAP (Public Peering))

Interconnection Type	Capacity	Cost / Capacity	Per M bit Cost
Transit	DS3	$26,000/45M	$578/M
	OC3	$72,000/155M	$464/M
Private Peering	OC3	($10,000~$20,000) /155M/2	$64/M ~ $129/M
	OC12	($20,000~$30,000) /622M/2	$32/M ~$48/M
Public Peering	DS3	$3,900/45M	$87/M
	OC3	$4,700/155M	$30/M

CONCLUSION

The Internet has become an important social and business tool. Furthermore, the market has become even more dynamic since it was privatized. Peering has emerged as a phenomenon that can at one time be beneficial to both parties while simultaneously discriminate against one of the peering partners. Professor Frieden calls it a "balkanization" in the Internet. If a new technology was introduced in this market, the internet providers with this technology would have a tendency from past practices to create their own peer groups to make money against the provider without this technology. On the other hand, ISPs are competitors and cooperators simultaneously: competitors for market share and cooperators for global connectivity. One ISP's decision has an influence on other ISP's decisions. Thus, they have a strong dependence on each other beyond just competitive factors.

In our paper, we believe that the two-tier, two layer market structure of both backbone and access is oligopolistic. This means, if a new technology is developed lowering costs, or increasing speed, or if some of them reach an agreement they could exercise their market power, maybe for the better for the consumer, maybe not. A policy maker's goal for the Internet industry is continuing growth and innovation. To achieve this goal, regulators need to continue to encourage competition and to give incentives for ongoing investment and in the development and deploying of new technologies, which will benefit consumers in the internet market. Therefore, it is a role of internet policy makers to make socially desirable competitive environments between higher tier ISBs and lower tier ISPs in the Internet industry.

REFERENCES

Carlton, D. & J. Perloff, (1999). *Modern Industrial Organization*, 3rd Edition. New York:Addison Wesley Longman.

Cawley, R.A. (1997). Interconnection, pricing, and settlements: Some healthy jostling in the growth of the Internet. In Kahin, B. & J. Keller (1997). *Coordinating the Internet* (pp. 346-376). Boston, Massachusetts: MIT Press.

Cremer, J., P. Rey, & J. Tiroel, (2000). Connectivity in the Commercial Internet. *The Journal of Industrial Economics, 40*(8), 433-472.

Erickson, T., (2001). *Introduction to the Directory of Internet Service Providers*, 13th Edition. Boardwatch Magazine, from http://www.ispworld.com/isp.

Frieden, R. (1998). Without public peer: The potential regulatory and universal service consequences of Internet Balkanization. 3 *Virginia Journal of Law & Technology* 8, from http: //vjolt.student.virginia.edu/.

Goldman, A., (2004). Top 23 U.S. ISPs by Subscriber: Q3 2004. *ISP-Planet magazine*, from http://www.isp-planet.com/research/ranking/usa.html

Greene, B.R., (2000). L2 Internet exchange Point (IXP) using a BGP Route Reflector. Draft version 0.4, from http://www.cisco.com/public/cons/isp/ixp.

Halabi, B., (1997). *Internet Routing Architectures*. Indianapolis, IN: Cisco Press.

Halabi, B., (2001). *Internet Routing Architectures* 2nd Edition. Indianapolis, IN: Cisco Press.

Jew, B. & Nicholls, R., (1999), *Internet Connectivity: Open Competition In the Face of Commercial Expansion*. 1999 Pacific Telecommunications Conference.

Kende, M., (2000). The Digital Handshake: Connecting Internet Backbones. FCC-OPP Working Paper No. 32, from http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp32.pdf.

Martin, L., (2001). Backbone Web Hosting Measurements. *Boardwatch Magazine*, from http://www.ispworld.com/isp/Performance_Test.htm.

Minoli, D. & A. Schmidt, (1999). *Internet Architectures*. New York, NY: Wiley.

Moulton, P., (2001). *The Telecommunications Survival Guide.* Upper Saddle River, NJ: Prentice-Hall.

Norton, W.B., (1999). Internet Service Providers and Peering, draft version 1.9, from http://www.nanog.org/mtg-0010/tree.doc.

OECD, (1998). Internet Traffic Exchange: Development and Policy, from http://www.oecd.org/dataoecd/11/26/2091100.pdf

Vara, V., (2004). High-Speed Surpasses Dial-up As Top Home Web Access in U.S., from http://www.wsj.com.

Weinberg, N., (2000). Backbone Bullies. *Forbes*, June 12 edition, 236-237.

Weiss, M., & S. Shin, (2004). Internet interconnection economic model and its analysis: Peering and settlement, *Netnomics, 6,* 43-57.

ENDNOTES

* Dr. Correa, an international scholar and professor in the Graduate School of Public and International Affairs at the University of Pittsburgh who passed away in the summer of 2004. We deeply appreciate his effort in this paper.

THE IMPACT OF GENERAL AND SYSTEM-SPECIFIC SELF-EFFICACY ON COMPUTER TRAINING LEARNING AND REACTIONS

Bassam Hasan, University of Toledo
Jafar Ali, Kuwait University

ABSTRACT

Recognizing the multilevel nature of the computer self-efficacy (CSE) construct, this study distinguishes between general and system-specific CSE and examines the effects of both levels of CSE on two key training outcomes: reactions and learning performance. Reactions were examined with respect to perceived ease of use and perceived usefulness, whereas learning performance was examined in terms of near-transfer and far-transfer learning. The results of a field experiment revealed that general CSE had significant effects on perceived ease of use and far-transfer learning. Conversely, system-specific CSE demonstrated significant effects on near-transfer and far-transfer learning. The results of this study provide several valuable research and practical implications for IS training and systems acceptance and usage.

INTRODUCTION

End users' computing skills and abilities have been correlated with individual performance, success of information systems, and organizational competitiveness (Byrd & Turner, 2001). As a result, businesses are increasingly trying to design and provide effective computer training in which end users can learn and master the skills necessary to use computer systems competently (Yi & Davis, 2001). The need for effective computer training is further exacerbated by the rapid changes in computer technologies and the growing diversity of end users who must use computers to perform their daily job activities. Thus, computer training has become a pivotal issue in information systems (IS) research and practice, and deserves further investigation and better understanding.

Computer training has received extensive research attention in recent years (Davis & Bostrom, 1993; Johnson & Marakas, 2000; Lu, Yu, & Liu, 2003; Simon & Werner, 1996; Simon, Grover, Teng, & Whitcomb, 1996; Yi & Davis, 2003; Yi & Davis, 2001). The impetus behind this increased research activity has been to understand how to design more effective computer training (Agarwal, Sambamurthy, & Stair, 2000; Bostrom, Olfman, & Sein, 1990; Simon & Werner, 1996; Simon et al., 1996; Yi & Davis, 2003). Among the myriad of variables investigated in past research as determinants of computer training performance, computer self-efficacy (CSE) - one's confidence in his or her computing skills - had shown positive effects on various outcomes associated with

computer training (Agarwal et al., 2000; Compeau & Higgins, 1995a; Gist, Schwoerer, & Rosen, 1989; Johnson & Marakas, 2000; Yi & Im, 2004).

However, a review of past research concerning the relationship between CSE and computer training outcomes reveals some important shortcomings. Firstly, although CSE has been shown to be a multilevel construct that operates at general and system-specific levels (Agarwal et al., 2000; Johnson & Marakas, 2000; Marakas, Yi, & Johnson, 1998; Yi & Davis, 2003), past research has predominantly focused on CSE as a one-dimensional, system-independent construct and little attention has been given to investigations of CSE at the system or application level. Secondly, whilst theories of learning distinguish between near-transfer and far-transfer learning, this distinction has not been adequately utilized in computer training contexts (Cormier & Hagman, 1987; Gururajan & Fink, 2002). Thus, it remains unclear how the two levels of CSE beliefs affect the two types of learning. Finally, while reactions represent an important criterion for evaluating training programs (Kirkpatrick, 1995), very few studies have considered reactions as a computer training outcome (Mahapatra & Lai, 2005). Consistent with recent studies (Bedard, Jackson, Ettredge, & Johnston, 2003), we investigate reactions with respect to perceived ease of use and perceived usefulness.

The present study seeks to address the shortcomings described above. Thus, this study proposes and empirically tests relationships among the following variables: general CSE, system-specific CSE, perceived ease of use, perceived usefulness, near-transfer learning, and far-transfer learning. In addition to the apparent research implications of addressing existing voids in the literature, this study provides interesting implications for practice. With the substantial resources that organizations are committing to end-user computer training and the vital consequences of such training, a better understanding of the factors affecting training outcomes will certainly be useful in planning, administering, and evaluating computer training programs.

RESEARCH BACKGROUND

Computer self-efficacy

Self-efficacy is a principal concept in Bandura's (1986; 1997) social cognitive theory (SCT). Self-efficacy refers to people's confidence in their abilities to organize and execute required skills, in order to perform a behavior successfully (Bandura, 1986). According to this definition, self-efficacy is not concerned with assessments or the sum of the actual skills that individuals may possess, but it is instead concerned with the level of one's confidence in his or her ability to use whatever skills he or she may have to perform a behavior successfully. Self-efficacy alters human behavior through its effects on the choices of behavior that people make, the amount of effort they put to exert on a behavior, the level of persistence they demonstrate to overcome difficulties, and the difficulty of goals they set for themselves.

Since self-efficacy is a dynamic construct that varies across domains (Bandura, 1986; Gist & Mitchell, 1992), the concept of self-efficacy has been applied to various domains, such as

mathematics, sports, and computing. Adapted from the general concept of self-efficacy, computer self-efficacy (CSE) refers to people's perceptions about their abilities in using a computer successfully (Compeau & Higgins, 1995a). The CSE construct demonstrated positive effects on many computer-related behaviors and outcomes (e.g. Venkatesh & Davis, 1996). For instance, CSE was found to have positive effects on perceived ease of use and computer use (Igbaria & Iivari, 1995; Ong, Lai, & Wang, 2004).

In the context of computer training, Compeau and Higgins (1995b) examined the impact of CSE on learning performance of two computer applications (i.e. Lotus and WordPerfect). Their results showed that CSE had a significant effect on learning the two applications, but its effect on learning of WordPerfect was stronger than that of Lotus. Similarly, subjects who demonstrated higher CSE beliefs prior to training attained higher learning performance than those who had lower CSE beliefs (Gist et al., 1989).

The seminal work of Marakas et al. (1998) show that CSE is a multidimensional construct that operates at a general (i.e. across systems) and a system-specific (i.e., application) level. System-specific CSE (SCSE) is defined as an individual's perception of self-efficacy in performing computing tasks related to a particular computer application or software package. Thus, unlike general CSE (GCSE) - which refers to a generalized and system-independent individual trait - SCSE refers to judgments of self-efficacy toward a specific software package or computing domain (Marakas et al., 1998).

The distinction between general and system-specific CSE is indispensable. First, this distinction parallels the malleable character of self-efficacy and Bandura's (1986) suggestion that evaluations of self-efficacy tailored to specific domains provide better predictors of behavior than omnibus assessments. Second, GCSE is considered a trait-oriented efficacy, while SCSE is a state-oriented efficacy that is easier to influence and manipulate (Hsu & Chiu, 2004). Finally, this distinction allows assessments of CSE to focus on computing skills and exclude evaluations of cross-domain skills necessary to perform a computing task (Marakas et al., 1998).

Despite the importance of SCSE to various IS-related outcomes, it is only recently that SCSE has begun to attract attention IS research. For instance, Agarwal et al. (2000) found significant support for the relationship between Windows 95 CSE and Lotus 123 CSE and perceived ease of use of the two applications. Likewise, SCSE beliefs were examined with respect to Access 97 (Martocchio, 1994) and a financial software package (Gist et al., 1989). Finally, Hsu and Chiu (2004) examined the impact of general Internet CSE and Web-specific CSE on behavioral intentions and actual usage of an Internet-based application. Their results revealed that Web CSE had positive effects on usage intention and actual system usage. In contrast, general Internet CSE demonstrated indirect effects on behavioral intention and usage through its direct effect on attitudes.

In computer training research, SCSE has been found to have a significant effect on learning of many computer tasks such as immediate tasks (as measured at the conclusion of the training session) and delayed tasks (as measured several days after training) (Yi & Davis, 2003). Johnson and Marakas (2000) found that both levels of CSE (GCSE and SCSE) have positive effects on

computer learning performance. However, there are some studies in the literature that have reported contradicting results. For instance, SCSE demonstrated a non-significant effect on computer learning performance (Yi & Im, 2004).

Training effectiveness

Kirkpatrick (1995) suggests that training effectiveness should be evaluated in terms of reactions, learning, behavior, and results. Reactions pertain to trainees' feelings and attitudes after training. Learning relates to whether participants learned and mastered the knowledge presented in training. Behavior refers to the extent to which trainees can apply the newly learned knowledge at their jobs. Finally, results pertain to evaluating the impact of training on the organization in quantitative terms, such as reduced costs and increased quality and quantity of work.

While Kirkpatrick's framework remains the most widely accepted and extensively used model for evaluating the effectiveness of various types of training programs, very few studies, if any, have attempted to use this model to evaluate computer-training outcomes (Mahapatra & Lai, 2005). As a result, this study attempts to apply Kirkpatrick's model to computer training.

In this study, as indicated earlier, only the first two groups of outcomes - reactions and learning - are broached. Based on past studies (e.g., Bedard et al., 2003) and given that computer training is often aimed at enhancing systems acceptance and utilization (e.g. Agarwal et al., 2000; Davis & Bostrom, 1993; Hu, Clark, & Ma, 2003; Noe, 1986; Yi & Davis, 2001; Yi & Hwang, 2003), reactions are examined with respect to perceived ease of use and perceived usefulness. Similarly, consistent with theories of learning (Ausubel, 1968), learning performance is assessed in terms of near-transfer and far-transfer learning.

RESEARCH HYPOTHESES

Perceived ease of use

The perceived ease of use (PEOU) construct was introduced as a principal component of the technology acceptance model (TAM) (Davis, 1989; Davis, Bagozzi, & Warshaw, 1989). PEOU refers to the degree to which one believes that using or learning a system would be effortless (Davis, 1989). According to TAM, the PEOU of a target system represents a primary determinant of subsequent system acceptance and use.

The relationship between CSE and PEOU has attracted a considerable attention, especially in systems acceptance and adoption research. For instance, Venkatesh and Davis (1996) found that CSE had a positive impact on PEOU before and after subjects received hands-on experience with a target system. While most studies found a significant, positive relationship between CSE and PEOU (Davis, 1989; Igbaria & Iivari, 1995; Ong et al., 2004), negative or non-significant relationships between GCSE and PEOU have also been reported (Agarwal et al. 2000). These mixed

results underscore a genuine need for more work to better understand the relationship between CSE and PEOU (Agarwal et al., 2000).

A very limited number of studies in the computer training literature have evaluated PEOU as a training outcome. For example, Davis and Bostrom (1993) examined the impact of system interface style and training approach on perceived ease of use and found that neither interface style nor training approach had any significant impact on perceived ease of use. Hu et al. (2003) found that GCSE had a positive impact on perceived ease of use both before training and after a four-week training program. Nonetheless, Gururajan and Fink (2002) did not find significant differences in PEOU between subjects who received training on an icon-based system and those who received training on a menu-based system. Hence, they suggested that other factors might have stronger effects on PEOU than interface style and training method (Davis & Bostrom, 1993).

Even less attention was given to the relationship between SCSE in past research. In studying the acceptance of Web-based applications, Yi and Hwang (2003) found that Internet CSE had a significant effect on PEOU before the start of the training and after a four-week training program on a Web-based system. Additionally, Agarwal et al. (2000) found that general and specific CSE had significant effects on the perceived ease of use of Windows 95. In a recent study, Bedard et al. (2003) found that GCSE and SCSE had positive effects on PEOU. Thus, general and system-specific CSE are expected to have significant effects on perceived ease of use.

H1a: General CSE will have a positive effect on perceived ease of use.
H1b: System-specific CSE will have a positive effect on perceived of ease of use.

Perceived usefulness

Perceived usefulness (PU) represents another key determinant of IS acceptance and use. PU refers to the degree to which a person believes that using a system would enhance his or her job performance (Davis, 1989). Thus, TAM suggests that the more an individual believes that using a system will bring such benefits such as increased performance, the more likely the individual will accept and use the system.

Investigations of the relationship between CSE and PU have been carried out mostly in IS acceptance and adoption studies. Lopez and Manson (1997) found that CSE had a positive impact on PU. Additionally, Agarwal and Karahanna (2000) examined the acceptance of the World Wide Web among undergraduate business students and found that CSE had a significant (albeit small) effect on PU. Finally, Ong et al. (2004) investigated the impact of CSE on PU of an e-learning system.

While the relationship between CSE and PU has been examined extensively in systems acceptance research, there is a clear paucity of studies examining this relationship in computer training literature. For instance, Compeau & Higgins (1995b) find mixed results regarding the relationship between GCSE and performance-related outcome expectations (similar to PU). Bedard

et al. (2003) found that users' perceptions of usefulness about a target system improved significantly after a computer training program. Consistent with these studies, we suggest that general and system-specific CSE beliefs will have positive effects on PU.

H2a: General CSE will have a positive effect on perceived usefulness.
H2b: System-specific CSE will have a positive effect on perceived usefulness.

Near-transfer and far-transfer learning

Learning transfer refers to the extent to which past or current learning is applied or adapted to similar or novel situations (Haskell, 2001). Therefore, learning that can be extended or applied to novel situations and tasks that are different from those in the training environment is known as far-transfer learning. Conversely, learning that cannot be transferred or applied to situations that are dissimilar to the learning context is known as near-transfer learning.

The theoretical basis of near- and far-transfer learning is rooted in the assimilation theory of learning (ATL) (Ausubel, 1968) ATL defines two types of learning: rote and meaningful learning. Rote learning occurs when the learner memorizes the new information without knowing how it relates to existing knowledge or a cognitive structure: thus, rote learning cannot be extended or transferred to situations that differ from the situation in which the knowledge was learned. Conversely, meaningful learning refers to the learning in which new knowledge is completely understood, and the learner is able to relate the new information to existing knowledge. As a result, meaningful learning can be extended, manipulated, and applied to unfamiliar situations that differ from the training setting, or to tasks that have never been encountered before.

While much research has focused on understanding how to enhance learning performance in computer training (Bostrom et al., 1990; Johnson & Marakas, 2000; Lu et al., 2003), only a few attempts have been made to distinguish between near- and far-transfer learning. In their study, Davis and Bostrom (1993) found that instruction-based training resulted in better performance in near-transfer tasks than exploration training. For far-transfer tasks, their results revealed non-significant differences between the two training approaches. Likewise, Simon et al. (1996) found that behavioral modeling training resulted in better performance in near-transfer and far-transfer tasks than other training approaches.

More recently, Martocchio and Hertenstein (2003) investigated the impact of SCSE on declarative knowledge (similar to near-transfer knowledge). They studied the influence of Access 97 self-efficacy on the learning of Access 97 skills following a training program to 96 clerical workers. Their results indicated that Access 97 CSE has a significant effect on declarative knowledge (i.e. specific knowledge about basic features of Access 97).

Depending on the quality and quantity of the learner's relevant knowledge and the amount of effort that he or she is willing to expend to integrate the new knowledge into the existing knowledge and cognitive structure, learning occurs on a continuum that ranges from rote learning

to highly meaningful learning (Novak, 2002). Thus, to accomplish higher levels of meaningful learning, an individual must make a sustained effort to connect new information with existing knowledge. Given that self-efficacy has a direct effect on the amount of effort and persistence that a person exerts to perform successfully (Bandura, 1986), individuals with higher CSE are expected to demonstrate better near-transfer and far-transfer learning of computer skills than those with lower-rated CSE beliefs. Accordingly, the following hypotheses are suggested.

H3a: General CSE will have a positive effect on near-transfer learning.
H3b: System-specific CSE will have a positive effect on near-transfer learning.
H3c: General CSE will have a positive effect on far-transfer learning.
H3d: System-specific CSE will have positive effect on far-transfer learning.

RESEARCH METHOD

Subjects and procedure

Data for this study were collected from 78 undergraduate business students enrolled in two elective computer information systems courses at a Midwestern American university. Subjects were either juniors or seniors, with a major or a minor in management information systems. Sixty-four percent of the subjects (n=50) were males and 36 percent were females (n=28). The mean age of the subjects was 23.06 years (SD=2.77), with a range of 20 to 37 years.

The subjects completed a survey questionnaire containing measures of general and system-specific CSE pertaining to UNIX, along with some background and demographic questions. Next, a behavior-modeling training (described below) was given to the subjects. The training session lasted about 60 minutes and at the end of the training session, subjects were given a short tutorial covering the material presented in training; they were encouraged to read the tutorial and to work its illustrative examples. Subjects were asked to take a computer learning test covering the skills and concepts covered in training, and to complete a questionnaire to measure perceived ease of use. Finally, based on the suggestion that perceptions of usefulness about a target system are better formed after some interaction with the system (Hong, Thong, Wong, & Tam, 2001; Venkatesh & Davis, 1996), perceived usefulness was measured two weeks after training.

Training content

The material presented in training was about using the directory and file structures in the UNIX environment. Examples of the skills presented in training include creating, renaming, copying, and deleting files and directories. Additional skills - such as listing contents of a directory or sub-directory, searching for a file or a directory, identifying the location of the current (working)

directory in the directory structure, and using absolute and relative directory paths - were also demonstrated and explained in the training session. Unix was selected for this study because subjects could use it in their courses and, thereby, evaluate its usefulness to them. Additionally, the lack of subjects' experience with Unix reduces the effect of their past experience with the technology on their perceptions of it.

Measurements

GCSE was measured by six items from a widely-used and well-validated instrument (Compeau & Higgins, 1995b). Items in this instrument asked subjects to rate their ability to perform unspecified computing task using unidentified software. Responses were recorded on a10-point interval scale starting with 1 (*not at all confident*) and ending with 10 (*totally confident*). SCSE was measured by six items adapted from Johnson and Marakas (2000). Since the examined technology was UNIX, the items asked subjects to indicate their agreement or disagreement with six statements related to their ability to manipulate UNIX files and directories. Consistent with previous research, responses to items on this instrument were recorded on a seven-point Likert-type scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

PEOU was measured by three items from Davis's (1989) perceived ease of use instrument. The three items asked subjects to indicate the extent of their agreement or disagreement with statements concerning the usefulness of UNIX. Similarly, four items from Davis's (1989) instrument were used to measure perceived usefulness: items on this instrument asked subjects to indicate the degree to which they agreed or disagreed with statements about the usefulness of UNIX. Responses to the perceived ease of use and perceived usefulness items were recorded on a seven-point Likert-type scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Simple or near-transfer tasks are usually used to measure rote or near-transfer learning (Davis & Bostrom, 1993; Simon et al., 1996). Near-transfer tasks are characterized as being simple tasks that are similar to the tasks presented in training and require minimal effort to perform successfully. Accordingly, a five-question learning test was used to measure near-transfer learning. The skills included in the learning test were identical to the skills highlighted in training. The specific near-transfer skills included on the learning test were: (1) renaming a file in the current directory, (2) navigating to a parent directory, (3) creating a new sub-directory in the current directory, (4) listing the contents of the current directory, and (5) deleting a file from the current directory. Each question was worth 3 points. Completely correct answers were given 3 points and completely incorrect answers received 0 points. Answers that were neither entirely correct nor completely incorrect received 2 points. As such, possible scores on near-transfer learning test have a possible range of 0 to 15.

Complex or far-transfer tasks are often used to measure meaningful and far-transfer learning (Davis & Bostrom, 1993). Simon et al. (1996) suggest that far-transfer tasks comprise two or more near-transfer tasks and it is left to the trainee to determine which and how the near-transfer tasks

should be combined to perform the given far-transfer task. Thus, far-transfer learning was measured by a four-item learning test. Similar to the tasks used by Davis and Bostrom (1993), far-transfer tasks included on the learning performance test were: (1) creating a multi-level directory structure, (2) deleting a non-empty subdirectory, (3) copying a file from one directory to another directory in different locations in the directory structure, and (4) renaming files in another directory. Because, performing a far-transfer task requires more mental and physical effort than performing a near-transfer task, each question was worth 6 points. Scoring of the far-transfer learning test followed the same procedure used in the near-transfer learning test with proportional points for partially correct answers.

RESULTS

Although validated and highly reliable instruments were used to measure the research variables, a confirmatory factor analysis was performed on all multi-item measures. All instruments demonstrated high internal consistency reliability (coefficient alpha), and the loading of each item on its underlying construct was higher than on any other construct.

Table 1: Means, standard deviations, and correlations

Variable	Mean	S.D.	1	2	3	4	5	6
General CSE	40.63	11.62	--					
System-specific CSE	28.43	9.32	0.62*	--				
Perceived ease of use	15.26	4.64	0.49*	0.34*	--			
Perceived usefulness	19.62	6.38	0.09	0.02	0.62*	--		
Near-transfer learning	10.82	2.74	0.20	0.56*	0.62*	0.03	--	
Far-transfer learning	15.79	5.45	0.53*	0.58*	0.30*	0.10	0.36*	--

* $p < 0.01$

The means, standard deviations, and correlations of the research variables are presented in Table 1. As Table 1 shows, subjects' GCSE beliefs were considerably higher than their SCSE (UNIX) beliefs. This was not surprising, since UNIX is not a technology that students commonly interact with or use very frequently. Additionally, as can be seen in Table 1, all the correlations are well below the 0.8 threshold for suspecting the presence of multicollinearity (Bryman & Cramer, 1994).

The hierarchical regression procedure was used to test the research hypotheses and evaluate the impact of each level of CSE on the four dependent variables. Based on empirical or theoretical

considerations, predictor variables are entered into the regression hierarchical model in stages. This approach allows assessing the unique explanatory contribution of each predictor variable in predicting the dependent variable. Hence, based on the established relationships between GCSE and computer training outcomes, GCSE was entered first in all regression models, followed by SCSE.

Table 2: Hierarchal regression results for PEOU

	GCSE	SCSE	R^2	ΔR^2	*Sig.*
Step 1	0.203*	0.247	0.247	0.001	
Step 2	0.170 **	0.066	0.256	0.010	0.345

* $p < 0.001$; ** $p < 0.01$

The results of the first hierarchical regression analysis which was conducted to examine the effects of GCSE and SCSE on perceived ease of use are presented in Table 2. As Table 2 shows, GCSE had a significant effect on PEOU and explained about 25% of the variability in PEOU ($\Delta R^2 = 0.247, p < 0.001$), supporting hypothesis 1a. In contrast, SCSE demonstrated a non-significant effect on PEOU and explained about an additional 1% of the variance in PEOU ($\Delta R^2 = 0.010, p < 0.345$). Thus, hypothesis 1b was not supported. The combination of GCSE and SCSE explained about 26% of the variability in PEOU.

Table 3: Hierarchal regression results for PU

	GCSE	SCSE	R^2	ΔR^2	*Sig.*
Step 1	0.090		0.008	0.008	0.451
Step 2	0.100	-0.016	0.008	0.000	0.920

Table 3 presents the results of the second hierarchical regression analysis which was conducted to assess the unique effects of GCSE and SCSE on PU. As can be seen in Table 3, GCSE had a non-significant positive effect on PU and explained less than 1% of the variability in PU ($\Delta R^2 = 0.008$, $p < 0.451$). Similarly, SCSE had a negative, albeit non-significant, effect on PU and explained less than variance perceived usefulness. Almost no additional variance in PUperceived usefulness was explained by SCSE ($\Delta R^2 = -0.000$, $p < 0.920$). Practically no amount of variance (less than 1%) in PU was explained by GCSE and SCSE. Accordingly, both hypotheses, 2a and 2b, were not supported by the data.

Table 4: Hierarchal regression results for near-transfer learning					
	GCSE	SCSE	R^2	ΔR^2	*Sig.*
Step 1	0.199		0.040	0.040	0.093
Step 2	-0.229	0.683*	0.322	0.282	0.001
* $p < 0.001$					

The third hierarchical regression analysis was conducted to assess the influence of GCSE and SCSE on near-transfer learning. As Table 4 illustrates, no support was found for hypothesis 3a. More specifically, GCSE had a non-significant effect on near-transfer learning and explained less than 3% of the variability in near-transfer learning ($\Delta R^2 = 0.040$, $p < 0.093$). In contrast, SCSE had a significant positive effect on near-transfer learning and explained about an additional 28% of the variance in near-transfer learning ($\Delta R^2 = 0.282$, $p < 0.001$), supporting hypothesis 3b. The combination of GCSE and SCSE explained about 30% of the variability in near-transfer learning.

Table 5: Hierarchal regression results for far-transfer learning					
	GCSE	SCSE	R^2	ΔR^2	*Sig.*
Step 1	0.540*		0.291	0.291	0.001
Step 2	0.259**	0.447*	0.412	0.121	0.001
* $p < 0.001$; ** $p < 0.05$					

The results of the final hierarchical regression analysis which was conducted to assess the unique effects of GCSE and SCSE on far-transfer learning are presented in Table 5. As can be seen in Table 5, GCSE had a significant positive effect on far-transfer learning and explained about 29% of the variability in far-transfer learning ($\Delta R^2 = 0.291$, $p < 0.001$). Similarly, SCSE had a significant positive effect on far-transfer learning and explained about an additional 12% of the variance in far-transfer learning ($\Delta R^2 = 0.121$, $p < 0.001$). The results provide support for hypotheses 3c and 3d. Taken in tandem, GCSE and SCSE explained about 41% of the variability in far-transfer learning.

DISCUSSION

This study sought to examine the impact of CSE as a multidimensional variable on key computer-training outcomes. Accordingly, GCSE and SCSE were hypothesized to have positive effects on perceived ease of use, perceived usefulness, near-transfer learning, and far-transfer learning. The results provided mixed support for hypothesized relationships.

The results indicate that GCSE has a significant positive effect on perceived ease of use, supporting H1a. This finding is consistent with the findings of previous studies (Agarwal & Karahanna, 2000; Hong et al., 2001; Hu et al., 2003; Igbaria & Iivari, 1995; Venkatesh & Davis, 1996), and further corroborates the positive relationship between GCSE and perceived ease of use. In contrast, H1b - which predicted that SCSE would have a positive effect on perceived ease of use - was not supported by the data.

The non-significant impact of SCSE on PEOU was surprising, and deviated from recent research findings that showed system-specific (Internet) CSE had a significant effect on perceived ease of use of a Web-based Blackboard system (Yi & Hwang, 2003). One plausible explanation for this disagreement may be due to the differences in the technologies examined in both studies. According to innovation diffusion literature, trialability and visibility have profound effects on user beliefs about IS (Karahanna, Straub, & Chervany, 1999; Moore & Benbasat, 1991). In their study, Yi & Hwang (2003) examined Blackboard, which is a popular course-management system that is increasingly being used in educational environments. Thus, it is plausible that students might have used Blackboard technology in one or more courses, making it more visible and available for trial by students than UNIX, which is not a commonly used technology. Another potential explanation concerns the complexity of the examined technology. Unlike UNIX, Blackboard is a user-friendly technology that can be learned sufficiently well within a relatively short period of time by individuals with varying levels of CSE.

The results also show that neither GCSE nor SCSE had any significant effect on PU, and thus both H2a and H2b were supported. Since in most previous studies CSE was found to have either a small (Agarwal & Karahanna, 2000; Ong et al., 2004) or a non-significant, negative effect on perceived usefulness (Chau, 1996), our results seem to be congruent with this overall pattern of findings. On the one hand, GCSE demonstrated a non-significant positive effect on perceived usefulness. On the other hand, SCSE exhibited a non-significant negative effect on perceived usefulness.

There are two potential explanations that could partially account for the results described above. First, given that perceptions of usefulness in this study were measured two weeks after training, it is not unreasonable to suspect that a two-week period was not long enough for subjects to fully explore the system and discover its potential to improve their course work. Thus, the usefulness of the examined technology might not have been fully apparent to the subjects within this limited period of time.

Alternatively, the availability of other functionally comparable applications that subjects could use to complete their work offers another possible explanation. Since the use of UNIX was not mandatory and students had access to other applications (such as integrated development environments, or IDEs) to complete their work, it is very likely that some of them may have used other applications. Thus, it is plausible that the usefulness of the examined technology was not apparent to those who had not used it.

H3a and H3b predicted that GCSE and SCSE would have positive effects on near-transfer learning. Only SCSE demonstrated a significant effect on near-transfer learning, supporting H3b only. These findings are consistent with the assimilation theory of learning (Ausubel, 1968), which indicates that near-transfer learning requires recalling concepts from memory and applying these concepts to similar situations, whereas far-transfer learning requires not only recalling concepts from memory but also manipulating and arranging these concepts to perform more complex tasks in novel situations. Since SCSE focuses on specific skills relevant to a computer application, it is likely that its effect is more dominant in situations that require recalling and using standard skills to perform a given computing tasks. GCSE, however, may play a stronger role in performing more complex tasks that require not only recalling standard skills, but also greater confidence in ability to manipulate these skills to perform a target behavior successfully.

As was hypothesized, H3c and H3d - which suggested that GCSE and SCSE would have positive effects on far-transfer learning - were both supported. Again, these findings are consistent with assimilative learning theory and self-efficacy theory. From the assimilative learning theory perspective, meaningful (i.e. far-transfer) learning occurs when the learner understands the new knowledge, and is able to recall and apply it in a correct sequence to complete a task. Likewise, self-efficacy theory suggests that competent and successful performance requires the presence of not only constituent skills, but also high efficacy beliefs to use these skills effectively. Furthermore, these findings provide support for Noe's (1986) assertion that trainees are more likely to transfer the learning they acquired in training when they have more confidence in the skills they learned (i.e. high self-efficacy beliefs).

IMPLICATIONS

A number of implications for research and practice can be drawn from the results of this study. In terms of research, this study examined CSE at the general and software-specific levels. The results provide empirical support for the multidimensionality nature of the CSE construct and provide support for earlier studies that reported a similar pattern of relationships concerning the multidimensionality of CSE (e.g. Agarwal et al., 2000; Johnson & Marakas, 2000).

In their seminal work on computer self-efficacy, Compeau and Higgins (1995a; 1995b) suggested that general and specific CSE beliefs may have different effects on computing-related behavior. Thus, by empirically examining these assertions and demonstrating that general and system-specific CSE demonstrate varying effects on computer training outcomes, this study empirically tested and validated some of those relationships.

This study took a broader approach to evaluating the effectiveness of computer training by employing two key training outcomes (reactions and learning performance) as identified in Kirkpatrick's (1995) model of training effectiveness and addressed a void computer training research (Mahapatra & Lai, 2005). Drawing on past research, reactions were examined with respect

to PEOU and PU. In addition, unlike most past studies, this study evaluated two types of learning (near-transfer and far-transfer) as indicators of learning performance.

In terms of practice, this study highlighted the important role that CSE plays in achieving different types of training outcomes. Accordingly, organizational attempts to boost trainees' GCSE and SCSE beliefs prior to training could be very useful in achieving successful computer training outcomes and fulfilling certain training objectives. Past studies have shown that GCSE and SCSE beliefs can be enhanced through behavior modeling (Gist et al., 1989), feedback and persuasion (Martocchio & Webster, 1993), organizational support (Igbaria & Iivari, 1995), induced conception of ability (Martocchio, 1994), and prior experience (Compeau & Higgins, 1995a).

Overall, the results show that GCSE has a stronger effect on reactions than learning performance, whereas SCSE has stronger effects on learning performance than reactions. This suggests that depending on the principal objective of training (i.e. improving reactions or maximizing learning), more emphasis on general or system-specific CSE beliefs prior to training may be highly beneficial to achieving the desired outcomes.

The study may also have some implications for enhancing specific types of learning performance. Cormier and Hagman (1987) suggest a trade-off between training for the rapid acquisition of skills and training for the transfer of skills. They maintain that training designed for rapid acquisition should involve more time spent on a few and similar examples. In contrast, training for skill transfer should focus on diverse tasks and situations in which the presented skills can be used. This study demonstrates how the two types of learning are influenced by the two levels of CSE, and thus these results could be useful in designing training to maximize the results specific types of learning.

In a broader context, the study offers implications for enhancing IS acceptance and usage. Perceived ease and perceived usefulness are considered primary determinants of IS acceptance. Although both levels of CSE had non-significant effects on perceived usefulness, GCSE was found to have a significant effect on perceived ease of use. Collectively, these results seem to imply that perceived usefulness may not be greatly affected by manipulating perceptions of CSE. This suggests that other factors may be more influential in affecting perceptions of usefulness, such as the type of technology in question and type of usage after training (eg. Internet Explorer vs. Photoshop). Therefore, emphasis on factors other than CSE may prove to be more useful in improving perceptions of usefulness. In contrast, GCSE demonstrated a significant effect on perceived ease of use. Given that enhancements to the interface style may not necessarily result in more favorable perceptions of ease of use (Keil, Beranek, & Konsynski, 1995), attempts to focus on other factors - such as GCSE - may be useful in improving perceptions of ease of use and thereby improve ultimate IS acceptance.

LIMITATIONS

Like all studies, this study has some limitations that should be recognized when interpreting the results. An obvious limitation involves the use of a student sample to test the research model. Although the use of student samples is widespread in studies of this nature, it is important for future studies to use more diverse samples in order to enhance the generalizability of the results to other populations of end users, especially in organizational environments. Moreover, this study used a test of learning performance rather than actual computing tasks. Thus, future research should consider using actual near-transfer and far-transfer computing tasks, and examine other technologies to enhance the validity of the results – hence increasing their generalization across technologies.

A possible limitation stems from the timeframe in which perceived usefulness was measured. Perceptions of usefulness were measure two weeks after the training. This relatively short period of time may not be long enough for subjects to fully explore the system functionality and determine its usefulness in their coursework. Thus, their perceptions of usefulness may not have been fully developed at the time of measurement. Although our approach was consistent with other studies (e.g., Yi & Davis, 2003), this issue deserves further investigation in future research to determine the appropriate time at which users can accurately determine system usefulness and form their perceptions of usefulness.

Another limitation concerns the use of two outcomes from Kirkpatrick's (1995) four-level model of training effectiveness. More specifically, trainee behavior and training results (the other two outcomes identified in Kirkpatrick's model) were not examined here. Because evaluating trainee behavior and training results involve assessing the impact of training on one's job and organization objectives, it was not possible to examine these two outcomes in the present study. Thus, to assess computer training effectiveness in a more comprehensive manner, future research should evaluate the four training outcomes suggested by Kirkpatrick's model rather than the first two categories of outcomes as was the case in this study.

FUTURE RESEARCH

Although this study shed some light on the impact of CSE on computer training outcomes, more research is undoubtedly needed to provide further insights into this important issue. First, the non-significant effects of general and system-specific CSE on perceived usefulness warrant additional investigation. Given that perceptions of usefulness were measured two weeks after training, a promising area for future research is to investigate the timeframe in which users' perceptions of usefulness about a target system are likely to be fully developed. If the time needed for end users to form their perceptions of usefulness can be accurately determined, this will clearly help IS managers and organizations make more timely decisions about the usefulness of a particular system, and, thus avoid premature or delayed decisions or courses of action. Moreover, further research is needed to examine methods that can be used by organizations and IS trainees that help

boost trainee's CSE beliefs prior to training which in return make computer training outcomes successful.

This study was by no means intended to be a comprehensive investigation of all the factors believed to affect computer training outcomes. On the contrary, it was made clear at the outset that the current research sought to examine the impact of multidimensional CSE on two key computer training outcomes. Clearly, other variables such as training method (Chou & Wang, 1999), system interface style (Davis & Bostrom, 1993), labeling of training as work or play (Martocchio & Webster, 1993), and individual learning style (Simon et al., 1996) – all of which have been found to have significant effects on computer training outcomes - were not examined here. Thus, to enhance the current understanding of factors affecting the effectiveness of computer training, it is imperative that future research explore how these and other variables affect learning performance, reactions to training and other training outcomes. Furthermore, further investigation is needed to measure effects of general and system-specific CSE on perceived usefulness for different types of technology trainings (e.g. Windows XP vs. UNIX).

REFERENCES

Agarwal, A. & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, 24(4), 665-694.

Agarwal, R., Sambamurthy, V. & Stair, R. (2000). The evolving relationship between general and specific computer literacy: An empirical assessment. *Information Systems Research*, 11(4), 418-430.

Ausubel, D.P. (1968). *Educational Psychology, A Cognitive View*, Holt, Rinehart and Winston, Inc, New York.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice-Hall, New Jersey.

Bedard, J.C., Jackson, C., Ettredge, M. L. & Johnstone, K. M. 2003. The effect of training on auditors' acceptance of an electronic work system. *International Journal of Accounting Information Systems* 4(4), 227-250.

Bostrom, R.P., Olfman, L. & Sein, M.K. (1990). The importance of learning style in end user training, *MIS Quarterly*, 14(1), 101-119.

Bryman A. & Cramer, D. (1994), *Quantitative data analysis for social scientists*, Routledge, New York.

Byrd, T.A. & Turner, D.E. (2001). An exploratory analysis of the value of the skills of IT personnel: Their relationship to IS infrastructure and competitive advantage. *Decision Sciences*, 32(1), 21-54.

Chau, P.Y.K. (1996). An empirical assessment of a modified technology acceptance model. *Journal of Management Information Systems*, 12(2), 185-204.

Chou, H.W. & Wang, Y.F. (1999). Effects of learning style and training method on computer attitude and performance in world wide web page design training. *Journal of Educational Computing Research*, 21(3), 323-342.

Compeau, D.R. & Higgins, C.A. (1995a). Application of social cognitive theory to training for computer skills. *Information Systems Research*, 6(1), 118-143.

Compeau, D.R. & Higgins, C.A. (1995b). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19(2), 189-211.

Cormier, S. & Hagman, J. (1987). *Transfer of learning: Contemporary research and applications*, Academic Press, Inc., San Diego: CA.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

Davis, F.D., Bagozzi, R.P. & Warshaw, P.R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982-1003.

Davis, S. & Bostrom, R. (1993). Training end users: An experimental investigation of the roles of the computer interface and training methods. *MIS Quarterly*, 17(1), 61-79.

Gist, M.E. & Mitchell, T.R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Science*, 17(2), 183-211.

Gist, M.E. , Schwoerer, C. & Rosen, B. (1989). Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology*, 74(6), 884-891.

Gururajan, R. & Fink, D. (2002). A study of influences of application interfaces on end user training outcomes, Cork, Ireland: Informing Science and IT Education Conference.

Haskell, R.E. (2001). Transfer of learning: Cognition, instruction and reasoning, Academic Press, San Diego.

Hong, W.Y. Thong, J.Y.L., Wong, W.M. & Tam, K.Y. (2001). Determinants of user acceptance of digital libraries: An empirical examination of individual differences and system characteristics. *Journal of Management Information Systems*, 18(3), 97-124.

Hsu, M.H. & Chiu, C.M. (2004) Internet self-efficacy and electronic service acceptance. *Decision Support Systems*, 38(3), 369-381.

Hu, P.J.H. , Clark, T.H.K. & Ma, W.W. (2003). Examining technology acceptance by school teachers: A longitudinal study. Information & Management, 41(2), 227-241.

Igbaria, M. & Iivari, J. (1995). The effects of self-efficacy on computer usage. *Omega*, 23(6), 587-605.

Johnson, R.D. & Marakas, G.M. (2000). The role of behavioral modeling in computer skills acquisition: Toward refinement of the model. *Information Systems Research*, 11(4), 402-417.

Karahanna, E., Straub, D. & Chervany, N. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly*, 23(2), 183-213.

Keil, M., Beranek, P.M. & Konsynski, B.R. (1995). Usefulness and ease of use: Field study evidence regarding task considerations. *Decision Support Systems*, 13(3), 75-91.

Kirkpatrick, D.L. (1995). Techniques for evaluating training programs. *Journal of the American Society of Training Directors*, 13(11-12), 3-26.

Lopez, D. & Manson, D. (1997). A study of individual computer self-efficacy and perceived usefulness of the empowered desktop information system. *Journal of Interdisciplinary Studies,* 10, 83-92.

Lu, J. , Yu, C. & Liu, C. (2003). Learning style, learning patterns, and learning performance in a WebCT-based MIS course. *Information & Management*, 40(6), 497–507.

Mahapatra, R. & Lai, V.S. (2005). Evaluating end-user training programs. C*ommunications of the ACM*, 48(1), 67-70.

Marakas, G.M., Yi, M.Y. & Johnson, R. (1998). The multilevel and multifaceted character of computer self-efficacy: Toward a clarification of the construct and an integrative framework for research. *Information Systems Research*, 9(2), 126-163.

Martocchio, J.J. & Hertenstein, E.J. (2003). Learning orientation and goal orientation context: Relationships with cognitive and affective learning outcomes. *Human Resources Development Quarterly*, 14(4), 413-434.

Martocchio, J.J. & Webster, J. (1993). The effects of feedback and microcomputer playfulness on performance in microcomputer software training. *Personnel Psychology*, 45(3), 553-578.

Martocchio, J.J. (1994). Effects of conceptions of ability on anxiety, self-efficacy, and learning in training. *Journal of Applied Psychology*, 79(6), 819-825.

Moore, G.C. & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3), 192-222.

Noe, R.A. (1986). Trainees' attributes and attitudes: Neglected influences on training effectiveness. *Academy of Management Review*, 11(4), 736-749.

Novak, J.D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education*, 86(4), 548-571.

Ong, C.S., Lai, J.Y. & Wang, Y.S. (2004). Factors affecting engineers' acceptance of asynchronous e-learning systems in high-tech companies. *Information and Management*, 41(6), 795-804.

Sein, M.K., Olfman, L., Bostrom, R.P. & Davis, S.A. (1993). Visualization ability as a predictor of user learning success. *International Journal of Man-Machine Studies*, 39(4), 599-620.

Simon, S.J. & Werner, J.M. (1996). Computer training through behavior modeling, self-paced, and instructional approaches: A field experiment. *Journal of Applied Psychology*, 81(6), 648-659.

Simon, S.J., Grover, G., Teng, J.T. & Whitcomb, K. (1996). The relationship of information system training methods and cognitive ability to end-user satisfaction, comprehension, and skill transfer: A longitudinal field study. *Information Systems Research*, 7(4), 466-490.

Venkatesh, V. & Davis, F.D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences*, 27(3), 451-481.

Yi, M.U. & Davis, F.D. (2001). Improving computer training effectiveness for decision technologies: Behavior modeling and retention enhancement. *Decision Sciences*, 32(3), 521-544.

Yi, M.U. & Davis, F.D. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Information Systems Research*, 14(2), 146-169.

Yi, M.U. & Hwang, Y. (2003). Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-Computer Studies*, 59(4), 431–449.

Yi, M.U. & Im, K.S. (2004). Predicting computer task performance: Personal goal and self-efficacy. *Journal of Organizational and End User Computing*, 16(2), 20-34.

A SOCIAL ENGINEERING PROJECT IN A COMPUTER SECURITY COURSE

Barbara Endicott-Popovsky, Seattle University
Diane L. Lockwood, Seattle University

ABSTRACT

A small private university began to offer undergraduate and graduate courses in computer security during the academic year 2002-2003 within the schools of computer science and business. In the introductory computer security course, a "social engineering" team project was included as a required assignment. This article briefly summarizes the social engineering literature, describes the project assignment and learning objective, provides actual student sample deliverables, and presents results of a follow-up student survey on the experience. The lessons learned from this effort should prove useful to other universities and instructors contemplating similar coursework.

INTRODUCTION

A woman, ostensibly from the human resources department, calls the company help desk and says she has forgotten her password. In a panic, she adds that if she misses the deadline to submit employee insurance applications online, all employees will be without health insurance until the problem can be corrected, adding that she might even be fired for this. The help desk worker feels sorry for her and quickly resets the password - unwittingly giving a hacker entrance into the corporate network. The hacker got the names of human resources employees from the company's recycling bin the previous night. This caper is known as social engineering.

Social engineering is basically pulling a con job to get information or access to systems that are normally only used by privileged users (Mitnick, 2002). Social engineering is the human side (i.e., "wetware" in hacker slang) of breaking into a corporate network. Organizations with elaborate firewalls, authentication processes, virus scan software, and network security monitoring technology are "still open to an attack if an employee unwittingly gives away key information in an email, by answering question over the phone with someone they don't know," by not shredding sensitive documents, or even talking about a project with coworkers at a restaurant (Gaudin, 2002b).

Kevin Mitnick, the famous convicted computer hacker, offered advice to businesses afraid that corporate spies and hackers may gain access to their internal systems using social engineering saying that "on the corporate side, as an employee, it all comes down to user awareness and education (Savage, 2003)."

Courses in computer security predominantly discuss the technical side of security (e.g., encryption, network security defenses, firewalls, software reliability, digital certificates, wireless eavesdropping, biometrics.), but often give short shrift to the human side of security - especially social engineering. The purpose of this article is to describe a social engineering student project that was undertaken to increase student awareness of this serious security vulnerability. The lessons learned from this effort should prove useful to other universities and instructors contemplating similar coursework (Vaughn & Boggess, 1999).

DESCRIPTION OF SOCIAL ENGINEERING ASSIGNMENT

Students in a graduate MBA business class on Computer Security were given a reading assignment from Kevin Mitnick's book, The Art of Deception (Mitnick, 2002), to learn what is meant by social engineering. With that background, they were asked to develop an exploit, using information gleaned from any open source (e.g., including telephone directories, dumpsters, waste baskets, online information, and any other publicly available information), against some specific target person on campus. They were prohibited from actually impersonating anyone like campus police since impersonating a law enforcement official is considered a criminal offense. They were also prohibited from contacting the target "mark" directly, or actually executing their exploit.

To bound and control this assignment, student activities were confined to local campus personnel and campus security was informed to prevent any misunderstandings. Students were instructed to carry a copy of their assignment (see Appendix A) at all times in the event they were confronted; however, they were warned that getting caught would result in a significant deduction of points! The "target mark" was not to be contacted about the exploit nor the information gained. In addition, all confidential information was to be deleted (lined out) from the final assignment and any confidential documents discovered were to be submitted along with their assignment so that the instructor could supervise proper disposal.

Please note that this assignment was set up as an object lesson in ethics, as well. The ethical dimensions of social engineering were discussed extensively before the experiment was introduced and students were asked to sign a lab policy statement that acknowledged their understanding of the "acceptable use" constraints established for this and other experiments conducted in class (See Appendix B). Students were informed that the campus security office was acquainted with the experiment and the constraints under which it would be conducted. Results were monitored and all materials discovered during this activity were collected and destroyed. The instructors took care that the experiment was under their control and supervision.

The deliverable for this assignment was an oral and written report that described the information obtained, from what source(s), and how they could use the information discovered through these covert means to exploit their target. To be certain that students stayed within the bounds of the assignment, the instructor required progress reports weekly until the due date for the assignment. They were given four weeks to complete the exercise. The major objective of the

assignment was to increase student awareness of the human side of computer security vulnerabilities. Two sample student team deliverables are presented in Appendix C.

SAMPLE ASSIGNMENTS

In one assignment, the student team interviewed administrative assistants on campus, claiming to be a sociology student doing a class assignment on how people constructed passwords. They expressly stated that they did NOT want participants to reveal their actual passwords; rather they were to describe the algorithm used to derive their passwords. For example, one respondent indicated that she used the abbreviations of various states, of course excluding any that were over six characters long. Obviously, it would not be hard for a hacker to subsequently crack their password using this information.

Another enterprising group of graduate students had an even more elaborate plan. They gained enough information on a professor in the department to steal her identity and devise a plan to apply for a tenured professor's position at a prestigious university, assuming her credentials. They found the "easy" documents (e.g., her professional resume, list of publications, etc. from the school's website), but more alarming was their ability to get official school transcripts, manufacture a letter of recommendation from her department head, and research sufficient confidential, personal information to pose successfully as her.

A third group raised a red flag when they reported that deeds of trust were recorded in our county, complete with social security numbers! That information about one of the school's faculty members, plus a traffic ticket that the target had discarded in a wastebasket-unshredded (traffic tickets contain the driver's license number of the person cited), is enough to get a replacement driver's license and to open credit card accounts, which is exactly what their exploit suggested they would do! Their specific exploit is described in Appendix C. Note that it is written in the first person, as the hacker! They have come up with three different exploits they could run with the information they have received, generally making Mr. X's life miserable!

The oral and written assignments that student gave in class were graded using the instructor evaluation forms found in Appendix D.

STUDENT FOLLOW-UP SURVEY RESULTS

Students in the class were asked two questions in a follow-up survey taken after the Social Engineering assignment was turned in.

1. To what extent did the social engineering assignment increase your awareness of the human side of computer security? (Circle One)

None	Very little	Somewhat	A lot
1	2	3	4

2a. To what extent has the social engineering assignment changed your behavior concerning computer security (e.g., " I change my password more frequently; I'm more careful about shredding sensitive documents," etc.)

None	Very little	Somewhat	A lot
1	2	3	4

2b. In what specific way(s), did you change your behavior (please describe):

In response to the first question, students indicated that the assignment increased their awareness of computer security "A lot." (avg, = 3.8). However, in response to the second question, the students indicated that the assignment actually changed their behavior concerning computer security between "A lot" and "Somewhat" (avg, = 3.5). The result that reported "behavioral" changes lagged "awareness" changes could be explained by the fact that a few students already had work experience or training in computer security and that they were already practicing secure "behaviors;" hence the assignment did little to change their subsequent behavior.

Among the ways students claimed they changed their behaviors included buying a shredder and shredding anything with their names on it before discarding; continuously checking for security updates for their PC software; taking more care with their computer passwords and computer access, and hardening their personal PC's with firewall, anti-virus and authentication software. A small number of students indicated they already were taking precautions with their computer systems at home, but for most, this exercise was eye-opening. One student even indicated that they had since initiated procedures at work to properly handle the computer of an employee who was terminated, another indicated they were much more sensitive to cold calls on the telephone from people interested in getting personal information over the phone.

SUMMARY AND CONCLUSIONS

This article described a social engineering assignment that can be used to increase student awareness of, and change their personal behavior toward, the human side of computer security vulnerabilities. A forthcoming paper (Part II) will provide a much more developed pre-post test questionnaire and analysis.

A similar social engineering assignment could also be given in actual computer security training programs within organizations that are trying to increase security awareness. After all, even Kevin Mitnick (2002) agrees that the best way to prevent such attacks is to increase employee awareness through training and education followed up by unannounced fire drills or penetration tests. Results of these tests should then be communicated to all employees and rolled into on-going training sessions on computer security.

REFERENCES

Arbaugh, B. (Feb 2002). "Security: technical, social, and legal challenges." *Computer,* v35n2,p109-11.

Castellvccio, M. (Dec 2002). "Social engineering 101." *Strategic Finance*, v84,n6, p. 57-58.

CERT Incident Note IN-2002-03 (March 19, 2002). "Social engineering attacks via irc and instant messaging." http://www.cert.org/incident_notes/IN-2002-03.html

"Electronic highwayman – a modern myth?" (June 1997). Anonymous author. *Accountancy*, v119n1246, p. 49.

Garceau, L. (Feb 1997). "The threat of social engineering." *Ohio CPA Journal*, v56n1, p. 42-44.

Gaudin, S. (July 24, 2002). "Tough computer crime bill clears hurdle." www.internetnews.com/bus-news/article.php/1107521

Gaudin, S. (July 24, 2002). "Social engineering: the human side of hacking." cin.earthweb.com

Hayes, F. (July 29, 2002). "Can't buy security." *Computerworld,* v36n31, p. 54.

Lemos, R. (July 24, 2002). "Mitnick teaches social engineering." ZDNet News US at http://news.zdnet.co.uk/story/0,,s2080227,oo.html

McClure, S. & Scambray, J. (May 3, 1999). "When looking for holes in network security, first give a call to your helpdesk and users." *Infoworld*, v21n18, p. 48.

Mitnick, K.D. & Simon, W.L. (2002). *The Art of Deception.* New York: Wiley Publishing.

Savage, M. (Apr 28, 2003). "Former hacker Mitnick details the threat of social engineering." *CRN*, no.1043, p. 58.

Schneier, B. (2000). *Secrets and Lies: Digital Security in a Networked World*. New York: Wiley Publishing.

Schwartau, W. (Feb 2, 1998). "Anatomy of a friendly hack." *Network World*, v15n5, p. 35-41.

Taubes, G. (Mar/Apr 1996). "The code snatchers." *Worldbusiness*, v2n2, p. 44-45.

Ulfelder, S. (Aug 11, 1997). "Spycatcher." *Computerworld,* v31n32, p. 80-81.

Vaughn, R. Jr. & Boggess, J. (Dec 30, 1999) "Integration of computer security into the software engineering and computer science programs." *Journal of Systems and Software*, v49,n2/3, p149-53.

APPENDIX A: SOCIAL ENGINEERING ASSIGNMENT
SCHOOL OF BUSINESS

ASSIGNMENT 1 SOCIAL ENGINEERING
ECIS-591 COMPUTER SECURITY
INSTRUCTOR: BARBARA ENDICOTT-POPOVSKY

ASSIGNMENT DESCRIPTION

You will be working in teams of 3 to 5 students. You will be gathering enough information to be able to impersonate some person on campus in a social engineering exploit:

Name:
Office number, office phone number:
Department, Department Manager:

Something personal like a meeting they attend this week, something that gives someone the sense that you are the person you are impersonating, etc. You can use any open source-telephone books, dumpsters, waste baskets, online information.

YOU MUST NOT:

Call or talk to anyone other than team members
Impersonate anyone to get this material (Impersonating a police office is a crime!)
Contact the individual you are targeting

APPENDIX B
ECIS591 COMPUTER LAB USE POLICY

1. The following Seattle University IT policy will be in effect, which can be found at: http://www.seattleu.edu/it

2. Seattle University acceptable computer use policy will be in effect, which can be found at: http://www.seattleu.edu/it/policies/default.asp

3. The ECIS591 General Responsibilities for All Students policy will be in effect can be found on the first page of this syllabus.

4. Due to the special nature of Computer Security certain additions to the above policies will be in effect, they are the following:

   Where appropriate, every experiment run in conjunction with this course will have certain rules and regulations regarding its conduct. These will be explained when the assignments are given and students are expected to comply with any additional restrictions.

   To the extent that classroom computers are used to stage attacks under controlled circumstances, they will be physically disconnected from all external networks. All student users must maintain this lack of connection and must verify this lack of connection (with instructor help) before running any malicious code or exploit.

Students may be allowed to attempt to run harmful software and obtain root access on classroom computers isolated from the network, as long as the students in question agree to fix any problems they cause (e.g. hardware damaging code).

Security flaws and other problems should be pointed out immediately to the instructor.

Any student running an exploit in connection with assignments in this class must file an Exploit Approval form with the instructor, before running any malicious code or attempting any exploit on any classroom computer.

Students are responsible for the consequences of any actions they take without the knowledge of the lab instructor.

I hereby certify that I have read and understand this policy and the relevant Seattle University policies referenced above regarding computer lab use and will abide by them.

Signature:
Printed Name:
Student ID No:
Date:

APPENDIX C
SOCIAL ENGINEERING EXAMPLE ASSIGNMENT

This is an example of several exploits that Group 3 developed after getting the social security number of the target. Note: this exploit was NOT actually carried out; only a frightening scenario of what could actually happen.

THE EXPLOIT

We, the social engineering experts of the MBA 591 Computer Security class exploited Mr. X in the following ways:

SCENARIO 1: IDENTITY THEFT

We applied for a credit card in the name of Mr. X, using his SSN, birth date, address, and a fabricated mother's maiden name. (It didn't matter that our request for credit was not approved; the information we submitted was added to Mr. X's credit report because the credit agencies merge everything into the database entry of the target person based on matching name and address).

With that and our dumpster diving we had enough information in place to obtain a credit report. We used the address and fax number of a non-franchise postal services shop called Mailmart, which has a physical address instead of a P.O. box, and had the clerk forward our mail on a weekly basis, as a precaution, to the address of a neighbor in the neighborhood who has a mailbox in a location that's not visible from the street. I know this person's work schedule and check it before the rightful owner does.

We got a credit-reporting agency to give us "our" credit report. We were actually able to obtain the credit report through a Web interface. We didn't even need our mailing address!

Right away we were able to examine the credit report. It has the target's actual ("primary") social security number and other information including other cards, accounts, assets and credit history. We were able to stock our meth lab and party for the rest of the month. Based on "his" purchase records Mr. X may expect a little visit from the Seattle Police Department - I would love to see him try to explain this one.

SCENARIO 2: THEFT & DEFACEMENT OF SCHOOL PROPERTY

We called the Help Desk and assuming the target's identity requested a change of password. The user IDs at SU are all first 7 characters of last name, first initial. I anticipated they might ask for my faculty ID number, (which I had after calling the Registrar's office, and asking for "my" faculty ID number), but they didn't even ask for that - they reset it to "password" for me - "Oh, and by the way," I said, "could you reset my SU Online password it was the same as my other one." They obliged and I was off to the races.

I logged into the faculty portion of SU online - I didn't change everyone else in Mr.X's class grades too dramatically, a B+ became an A-, a C became a B, but my C became a straight A for "Access".

SCENARIO 3: DEFAMATION OF CHARACTER

Let's just say I don't agree with Mr. X's politics. I successfully changed "my" password for the second time a couple weeks later at the Help Desk; this time they double-checked by asking for my SSN, which of course was no problem. I logged into "my" email account, and sent a bulk email to all the members of the 12th Avenue Parking Workgroup. I didn't want it to be deleted as spam so I started out in a way that would be mistaken for Mr. X but that got more and more insulting and vulgar.

I didn't have to imagine the shock and dismay of the 12th Avenue Parking Workgroup because many responded to my email right away!

Good luck with the re-election Mr. X!

APPENDIX D
ORAL AND WRITTEN GRADING CRITERIA

Student Social Engineering Assignments were graded according to rubrics presented in this Appendix. The written report earned a percentage grade as follows:

Thoroughness of preparation	30%	
Depth of research	25%	
Quality of exploit	25%	
Class presentation	10%	
Documentation	10%	
TOTAL		100%

The oral presentations were graded using this additional rubric:

Clear grasp of major issues	20%	
Quality of presentation	20%	
Appropriate analysis, evaluation	20%	
Demonstrated ability to employ course concepts	20%	
Organization and logical flow	20%	
TOTAL		100%

ECONOMIES OF IT SYSTEMS AT WAL-MART --- AN HISTORICAL PERSPECTIVE

Jianfeng Wang, Indiana University of Pennsylvania

ABSTRACT

This paper takes a retrospective look over Wal-Mart's IT systems. Major types of IT systems used by Wal-Mart are listed chronologically. The paper then explains the economies of IT systems at their application areas and how such economies have contributed to Wal-Mart's everyday low price strategy. The paper explores how distribution centers and IT systems function strategically at Wal-Mart, how Wal-Mart uses point-of-sale data and data warehouse technology, how Wal-Mart has used Retail Link in inventory management, and how economies of scale and scope and external economies are realized in the implementation of IT systems at Wal-Mart Stores, Inc. The paper concludes with the challenges Wal-Mart is facing.

INTRODUCTION

Researchers have attempted to reveal the importance of the relationships between retailers and suppliers using Wal-Mart as a case (Bloom et al., 2001; Bloom and Perry, 2001; Kumar, 1996). Stalk et al. (1992), on the other hand, have advocated capability-based competition by comparing Wal-Mart and Kmart. They have suggested that Wal-Mart's strategic vision was fully expressed in Wal-Mart's cross-docking system. None of the previous studies, however, have ever tried to take a comprehensive overview of Wal-Mart's IT systems and reveal how its IT systems have helped Wal-Mart in realizing its everyday low price strategy. This case study takes a retrospective look over Wal-Mart's IT systems. Major types of IT systems used by Wal-Mart are listed chronologically. The paper then explains the economies of IT systems at their application areas and how such economies have contributed to Wal-Mart's everyday low price strategy. IT systems at Wal-Mart are applied in major areas such as inventory management, administrative management, customer management and supplier management, etc. Efficiency and performance of IT systems in these areas apparently affect each other. Wal-Mart seeks to achieve not only the efficiency of individual systems, but also that of the integrated systems. Finally, the paper clarifies the existing challenges that Wal-Mart is facing. Cost control and inventory management are one of major concerns for Wal-Mart, though cost control goes beyond inventory management. Wal-Mart also has to control and reduce its cost of operation and administration and cost of sales.

The case is divided into a few sections. Section one is a brief overview of Wal-Mart's IT systems; section two explain the case study approach and data; section three discusses the strategic

role of Wal-Mart's distribution centers; section four describes how Wal-Mart uses Point-of-Sale systems to collect customer data and uses data warehouse and datamining to analyze customer behavior; section five is about Electronic Data Interchange (EDI) and Retail Link systems at Wal-Mart; section six discusses productivity improvement from IT systems; section seven shows the economies of scale and scope at the distribution centers with IT systems; section eight is about the business process standards and external economies with IT systems. Section nine covers Wal-Mart's integrated structure and major challenges Wal-Mart is facing. The final section concludes the case study with a summary.

THE CASE STUDY APPROACH

The case study approach in this paper follows Yin (1994). Yin (1994) gives very detailed instructions of how to conduct a case study. According to Yin (1994),

> A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident. The case study inquiry copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result benefits from the prior development of theoretical propositions to guide data collection and analysis (p.13).

This paper explores the economies of IT systems at Wal-Mart from the perspectives of transaction cost theory, information asymmetry, information sharing, economies of scale and scope, and externalities are useful in explaining the economies of IT systems at Wal-Mart. The author collected data about Wal-Mart from a variety of resources: SEC filings, Wal-Mart annual reports, news reports, available interviews with Wal-Mart officers, and publications about Sam Walton and Wal-Mart.

Wal-Mart annual reports often provide valuable information about information systems or technologies used in prior years. Especially when some information systems were successfully implemented and applied, the annual report for a given year often tells stories and gives comments by Wal-Mart officers. News reports about Wal-Mart are scattered around numerous sources. They also provide very important evidence about Wal-Mart IT systems. The author did not have the opportunities to interview the Wal-Mart officers, whose comments are cited in this paper, and thus must rely on the sources for veracity. Editors or journalists of many magazines such as Fortune, Information Week, Computer World, CIO, etc, interviewed Wal-Mart executives at different times. This paper incorporates those interview reports. The Wal-Mart Corporate Office provided all the annual reports and answered some of the author's questions.

A BRIEF OVERVIEW OF WAL-MART IT SYSTEMS

Accounting systems were first introduced into Wal-Mart in the late of 1960s. In 1972, Wal-Mart reported to have "considerable additions in Accounting, Data Processing and other management areas. A number of new programs and control systems were activated in 1972, both at store and management levels." (Annual Report, 1972)

In 1973, Wal-Mart developed a complete vendor system for its distribution center rebuyers and converted from an IBM 360/20 to an IBM 370/125 computer. That required physically changing all the existing programs (Annual Report, 1973).

In 1974, Wal-Mart reported that the Company realized significant freight savings through the Traffic Department's ability to obtain the most economical landed cost on merchandise purchased and to utilize the most economical modes of transportation (Annual Report, 1974).

In 1975, Wal-Mart leased an IBM 370/135 computer system, which was utilized to maintain inventory control on an item basis for all merchandise in the warehouse and distribution centers and on a classification basis for each Wal-Mart store, and to prepare income statements on a store-by-store basis. Singer electronic cash registers in 64 Wal-Mart stores and NCR mechanical and electronic registers in 61 Wal-Mart stores recorded point of sale data used to maintain inventory control (Annual Report, 1975).

In 1977, Wal-Mart built a company-wide computer network. With this communication network, messages pertaining to any phase of its operations could be sent to and from the stores immediately. The system was also used by the stores to place orders for merchandise, which expedited processing. Wal-Mart, as well, installed the cross-docking system in the Searcy Distribution Center. With the cross docking systems, merchandise in incoming trucks is unloaded at an entrance, processed, and conveyed automatically onto trucks on an exit, which are bound to branches stores (Annual Report, 1977).

Wal-Mart also developed a system for management of the corporate payroll. This system allowed store management to know, on a daily basis, their exact payroll costs and also permitted the stores to forward their payroll data to the general office without delay (Annual Report, 1977).

In 1978, the fashion distribution center utilized a computer-programmed split-ticket system of marking, where over a half a million fashion tickets could be received in an average week and quickly evaluated by buyers. Items were pre-marked and inspected for quality before leaving the distribution centers to ensure the condition of merchandise and uniform low prices in all stores. Buying decisions, reorders on hot-selling items, markdown percentages, and the automatic replenishment program could be calculated rapidly and with more accuracy (Annual Report, 1978).

Because of rapid growth in its computing power, Wal-Mart built a computer center in 1979. The first installed store computer terminal was the IBM 3774 (Annual Report, 1979).

Wal-Mart's computer development group set priorities on projects and studied the constant changes in computer technology. The aim of this department, in addition to rapid company communication, was to refine the information acquired and render it more useful and easier to read.

Computer specialists and management constantly evaluated programmed data for accuracy, simplicity, and usefulness. As a result,

Outdated reports can be discontinued immediately and necessary adjustments can be made to improve existing computer programs. The financial savings and the number of personnel hours saved daily by using the computer center is incalculable---even by the computer. (Annual Report, 1978)

In 1981, Wal-Mart implemented a new purchase order system and successfully tested a point-of-sale scanning system (Annual Report, 1981).

In 1983, Wal-Mart finished a two-year project to upgrade its in-store, backroom computers in all stores. Wal-Mart began to use Uniform Product Code as its product barcode in electronic scanning of point-of-sale (POS) data (Annual Report, 1983).

As the number of stores continued to increase, the timeliness of data became more critical. In 1984 Wal-Mart decided to establish a satellite communication network to allow the simultaneous sending and receiving of information to all stores (Annual Report, 1984). As the Wal-Mart Annual Report of 1985 noted,

The network will assure direct store-to-store and store-to-general office voice and data communication at a fixed future cost in a general environment of escalating cost. Communications will be enhanced further by WSN's capacity to beam live video presentations from the general office to the balance of the system.

Wal-Mart's Satellite Network (WSN) was successfully set up in 1987.

In 1984, Texlon handheld terminals were used as a direct assistance to store associates in re-ordering merchandise. Upon scanning a shelf label, the unit provided a description of the merchandise and information on prior quantities ordered, cost and retail of the merchandise item and the extended cost and retail of quantity being ordered. This device and accompanying system saves significant time in the ordering cycle, which in turn improves services to Wal-Mart customers (Annual Report, 1984).

In 1987, the Universal Product Barcode (UPC) was applied to all the stores and distribution centers. A check-in system designed to take full advantage of container bar code labeling, was also in the backroom of every Wal-Mart store in summer 1987. The use of the UPC brought new convenience in inventory control and distribution management: more accurately sorted shipments and immediate paperless billing; cost efficiencies in the stores' backroom freight processing and automated receiving (Annual Report, 1987).

In 1988, Wal-Mart deployed EDI systems.

In 1990, a prototype of data warehouse system was created at Wal-Mart, which stored historical sales data. In 1992, as a response to the suppliers' request for sales data sharing, Wal-Mart deployed the Retail Link system to further their partnership by moving beyond electronic data sharing. With the Retail Link system, Wal-Mart provides its vendors quality information concerning sale trends and inventory levels to facilitate genuine partnering. The actual system capitalizes on UPC and satellite communications capacity to bring its suppliers closer to its individual stores.

Wal-Mart invested $2.5 billion in capital expenditures in 1990 and $2.8 billion in 1993 for building and maintaining systems, distribution, and transportation infrastructure capacity. The goals are to not just sustain growth, but improve its productivity and reduce expense in the existing operation (Annual Report, 1993).

In 1996, Wal-Mart made Retail Link and EDI available through the Internet and began to use the Internet as an application platform. As one report estimated, up to the year 2001, Retail Link cost Wal-Mart about $4 billion and its suppliers $40 billion (Technological Review 03/2002). Wal-Mart, in 2001, further invited Atlas Commerce to add to Retail Link functions of a private online marketplace.

Wal-Mart and Sam's Club went online in June 1996. The online sites use Secure Socket Layer and Secure Transaction Protocol for secure online card transactions. In November 2004, Wal-Mart Stores launched one-hour digital photo service online. Wal-Mart customers and Sam's Club members can now upload digital photos online and pick up prints at local stores and clubs after one hour.

In 1997, Wal-Mart further increased the size of its data warehouse to 24 terabytes with 30 decision support applications. The data warehouse could process up to 50000 queries per week. In the same year, Wal-Mart installed data mining software developed by NeoVista Software, Inc. in Cupertino, Calif. (now part of the JD software group). NeoVista data mining system allows Wal-Mart to analyze point-of-sale data from each store down to the item level and assists automated product ordering and replenishment systems. Rob Fusillo, director of replenishment systems at Wal-Mart, explained that the system can further help Wal-Mart reduce inventory costs. In 2004, the storage of data warehouse amounted to 423 terabytes.

In 2002, Wal-Mart chose Internet Protocol for electronic data exchange with thousands of its suppliers in the world. With the use of the Electronic Data Interchange-Internet Integration Applicability Statement (EDIINT AS2) protocol can lower costs for the company and its suppliers. The adoption of EDIINT AS2 means that Wal-Mart suppliers will not need to resort to deploying value-added networks, which are usually expensive to use.

In 2004, Wal-Mart started to deploy RFID (Radio Frequency Identification) tags or electronic product codes. In April, it first tested the RFID tag in seven stores and a regional distribution center in Texas. Wal-Mart required its top 100 suppliers to have cases and pallets RFID tagged starting January 2005. Wal-Mart has plans to implement the electronic product code in its stores and distribution centers one after another. And Wal-Mart is urging its top suppliers to follow suit. Though there are debates about the functionality and cost problems of the RFID tag, Wal-Mart top management believe that the electronic product code is going to bring huge benefits to Wal-Mart and its suppliers and further cut their inventory and supply chain costs. Electronic product code through RFID tag will allow more data to be packaged and collected and speed up checkout for customers and provide better visibility in inventory management. RFID, together with EDI and UCCnet, can provide a huge convenience in synchronizing the supply chain[1].

The Wal-Mart Information System Division agenda includes projects such as revamping supply chain processes, synchronizing product data with suppliers using the UCCnet standard, and improving E-commerce platform, developing talent, and fostering regulatory compliance across the globe.

DISTRIBUTION CENTERS

Wal-Mart was not the first organization to set up a distribution center. In 1969 it built a distribution center, following Kmart and WoolCo's[2] lead. Before 1969, Wal-Mart bought most of its merchandise from wholesalers and some directly from suppliers. The distribution centers changed the way the merchandise was purchased and distributed. The distribution centers internalized services that were previously provided by external suppliers: manufacturers, trucking companies, and wholesalers.

A distribution center is usually equipped with many advanced information systems and robotics. Advanced systems in a distribution center includes: a system for product barcode, a cross-docking system, a stock location system, systems for tachograph analysis, vehicle routing and scheduling, systems for information processing and flow between a distribution center and the Merchandising Division such as ordering and accounting systems. Wal-Mart has been upgrading its distribution centers with new advanced systems. McKinnon (1990) analyzed four ways through which distribution centers can improve inventory control at levels of both distribution centers and stores and thus reduce inventory costs. In the case of Wal-Mart, the benefits of distribution centers include:

1. Volume buying from manufacturers brings huge savings. As its Annual Report of the year 1973 read, "The Distribution Center does not carry any item that would not result in at least a 5 percent savings, unless there were other reasons involved." Later annual reports have further confirmed this principle.
2. Better scheduling control of its own trucks dramatically reduces freight costs: "Approximately 60 percent of the trucks returning to the distribution centers from store deliveries are used to backhaul merchandise from suppliers" (Annual Report, 1973). The year 1977, 1980, and 1997 Annual Reports emphasized that Wal-Mart private truck fleets brought in large savings in freight costs, and was a key source of competitive advantage.
3. Reduction in inventory cost because of frequent and quick replenishment from distribution centers to local stores instead of slow replenishment from suppliers to local stores. All stores are one-day's drive away from the distribution centers. The year after building the first distribution center, 66 percent of merchandise was distributed to local stores through the

distribution center. In 1984, this percentage rose to approximately 77 percent of the merchandise sold in Wal-Mart stores. Now this percentage rate is kept around 80 percent.

4. More effective quality control results in less shrinkage loss and reduces return service from local stores to manufacturers, saving money.

Distribution centers internalize a variety of services that were previously provided by manufacturers, wholesalers, and outside carriers. Efficient managerial hierarchies and smooth information flow enabled by IT systems have replaced market mechanisms in coordinating distribution and logistics activities. Building a distribution center, however, needs a large amount of investment. Only when the business scale is big enough will this be possible and justifiable. Without adequate economies of scale, a distribution center can be prohibitively costly. Information systems and automations from robotics almost fix the costs of running a distribution center when the merchandising volume increases, because marginal cost of handling an additional unit of commodities or additional transaction is very small or insignificant in the short term.

Comet, a British retailer, has also reported that centralized distribution centers save costs by delivering merchandise from its centers to branch stores (McKinnon, 1990). It should be noted that small and medium sized retailers have to buy distribution service from wholesalers.

Wal-Mart has described the importance of distribution centers in its annual reports (1974, 1980, 1984, 1988, 1992, and 1998):

Distribution is a key in our ability to remain competitive. Logistics, distribution centers and transportation-the Wal-Mart distribution team is a key in our ability to remain competitive. Our 22 centers, averaging almost one million sq ft, received and shipped more than 769 million cases to our stores this past year. Our private fleet enables customized cost-efficient delivery to our stores, accommodating peak seasonal periods, night deliveries, and accelerated delivery. Our 2500 drivers and 16 000 distribution associates hard work and commitment to continuous improvement make this investment in centers and equipment pay by improving the in-stock position of our stores and making just-in-time inventory management a reality for us and our vendors. (Annual Report, 1992)

Combine these information systems with our logistics—our hub-and-spoke system in which distribution centers are placed within a day's truck run of the stores—and all the pieces fall into place for the ability to respond to the needs of our customers, before they are even in the store. In today's retailing world, speed is a crucial competitive advantage. And when it comes to turning information into improved merchandising and service to the customer, Wal-Mart is out in front and gaining speed. (Annual Report, 1998)

COLLECTION AND MANAGEMENT OF POS DATA

Wal-Mart recognizes that if it can control the sources of information and intelligently analyze and process information, it can influence those entities, which need these types of

information for their decision making. Wal-Mart, through its collection and analysis of POS data, has tried to develop information asymmetry over its suppliers and competitors. Though economies of information asymmetry may vary with respect to specific situations, such asymmetry may enable Wal-Mart to extend its business areas over production processes, devote its resources for product development, and create a series of own-label products.

At Wal-Mart, a POS data table usually consists of: store number, item number, department number, activity sequence number, selling unit quantity, selling amount, selling cost, Monday unit quantity, Tuesday unit quantity, Wednesday unit quantity, Thursday unit quantity, Friday unit quantity, Saturday unit quantity, and Sunday unit quantity (Westerman, 2001). Data are recorded with product bar codes, through which product and supplier information is coded and can be read and tracked. POS data can be analyzed for the studies and forecasts of customer demand changes, processed to display instant inventory movement down to the item level, and mined to find consumption patterns of customer and desired patterns of product display. Today all the Wal-Mart stores have POS systems connected with the Wal-Mart data center and integrated with the Wal-Mart data warehousing systems.

Back in 1984, Wal-Mart found that the data amount became too large and it became very burdensome to analyze without the aid of more advanced information systems. To simplify the use of POS data, the Data Collect system was developed in 1985 to speed data gathering and provide decision support for people who needed to analyze these data. It provides accurate tracking of merchandise sales by item for sale management on a daily basis. The system creates opportunities to maximize sales on "hot items" by maintaining proper in-stock position. Scanning also serves to reduce shrinkage by providing control over the capture and recording of markdowns (Annual Report, 1985).

Wal-Mart's IT department created an Executive Information System (EIS) in 1988 (Westerman, 2001). The EIS incorporated information from POS systems. Most of the data were about store operations, such as yearly or monthly sales, some information down to store and department level and article level. IBM's earlier version of SQL could be run in the system. Data were usually updated nightly, but there was a contest for computing sources between functional analysis and transactional processing.

To relieve the conflict, a data warehousing project was suggested and approved in 1990 to support analytical processing. The prototype system was first tested at the Merchandising Division where buyers and their assistants liked to do time series analysis over data concerning orders, receiving, and articles. The prototype system was very welcomed and proved to have a high return of investment (Westerman, 2001). POS systems were the data sources for the data warehousing system and were soon integrated with the data warehousing system. Two years later the data warehousing system was expanded and became a company-wide system, becoming an integral and fundamental block of the Wal-Mart information system network. With data warehouse available, data mining was also started in 1990 (Annual Report, 1998).

The control and analysis of POS data give Wal-Mart powerful sources to better understand its customers, enable more delicate decision-making processes, and motivate Wal-Mart to integrate some activities and risks that are originally undertaken by manufacturers and suppliers. Lee Scott, the current CEO and President of Wal-Mart, marked in the Annual Report 2003 that Wal-Mart had made big strides in internal product development. He noted, "our product development team, working with Wal-Mart buyers, is driving significant improvement in key product categories such as apparel, domestics and electronics." Wal-Mart now owns a rich series of private label products, and makes decisions on its own on how brand-name products should be presented and displayed.

Before 1992, there were few own label products at Wal-Mart. Own label products consisted of a very small portion of its revenue. In 1992, a series of Sam's American Choice was introduced as private label products. Actually Sam Walton began to plan such private labels in 1985 when Data Collect system was installed. Today Wal-Mart Stores' own label products include: OI' Roy; Great Value; Equate; Spring Valley; White Cloud; Glory; Sam's American Choice; Neighborhood Market; EverActive; Member's Mark; Special Kitty. Wal-Mart is deeply involved with private label products design, production, packaging, etc. Own label products are more profitable than national brand products because own label products do not need as large of advertising fees as national brands.

In 1997, Rob Fusillo, Director of Replenishment Information Systems at Wal-Mart, explained that Wal-Mart had started a data mining application that analyzed sales of individual items "at a lower level than we could ever have done previously."

Limited sharing of POS data with its supplier also enables new type of collaboration and coordination and extends Wal-Mart's control into the boundaries of its suppliers' business, such as production planning, forecasting, and package size and design decisions. Information sharing is limited to the sales data of suppliers' products. A supplier can only know the inventory movement of their own products. Wal-Mart discontinued the sharing of its POS data with retailing consultancy companies in 2001 (retailinfo.com 2001). Wal-Mart never shares any of its proprietary information with its suppliers, such as Wal-Mart's cost information. That is, no information sharing is intended to affect Wal-Mart's bargaining power when its buyers negotiate with its suppliers. Information sharing is often designed to increase mutual understanding between Wal-Mart and its suppliers and help its buyers convince its supplier representatives to yield and shorten negotiation processes.

INVENTORY CONTROL WITH EDI AND RETAIL LINK

In the early 1970s, Wal-Mart used computer systems to keep track of inventory down to the item level. It also developed a vender system to aid its distribution center rebuyers. To further reduce lead time, Wal-Mart installed the purchase ordering system in 1981. EDI was deployed to improve information coordination and processing with its suppliers in 1988. In 1983, Wal-Mart reported an obvious reduction in its lead time between ordering and receipt of merchandise. In 1984 Wal-Mart associates could use Texlon handheld to reorder merchandise and get information about any item inventory. In 1986 UPC was applied in inventory management. In 1987 the WSN project

was finished. This made both inside and outside coordination quicker and smoother. Simultaneous information sharing was enabled across the company and stores.

In 1992, Wal-Mart started to deploy Retail Link. Retail Link has relied on Wal-Mart data warehousing systems since it was first developed (Westerman, 2001). Retail Link allows suppliers to access their products' sales and inventory data at Wal-Mart Stores. Information sharing through Retail Link significantly reduces negotiating time and lead time between Wal-Mart and its suppliers (Westerman, 2001; Wal-Mart Annual Report, 1998).

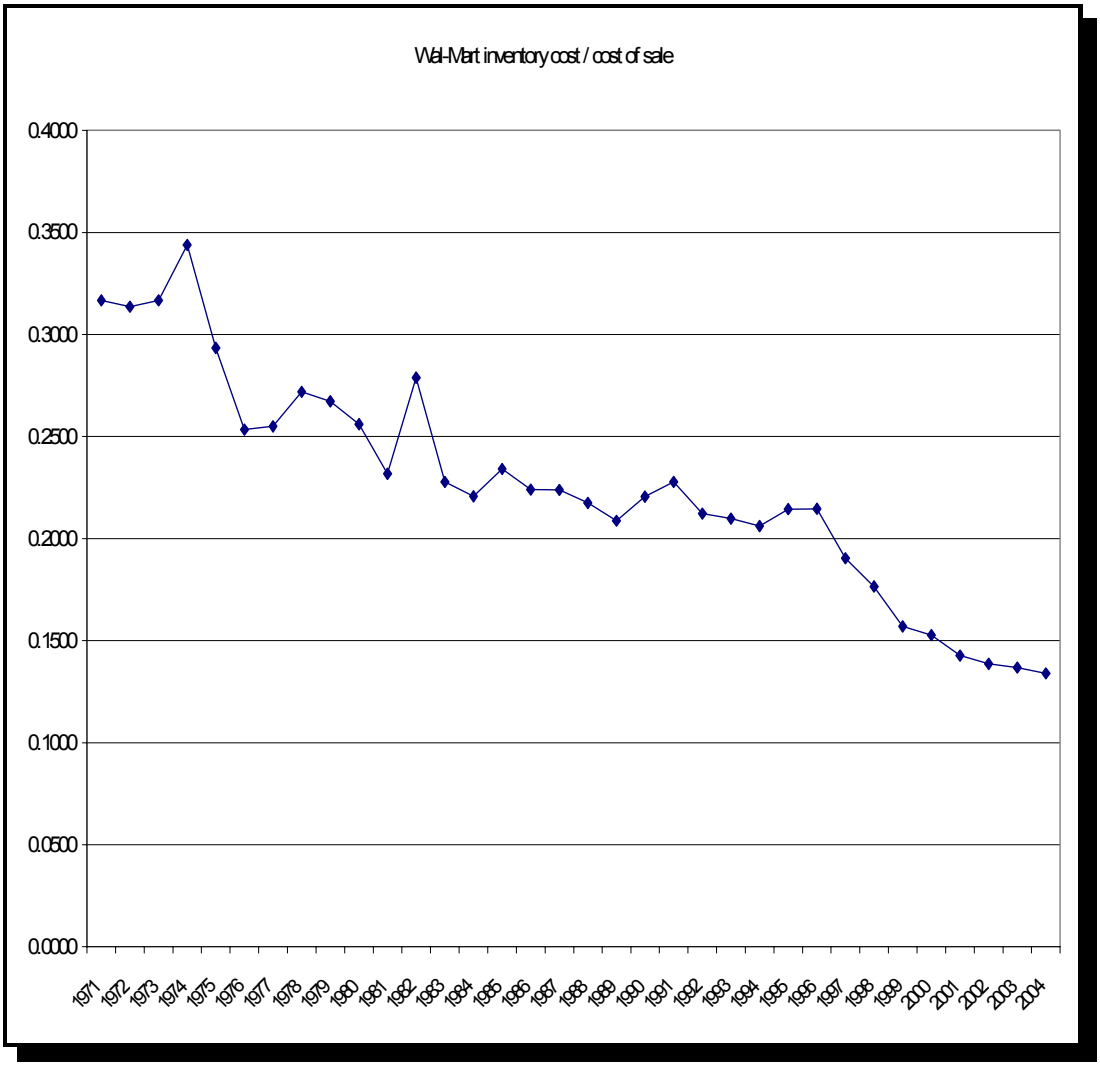
Data warehousing and data mining systems enable Wal-Mart to better understand and forecast its customers' demands and increase the probability of buying the right merchandise for its customers at the right amount and at the right time.

These systems contribute to reducing the inventory level and increasing inventory turnover as described by the CIO and other executives of Wal-Mart in the Wal-Mart Annual Reports of 1997, 1998, 1999 and 2001. As we can see from Figure 1 and 2, there are apparent increases in inventory turnover and obvious decreases in the ratio of inventory cost over cost of sale.

Figure 1: Wal-Mart Inventory Turnover 1971-2004

Data Source: Wal-Mart Annual Reports 1971--2004

Figure 2: Wal-Mart Inventory Cost/ Cost of Sale



Data Source: Wal-Mart Annual Reports 1971--2004

Retail Link has cost Wal-Mart about $4 billion up to 2001. According to Kevin Turner, then-CIO and Vice Executive President of Wal-Mart:

> Retail Link is the business leader for supplier collaboration via the Internet and has been a source of competitive advantage for Wal-Mart and SAM'S Club since 1991.

> Retail Link is an Internet application that is a free offering to our suppliers so that our suppliers may collaborate with our buyers to make better decisions that result in a lower cost of goods for our customers. Retail Link allows Wal-Mart suppliers and Wal-Mart merchants to view, manipulate and access 104 weeks of on-line, real-time, item level data that is kept at the lowest level of detail.

Since there is instant information sharing between Wal-Mart and its suppliers, the suppliers can use better data of inventory change at Wal-Mart when they schedule their production. Wal-Mart leads the retail industry with its version of a "just-in-time" supply system in which "computers track every product and automatically alert warehouses when it's time to restock the shelves" (Wal-Mart Annual Report, 1997). On the one hand, Wal-Mart suppliers are able to better plan production and reduce their inventory level though Wal-Mart may just be one of their product outlets. On the other hand, Wal-Mart itself is able to reduce its inventory level through timely replenishment of more wanted products at more accurate amounts or by simply asking its suppliers to take care of their products inventory management at Wal-Mart stores.

Wal-Mart is also testing new systems. It is implementing and applying CPFR (collaborative planning, forecasting and replenishment) and CTM (collaborative transportation management) with P&G.

Bob Connolly, Executive Vice President of Merchandising in 1997, said that there were four keys to the improvement in inventory management: 1) Systematic reduction of unproductive inventory; 2) Reduction of orders by 15 percent, enabling stores to manage their own inventory; 3) Reduced pack size across many categories; 4) Timely mark-downs. Rather than blindly slashing inventory, Wal-Mart has used the data gathered by technology to make more inventory available of the key items that customers want most, while reducing inventories overall (Annual Report, 1997).

The deployment of the electronic product code is simply to further improve the efficiency of inventory control and supply chain management. Linda Dillman, current CIO and Executive Vice President of Wal-Mart, believe that RFID will bring $7 billion benefits to Wal-Mart and its suppliers in year 2005 (InformationWeek.com, 2004).

PRODUCTIVITY IMPROVEMENT

In 1980 Wal-Mart reported that:

> Merchandise productivity is enhanced by the Distribution and Transportation Divisions. Merchandise flows from the manufacturers to the Company's distribution centers. Distribution center facilities efficiently sort the large quantities received into outbound shipments to each store. Deliveries are made on Wal-Mart trucks which backhaul other merchandise to the warehouses, eliminating as many miles traveled with empty trailers as possible. . . . Merchandise productivity is improved by the

> utilization of the store terminal network systems, which provides a means for replenishment as well as tracking of item movement and changing sales trends. Having information immediately available enables Management to respond quickly to any problems or opportunities . . . Wal-Mart continues to explore all areas for productivity improvement possibilities (Annual Report, 1980).

"Make technology pay" is frequently cited within Wal-Mart stores as new equipment, software, and communications are applied to reduce costs and improve productivity. As the Annual Report for 1992 stated Wal-Mart's "aim is the simplification of what we do, elimination of waste, and access to more meaningful information." Wal-Mart in its 1996 Annual Report revealed some information of its IT strategy:

> With an annual technology and communication budget of $500 million and an information system staff of 1200, Wal-Mart leads the industry in information technology—and we're not slowing down. We know our future earnings growth has to come not just from increased market share, but also from increased productivity. (Annual Report, 1996)

McKinsey Global Institute reported in 2002 that:

> In 1987 Wal-Mart had just 9 percent market share but was 40 percent more productive than its competitors. By the mid-1990s, its share had grown to 27 percent while its productivity advantage widened to 48 percent. Competitors reacted by adopting many of Wal-Mart's innovations, including . . economies of scale in warehouse logistics and purchasing, electronic data interchange and wireless bar code scanning. From 1995 to 1999, competitors increased their productivity by 28 percent while Wal-Mart raised the bar by further increasing its own efficiency another 20 percent. (Technology Review, March 2002)

ECONOMIES OF SCALE AND SCOPE

Economies of scale and scope are extremely important to any retailers. Economies of scale mean a decreasing marginal cost for processing an additional unit. In system application, it means applying systems to their reasonable maximum scale such that marginal cost for additional unit of information processing can be minimized or even ignored. Economies of scope mean a decreasing marginal cost for an additional unit of a relevant service or product. In system application, economies of scope mean that systems can be applied to different business processing with a decreasing additional cost.

Construction and operation of distribution centers reveal how Wal-Mart has looked for economies of scale and scope. The information systems can be used to handle information processing and distribution for Wal-Mart Stores, SuperCenters, Sam's Clubs, and Neighborhood Markets. They are used to process information for retailing of both grocery and general merchandise.

A distribution center usually has to handle a large variety of commodities. Before 1992, Wal-Mart was just a general merchandise retailer. Now a Wal-Mart discount store has about 36,000 different kinds of commodities. About 80 percent of merchandise is delivered from a distribution center. So a distribution center should be able to handle about 28,800 kinds of commodities. After Wal-Mart Supercenters and Sam's Clubs are opened, distribution centers have to deal with the double number of kinds of commodities, including general merchandise and grocery. A Supercenter usually has about 72,000 items of commodities; and a Sam's Club has more to cope with. So systems must be designed with capabilities to handle a large variety of merchandise. The growth of Wal-Mart International increases the complexity of information systems. Nevertheless, Wal-Mart Information System Division has been developing systems that can be used across different countries since Wal-Mart acquired a Canadian retailer in 1994.

That is why UPC is so important. It is used to integrate information from different terminals and systems. Information about merchandise can be traced by using UPC, regardless of the physical shape or the physical property of that product. Because of UPC, Wal-Mart merchandising systems can be applied to manage merchandise flow for different retailing channels. The implementation of the electronic product codes will lift system efficiency to a higher level.

The economies of scale and scope in system use are tremendous. The setup cost or fixed cost of building a system is high. But the marginal cost for processing an additional unit of information is very low--just some utility fees. Depreciation in the short term can almost be ignored. So the cost efficiency of large-scale IT systems requires a larger scale whenever reasonable and possible. Wal-Mart is strongly motivated to increase its scale to reduce system costs amortized to any additional dollar sale.

In 1973, the Bentonville Distribution Center distributed goods for 64 Wal-Mart discount stores. In 1975, the Bentonville Distribution Center was responsible for 104 stores with the sales of $236.2 million. In 1992, the 22 distribution centers provided service to 1928 Wal-Mart stores and Sam's Clubs with revenues of $43.9 billion. In 2001, the 33 distribution centers distributed goods to 3719 Wal-Mart discount stores, Supercenters, Sam's Clubs, and Neighborhood Markets in the USA, Mexico, and Canada with sales revenues at $191.3 billion. In 1977, the only Distribution Center at Bentonville handled about 70 percent of $478.8 million goods. In 2001, each distribution center on average had to cope with 80 percent of $4.72 billion goods and more kinds of goods. See Table 2 and 3 and Figure 3 for the detail.

Table 1
Sales Supported by Bentonville Distribution Center

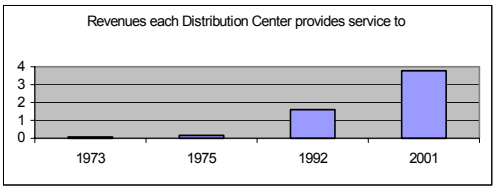
	1973	1975
Bentonville Distribution Center (236, 800 sq ft)	64 stores	104 stores
Sales supported by the distribution center	55% of $124.9 ($ million)	65% of $236.2 ($ million)
Sales supported by per sq ft at a distribution center	$290.1	$648.3

Table 2
Sales Supported by an Average distribution center

	1992	2001
Number of Distribution Centers with average size of about one million square feet	22	33
Revenues	$43.9 billion	$155.8 billion*
Average sales supported by the distribution center	80% of $1.99=1.592 ($ billion)	80% of $4.72=3.77 ($ billion) **
Sales supported by per sq ft of a distribution center	$1592	$3776

*Wal-Mart in 2001 revenue was $191.3 billion. Sales outside North America were $35.5 billion. Data are derived from Wal-Mart annual reports.
** In 1973, only 55 percent of its merchandise was distributed through its distribution center. The ratio rose to 65 percent in 1978. Since late of 1980s, the ratio has been kept around 80 percent.

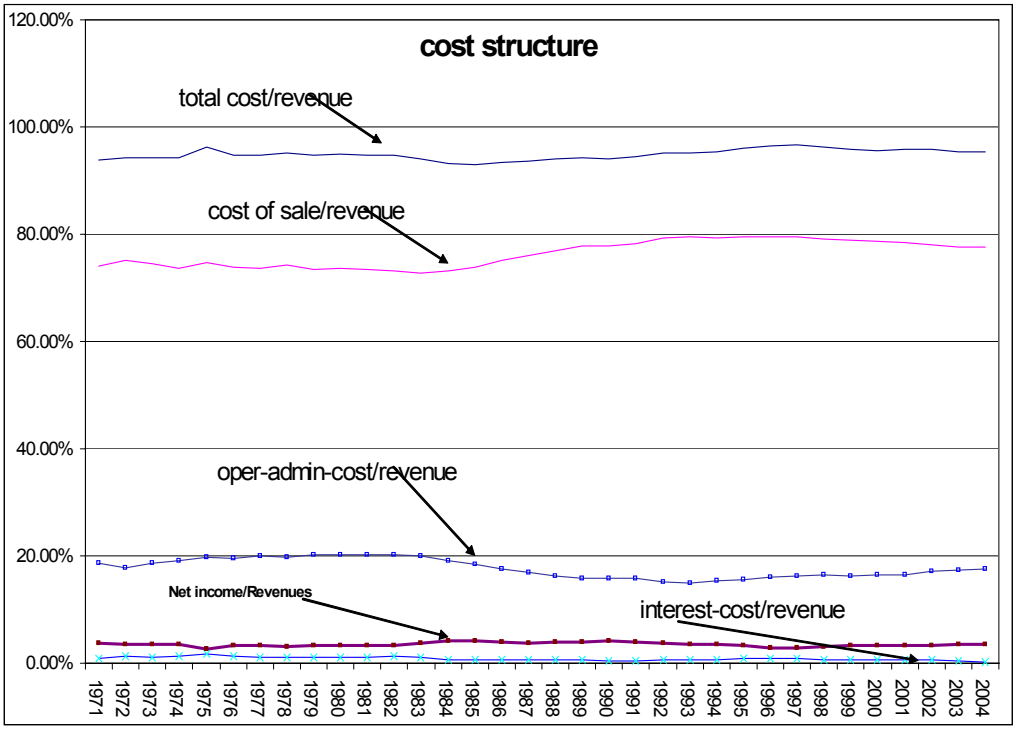
Figure 3: Average Sales Supported by A Distribution Center ($ billion)



Data Source: Wal-Mart Annual Reports 1971--2001

As IT systems contribute to the economies of scale and scope in Wal-Mart's merchandising processes, many people have expected that such a firm would have a declining ratio of administration cost over revenue. In Figure 4, we can see the operation-and-administration-cost/revenue curve declined from 1982 to 1993, but has slightly increased since 1994.

Figure 4: Wal-Mart Cost Structure 1971-2004



Data Source: Wal-Mart Annual Reports 1971--2004

BUSINESS TECHNOLOGY STANDARDS AND EXTERNAL ECONOMIES

Worcester (1969) found that external economies might enable a company to grow into a monopoly. External economies of highway, telecommunication network and technologies have been widely discussed and agreed by economists (Worcester, 1969). External economies of software technologies, termed "network externalities" by Katz and Shapiro (1985), also have been discussed. External economies of prevailing use of a software system among the public are well acknowledged though some authors would prefer to use other terms and think that network externalities are actually economies of scale in the use of software (Liebowitz and Margolis, 1994).

Technological development in telecommunication, computing, and networking have brought down the cost for information and communication processing. Prevailing use and thus wide availability of these technologies enable firms to enjoy the external economies such that marginal cost for information and communication processing has been decreasing. Wal-Mart has enjoyed external economies from the adoption of general IT technologies such as POS, UPC, telecommunication network, EDI, and Internet.

From 1983 to 1987, Wal-Mart changed its product barcode to the UPC system, resulting in many improvements in coordination consistency and convenience and big cost savings (Annual Reports 1983, 1984, 1985, 1986 and 1987). In 1988, Wal-Mart adopted EDI systems to further its effort for paperless document processing, making distribution centers online. By the early 1990s, Wal-Mart had set up arrangements with about 1810 of its approximately 5000 suppliers, making it the nation's largest user of the technology. The EDI system combined with the improved ordering system provides a more rapid replenishment of merchandise at a reduced cost (Vance and Scott, 1994).

In the mid 1990s, Wal-Mart upgraded EDI and Retail Link and made them available through the Internet. Efforts are being made to upgrade Retail Link to be XML-based, which enables compatibility between different transmission systems. Wal-Mart IT division carefully controls the software cost as the hardware cost dramatically decreases.

But Wal-Mart has hesitated in some areas about whether it should adopt any standards created by standard bodies or follow any technology vendors. In an April 2002 interview with *CIO.com* Editor-in-Chief Abbie Lundberg, Kevin Turner, CIO and Executive Vice President of Wal-Mart, said that Wal-Mart rarely buys any packaged software to avoid being tied by any software vendors. As he emphasized, Wal-Mart doesn't "have to run at the pace the software company wants to." Wal-Mart IT department does everything internally to, "take the best parts of software programs, get rid of the worst and customize to fit Wal-Mart's goals, and benchmark its progress against outsourcers and software developers.[3] That has been Wal-Mart's tradition in the management of IT system development.

Retail Link actually has had some features as a private online exchange since its start in 1990. There are retailing online exchanges (B2B) for suppliers and retailers, such as Transora and WorldWide Exchange (WWRE). But according to a report by Forrester Research, because Transora and WWRE are each seeking exclusive rights to retailer-supplier transactions, retailers are hesitant to invest resources in either one for fear of backing the loser in the contest. As a result, members have limited their participation to low-level activities like auctions, preventing any growth in exchange adoption. So Wal-Mart has decided to build its own B2B exchange based on Retail Link. Wal-Mart has invited Atlas Commerce to join its efforts to build a Web exchange. Retail Link actually set standards for supply chains when many system vendors developed their systems to be compatible with Retail Link and now connects Wal-Mart with its more than 10,000 suppliers. Vendors cannot ignore its existence when they try to develop a supply chain system.

Though UCC and ARTS, two technical organizations under VICS[4], are also developing some standards, whether retailers and suppliers would adopt the standards is uncertain. Ken Harris, SVP/CIO at GAP Inc., has doubted the viability of industry-wide standards for supply chains. As he argues, standards bodies cannot accomplish their objectives because leading companies will join in consortiums to create rival camps that will set standards for their own efforts. Actually, Wal-Mart Retail Link has a tendency to be the standard of supply chain coordination in the retailing industry.

A HIGHLY CENTRALIZED AND INTEGRATED STRUCTURE

Wal-Mart has been centralized and integrated in many areas from the start, and has kept a tight control over IT development all along. As Kevin Turner says, Wal-Mart's IT department follows three rules in system development: 1) Maintain a centralized information system infrastructure, 2) Standardize systems and platforms, 3) Be merchants first and technologists second. Wal-Mart has been widely recognized as one of the best places for IT people to work because of its productive and efficient management of system development. Wal-Mart's organizational structure does not conflict in any way with its IT policy, as Wal-Mart itself has been very centralized and integrated. There have been few changes in Wal-Mart's centralized organizational structure since the adoption of IT systems.

IT systems are tightly integrated to assure economies from system security, compatibility, and integrity. The system security, compatibility, and integrity provide the technological foundation for economies of scale and scope.

Tightly integrated IT systems do not necessarily pre-require an integrated business structure or ownership control or centralized administrative structure. With the support of IT systems, a centralized structure can also be very flexible and responsive to internal or external changes. The major problem facing Wal-Mart is the increased cost of sales and the slight but steady increase in the cost of operation and administration.

From Figure 4, we can see that the cost of sale has been at a higher level since 1992. The ratio of interest cost over revenues has been kept at a pretty low level because of Wal-Mart's successful inventory management and category management to provide continuous cash flow. Wal-Mart's major costs include cost of sales and cost of operation and administration. Cost of sales includes actual product cost, change in inventory, buying allowance from suppliers, the cost of transportation to the company's warehouse from suppliers, the cost of transportation from the company's warehouses to the stores and clubs, and the cost of warehousing for the Sam's Club segment. When asked why the ratio of cost of sale over revenue was at a higher level, Wal-Mart's Corporate Office responded, "Because Wal-Mart is not willing to increase its prices." Wal-Mart could reduce the ratio by marking up prices. But then customers would visit Wal-Mart's competitors more frequently. On the other hand, there are limits to reducing suppliers' charges. In this case, what can Wal-Mart do to reduce its increasing cost of sales?

Another issue is whether Wal-Mart has applied IT systems to their potential limits, or how much further can Wal-Mart use IT systems to reduce its costs. This issue is raised because of the steady increase in the cost of operation and administration. The ratio of cost of operation and administration over revenues decreased from 1983 to the early 1990s. Since 1992, this ratio has been slightly but steadily increasing. It is worth asking whether or not Wal-Mart has applied IT systems to their potential limits. Are IT systems at Wal-Mart facing a diminishing return problem with Wal-Mart's current size of business? These are questions waiting for further answers.

CONCLUSION

The economies of IT systems at Wal-Mart are prominent from all the perspectives. IT systems at Wal-Mart reduce costs of transaction at the distribution centers and support Wal-Mart's long-term strategy of owning the distribution centers. While many studies have indicated that IT systems reduce transaction costs and lead to smaller companies, Wal-Mart obviously is a clear case that shows the opposite.

The literature on information asymmetry and IT systems suggests that IT such as Internet can reduce information asymmetry. The POS systems and data warehousing and mining systems definitely work in a different way. While POS systems collect data for Wal-Mart and increase Wal-Mart's visibility about customers and markets, Wal-Mart's suppliers have to rely on whether Wal-Mart is willing to share information with them. There is asymmetry in understanding customers and markets between Wal-Mart and its suppliers.

Improvement in productivity due to IT systems is dramatic at Wal-Mart. Information sharing through the Retail Link between Wal-Mart and suppliers improve inventory control of Wal-Mart and suppliers. But probably we will never be able to get a complete picture of how Wal-Mart is managing and doing those things.

By using standard technologies, Wal-Mart has positive externalities. Standard technologies also provide technical foundations to realize the economies of scale. But with the size it has, Wal-Mart keeps its own standards in some areas such as supply chain management using the Retail Link system.

The economies of IT systems at Wal-Mart are evident and unique. Yet we have to investigate further whether Wal-Mart IT systems are facing a diminishing return problem. This will need further research in the future.

ENDNOTES

[1] For benefits of RFID, please refer to Fugerson (2002), Jilovec (2004), and Sheffi (2004). UCCnet is a subsidiary of the Uniform Code Council, Inc. For details about UCCnet, please visit http://www.uccnet.org/.

[2] WoolCo was a Canadian discount retailer. It was acquired by Wal-Mart in 1994.

[3] See http://www2.cio.com/conferences/april2002/coverage57_content.html

[4] VICS, Voluntary Interindustry Commerce Standard Association. Please visit http://www.vics.org/home for details. UCC, Uniform Code Council, INC. Details about UCC can be found at http://www.uc-council.org/ean_ucc_system/stnds_and_tech/vics_edi.html. ARTS, the Association for Retail Technology Standards. Please visit http://www.nrf-arts.org/ for details.

REFERENCES

Amit, R. and Schoemaker, P. (1993). Strategic assets and organizational rent. *Strategic Management Journal*, 14: 33-46.

Arnold, S. and Fernie, J. (2000). Wal-Mart in Europe: Prospects for the UK. *International Marketing Review*, 17 (4/5): 416-432.

Bakos, J. Y. and Treacy, M. E. (1986). Information technology and corporate strategy: A research perspective. *MIS Quarterly*, 10(2): 107-119.

Besen S. and Farrell J. (1994). Choosing how to compete: Strategies and tactics in standardization. *Journal of Economic Perspective*, 8 (2): 117-131.

Bloom, P., and Perry, V. (2001, fall). Retailer power and supplier welfare: The case of Wal-Mart. *Journal of Retailing*, 77 (3): 379-96.

Brynjolfsson E., Malone T., Gurbaxani S., & Kambil A. (1994). Does information technology lead to smaller firms? *Management Science*, 40(12): 1628-44.

Carlton, D. (1979). Vertical integration in competitive markets under uncertainty. *Journal of Industrial Economics*, 27(3): 189-209.

Casadesus-Masanell R. and Spulber D. (2000). The fable of Fisher Body. *Journal of Law and Economics*, 40(1): 67-103.

Chandler, A., Jr. (1977). *The* visible *hand*. Cambridge: Harvard University Press.

CIO. http://www.cio.com.

ComputerWorld. (2003). http://www.computerworld.com,

Clemons, E. and Row M. (1992). Information technology and industrial cooperation: The changing economics of coordination and ownership. *Journal of Management Information System*, 9(2): 9-28.

Coase R. H. (1937). The nature of the firm. *4 Economica N.S,* 386-405.

Coase, R.H. (1988, spring). Lecture. Journal of Law, Economics, and Organization.

Coase, R.H. (2000). The acquisition of Fisher Body by General Motors. *Journal of Law and Economics*, 40(1): 15-31.

Cooke James. (1996, April). The retail revolution is coming! *Logistics Management*: 48-50.

Cooper R. and Slagmulder R. (1999). Supply chain development for lean enterprise. The IMA Foundation for Applied Research, Inc., Montvale, New Jersey.

Crocker, K.J. (1983). Vertical integration and the strategic use of private information. *Bell Journal of Economics*, 14: 236-248.

Demsetz, Harold (1995). *The economics of the business firm*. New York: Cambridge University Press.

Fernie J. and Sparks L. (1999). Logistics and Retail management. San Francisco: CRC Press.

Fernie J. (1990). Retail distribution management. London: Kogan Page.

Foote, P., and Krishnamurthi, M. (2001, fall). Forecasting using data warehousing model: Wal-Mart's Experience. *Journal of Business Forecasting Methods and Systems*, 20 (3):13-17.

Fugerson, Glover. (2002, June). Have your objects call my objects. *Harvard Business Review*: 138-144.

Grant, RM. (1987). Manufacturing-retailer relations: The shifting balance of power. In *Business Strategy and Retailing*. Ed. G Johnson. Chichester: Wiley.

Gurbaxani, V.and S. Whang. (1991). The impact of information systems on organization and markets. *Communications of the ACM*, 34(1): 59-73.

Jilovec, Nahid. (2004, November 10). EDI, UCCnet & RFID: Synchronizing the supply chain. *29th Street Press*.

Sheffi, Yossi. (2004). RFID and innovation cycle. *International Journal of Logistics Management*, 15 (1): 1-11.

Katz M. and Shapiro C. (1994). Systems competition and network effects. *Journal of Economic Perspective*, 8 (2): 93-115.

Lieberman M. (1991, September). Determinants of vertical integration: An empirical test. *Journal of Industrial Organizations*: 451-65.

Liebowitz S. and Margolis S. (1994). Network externality: An uncommon tragedy. *Journal of Economic Perspective*, 8(2): 133-150.

Levy M. (2001, June). *Retailing management.* New York: McGraw Hill.

Lowson B., King R., & Hunter A. (1999, July). *Quick response: Managing the supply chain to meet consumer demand.* New York: John Wiley & Sons.

Malone T., Yates J., & Benjamin R. (1987). Electric markets and electronic hierarchies. *Communications of the ACM*, 30(6): 484-497.

McKinnon, A. (1986). The physical distribution systems of supermarket chains. *Service Industries Journal*, 5(2): 226-38.

McKinnon, A. (1990). The advantage and disadvantage of centralized distribution. In *Retail distribution Management.* Ed .J Fernie. London: Kogan Page.

Moore, J. (1993, May-June). The evolution of Wal-Mart: Savvy expansion and leadership, *Harvard Business Review*: 82-3.

Myer, Randy. (1989, Nov-Dec). Suppliers—manage your customers. *Harvard Business Review*: 160-8.

Perry M. (1989). Vertical integration. In *Handbook of Industrial Organization*. Ed. by Richard Schmalensee and Robert Willig. London: Science Publishers B.V.

Porter, Michael. (1985). *Competitive strategy*. New York: Free Press.

Reid M. (1995, March 4). Retailing survey. *The Economist*.

Retail infosystems. (2003). http://www.retailinfo.com.

Reithel Brian et al. (2001). Competitive force/marketing mix framework. In *Strategic information technology: Opportunities for competitive advantage.* Ed. by Raymond Papp. Idea Group Publishing.

Shrage Michael. (2002, March). Wal-Mart trumps Moore's law. *Technology Review*: 21.

Stalk G., Evans P., & Shulman L. (1992 March-April). Competing on capabilities: The new rules of corporate strategy. *Harvard Business Review*: 57-69.

Vance S. and Scott R. (1994). *Wal-Mart: A history of Sam Walton's retail phenomenon*. Boston: Twayne Publishers.

Varley R. and Gillooley D. (2001, February 15). *Retail product management: Buying and merchandising*. New York: Routledge.

Wal-Mart Stores, Inc. Annual reports and SEC filings, 1972-2001.

Westerman, P. (2001). *Data warehouse: Using Wal-Mart model.* New York: Morgan Kaufmann Publishers.

Wetherbe J. and Frolick M. (2000, May). Cycle time reduction: Concepts and case studies. C*ommunications of Association for Information Systems,*3 (13).

Williamson Oliver. (1985). *The economic institutions of capitalism*. New York: Free Press, 1985.

Williamson, Winter S. (1991). *The nature of the firm: Origins, evolution, and development*. New York: Oxford University Press.

Worcester D. (1969). Pecuniary and technological externality, factor rents, and social cost. *American Economic Review*, 59(4): 873-885.

Yin, Robert, (1994). *Case Study Research*, 2nd edition. Sage Publications, 1994.

EXAMINING THE DIFFERENCES IN GENDER PERCEPTION IN THE USE OF SPEECH RECOGNITION AS A TOOL IN GROUP SUPPORT SYSTEMS

Carl M. Rebman Jr., The University of San Diego
Casey G. Cegielski, Auburn University
Brad Prince, Auburn University

ABSTRACT

Group Support Systems (GSS) technology is recognized as a tool with advantages for facilitating, improving, and speeding communication between the different members in a meeting session. One innovation that could enhance the efficiency and effectiveness of GSS technology is speech recognition. While the concept of human speech interaction with computer-based information systems is not novel, Speech Recognition (SR) technology presents an opportunity to reduce the challenges in human-computer interaction. One of the advantages that Speech Recognition offers is the ability to capture a larger amount of text/data over traditional keyboard entry. This is of interest as one of the limitations of GSS performance and end user satisfaction has been attributed to keyboard-based comment entry challenges. Simply put, people talk and think faster than they can type. This paper reports the results of an experiment using a prototype voice-based Group Support System. Fourteen groups of five subjects each submitted speech-synthesized comments into the system and read others' comments visually on computer screens. This prototype seeks to maximize the relative efficiencies of speaking and reading to generate the most ideas in the shortest amount of time while maintaining accurate transcriptions. Results of this study illustrated some interesting differences between genders. First, male subjects reported spending too much time on idea generation while female subjects reported a higher level of understanding of decisions made. Females also reported a higher level of satisfaction with both the decision outcome and the process. These results provide evidence that Speech Recognition may be a viable tool for decision-making processes where gender issues play an important role.

INTRODUCTION

As organizations seek to engage in methods to improve collaboration and communication methods they must also analyze societal factors that can either advance or impact end-user adoption. One such factor that can play a major role in end-user adoption is gender. According to Nicovich,

et. al, "the role of gender in reaction to communication stimuli has been of interest to both academics and practitioners for some time (Darley, 1995; Holbrook, 1986; Meyers-Levy & Maheswaran, 1991; Meyers-Levy & Sternthal, 1991; Prakash & Flores, 1984). Gender is often used as a segmentation variable in communication strategies, and interesting differences have been noted in academic literature over the past few decades." Furthermore, according to Wilson and Howcraft, "Within the IS literature, the issue of gender is largely under-theorized, partly due to the belief in the gender neutrality of technology" (Knights & Murray, 1994). Lastly, because there have been some reports of negative experiences with technology by women (Adam, 1997) it is important that research be conducted in the introductory stage of any new technology. This paper seeks to examine the viability of speech recognition as a new tool for current group support systems and more particularly its effects on gender acceptance.

OVERVIEW OF GROUP SUPPORT SYSTEMS

Group Support Systems (GSS) were originally envisioned by Huber (1984a), who extended the decision support systems (DSS) concept to group decision settings (Rebstock, et al., 1997). A GSS is a computer-based information processing system designed to facilitate group decision-making. It is a group technology, which represents a networked computer-based system that combines computing, communication, and decision support technologies (Aiken & Chrestman, 1995). A GSS supports a group of people involved in a common task through a shared environment. Its design is centered on groups and group productivity through idea generation, preference, and opinion exchange (Aiken & Chrestman, 1995; Aiken and Hassan, 1996). Group Support Systems allow for generating new ideas, developing new plans, and facilitating new research efforts (Anson, 1996). Overall group members in electronic meetings participate more, save more time, and are more satisfied than participants in traditional, verbal meetings (McLeod 1992).

A GSS typically includes: (1) a communicative interface component providing computer-mediated channels, (2) a decision modeling component-suggesting solutions, and (3) a documentation component recording the decision process (Pollack and Kanachowski, 1993; Alavi and Joachimsthaler, 1992, Huber, 1984b). In the past decade, numerous group support systems have been developed to support planning, brainstorming, negotiating, problem solving, and other decision making processes utilizing a variety of technologies (Teng and Ramamurthy 1993; Dennis et al., 1988; Huseman and Miles, 1988; Johansen, 1988; Kraemer & King, 1988; DeSanctis and Gallupe, 1987; Huber, 1984b). Improved computer-based information processing technologies and increased computer literacy have led to a constant impetus for further GSS development (Broome and Chen, 1992; Huber, 1984a). Consequently, GSS systems vary greatly today as a GSS can mean anything from electronic mail systems to sophisticated decision rooms and nearly anything in between (Rebstock, et al., 1997).

CLASSIFICATION OF GROUP SUPPORT SYSTEMS

Group Support Systems facilitate group decisions by allowing groups to meet in a single room or at remote sites, and the computer replaces chalkboards and projected video images. The interactive computer system has specifically designed hardware and software, which are designed to be easy to learn and use (Aiken, 1992). There are four possible GSS settings (Turban, 1993):

Same Time-Same Place: Participants meet in the decision room. They are face-to-face, in one place at the same time. A brainstorming or voting session demonstrates this type of setting.

Same Time-Different Place: Participants are communicating at the same time, but they are in different places. This setting describes a meeting where participants are indifferent locations in the city.

Different Time-Same Place: Participants work in shifts in this type of setting. International trading is an example of this setting.

Different Time-Different Place: Participants are in different time zones or countries. A company may be located in Germany and its subsidiary is in Canada.

There are also three different types of GSS environments (Turrof et al., 1993):

Single-computer systems are the simplest types of group support systems designed primarily for single users that are dispersed geographically. Each group member uses a video display system through which communication with other group members takes place. Members can observe how other group members approach problems and participate by presenting their ideas. Single-computer systems are relatively inexpensive, portable, and suitable for problems whereby confidentiality is not important (Gupta, 1996). One problem suitable for single-computer systems is making investment decisions based on standard weighted criteria. Group members can assign individual weights to the criteria, the system ranks the various alternatives, and members rank the criteria. Changes are suggested, while assigning and prioritizing the weights takes place until consensus is attained (Gupta, 1996).

Keypad-response systems are systems where group members communicate with each other using hand-held keypads. In this type of system, group members are linked by networked personal computers. They are in a room referred to as the decision room. A projection screen at the front of the room is visible by all members

in the decision room. All inputs, comments, or ideas are displayed on the screen. Each member uses the keypad to coordinate and communicate ideas with each other. The keypad has customized rating scales capable of processing, analyzing, ranking, and displaying the output, in text or graphical form. The advantages of keypad systems are anonymity, portability, and instantaneous summarization of input (Gupta, 1996).

Full-keyboard workstation systems are similar to the keypad-response systems, where a full-keyboard workstation system is also built in a single room containing a projection screen and a series of personal computer workstations networked through a facilitator workstation. Each workstation has a private display monitor. Because group member responses are displayed on the projector screen, as well as on the participant's monitors, anonymous communication is maintained. The system design is adaptable to a wide array of uses and informational requirements. The tools utilized for generating, evaluating, and ranking ideas result in efficient group sessions because time is not lost in tabulations, and because participants can monitor the progress of their work (Gupta, 1996; Aiken and Chrestman, 1995).

SPEECH RECOGNITION

Automated speech recognition technology has been around since the 1950's, but the trade press has begun only recently to focus attention on it because of recent advances in the quality of software, and personal computers have become fast enough to process it in real time (Lindquist, 1999; Miatkowski, 1999). SR technology digitizes spoken words, identifies individual sounds (phonemes), and uses mathematical models to selected discrete words or complete phrases.

Speech recognition can be used to enhance dictation by eliminating keyboard input and increasing productivity. Similarly, it can be used in telephony to replace cumbersome touch-tone menus or to reduce the number of screen prompts required to reach the desired person. Lastly, SR systems can be used to control machinery or recording inspection, allowing hands-free operation and greater concentration on more important tasks.

SPEECH RECOGNITION TECHNOLOGY

Speech Recognition can be categorized in terms of speaker input and speech mode (Rodman, 1999).

Speaker Dependent or Independent Speaker dependent SR technology requires a user to train the program to recognize his or her voice (the process is referred to as "enrollment"), but this type of SR may be better for those with non-standard

speaking (e.g., patterns, dialects, or foreign accents) (Rabiner, 1993). Speaker-independent programs are designed to interpret any user's voice with no enrollment, but this type of software is usually less accurate. If the software is speaker-independent, a default set of discrete sounds or phonemes is provided, otherwise, user enrollment user creates a personalized set of phonemes for improved accuracy.

Continuous or Discrete Speech Continuous speech SR allows the user to talk normally, in complete sentences, with no breaks between words, while discrete speech SR requires speakers to pause after each word. Continuous speech usually is considered to be more natural, less frustrating, and faster. In addition, being more complex, continuous speech SR can recognize individual words (discrete speech) as well as entire phrases. Although some studies have suggested that discrete speech SR is more accurate, others have shown the opposite to be true (Klevans, 1997; Niccolai, 1997).

The speech recognition process follows five steps (Markowitz, 1996):

Audio Input The human voice is transmitted through a microphone connected to a PC with a standard sound card.

Acoustic Processor The acoustic processor filters out background noise and converts captured audio into a series of phonemes.

Word Matching The software attempts to match the sounds to the most likely words in two basic ways. First, it uses acoustical analysis to build a list of possible words that contain similar sounds. Then, it uses language modeling (the likelihood that a given word will appear between those coming before and after it) to narrow the list and come up with the best candidates. In addition, the word matching process draws on the domain (the set of vocabularies, pronunciations, and word-usage models, as well as a model of the user's speech and words) employed by the application. The user can extend the domain by adding new words or even creating multiple domains for different applications. Finally, continuous speech SR looks at contextual information to predict what words should come next. This also helps the system to distinguish homonyms.

Decoder The decoder selects the most likely word based on the rankings assigned during word matching and assembles the words in the most likely sentence combinations. It then transfers the sentence to the word processing application.

Text Output Some SR programs include their own word processors, but many also will allow text generation directly into a word processor or a text box in another application.

Several SR systems have been compared in the literature, and accuracies ranging from 80% to 99% have been reported (Alexander, 1999; Bethoney, 1999; Jang & Hauptmann, 1999; Wactlar, et al., 2000). Tests of Dragon Systems' *Naturally Speaking*, one of the leading SR program for personal computers, showed an accuracy of 89.1% at 74 wpm as compared to 87.8% at 69 wpm with IBM's *ViaVoice* SR software. Other tests of *DragonSystems* showed that text could be generated at 120 wpm with word accuracies of 80% to 95% (Pallett, Garofolo, & Fiscus, 2000).

In a comprehensive study of *DragonSystems* (Aiken, Wong, & Vanjani, 2001), the absolute accuracy was about 79% while an adjusted accuracy accounting for reader comprehension was 91%. That is, even though some words were transcribed incorrectly, an independent reviewer could correctly guess what was meant for many of the errors.

RESEARCH QUESTIONS

This study examines five issues related to the successful implementation of SR technology in a GSS with a particular focus on gender adoption. Specifically, the issues addressed in this research are: 1) background noise, 2) user training, 3) comment submission, 4) comment quality, and 5) transcription accuracy.

How does SR technology impact comment submission between genders in a GSS setting?

Many executives, in particular, are poor typists, and therefore, their productivity may suffer when they are restricted to using a keyboard (Aiken, et al., 2000). Given that many executives are male this question seeks to determine if SR technology can increase the ease of entering comments into the GSS application.

H1: Male subjects using a SR GSS system will report a higher ease-of-use than females using a GSS system.

How does SR technology impact GSS meeting satisfaction between genders?

Groups in GSS meetings utilizing keyboard input have reported higher satisfaction than group in verbal meetings. This question seeks to determine if SR technology can produce even higher satisfaction.

H2: Male subjects using an SR GSS will report higher satisfaction than their female counterparts.

Does SR Technology help females to express themselves better?

Most people speak faster than they can type. Past GSS studies have shown a higher number of total comments generated over verbal meetings. This question seeks to investigate the impact of SR input on total comment generation in the GSS application.

H3: Male subjects using a SR GSS system will generate more comments than female subjects.

How does SR technology impact conflict/non-task behavior between genders in a GSS Setting?

One of the drawbacks to any group decision-making process is the amount of irrelevant or flaming comments made that are not germane to the topic at hand. Prior research has shown that in an anonymous keyboard GSS session flaming increases over that of a face-to-face meeting. This question seeks to determine if SR technology can decrease the number of irrelevant comment submission and increase comment quality.

H4: Male participants using a SR GSS system will generate less irrelevant and flaming comments than female participants using a keyboard GSS system.

How does SR technology impact comment accuracy in a GSS setting?

One of the challenges with any transcription method is the number of errors or misspelled words that limit the comprehension of the comment. This question seeks to determine if SR technology can more be more accurate in the elimination of misspelled words.

H5: Male subjects using a SR GSS system will generate less misspelled words and have an overall higher accuracy than their female counterparts.

EXPERIMENTAL STUDY

A group of 70 Business school students participated in the experiment for extra credit. Five groups of seven each used electronic gallery writing to type comments while five groups of seven each used *DragonSystems Naturally Speaking* SR software to generate the text. The *DragonSystems* software is a speaker-dependent program that accepts continuous and discrete speech and increases its speech recognition with each use of the software. An a-priori power test indicated a level of .80

for the study (Cohen, 1988). Of the total 70 subjects, 44 were male, 26 were female, 16 were seniors, 41 were juniors, and 13 were sophomores. All participants were trained in a specialized computing facility that minimized background interference.

All subjects were provided with an explanation of the purpose of the study and each participant spent 20 minutes in the training phase of the software. The training time was determined based on the results of previous research (Rebman, 2001). The subjects also spent 10 minutes discussing promoting tourism, a topic used in many prior experiments. All comments were automatically recorded in a separate file with user number, group number, and time written. In addition all comments were anonymous save for the possibility of hearing others speak. Upon completion of the training exercise each participant was asked to vote and select the format of his or her final exam. Following the meeting the subjects completed a questionnaire that was adopted from previous voice-GSS studies (see Appendix A).

RESULTS

Table One presents the summary statistics between the male and female participants. Overall, there were very little significant differences in gender. Actually, there were no significant differences between males and females in the keyboard session, yet there were a few noticeable differences between males and females in the voice sessions. It appeared that both males and females had approximately the same perceived efficiency, and ease of use of the GSS application. The few items that came close to being statistically significant were keyboard experience (females reported a higher number of hours of keyboard experience), and females found the voice GSS system a little easier to submit comments. Two items that were statistically significant were satisfaction of the process and outcome, and importance of decision, which is illustrated in Table 2. Interesting enough, females in the voice groups reported that not enough time was spent on generating ideas. T-Tests were conducted on 46 variables in both voice and keyboard groups and only 6 variables were found to be statistically significant. These differences are only applicable to male and female subjects in the voice sessions and they are listed and discussed below; there was no significant difference between male and female keyboard participants.

Idea generation – Male subjects reported feeling that they had spent too much time on idea generation during the group discussion as opposed to the female participants.

Group understanding of decisions made – Females reported a higher level of understanding of the decisions made.

Consensus on Decisions made – Females reported a stronger impression that they felt the group decision were made by consensus as opposed to some other type of decision-making style.

Satisfaction with Outcome – Females reported a much higher satisfaction level (5.88) then their male counterparts (4.72).

Importance of Decision – Females apparently felt the decision they were engaged in was more important than the males.

Satisfaction with Process – In addition to satisfaction with the outcome, females also reported higher satisfaction with the decision-making process.

Table 1

Variable	Gender	Mean	Std. Deviation
Ease of Use	Male	3.37	1.24
	Female	3.79	1.44
Tkcompar	Male	5.24	1.35
	Female	5.29	1.12
Peffic	Male	3.57	1.46
	Female	3.88	1.03
Keyexp	Male	1.96	1.21
	Female	1.54	0.66
Afraid	Male	1.76	1.35
	Female	1.63	1.31
Submit	Male	3.41	1.82
	Female	4.33	2.24
Knowgroup	Male	3.43	2.09
	Female	3.17	2.1
Express	Male	4.72	2
	Female	5.08	1.82
Prefer	Male	2.06	1.79
	Female	2.92	1.59
Accuracy	Male	2.87	1.5
	Female	2.79	1.44

Table 1			
Variable	Gender	Mean	Std. Deviation
Reak ok	Male	4.41	1.48
	Female	4.75	1.67
Ease of comm	Male	2.80	1.36
	Female	3.08	1.61
Satisfied	Male	3.41	1.64
	Female	3.88	1.65
Work in future	Male	5.26	1.48
	Female	5.17	1.40
Importance of Dec	Male	4.59	1.51
	Female	5.29	1.16
Satisfied w/Dec	Male	5.61	1.39
	Female	6.21	1.18
Satified w/Process	Male	4.52	1.64
	Female	5.58	0.97

Table 2		
Variable	t stat	Sig. (2-tailed)
Ease of Use	-1.22	0.23
Tkcompar	-0.17	0.86
Peffic	-1.03	0.31
Keyexp	1.86	0.07
Afraid	0.41	0.69
Submit	-1.74	0.09
Knowgroup	0.51	0.61
Express	-0.77	0.44
Prefer	-0.74	0.46
Accuracy	0.21	0.83

Table 2		
Variable	t stat	Sig. (2-tailed)
Reak ok	-0.83	0.41
Ease of comm	-0.72	0.47
Satisfied	-1.11	0.27
Work in future	0.26	0.79
Importance of Dec	-2.17	0.03*
Satisfied w/Dec	-1.90	0.06
Satified w/Process	-3.39	<.001*
df=65, male=46, female=24		
*significant at 0.05		

LIMITATIONS

This study made attempts to analyze gender differences, however the study was conducted with unequal sample sizes. Although, the sample sizes were different, it did become evident that gender differences did exist in the speech-computer interface. More research is required to determine the full depth and implications of these differences. Likewise, research is necessary to determine other demographic differences stemming from ethnicity and age.

CONCLUSIONS

For several decades, system developers intensely explored opportunities to integrate speech recognition into the human-computer interface. Due to previous inefficiencies in the technology, the promise of SR in the human-computer interface has yet to be fully realized. However, evolutions in speech recognition software during the last 10 years again have prompted systems developers and researchers alike to examine the feasibility of speech recognition technology as a method of human-computer interface. Although the latest evolutions of SR technology provide greater technical opportunities for system integration, there are still many human user variables associated with the technology that remain to be examined. We offer the findings presented in this paper as an initial point of investigation from which researchers may begin to explore the human user variables associated with speech recognition. Gender, as used in this paper, is only one user variable that requires examination in the context of speech recognition-enabled systems. Additionally, researchers

should consider the larger scope and dynamic relationship among other human user variables both on an individual and a group level and examined these variables in similar contexts used in other studies. Speech recognition technology and SR research must move beyond the technical assessments of the systems and focus upon the users of the technology.

REFERENCES

Adam A (1997) 'What should we do with cyberfeminism?' *Women in Computing* Intellect Books, Exeter

Aiken, M.; Rebman, C.; and Paolillo, J.; "Lessons Learned with a Voice-based Group Support System: *Proceedings of the 32nd Annual Decisions Sciences Institute*, November 2001, San Francisco, California.

Aiken M.; Rebman, C.; and Vanjani, M.; "A Voice-based Group Support System," *Proceedings of the 32nd Annual Southwest Decision Sciences Institute Conference*, New Orleans, Louisiana, February 2001,

Aiken, M., Wong, Z., and Vanjani, M. (2001). An analysis of errors using a speech recognition system. *Proceedings of the 32nd Annual Southwest Decision Sciences Institute Conference*, February 2001, New Orleans, LA.

Aiken, M., Sloan, H., Paolillo, J., & Motiwalla, L. (1997). The use of two electronic idea generation techniques in strategy planning meetings. *Journal of Business Communication*, 34(4), October 1997, 370-382.

Aiken, M.; and Hassan, B. (1996). Total Quality Management: A group Decision Support System Approach, *Information Systems Management*, 13(1), January 1996, 73-76.

Aiken, M.; and Chrestman, M. (1995). Electronic Meeting Systems" Journal for Quality and Participation, 18(4), April 1995, 98-102.

Aiken, M., Vanjani, M., and Krosp, J. (1995). Group decision support systems. *Review of Business*, 16(3), Spring 1995, 38-42.

Aiken, M. Kim, D., and Singleton, T. (1994). Future developments of group decision support systems. *1994 Southeast Decision Sciences Institute Conference*, Williamsburg, VA, March 1994.

Aiken, M. (1992). Using a Group Decision Support System as an Instructional Aid: An Exploratory Study. *International Journal of Instructional Media*, 19(4), April, 328-329.

Aiken, M., Liu Sheng, O., and Vogel, D. (1991). Integrating Expert Systems with Group Decision Support Systems. *ACM Transactions on Information Systems*, 9(1), January, 75-95.

Alavi, M.; and Jachimsthaler, E.(1992). Revisiting DSS implementation research: A meta-analysis of literature and suggestions for researchers, *MIS Quarterly*, 16(1) March, 95-116.

Alexander, S. (1999). Speech recognition. *Computerworld* 33(45), November 8, 1999, 65.

Anson, A. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams*Academy of Management Journal*, 39(1) January 1996, 123-148.

Bethoney, H. (1999). Speech products: The talk of the town. *PC Week,* August 16, 1999, 5.

Broome, B.; and Chen, M. (1992). Guidelines for computer-assisted group problem solving *Small Group Research*, 23(2) February 1992, 216-236.

Darley, W. K., & Smith, R. E. (1995). Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. *Journal of Advertising,* 24 (1), 41-56.

Dennis, A.; George, J.; Jessup, L.; Nunamaker, J.; and Vogel, D. (1988). Information technology to support electronic meetings. *MIS Quarterly* 12(4) December 1988, 591-624.

DeSanctis, G.; and Gallupe, B. (1987). A foundation for the study of group decision support systems *Management Science*, 33(8) August 1987, 589-609.

Gupta, U. (1996). *Management Information Systems*. St. Paul, MN: West Publishing Company.

Holbrook, M. B. (1986). Aims, concepts, and methods for the representation of individual differences in esthetic responses to design features. Journal of Consumer Research, 13 (3), 337-347.

Huseman, R.; and Miles, E.(1998). Organizational communication in the information age: Implications of computer-based systems *Journal of Management*, 14(4), April, 181-204.

Huber, G. (1984a). A theory of the effects of advanced information technologies on organizational design, intelligence, and decision making Academy of Management Review (15(1) January 1984a, 47-71.

Huber, G. (1984b). Issues in the design of group decision support systems *MIS Quarterly*, 8(3) September 1984b .195-204.

Jang, P. and Hauptmann, A. (1999). Learning to recognize speech by watching television. *IEEE Intelligent Systems*, 14(5), May 1999, 51-58.

Johansen, R. *Groupware: Computer support for business teams*. New York, NY. Free Press.

Jones, D.; Hapishi, K.; and Frankish, C. (1991). Automated Speech Recognition in Practice *Behavior and Information Technology*, March/April 1991, 47-52.

Klevans, R.; and Rodman, R. (1997). *Voice Recognition*. Boston, Artech House.

Knights D & Murray F (1997) Markets, managers, and messages: Managing information systems in financial services In: B P Bloomfield, R Coombs, D Knights & D Littler (eds.) *Information Technology in Organizations: Strategies, Networks, and Integration* Oxford: Oxford University Press

Kraemer, K.; and King, J. (1988). Computer based systems for cooperative work and group decision making *ACM Computing Surveys*, 20(2) February1988, 115-146.

Lindquist, C. (1999). Speak easy *PC World*, October 1999, .185-195.

Markowitz, J. (1996). *Using Speech Recognition*. Upper Saddle River NJ, Prentice Hall

McLeod, P. (1992). An assessment of the experimental literature on electronic support of group work: Results of a meta-analysis *Human-Computer Interaction*, July1992, 257-280.

Miastkowski, S. (1999). Latest speech software gets you up and running faster. *PC World*, November 1999, . 63-66.

Meyers-Levy, J., & Maheswaran, D. (1991). Exploring differences in males' and females' processing strategies. *Journal of Consumer Research*, 18 (1), 63-70.

Meyers-Levy, J., & Sternthal, B. (1991). Gender differences in the use of message cues and judgments. *Journal of Marketing Research,* 28 (1), 84-96.

Niccolai, J. (1997). First speech-recognition email announced *InfoWorld*, November 24, 1997, . 57.

Nicovich, S. G., Boller, G. W., and Cornwell, T. B. (2005). Experienced presence within computer-mediated communications: Initial explorations on the effects of gender with respect to empathy and immersion. *Journal of Computer-Mediated Communication*, 10(2), article 6.

Nunamaker, J.F., Dennis, A.R., Valacich, J.S., Vogel, D.R., and George, J.F. "Electronic Meeting Systems to Support Group Work," *Communications of the ACM,* (34:7), July 1991, pp. 40-61.

Pallett, D.S., Garofolo, J.S., Fiscus, J.G. (2000). "Measurements in Support of Research Accomplishments". *Communications of the ACM.* (43:2_, February 2000, pp. 75-79.

Pollack, C.; and Kanachowski, A. "Application of theories of decision making to group support systems (GDSS)" *International Journal of Human-Computer Interaction*, (5:1) January 1993, pp. 71-94.

Pollard, C. Electronic meeting systems: Specifications, potential, and acquisition strategies. *Journal of Systems Management*, (33:3) May/June 1996, 33, pp. 22-28.

Prakash, V., & Flores, R. C. (1984). A study of psychological gender differences: Applications for advertising format. Advances in Consumer Research, 12 (1), 231-237.

Rabiner, L.; and Biing-Hwang, J. *Fundamentals of Speech Recognition.* Englewood Cliffs NJ, Prentice Hall, 1993.

Rebman, C. "An Exploratory Study of the Impact of Training Times on User Acceptance of Speech Recognition Systems," *Proceedings of the 32nd Annual Southwest Decision Support Institute Conference*, New Orleans, Louisiana, February 2001, pp.54-57.

Rebstock, S.; Williams, S.; and Wilson, R. "Group support systems, power and influence in an organization: A field study" *Decision Sciences*, (28:4), December 1997, pp. 911-937.

Rodman, R. *Computer Speech Technology*. Boston, Artech House, 1999.

Teng, J., and Ramamurthy, K. "Feedback as a source of control in decision support systems: Clarifying the concept and establishing a functional taxonomy" *INFORMS*, (41:3), March 1993, pp.166-185.

Turban, E. *Decision Support and Expert Systems: Management Support Systems,* 3rd Edition. New York, NY: MacMillan Publishing, 1993.

Turoff, M., Hiltz, R.S., Ahmed, N.F., and Bahgat, A.R. Distribution Group Support Systems. *MIS Quarterly*, (17:2), June 1993, pp. 399-405.

Wactlar, H., Hauptmann, A., Christel, M., Houghton, R., and Olligschlaeger, A. Complementary video and audio analysis for broadcast news archives. *Communications of the ACM*, (43:2), February 2000, pp. 42-47.

Wilson, M., and Howcraft, D. "Gender and User Resistance in Nursing Information Systems Failure" Working Paper Manchester School of Management *http://www.sm.umist.ac.uk/wp/Papers/wp2013.htm*

APPENDIX A EXPERIMENTAL QUESTIONNAIRE

Overall ease of use

1. In general using the GSS program was

1 Extremely difficult
2 Difficult
3 Fairly Difficult
4 Neither easy or difficult
5 Fairly easy
6 Very easy
7 Extremely easy

2. Overall, compared with traditional meetings (verbal), using the GSS program appears to be

1 Very much easier
2 Easier
3 A little easier
4 About the same
5 A little harder
6 Harder
7 Very much harder

Perceived Efficiency

1. My impression of the GSS program is that it is

1 Extremely inefficient
2 Very inefficient
3 Inefficient
4 Neither efficient nor inefficient
5 Efficient
6 Very efficient
7 Extremely efficient

Previous Keyboard Experience

1. I estimate that before today's session my keyboard experience was

1 More than 1,000 hours
2 500 to 1,000 hours
3 250 to 499 hours
4 100-249 hours
5 50 to 99 hours
6 1 to 49 hours
7 0 hours

GSS Questions

Please Circle your answers below.

1. I was afraid that others in my group would criticize my comments.

1 2 3 4 5 6 7
Disagree Neutral Agree

2. It was easy to submit and read comments.

1 2 3 4 5 6 7
Disagree Neutral Agree

3. I know the members of my group well.

1 2 3 4 5 6 7
Disagree Neutral Agree

4. I was able to express my opinion easily.

1 2 3 4 5 6 7
Disagree Neutral Agree

5. I prefer this type of meeting to a traditional, verbal meeting.

1 2 3 4 5 6 7
Disagree Neutral Agree

6. The program was able to accurately print out(on the screen) my comments.

1 2 3 4 5 6 7
Disagree Neutral Agree

7. I was able to read and understand everyone's comments.

1 2 3 4 5 6 7
Disagree Neutral Agree

IMPACT OF DATA INTEGRATION ON CRM IN THE ELECTRONIC COMMERCE OF SMES

Assion Lawson-Body, University of North Dakota
Abdou Illia, Eastern Illinois University
Jeanne Jimenez, University of Montpellier - Campus de Poitiers

ABSTRACT

Customer Relationship Management (CRM) has been discussed in literature as a data integration system that can provide a strategic advantage via customer loyalty to small and medium-sized enterprises (SMEs). Although CRM promises increased revenues and profits, SMEs face potential failure because of complex relationship issues.

This study is a theoretical and empirical examination of the impact of data web integration on the relation between CRM (which includes customer prospecting, empowerment and understanding customer expectations) and customer loyalty. Data collected from 224 SMEs in the USA and processed with Partial Least Square (PLS) show that the use of data web integration reduces the impact of customer prospecting on customer loyalty. In contrast, the use of data web integration increases the impact of empowerment and understanding customer expectations on customer loyalty. The implications of the results for the study are discussed.

INTRODUCTION

Customer Relationship Management (CRM) was easier in earlier times when it was based on relationship marketing (RM) (Goodhue et al., 2002; Shin, 2003). RM has established itself as an information system (IS) reference discipline. The four Ps (product, price, place and promotion) of the marketing mix have become the universal marketing model or even theory (Gronroos, 1994). The four Ps of the marketing mix are a clinical approach, which makes the seller the active party and the buyer passive (Gronroos, 1994). No personalized relationship between the seller and the buyer is supposed to exist (Gronroos, 1994). The concept of RM has emerged because (1) it personalizes the relationship between the seller and the buyer, and (2) it relies on a two-way dialogue between a company and a customer to develop a deep relationship. Unfortunately, this two-way dialogue proved to be labor intensive and thus had to be limited to a small subset of customers.

Today, the use of information technology (IT) has modified RM into CRM, thereby making CRM feasible across a wide range of customers (Goodhue et al., 2002). CRM is driven by the availability of data web integration and advances in business-to-business (B2B) electronic commerce. In particular, IT based on data web integration has made CRM efficient and effective.

The meaning of CRM refers to relationships between small and medium-sized enterprises (SMEs) and their customers. There has been a large increase in SMEs in the world of B2B electronic commerce (U.S. SBA, 2000). The transaction value of B2B electronic commerce over the Internet is expected to be $2.0 trillion by 2003 and an additional $780 billion in purchases will be made over private networks using EDI (Electronic Data Interchange) (U.S. SBA, 2000). Although EDI concepts can be traced all the way back to the 1970s with American Airlines (SABRE) and American Hospital supply, the move away from proprietary systems towards using the web and data web integration technologies has reduced the cost and increased the number of SMEs using B2B systems (Poon, 2000). The growth of Internet and data web integration technologies has opened up numerous opportunities for SMEs to penetrate new markets and forge a new CRM (Li and Williams, 1999). Since a number of authors have suggested that loyalty is a relational phenomenon, our intention was to link CRM between SMEs and their loyal customers.

Although CRM is on the rise and holds tremendous promise for building mutually beneficial relationships with customers, SMEs often struggle with their CRM efforts. Many SMEs with CRM projects are either experiencing difficulties or close to failure (Goodhue et al., 2002). We believe SMEs can minimize their risks of failure by first having data web integration as a CRM enabler and then linking it to customer loyalty. In addition, there is a little existing research that has empirically tested the impact of data web integration on CRM which leads to customer loyalty.

The primary objective of this study is to examine the moderating effects, on a theoretical and an empirical basis, of data web integration on the relation between CRM and customer loyalty.

This article is organized as follows: In Section 2, we present the literature review. In Section 3, we present the research model and hypotheses. In Section 4, we introduce the methodology. In Section 5, we present analysis and results. In Section 6, we outline the discussion. In Section 7, we draw conclusions, identify the limitations and specify future directions.

LITERATURE REVIEW: DATA WEB INTEGRATION, CRM, AND CUSTOMER LOYALTY

Customer Loyalty as a relationship phenomenon

Many authors, such as Yoon & Kim (2000), Rowley & Dawes (2000) and Dick & Basu (1994) distinguish between two possible properties that underlie loyalty. The first property is behavior, which they define as the overt act of selective repeat purchasing based on evaluative psychological decision processes. The second property is attitude, which refers to underlying predispositions to behave in a selective fashion. Dick & Basu's investigation (1994) into customer loyalty explores the antecedents of attitude. They found that loyalty is determined by the strength of the relationship between relative attitude and repeat patronage, a customer behavior.

It has been found that loyal customers are less price sensitive and lower costs are incurred by the provider as the expense of pursuing new customers is reduced (Birgelen et al., 1997; cited in

Rowley & Dawes, 2000). Much of the literature on customer loyalty has looked at brand loyalty, but there is a range of conceptualizations of loyalty from repeat purchase to a lifetime relationship (Rowley & Dawes, 2000). Research also has shown that there are clear linkages between customer loyalty and organizational profitability (Rowley and Dawes, 2000). Customers can demonstrate their loyalty in several ways. Clearly, they may choose to stay with a provider, whether this continuance is defined as a relationship or not, or they may increase the number of purchases or the frequency of their purchases or even both (Rowley & Dawes, 2000). The enduring desire to maintain a valued relationship should, in turn, impact loyalty (Chow & Holden, 1997). Not only do loyal customers generate more revenue for more years, but it also costs less to maintain existing customers than to acquire new customers (Berry, 1995).

The enthusiasm for customer loyalty or the idea of a company developing a relationship with its customer is not new (Sharp, 1997). Fournier & Yao (1997) used the perspective of interpersonal relationship theory to critically examine the notion of loyalty. They found that not all interpersonal relationships possess characteristics of loyalty. Following the customer loyalty conceptualization of Chow & Holden (1997), our intention was to examine the theoretical and empirical impact of CRM on customer loyalty.

CRM from SME perspectives

The success of a CRM depends on its participants (Poon, 2000). Our focus was on SMEs because resource scarcity compels them to work closely with others (customers, suppliers and partnering firms) to achieve business goals. SMEs tend to form alliances (or networks) in order to pool limited resources for large projects. According to the manager of one SME, "the reason we have been able to grow so fast for so long is that we aren't really one big company. We are really a confederation of small businesses as a network of entrepreneurial partnerships" (Reichheld, 2001). Network members often have complementary resources so the total benefit is larger than the sum of individual parts (Poon, 2000). When branches grow large, their customer and employee satisfaction tend to suffer (Reichheld, 2001). So when an enterprise branch grows to a specified size, it is split in two and a new manager is put in charge of the new branch (Reichheld, 2001). Members of the SME network may include partner firms and competitors in the same industry sector and suppliers and customers in the value system (Poon, 2000).

There are key phases SMEs may consider before putting a CRM program to work: identify customers, differentiate among them by understanding their needs, and empower employees to better interact with them (Peppers et al., 1999).

CRM based on traditional database systems

CRM takes a broad view of a company's customers by including both current and prospective customers as well as trading partners in the supply chain. To optimize interactions with these customers,

it is necessary to collect, store, and manage data on every interaction with them, no matter where the data come from (Goodhue et al., 2002).

According to Goodhue et al, (2002), the CRM technical architecture can include many applications, performing both analytical and operational functions. On the analytical side, a data warehouse typically maintains historical data that supports generic applications, such as reporting, queries, online analytical processing (OLAP), and data mining, as well as specific applications, such as campaign management, churn analysis, propensity scoring, and customer profitability analysis. The data warehouse entails a high degree of complexity in the system's architecture because its data sources comprise transactional databases, operational data stores, purchased data, and external data (Shin, 2003). On the operational side, data must be captured, integrated, and stored from all in-bound touch points. That data may be augmented with external demographic data. Current data can be maintained in an operational data store that supports operational applications, such as customer support.

Goodhue et al., (2002) also introduced the concept of data infrastructure. The authors distinguish between data architecture and data infrastructure because the terms can be confusing. The intent of data architecture is to ensure that the data used in one CRM application can be shared with all the other CRM applications. Data infrastructure, on the other hand, refers to the actual state of the existing databases, and the existing state of the data. The critical characteristic of data infrastructure is the degree to which existing data and databases can be used to support new applications, without incurring data problems.

CRM, Data web integration and Customer Loyalty

The newly available ability to access comprehensive customer information presumably makes possible new ways for SMEs to interact with their customers. Effective access to comprehensive customer information can be built with data web integration. Peterson (1995) and others present the Internet as a business channel that offers a company the possibilities of storing vast amounts of data at different virtual locations, permitting customers of the value chain to access databases via password-protected extranets (Joseph et al., 2001).

Before integrating their database into their web sites, some SMEs' customer service representatives handled thousands of interactions per month, including inquiries about price, product information, and availability (Goodhue et al.., 2002). All of the requested information resided in SMEs' back-end databases. Management reasoned that by giving the customers real time, online access to the database, it could minimize this labor-intensive interaction with customer service representatives. This resulted in reduced cost, increased profit margins, and higher customer loyalty.

Web sites integrated with database systems have significantly improved the efficiency and speed of record keeping and the ability to develop CRM. They are used to store and maintain a record of all business transactions (Orman, 2001). Via web sites, the data warehouse can enhance business performance in various ways, including better-targeted products and improved CRM (Shin, 2003). For instance, automated and integrated information delivered from the data warehouse via a web site

interface may substantially free up managers' time and efforts, thereby increasing their availability for other tasks. Venkatraman (1991) proposes the concept of electronic integration to promote relationships between companies and their customers (Grover et al., 2002). This integration could involve automation of exchange procedures and documents and the sharing of applications and databases (Grover et al., 2002). Explicit relations between the initiator of a web information system and its customers are also supported through the use of databases, application interfaces and cookies (Ghosh, 1998). If customers are satisfied by the information available on the web site that is integrated with a database, they will be more involved in the relationship which should lead to loyalty (Ghosh, 1998). SMEs know that customer loyalty is important because in addition to purchasing more, loyal customers also regularly refer new customers to a supplier, providing another rich source of benefits (Reichheld & Schefter, 2000). Via Web stores, SMEs can sell an increased range of products and services to loyal customers, broadening as well as deepening CRM over time (Reichheld & Schefter, 2000).

RESEARCH MODEL AND HYPOTHESES

The key variables of the research model, identified through the literature review are as follows: the dependent variable will be customer loyalty, the independent variables will be drawn from the CRM and the moderating variable will be the data web integration. Each is discussed below.

Customer loyalty

In this paper, loyalty is defined as building and sustaining a relationship with customers that leads to the customers' repeated purchases of products or services over a given period of time (Chow & Holden, 1997). In the literature, customer loyalty has been conceptualized as an interaction of attitude, behavior and relationships. This research is focused on the relationship aspects of customer loyalty.

Customer Relationship Management (CRM)

In this study, CRM is defined as any initiative designed to help an SME optimize interactions with customers or prospects via one or more touch points such as a call center, salesperson, or distributor for the purpose of retaining customers (Goodhue et al., 2002).

The primary goal of CRM is to let SMEs achieve a competitive advantage via mutual loyalty. That goal is met by (1) enabling an SME to better prospect customers, (2) ensuring that employees satisfy customer needs, and (3) understanding customer expectations.

Customer Prospecting

Customer prospecting refers to the ways businesses track, locate, and attract new customers (Shultz, 1995). Thomas (2001) has found that customer acquisition (customer prospecting) affects customer retention (customer loyalty). Payne (1994) has conceptualized CRM as loyalty ladder leading from prospect to customer, client, supporter, and partner. According to Payne (1994), customer prospecting is the beginning of the CRM process. Thomas (2001) studied a methodology for linking customer acquisition (prospecting) to customer retention (loyalty). He found that customer acquisition and retention are not independent processes. Using the model elaborated in his study, Thomas (2001) shows the financial impact of not accounting for the effect of acquisition on customer retention. Thus we propose the hypothesis below:

H1: Customer prospecting will have a positive effect on customer loyalty

Empowerment

Empowerment is defined as the process SMEs adopt to encourage and reward employees who exercise initiative and make valuable creative contributions (Evans & Laskin, 1994), or do everything that is possible to help customers solve their problems. Reichheld (2001), reports that he has yet to encounter an SME that has achieved extremely high customer loyalty without fostering similarly high loyalty among employees. Leaders who are dedicated to treating people right drive themselves to deliver superior value to customers which allows them to attract and retain the best employees (Reichheld, 2001). That's partly because higher profits result from customer loyalty, but, more important, it's because providing excellent service and value generates pride and a sense of purpose among employees (Reichheld, 2001). Collectively, we hypothesize that:

H2: Empowerment will have a positive effect on customer loyalty.

Understanding customer expectations

This concept is defined as the ability to identify customers' needs and to supply to those customers products and services which meet their expectations (Evans & Laskin, 1994). Some studies show that SMEs can benefit from electronic commerce but specific conditions need to be satisfied in order to achieve any benefit (Poon, 2000). For example, the management of expectations is important because a lack of benefit may be due to unrealistic expectations rather than unrealized benefit (Poon, 2000). Yoon & Kim (2000) describe understanding customer expectations as the strategy adopted by firms to generate more knowledge of customer expectations and needs and to provide customers with the best services in order to win their loyalty. Hence, we hypothesize:

H3: Understanding customer expectations will have a positive effect on customer loyalty.

The moderator role of Data web integration

A third variable plays the role of moderator when it partitions a focal independent variable into subgroups that establish its domains of maximal effectiveness in regard to a given dependent variable (Baron & Kenny, 1986). In general terms, a moderator is a qualitative or quantitative variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable (Baron & Kenny, 1986; Chin et al., 1996). In this article the data web integration is interpreted as a moderator of the relation between CRM and customer loyalty. The data web integration has functions of a third variable therefore it is supposed to affect the strength of the relation between CRM's variables (independent variables) and customer loyalty (dependent variable). The data web integration may change the direction of the relation between CRM and customer loyalty from positive to negative or vice versa. In this study, moderation implies that the causal relation between CRM and customer loyalty changes as a function of the moderator variable of the data web integration.

At the level of statistical analysis, a basic moderator effect can be represented as an interaction between a focal independent variable and a factor that specifies the appropriate conditions for its operation (Baron & Kenny, 1986). Therefore we will represent the moderator effect as an interaction or product of the data web integration and CRM variables. In addition, we will test the interaction or product of the data web integration and CRM on customer loyalty. The moderator hypothesis is supported if the interaction (path) is significant.

Data web integration

The data web integration refers to the services that connect different buyers and sellers through a shared data/information resource and provide tools for searching data in order to match product offerings and buyer desires to define mutual loyalty criteria (Kollman, 2001). CRM is a powerful tool for cross-organizational and analytical databases (Swift, 2002). This variable indicates that the web site should be integrated into the SME's database system. In many SMEs, the web browser is the universal user interface to their corporation databases where information is now accessed via corporate web-servers (Poon, 2000). Perhaps the greatest concern about the value of the database is if a one-way relational system is developed to favor the seller and then becomes a disincentive to the buyer (Sandeep, 2001). Therefore, it is stressed that the CRM and integrated database should work for both the buyer and the seller for mutual loyalty.

Many SMEs have developed databases that contain detailed interaction data on prospects as well as customers (Thomas, 2001). Many of the databases that exist tend to have data only on existing customers. The practice of collecting and analyzing data only on existing customers is also prevalent in established businesses that are just starting to adopt a CRM. Maintaining a customer database, which

means having one system integrate seamlessly with another, tracking each customer's contacts with the company, all of those activities require web site integrated with the company database system. A customer can inquire about the pricing, the tariff and availability of a particular product, or request information from internal databases from the intranet of its supplier. Therefore, we offer the following hypothesis:

H4: The data web integration will have a positive effect on the relation between customer prospecting and customer loyalty.

Ozer (2000) points out that some SMEs already integrate their databases to their web site for future projects among their employees, such as client proposals, prospects, repositories, presentations, deliverables, internal training curriculum, contact management, regional news, methodologies, and external information sources. Some SMEs maintain a database for their support staff that provides stored solutions to common problems. The problems and solutions can then be tracked using the Web. SMEs that use databases integrated to their web sites to support their employees should increase customer loyalty. So, we formulated hypothesis H5 as:

H5: The data web integration will have a positive effect on the relation between empowerment and customer loyalty.

Today, many companies are trying to go back to "the good old days" of knowing their customers well by capturing the available wealth of internal and external data, analyzing that data to better understand customer needs, preferences and loyalty, and then leveraging that knowledge in every customer contact (Goodhue et al., 2002). Some SMEs are integrating their customers' databases into their web sites to improve their understanding of customer needs (Chatterjee et al., 2002), so they can tie their customers' inventory and procurement systems to their systems. This will create better understanding of their customers' needs, requirements and expectations. The cost of switching will be high which will lead to loyalty (Reinartz & Kumar, 2002). We formulated hypothesis H6 as:

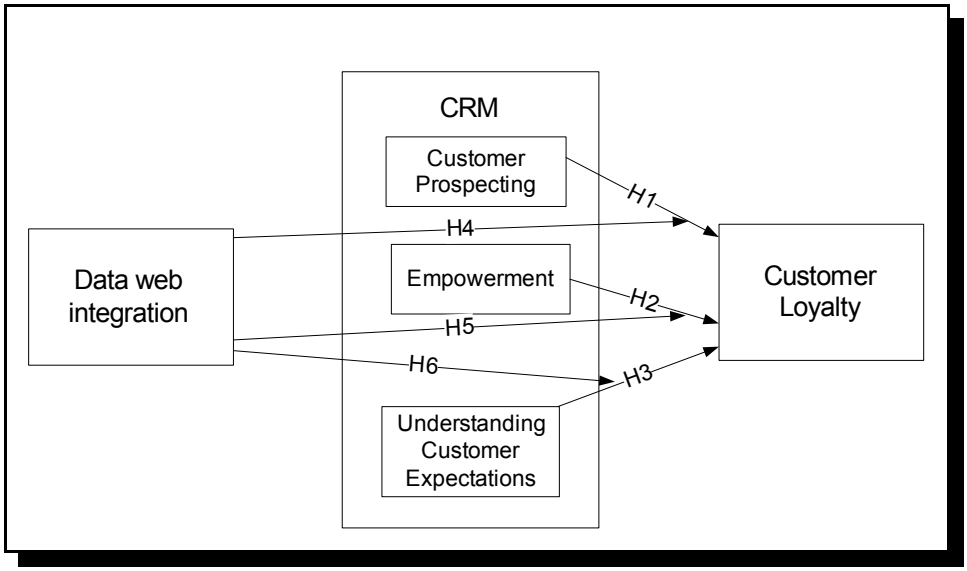
H6: The data web integration will have a positive effect on the relation between understanding customer expectations and customer loyalty.

The Research Model

The choices that have been made regarding the definition of each variable, and the arguments used to support them, have led to the redesigning and redefining of the research model below. In the research model, the CRM variables have a direct impact on the customer loyalty variable. The data web integration (moderating variable) can support the formation and maintenance of CRM because it facilitates the way SMEs prospect, understand customer expectations, and empower to create loyal

customers. In other words, SMEs can use the data web integration to build tight relationships with their trading customers who become loyal to them, rather than to identify customers on a transaction-by-transaction basis from a large pool of non-loyal customers. The quality of web site infrastructure can directly influence the type of relationships developed among businesses desiring reciprocal loyalty. The research model is shown in Figure 1.

Figure 1. Research Model



METHODOLOGY

The study sample

The sample consists of 1000 SMEs in the United States, each having a website and an e-mail address. Company size is measured by number of employees (Chow & Holden, 1997). In this study, an SME is one with fewer than 500 employees. The senior sales representative, company executive, or president of each of the above companies was sent a cover letter through the Internet (by e-mail) along with the URL of the web site containing the research instrument (questionnaire). The cover letter was on university letterhead. As an incentive, respondents were told that a summary of the results would be sent at their request. A total of 224 companies responded, producing a 22.4% response rate.

Measures

All of the measures were selected from the survey instrument used by Lawson-Body & Limayem (2001) in their study. After slight modification, the questions measuring all of the variables, except loyalty, used a scale of 1 (never) to 5 (always). The scale used to measure loyalty was 1 (decreasing sharply) to 5 (increasing sharply). Since the instruments have been slightly revised from the originals, reliability coefficients have been calculated.

Table 1 presents the reliability coefficients Rho. Its first column presents the independent and dependent variables of the research model. The second column of Table 1 presents the indicator of the reliability of a measure which is the Rho coefficient. Aubert & Rivard (1994), report that the guidelines established by Nunnally (1978) for the interpretation of Cronbach's alpha also apply to the Rho coefficient. These guidelines estimate that acceptable reliability coefficients must be higher than 0.6. It can be seen that all Rho coefficients are ranged between 0.701 and 0.857. This is considered very satisfactory.

Table 1: Reliability Coefficients RHO

Variables	Coefficients Rho
Empowerment	0.802
Understanding customer expectations	0.857
Customer Prospecting	0.701
Loyalty	0.761

To measure the data web integration variable, the evaluation grid, mounted according to the guidelines offered by Kassarjian (1977) and found in the study of Lawson-Body & Limayem (2001) was used. Data web integration was evaluated by two judges: the researcher and a graduate student. The inter-judge reliability coefficient is 83%. Berelson (1952; cited in Kassarjian, 1977) claimed a range located between 66% and 95% with a concentration at 90% for acceptable inter-judge reliability coefficients. The ratio of 83% appeared to be satisfactory.

ANALYSIS AND RESULTS

Characteristics of Participating SMEs

The respondents were spread across 13 different industry sectors. However, as shown in Table 2, seventy-four percent (74%) of the respondents were primarily involved in the Manufacturing,

Information Technologies Hardware/Software, Consulting Firms, Aerospace and Defense industries. In terms of annual sales volume, 85.56% of the sample had annual sales of less than $10 million, while 12.71% sold less than $100 million. Only 1.69% sold more than $100 million. About half of the SMEs had fewer than 50 employees, while about 85% had fewer than 300 employees. About 14% of the SMEs possess between 300 and 500 employees.

Most of the participating SMEs (about 92%) have been in business for more than 10 years. A total of 44.72% of the respondents claimed to have the duration of firm web site existence between 1 and 3 years, while 42.28% of SMEs claimed to have the number of years online between 4 and 6. A total of 88% of the SMEs have the average duration of sale cycle between 1 and 8 months. This information indicated that the responses were experience-based.

Table 2: Characteristics of the companies sampled

Industry sector	Percentage
Aerospace & Defense	10.48%
Automobile and Transportation	5.64%
Chemicals	0.80%
IT (non-internet) hardware and software	19.35%
Internet Access and Service Providers	1.61%
Manufacturing	28.22%
Financial Institutions (Banks, Insurances, etc.)	4.83%
Life Sciences	1.61%
Health Industry	3.22%
Ocean Industries	0.80%
Consulting Firms	16.93%
Government Agency	2.41%
Airline Industry	4.03%
The number of years firms have been in business	Percentage
Less than one year	0.81%
1-3 years	0%
4-6 years	3.25%
7-9 years	4.07%
10 years or more	91.87%
The size of firm (persons)	Percentage
1-50 employees	45.53%
51-100 employees	19.51%

Table 2: Characteristics of the companies sampled	
Industry sector	Percentage
101-200 employees	6.50%
201-300 employees	14.63%
301-500 employees	13.82%
The average duration of sale cycle of firms	Percentage
Less than one month	36.89%
Between 1 and 4 months	31.97%
Between 5 and 8 months	18.85%
Between 9 and 12 months	7.38%
13 months or more	4.92%
The revenue (turnover) of firms	Percentage
Less than $10 million	85.56%
$11-$100 million	12.71%
$101 million or more	1.69%
The duration of firm web site's existence	Percentage
Less than one year	6.50%
1-3 years	44.72%
4-6 years	42.28%
7-9 years	1.63%
9 years or more	4.88%

Procedures for Testing Hypotheses

In this research, Partial Least Squares (PLS), a second generation multivariate method, was used to process and analyze the data. The PLS method simultaneously evaluates both the measurement model and the theoretical model. It adjusts the relationships among the variables accordingly (Aubert et al., 1994). PLS was selected in this research because it presupposes no distributional form on the data. PLS simultaneously models the structural paths (i.e., theoretical relationships among latent variables) and measurement paths (i.e., relationships between a latent variable and its indicators). In this sense, PLS is preferable to techniques such as regression, which assume error-free measurement (Chin et al., 1996).

Hypothesis Testing (H1, H2, H3): The test of Hypotheses H1, H2, and H3 on the sample of the 224 respondents was carried out with a statistical tool named PLS-GRAPH. Table 3 shows the Student's T (t value) of impacts of customer prospecting (2.31), empowerment (1.85), and understanding customer expectations (2.78) on customer loyalty are higher than 1.65 ($P<=0.05$). This first hypothesis test shows that these variables of CRM have a positive and direct impact on customer loyalty. In other

words, customer loyalty is increased (R^2 = .67) by customer prospecting, by empowerment and by understanding customer expectations, supporting, H1, H2 and H3 respectively.

Table 3: Path Coefficients and Student's T (T values)		
	Loyalty (R^2 = .67)	
Variables	Path Coefficient	T-Statistic
Customer Prospecting	2.108	2.3195*
Empowerment	2.157	1.8512*
Understanding customer expectations	2.612	2.7811*
*T-Student significant at 1.64 (P<= 0.05)		

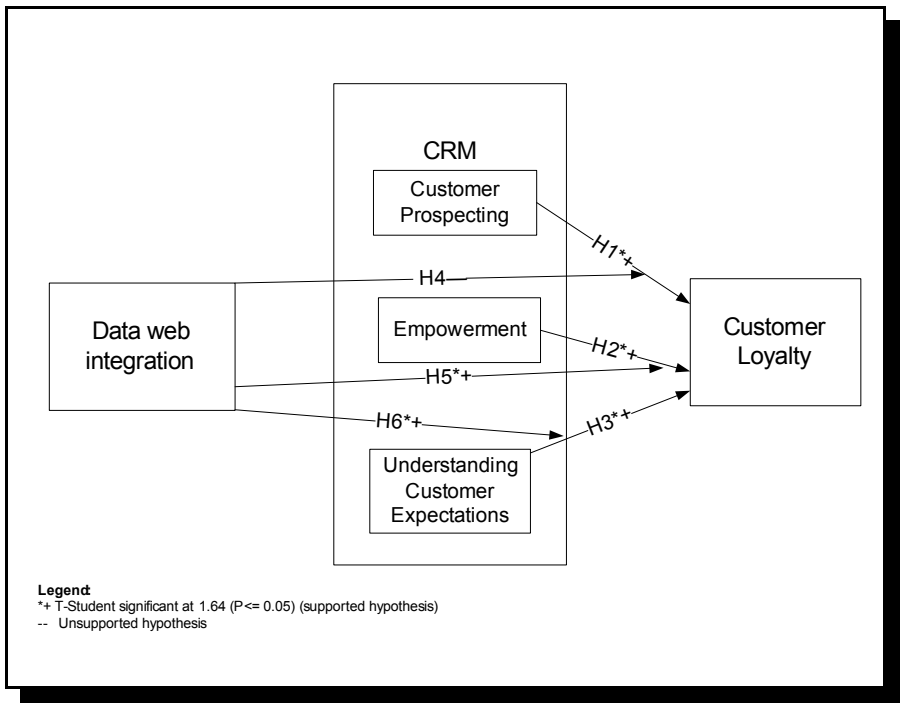
Hypothesis Testing (H4, H5 and H6): In order to determine the impact of the moderating variable (data web integration), the interaction tests of hypotheses H4, H5, and H6 were carried out with PLS-GRAPH. Table 4 presents the results of this test of the interaction effect and the size of the interaction effect, including the statistical values of the Student (T-statistic).

Table 4: Path Coefficients and Student's T (T values) for the Interaction Effects		
	Loyalty (R^2 = .76)	
Variables	Path coefficient	T-Statistic
Customer Prospecting	0.2103	2.4107*
Empowerment	0.2193	2.1270*
Understanding customer expectations	0.1972	3.4021*
Data web integration X Customer Prospecting	0.1872	1.0485
Data web integration X Empowerment	0.259	1.9140*
Data web integration X Understanding customer expectations	0.2477	2.1099*
*T-Student significant at 1.64 (P<= 0.05)		

As shown in Table 4, the path coefficient with interaction is lower for customer prospecting; this indicates a lower size of the interaction effect. For example, the beta for the impact of customer prospecting on customer loyalty is 0.2103. The interaction beta for the impact of data web integration on the relation between customer prospecting and customer loyalty is 0.1872. In other words, the use of data web integration reduces the impact of customer prospecting on customer loyalty. However, the

interactions between data web integration and empowerment, and data web integration and understanding customer expectations cause gains in customer loyalty. The betas for the impact of empowerment on loyalty and for the impact of understanding customer expectations on loyalty are 0.2193 and 0.1972 respectively. The interaction beta for the impact of data web integration on the relation between empowerment and loyalty is 0.2590. The interaction beta for the impact of data web integration on the relation between understanding customer expectations and loyalty is 0.2477.

Figure 2: Results of PLS analysis



DISCUSSION

Results from the survey of 224 SMEs in the USA show the use of data web integration reduces the impact of customer prospecting on customer loyalty. In contrast, the use of data web integration increases the impact of empowerment and understanding customer expectations on customer loyalty.

Impact of CRM on customer loyalty

The results of the hypothesis tests show that empowerment has a positive impact on customer loyalty. This finding corroborates the speculations of Berry (1995). Berry (1995) identifies

empowerment as one of the major emerging perspectives in building relationship, which leads to customer loyalty. Empowerment means workers can strive to meet customer requirements and needs and resolve problems; that can lead to spontaneous, creative rule-breaking to turn a potential frustrated customer into a loyal one. High employee turnover discourages management from investing much in hiring, training, and other relationship-building activities; this, in turn, is ineffective in creating customer loyalty. High employee turnover negatively affects customer retention or loyalty. Therefore, this result confirms some of the arguments found in the literature that employee retention is an antecedent of customer loyalty and low employee turnover positively impacts customer loyalty.

The positive impact of customer prospecting on customer loyalty was proved. But this finding was particularly surprising because it is four times less expensive to do business with existing customers than to attract new ones, if the SME pursues customer loyalty as an objective (Rosenberg & Czepiel, 1983; cited by Dekimpe, Steenkamp, Mellens, & Abeele, 1997). As customer prospecting is the cornerstone for attracting new customers, it is not normally considered as a form of CRM input which is conducive to customer loyalty. But this finding supported the results of Thomas' (2001) study. He found that customer acquisition and retention are not independent processes.

The hypothesis that understanding customer expectations will have a positive impact on customer loyalty was accepted. This result does not match one that was reached by Evans & Laskin (1994) to the effect that the CRM input, in terms of understanding customer expectations, did not have a positive impact on customer loyalty. These authors added, however, that this did not mean that understanding customer expectations was not an important part of the CRM input. The result in this study revealed that, understanding customer expectations remains a useful CRM component. This result also confirmed the importance attached to this concept by Szeinbach, Barnes, & Garner (1997) who described understanding customer expectations as the strategy developed by SMEs to collect more knowledge about customer needs and desires in order to be able to provide the best services in the hope of acquiring loyal customers.

Impact of data web integration on the relation between CRM and customer loyalty

The hypothesis that data web integration has a positive effect on the relation between empowerment and customer loyalty is significant. SMEs' contact employees are able, via data web integration, to access information about their company and its products and services. A web site integrated with an SME's database system, then, is a tool for employee empowerment. This employee empowerment has beneficial effects which affect external customers and ultimately lead directly to their loyalty. Once the company data web integration is completed and used by employees, the customers will most likely be satisfied with their services and loyalty will develop. The SME's data web integration protected by a secured firewall can be used to enhance the communication and collaboration among authorized employees. When the secured firewall protects the SME's data web integration, it becomes an Intranet. The firewalls protect the Intranets from unauthorized outside access. Only authorized employees are able to use them because they are limited to data pertinent to the company and contain

exclusive and often proprietary, sensitive and strategic information. An SME that uses the Intranet (the secured private Internet) to support their employees will increase their customer loyalty.

Although SMEs may articulate various reasons for their emphasis on customer loyalty, data limitations are a practical factor that motivates firms to disproportionately emphasize customer loyalty as opposed to customer prospecting. Because of data web integration limitations, electronic customer prospecting is frequently based only on an analysis of acquired customers. This study concludes that data web integration as well as the traditional databases that exist tends to have data only on existing customers. The findings of this research support that the practice of collecting and analyzing data on the web only on existing customers is prevalent in established businesses that are just starting to adopt CRM. Finally, one of the possible explanations for this phenomenon is that customer prospecting is the concept which is most appropriately supported by Internet tools but the customer prospecting strategy is not appropriate as a component of CRM which is supported by data web integration.

The acceptance of the hypothesis, stating that data web integration reinforces the relation between understanding customer expectations and customer loyalty seems to be totally justified because it supports the arguments developed by Goodhue et al., (2002) and Chatterjee et al., (2002). In fact, those authors mention that many companies believe that the availability of a wealth of internal and external data allows them to better understand their customer expectations. Therefore they do not hesitate to integrate their database system to their web site in order to track data about their customer behaviors and needs. The use of online customer data will create a high switching cost for the customers who are well-known by their suppliers. Those customers will then be loyal to their suppliers and vice-versa. There is another potential offered by the Internet to develop surveys and market research studies for collecting data. This potential can be enhanced by the use of forums and news groups, which can be vehicles for the collection, analysis and sharing of data as well as engaging in a dialogue with customers which leads to their loyalty.

CONCLUSIONS

Little of the existing research has empirically tested the effect of data web integration on CRM which leads to customer loyalty. This study has overcome these shortcomings with data collected from 224 SMEs in the USA and processed with PLS showing that the use of data web integration increases the impact of empowerment and understanding customer expectations on customer loyalty. However, the use of data web integration reduces the impact of customer prospecting on customer loyalty.

SMEs should initiate, maintain, and develop relationships with loyal customers by using, for example, forms available on their web sites. SMEs could integrate databases to their web sites containing those forms. Therefore, SMEs would be able to use online mechanisms to insert their customer data into their databases. SMEs will have the ability to use those data to understand their customer expectations, to analyze their customer profitability, and to make decisions related to loyalty

objectives. Employees should be empowered in their tasks to generate repeat purchases by the use of a data web integration system.

Implications for practice

The results of this study will allow Internet and information systems experts to better inform SMEs about the impact of data web integration use on customer loyalty. These experts need to be able to show SMEs reliable methods of using the management of database systems integrated to the Internet to build customer loyalty that involve increasing the level of CRM.

This study indicates that the research model introduced can become a good framework in understanding the success of data web integration. The ability to locate data through the web site will not pose a significant roadblock to SMEs that are willing to increase their customer loyalty via CRM. The findings of this research will help SMEs identify the CRM factors which they should emphasize when data web integration is used to augment customer loyalty. At a minimum, CRM, supported by data web integration in order to increase customer loyalty, should center on these factors: understanding customer expectations and empowerment. The results of this study also support the assumptions that even the opportunity offered by the integration between database systems and web sites continues to present data limitations for SMEs seeking to benefit from electronic customer prospecting. The findings of this study will help practitioners understand that the unique use of sophisticated data web integration does not automatically augment customer loyalty. It is rather through a combination of sophisticated data web integration and effective CRM that customer loyalty will grow.

Implications for research

The objective of this paper is, in part, to examine the theoretical foundation of electronic commerce. In fact, the findings of this study should be used to enhance transaction cost theory (TCT) because the use of data web integration technologies should reduce the cost and increase the number of SMEs using B2B electronic commerce systems. The findings of this study show that we built on advances in RM theory to establish a CRM construct and discussed its theoretical impact on customer loyalty.

The findings of this study let us understand that data web integration may have a direct impact on CRM or it may have an indirect impact on customer loyalty. These findings show the simultaneous theoretical relationship among data web integration, CRM, and customer loyalty. Therefore, the theoretical advance in this research is that loyalty is the result of the effective combination between CRM factors and IT facilities such as data web integration.

Further investigation is necessary for improved understanding of CRM and customer loyalty. In fact, continual empirical research is needed because Internet technology evolves so rapidly. Additional research should also expand the range of the data web integration variable and examine its effects on the relation between CRM and organizational performance. Further research also could

involve application of this research model to large businesses or use of the model to compare industries and companies of different sizes.

A multitude of research questions about the theoretical foundation of electronic commerce dimensions exists. The result of this research gives one way to organize thoughts about such questions. Another way of identifying and characterizing potential research is to consider the implications and problems that data web integration, CRM, and customer loyalty present to each of the traditional fields of business and IS. Many exciting research issues are being addressed with the results of this study and some are yet to be addressed. We hope that this paper inspires others to do future research by using concepts or variables from this research model.

Clearly, there is much uncertainty in the emerging world of electronic commerce. This uncertainty spans a variety of areas: marketing, organizational economics, technology, etc. All participants in the IS research community need to wake up to, understand, and adapt to the theoretical foundation of electronic commerce. This adaptation can help them redefine their research philosophies and approaches.

Finally, it is important to point out the contribution of data web integration as a moderating variable in a research model of electronic commerce for SMEs. The use of the statistical PLS software to process data is another distinctive element of this research. The authors encourage researchers to use the research model for various studies on electronic commerce.

Limitations

There were two notable limitations with the study. First, SMEs are in general not good at keeping precise business statistics as in the case of large organizations. Consequently, it is often fruitless to ask for such statistics from SMEs. Even if a small percentage of SMEs maintain such information, it is unlikely they are readily available to researchers due to suspicions about the intended use. Whether or not a small business is successful in electronic commerce on the Internet, there is little incentive to provide information to researchers. If an SME is successful, the knowledge of how the success was achieved is valuable and they may not want to reveal it. If an SME has not been successful, then it is very likely that they will have little to tell. Second, there were limitations of the questionnaire (survey) used, in particular, regarding the profile of respondents. From certain statements, we became aware that some respondents were webmasters who did not necessarily have a general background in their SMEs. The survey also did not take into account the measurement of the amount of the customers' interaction with the SMEs' databases.

REFERENCES

Aubert, B., S. Rivard & M. Patry (1994). Development of Measures to Assess Dimensions of IS Operation Transactions. Proceedings of the International Conference on Information Systems, 13-26.

Baron, R. M. & D.A. Kenny (1986). The moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Berry, L. L. (1995). Relationships Marketing of Services- Growing Interest, Emerging Perspectives. *Journal of Academy of Marketing Science, 23*, 236-245.

Chatterjee, D., R. Grewal & V. Sambamurthy (2002). Shaping up for E-commerce: Institutional enablers of the organizational assimilation of web technologies. *MIS Quarterly, 26*, 65-89.

Chin W. W., B. L. Marcolin & P. R. Newsted (1996). A Partial Least Squares Latent Variable Modeling Approch for Measuring Interaction Effects : Results from a Monte Carlo Simulation Study and Voice Mail Emotion/Adoption Study. Proceedings of the International Conference on Information Systems, 17-25.

Chow S. & R. Holden (1997). Toward An Understanding of Loyalty: The Moderating Role Of Trust. *Journal of Managerial Issues, 9*, 275-298.

Dekimpe, M., G. Steenkamp & M. Abeele (1997). Decline and variability in brand loyalty. *International Journal of Research in Marketing, 14*, 405-420.

Dick, A. S. & K. Basu (1994). Customer Loyalty: Toward an Integrated Conceptual Framework. *Journal of the Academy of Marketing Science 22*, 99-113.

Evans, J. R. & R. L. Laskin (1994). The Relationship Marketing Process : A Conceptualization and Application. *Industrial Marketing Management, 23*, 439-452.

Fournier, S. & J. L. Yao (1997). Reviving brand loyalty: A reconceptualization within the framework of consumer-brand relationships. *International Journal of Research in Marketing, 14*, 451-472.

Ghosh, S. (1998). Making Business Sense of the Internet. *Harvard Business Review. 2*, 126-135.

Goodhue, D. L., B. H. Wixom & H. J. Watson (2002). Realizing benefits through CRM: hitting the right target in the right way. *MIS Quarterly Executive, 1*, 79-94.

Grönroos, C. (1994). From Marketing Mix to Relationship Marketing: Towards a Paradigm Shift in Marketing. *Management Decision, 32*, 4-20.

Grover, V., J. T.C. Teng & K. D. Fiedler (2002). Investigating the Role of Information Technology in Building Buyer-Supplier Relationships. *Journal of the Association for Information Systems, 3*, 217-245.

Joseph, V. B., R. W. Cook & R. J. Javalgi (2001). Marketing on the Web: How Executives Feel, What Businesses Do. *Business Horizons, 44*, 32- 40.

Kassarjian, H. H. (1977). Content Analysis in Consumer Research. *Journal of Consumer Research, 4*, 8-18.

Kollmann, T. (2001). Measuring the Acceptance of Electronic Marketplaces: A Study Based on a Used-car Trading Site. *Journal of Computer Mediated Communication, 6*(2), Retrieved November 4, 2002, from http://www.ascusc.org/jcmc/vol6 /issue2/

Lawson-Body, A. & M. Limayem (2001). Business-to-Business Electronic Commerce: Impact of web site characteristics on Relationships Marketing Efficiency. *Management Science* (Science de Gestion), 5-6, 113-158.

Li, F. & H. Williams (1999). Interfirm collaboration through interfirm networks. *Information Systems Journal, 9*, 103-115.

Orman L. V. (2001). Database audit and control strategies. *Information Technology and Management, 2*, 27-51.

Ozer, M. (2000). Information Technology and New Product Development. *Industrial Marketing Management, 29*, 387-396.

Payne, A. (1994). Relationship Marketing Making the Customer Count. *Managing Service Quality, 4*, 29-31.

Peppers, D., M. Rogers & B. Dorf (1999). Is Your Company Ready For One-to-One Marketing. *Harvard Business Review, 1,* 151-160.

Poon, S. (2000). Business environment and internet commerce benefit---a small business perspective. *European Journal of Information Systems, 9,* 72-81.

Reichheld, F. F. (2001). Lead for Loyalty. *Harvard Business Review, 4*, 76-84.

Reichheld, F. F. & P. Schefter (2000). Your Secret Weapon on the Web. *Havard Business Review, 78*, 105-113.

Reinartz, W. & V. Kumar (2002). The Mismanagement of Customer Loyalty. *Harvard Business Review, 80*, 86-94.

Rowley J. & J. Dawes (2000). Disloyalty: a closer look at non-loyals. *Journal of consumer marketing, 17*, 538-549.

Sandeep, K. (2001). A Comprehensive Analysis of Permission Marketing. *Journal Computer Mediated Communication, 6*(2), Retrieved February 10, 2003, from http://www.ascusc.org/jcmc/

Sharp, B. (1997). Loyalty programs and their impact on repeat-purchase loyalty patterns. *International Journal of Research in Marketing, 14*, 473-486.

Shin B. (2003). An Exploratory Investigation of System Success Factors in Data Warehousing. *Journal of the Association of Information Systems, 4,* 141-170.

Shultz, D. H. (1995). Marketing Consuling Services on line, *Journal of Management Consuling, 8*, 35-42.

Swift, R. S. (2002). Executive response: CRM is changing our eras, the information we require, and our processes. *MIS Quarterly Executive, 1,* 95-96.

Thomas J. S., (2001). A Methodology for linking Customer Acquisition to Customer Retention. *Journal of Marketing Research, 38*, 262-268.

U.S. SBA, Office of Advocacy, (2000). Small Business Expansions in Electronic Commerce. Retrieved January 12, 2001, from Http://www.SBA.GOV/ADVO/STATS/

Yoon, S-J. & J-H Kim (2000). An empirical validation of a loyalty model based on expectation disconfirmation. *Journal of consumer marketing, 17*, 120-136.

A DESIGN OF SELF-QUESTIONING MECHANISM FOR INFORMATION REQUIREMENT SPECIFICATION

I-Lin Huang, Langston University

ABSTRACT

Incorrect requirement specifications are widely recognized as the major cause of information system failures. In order to improve the correctness of requirement specifications, various requirement specification techniques such as Data Flow Diagram, and Object Model, have been invented to help information analysts capture, understand, and represent information requirements. However, information analysts' cognitive abilities are still the most important determinant for the correctness of requirement specifications. Empirical studies have showed that novice information analysts cannot use requirement specification techniques effectively; and hence their performance of specifying information requirements is significantly lower than that of expert information analysts.

Self-questioning has long been recognized in the field of learning research as a strategy that can improve students' cognitive abilities on reading comprehension and problem solving. In order to improve the cognitive abilities of novice information analysts, this research argued that novice information analysts should be trained to incorporate self-questioning mechanism into their requirement specification process. On the basis of the theories on human cognition, this research also proposed several design strategies for self-asking questions that can guide novice information analysts to make more effective model-based reasoning and hence to achieve a higher correctness of requirement specifications.

INTRODUCTION

Specifying information requirements for the target information system under development is an important step in information systems development. The major output of specifying information requirements is a set of information requirement specifications (or simply requirement specifications) that state the desired functional and performance characteristics of the target information system (Roman, 1985). Basically, there are three purposes of requirement specifications: (1) facilitating an understanding of the target system, (2) guiding the process of information system design, and (3) serving as a basis for all communications concerning the information system being developed (Hsia, Davis & Kung, 1993; Schemer, 1987).

Incorrect requirement specifications are widely recognized as the major cause of information system failures (Dardenne, van Lamsweerde & Fickas, 1993; Davis, 1988; Dorfman, 1990; Greenspan & Mylopoulos, 1982; Scharer, 1981; Vessey & Conger, 1994). It has been reported that about two-thirds of information system failures can be attributed to the mistakes made in requirement

specifications (Fraser, Kumar & Vaishnavi, 1991; Shemer, 1987). According to an estimation, incorrect requirement specifications may cost fifty to one hundred times more than what would have been required if the errors are not discovered until system implementation (Roman, 1985; Shemer, 1987).

In order to improve the correctness of requirement specifications, various requirement specification techniques such as Data Flow Diagram, and Object Model, have been invented to help information analysts capture, understand, and represent information requirements (Couger, Colter & Knapp, 1982; Davis, 1988; Wieringa, 1998). However, the cognitive abilities of information analysts are still the most important determinant for the correctness of requirement specifications (Schenk, Vitalari & Davis, 1998). It has been reported that expert information analysts can use requirement specification techniques effectively by retrieving and applying more relevant modeling principles, making more critical testing of hypotheses and finally achieving requirement specifications with higher correctness (Allwood, 1986; Koubek, et al, 1989; Vitalari & Dickson, 1983). On the other hand, novice information analysts cannot use requirement specification techniques effectively because they have difficulties in correctly identifying important concepts in problem statements. The inadequacy of cognitive abilities makes novice information analysts unable to perform model-based reasoning effectively, and therefore leads to more errors in requirement specifications (Batra & Davis, 1992; Batra & Sein, 1994; Sutcliffe & Maiden, 1992).

It is believed that there are two cognitive characteristics that account for the difference of cognitive abilities between novice and expert information analysts: (1) reasoning processes, and (2) the knowledge organizations that support the reasoning processes. In addition, the learning process is slow for novice information analysts to reach the expert level of reasoning processes and knowledge organizations (Huang & Burns, 2000; Schenk, Vitalari & Davis, 1998). Therefore, the research question for this research is how to help novice information analysts achieve more correct requirement specifications even with relatively inadequate reasoning processes and knowledge organizations.

In the field of learning research, self-questioning is regarded as a cognitive strategy that can help students focus attention, organize new material, and finally integrate the new information with existing knowledge (Doerr & Tripp, 1999; Glaubman & Ofir, 1997; King, 1989, 1992; Wong, 1985). Empirical studies also showed that self-questioning can improve students' abilities of reading comprehension and problem solving (Doerr & Tripp, 1999; King 1992; Wong, 1985). Therefore, this research argued that novice information analysts should be trained to incorporate self-questioning mechanism into their cognitive process of information requirement specification. When novice information analysts can ask right questions to themselves during requirement specification, they will be able to learn how to think like an expert, organize their knowledge like an expert, and finally specify information requirements like an expert.

In order for the self-questioning mechanism to effectively guide novice information analysts to specify correct information requirements, a set of self-asking questions should be available to novice information analysts and meet at least two criteria: (1) can lead to correct requirement specifications, and (2) can fit the cognitive behavior of novice information analysts naturally. Currently, the common practice of requirement specification is focused on deriving information requirements from the answers

of the generic questions for the constructs of requirement specification techniques. However, the generic questions do not fit the cognitive behavior of novice information analysts naturally. Consequently, the generic questions may induce from novice information analysts the answers that lead to incorrect requirement specifications. Therefore, on the basis of the theories in human cognition (Gernsbacher, 1990; Graesser, 1995; Kintsch, 1988; Ortony, 1978), this research proposed that basic objects, coherence, and systematicity are the three cognitive principles that should be the guides for designing effective self-asking questions. For demonstration, a set of self-asking questions deriving from the three cognitive principles were designed to help specify information requirements by Data Flow Diagram.

The rest of this article will be organized into four sections. First, the self-questioning as a strategy to support human reading comprehension and problem solving will be discussed. Second, basic objects, coherence, and systematicity as guides for designing self-asking questions will be elaborated. In addition, Data Flow Diagram will be used to demonstrate the design of self-asking questions. Third, a brief example will be used to illustrate on how the self-asking questions support the process of specifying correct requirements. Finally, the conclusion will be made in the final section.

SELF-QUESTIONING AS A LEARNING STRATEGY

In learning research, self-questioning has been promoted as a strategy to improve students' abilities of reading comprehension and problem solving (Chin & Chia, 2004; Doerr & Tripp, 1999; Wong, 1985). When applying the strategy in reading comprehension or problem solving, the students are instructed to ask themselves questions before, during or after reading a text (for reading comprehension) or a problem statement (for problem solving). In this section, the discussion will be focused on three aspects of self-questioning in learning research: the benefits, the timing, and the questions of self-questioning.

The Benefits of Self-questioning

The benefits of self-questioning have been explored in learning research from three perspectives: active processing theory, metacognitive theory, and schema theory (Wong, 1985). Firstly, active processing theory suggests that it is the quantity, rather than the quality of self-asking questions that improves students' performance of reading comprehension or problem solving. The theory argues that actively asking questions during reading comprehension or problem solving can help students focus their thinking on the reading material, and hence improve their performance of reading comprehension and problem solving. The argument of active processing theory is supported by the evidence that the questions asked by students themselves facilitate understanding better than those asked by the instructors. Even further, more questions asked during reading or problem solving result in better comprehension and retention.

Secondly, metacognition theory suggests that the function of self-questioning is to help students monitor their understanding of the reading material (King, 1989; Nolan, 1991; Ozgungor & Guthrie, 2004). Therefore, the right self-asking questions as a metacognitive strategy should be able to help students focus on the important aspects of the material they read. Through a questioning--answering--questioning cycle, students can effectively analyze the content, relate it to their prior knowledge, and finally evaluate it and reassign their cognitive resources accordingly. This comprehension-monitoring process can not only let students know what they have learned, but also make them become aware of what they have not yet understood. When students fail to answer the questions they themselves post, they can take remedial action by asking themselves related questions, or asking questions of other information sources.

Finally, schema theory believes that students' reading comprehension or problem solving depends on their ability of activating prior knowledge (or called schema) in their minds (Chin & Chia, 2004; Wong, 1985). Therefore, the right questions asked by students themselves should be able to trigger the related concepts and experiences in students' minds, help students integrate their knowledge with the ideas in the reading material, and as a result, achieve better reading comprehension or problem solving.

The Timing of Self-questioning

It is suggested that self-questioning can enhance students' reading comprehension and problem solving no matter it is performed before, during, or after reading a text. Normally, students are advised to ask questions during reading a text. The function of self-questioning under this situation is mainly seeking information to (1) remove an obstacle in a plan or problem, (2) resolve a contradiction between ideas, (3) explain an unusual or anomalous event, (4) fill an obvious gap in students' knowledge base, and finally (5) make a decision among a set of alternatives that are equally likely (Graesser & Person, 1994).

Asking questions of predicting the ideas of a text before reading it can also facilitate reading comprehension (Nolan, 1991; Osman & Hannafin, 1994). It is believed that prediction can activate a cognitive plan to guide the students to understand the text. In addition, prediction can motivate students to read the text to confirm their prediction (Kletzien & Bednar, 1988).

Finally, self-questioning after reading a text can be helpful for students' reading comprehension because it provides students with a way to test their understanding; that is, it helps them to evaluate how well they understand what they are studying (King, 1992).

The Questions of Self-questioning

The main function of self-questioning is supposed to facilitate students' comprehension by guiding them in focusing attention (active processing theory), identifying important information (metacognition theory), and integrating the important information with existing knowledge (schema

theory). In order to achieve the function effectively, two types of questions have been proposed: domain-specific and domain-independent questions. Firstly, domain-specific questions are particularly useful in activating the prior content knowledge of the target text in students' minds. Therefore, domain-specific questions are very efficient in facilitating the integration between the new ideas and the students' prior knowledge. However, this approach will become very ineffective in situations that students have little or no background knowledge (Osman & Hannafin, 1994).

On the other hand, self-questioning with domain-independent concepts can improve students' comprehension in different domains. For example, when reading narratives, students can be taught to incorporate general story grammar elements such as leading character, action, obstacle, and outcome, into self-asking questions (Singer & Dolan, 1982). In this approach, students answer a general story grammar question for each element included in the story. Examples of general questions are as follows (Swanson, 1998): (a) Who is the leading character? (b) What is the leading character trying to accomplish? (c) What obstacles does the leading character encounter? Using the general questions as models, students create and answer their own specific questions based on the particular story they are reading, and hence reach a better understanding of the story.

In the next section, this research will discuss the three cognitive principles that can guide the design of effective self-asking questions for improving novice information analysts' ability to specify correct requirement specifications.

DESIGN STRATEGIES FOR SELF-QUESTIONING MECHANISM

During information requirement specification, users often provide information analysts with a verbal or written problem statement concerning the users' information requirements for a particular problem domain. The problem statement reflects the users' cognitive model of the problem domain. The task of the information analysts is to capture, understand, and represent the users' information requirements from the problem statement with the help of requirement specification techniques. From the perspective of human cognition, the cognitive behavior of the information analysts can be viewed as a reading comprehension process (Huang & Burns, 1998). Empirical studies on the modeling behavior of information analysts show a strong association among the activities of gathering information, identifying relevant facts, and conceptual modeling (Batra & Davis, 1992; Sutcliffe & Maiden, 1992). This strong association reflects that information requirement specification is basically an understanding process.

Because self-questioning is proven to be an effective strategy for improving reading comprehension and problem solving, this research argues that novice information analysts should be trained to incorporate self-questioning mechanism into their cognitive behavior of specifying information requirements. Even further, this research argues that effective self-asking questions should meet at least two criteria: First, they should promote the answers that can lead to correct requirement specifications. And second, they should be compatible to human cognition in specifying information requirements. Current practice of requirement specification is focused on the first criterion and tries to

induce correct requirement specifications directly from the generic questions for the constructs of requirement specification techniques. However, due to the inadequacy of reasoning processes and knowledge organizations, novice information analysts are often unable to come up with right answers for the generic questions, resulting in incorrect requirement specifications. Therefore, the design strategies that meet the above two criteria for effective self-asking questions should be based on the cognitive principles that people follow in reading comprehension and problem solving. Specifically, this research proposes that the design of self-asking questions should follow the principles of basic concepts, coherence, and systematicity because they are the effective guides for human cognitive behavior in reading comprehension and conceptual modeling. The rest of this section will discuss the four design strategies for self-asking questions including the generic question approach, and the three approaches based on the principles of basic concepts, coherence, and systematicity. In this research, Data Flow Diagram is used as an example of requirement specification techniques for demonstrating the design of self-asking questions.

Generic Questions

Generic questions are the questions that directly map the constructs of the intended requirement specification technique onto the concepts in the problem statement. For example, Data Flow Diagram is a requirement specification technique with four constructs: external entity, data flow, process, and data store. The generic questions for specifying information requirements on the basis of Data Flow Diagram are like those in table 1.

Table 1: The Generic Self-asking Questions for Data Flow Diagram
What are the external entities in the problem statement?
What are the data flows in the problem statement?
What are the processes in the problem statement?
What are the data stores in the problem statement?

By asking these generic questions, information analysts try to identify the relevant concepts in the problem statement and match them to the constructs of the intended requirement specification technique. Without rich experience and knowledge in requirement specification, novice information analysts basically use the definition of the constructs as shown in table 2 to guide them in identifying and mapping relevant concepts in the problem statement. Although this approach can focus novice information analysts on the relevant concepts, incorrect requirements or missing mappings may happen very often due to the inadequacy of novices' reasoning processes and knowledge organizations.

Table 2: The Definitions of the Four Important Constructs in Data Flow Diagrams (Satzinger, Jackson & Burd, 2002)	
Construct	Definition
External entity	External entities are sources or destinations of data flows which are outside the system under development. They represents users or other systems
Process	Processes transform inputs into outputs. They are the only active elements in data flow diagrams.
Data flow	Data flows are composite data items flowing from an element to a process (input dataflow) or from a process to an element (output dataflow) .
Data store	Data at rest. Data are stored for later. Corresponding to entities in Entity-relationship Model. Often implemented by databases

Basic Concepts

1. The principle

According to the research in concept formation (Komatsu, 1992; Rosch, et al., 1976), humans can recognize objects at different levels of abstraction. However, there is one level of abstraction at which the basic concepts of the objects are made. From the perspective of human cognition, the basic concepts are the concepts that carry the most information for human reasoning and are easiest to understand. For example, the concept of "chairs" can be viewed as "furniture" at a higher-level of abstraction or as "kitchen chairs" at a lower-level of abstraction. However, when we see a kitchen chair, most of us would recognize it as a chair rather than a piece of furniture or a kitchen chair because most of us view the concept of "chairs" as the basic concept. We recognize a kitchen chair as a chair rather than a kitchen chair because it is difficult to decide whether a chair is a kitchen chair or not due to few minor differences between kitchen chairs and other chairs. On the other hand, although we can easily classify a chair into furniture, we do not recognize a chair by the concept of furniture because the concept of chairs carries more information for human reasoning than furniture. The instances of furniture share fewer properties than those of chairs. Many objects like lamps can also be classified into furniture even though lamps are very different from chairs. As a result, there are fewer inferences can be made by the concept of furniture than by the concept of chairs

In order to manage infinite phenomena happening in the real world, we have to generate concepts to abstract the phenomena into groups, with the instances of each group sharing similar important properties. By abstraction, we can use a few concepts to handle infinite phenomena in the real world. However, by abstraction we also lose the information about the differences among the instances in the same group. Therefore, we recognize objects by the basic concepts because they are the concepts at the best level of abstraction that include the most phenomena but lose the least information. According to Komatsu (1992), the basic concepts are the most useful concepts for us

because they are those formed during perception of the environment, those formed during our childhood, and those to be most codable, most coded, and most necessary in language.

2. The design strategy based on basic concepts

For analyzing business information systems, business concepts are the basic concepts for the business users and hence easier to identify and understand than the information concepts such as objects, and processes used by requirement analysis techniques. For example, customers as a business concept are more natural for users' cognition than objects in Object Model, or external entities in Data Flow Diagram because users can generate more inferences from the concept of customers than from that of objects, or external entities. Therefore, on the basis of the theories on concept formation, the self-asking questions that are focused on business concepts are easier to answer correctly by novice information analysts than those on information concepts that are used as constructs for requirement specification techniques.

Assume that an information analyst is reading a problem statement and feeling that the information requirements are concerned with order processing. And if the information analyst has the domain knowledge about order processing as shown in Figure 1, she or he can ask questions based on business concepts during requirement specification as in table 3:

Figure 1: the Data Flow Diagram for Order Processing

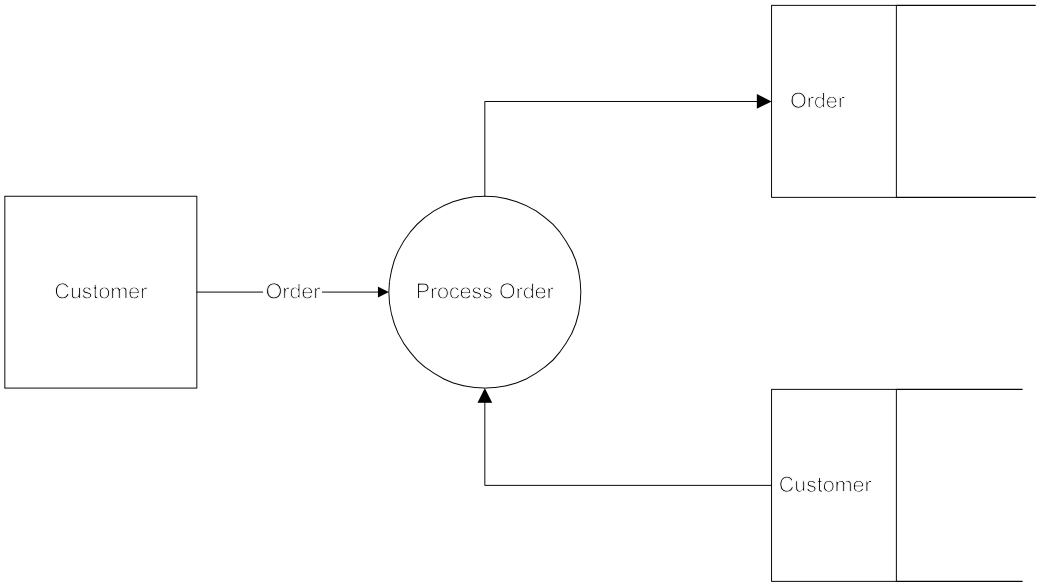


Table 3: Self-asking Questions Based on the Basic Concepts from Order Processing System
Are there people processing orders?
Is there an order file?
Is there a customer file?
Are there customers outside of this system?
Are there orders flowing in the system from the customers?

It is obvious that the above questions about business concepts are easier to answer than those about information concepts. However, business concepts are domain-specific concepts and will become difficult to identify and model if there is no pre-exist template of requirement specifications, or if the information analysts have little or no background knowledge.

Coherence

1. The principle

Coherence is well accepted as the goal of human cognition on reading comprehension. Specifically, reading comprehension is a process of building a coherent situation model of the text or problem statement being comprehended (Gernsbacher, 1990; Graesser, 1995; Kintsch, 1988; Ortony, 1978). A situation model is a mental representation of the concepts that are explicitly mentioned in the problem statement and that are derived by the reader's knowledge (Graesser, Singer, and Trabasso, 1994; Kintsch, 1988; Mckoon & Ratcliff, 1992). People construct situation models in order to be able to answer the questions related to the problem statement. If people can answer the questions related to the problem statement by searching the structure of the situation model, then people feel that they have understood the material in the problem statement (Gernsbacher, 1990; Graesser, 1995; Kintsch, 1988; Ortony, 1978). Coherence is "a state or situation in which all the parts or ideas fit together well so that they form a united whole" (Collins Cobuild English Dictionary, 1995). The coherence of a situation model for a problem statement determines how well the reader can answer questions about the problem statement, and remember, summarize, or verify the statement (Kintsch, 1988; Trabasso, 1989). In general, the more coherent the situation model of a problem statement, the better the comprehension and usage of the problem statement (Trabasso, 1989).

From the perspective of reading comprehension, the task of information analysts is to make goal-directed inferences to achieve coherent requirement specifications from problem statements constrained by the predetermined constructs and cognitive structures of requirement analysis techniques. While information analysts analyze a problem statement on the basis of a particular requirement specification technique, they may find that the problem statement can not be fitted into the conceptual model very well. Information analysts then perform intensive coherence inferences to resolve the discrepancies. Although correctness may be the goal for information requirement

specification, information analysts believe that understanding is achieved if the requirement specifications are coherent.

2. The design strategy based on coherence

With coherence as the goal for requirement specification, we can say that we have specified requirements for a problem statement accurately only if we can construct a coherent requirement model for the problem statement. On the other hand, if there is still any disconnected concept, or called incoherent concept, in our requirement model, we will feel that we fail to specify the requirements completely and accurately. In this situation, self-questioning can be invoked to ask questions that will be answered by our domain knowledge or the experts who have domain knowledge about the incoherent concept. If the answers can provide coherent connections for the incoherent concept, then the self-questioning serves its purpose and we have better chance to model the requirements completely and accurately.

Therefore, from the perspective of coherence, the purpose of self-questioning is to derive answers that can provide coherent structures for a concept. Coherence structures can be at least divided into two levels: local coherence and global coherence (Long, Oppy & Seely, 1997; Mckoon & Ratcliff, 1992). First, local coherence can be defined as "a small set of adjacent sentences that makes sense on its own or in combination with easily available general knowledge" (Mckoon & Ratcliff, 1992, p.444). From the perspective of local coherence, the coherent structures in Data Flow Diagram are the acceptable connections through data flows as shown in table 4. Derived from the coherent structures in table 4, a list of the possible self-asking questions that can be used to derive the locally coherent connections is thus shown in table 5.

Table 4: Legal Conceptual Connections for Local Coherence

		Data Moved To		
	Data Flow	External Entity	Process	Data Store
Data Moving From	External Entity		**A**	
	Process	**B**	**C**	**D**
	Data Store		**E**	

Second, global coherence involves constructing a requirement model that reflects a bigger unit of problem statement like a paragraph, or even the whole problem statement (Long, Oppy & Seely, 1997). Two examples of the generic coherent structures defined at the global level are process decomposition structure (Martin, 1989), and event partition structure (McMenamin & Palmer, 1984). The coherent structure of event partition model is shown in figure 2. The questions that can drive the concepts in the event partition structure are shown in table 6.

Table 5: Local Coherence-driven Questions for Data Flow Diagram Modeling	
Construct	Question
External Entity	A-structure: (external entity---request---process) What is the request of the external entity? To what process? B-structure: (process---report---external entity) What is the report for the external entity? Provided by what process?
Process	A-structure: (external entity---request---process) What initiates the process? From which external entity? B-structure: (process---report---external entity) What report is generated by this process? To what external entity? C-structure: (process---intermediate---process) What immediate result is generated by this process? To what process? D-structure: (process---save---data store) What data store save the result of the process? E-structure: (data store---retrieve---process) What data store provides information for the process? What data are provided?
Data store	D-structure: (process---save---data store) What data in the data store are saved? From what process? E-structure: (data store---retrieve---process) What data in this data store are retrieved? By what process?
data flow	A-structure: (external entity---request---process) What external entity is the sender of the request? To what process? B-structure: (process---report---external entity) What external entity is the receiver of the report? From what process? C-structure: (process---intermediate result---process) What process generates the immediate result? What process will do the further processing? D-structure: (process---save---data store) What is the data store saving the data? What is the process generating the data? E-structure: (data store---retrieve---process) What is the process retrieving the data? From which data store?

Figure 2: Event Partition Structure (Satzinger, Jackson & Burd, 2002)

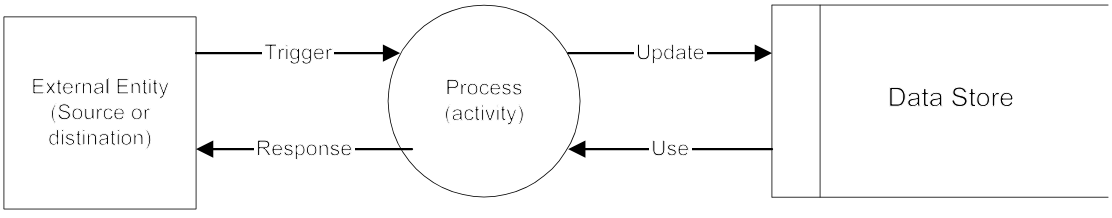


Table 6: Global Coherence-driven Questions for Event Partition Structure	
Construct	Question
Event	What is the event?
Source	What is the source?
Destination	What are the destinations?
Trigger	What is the trigger?
Response	What are the responses?
Process	What is the process?
Data Store	What are the data stores?
Update	Does the process update any data stores?
Use	Does the process use any data stores?

Systematicity

1. The principle

When selecting a base structure (e.g. a conceptual structure from a particular requirement specification technique) to be mapped onto a target structure (e.g. a conceptual structures from a problem statement), the base structure with higher-order relation will be more likely to be imported into the target structure than is that with an isolated relation or object-attribute. It is called the principle of systematicity (Gentner & Markman, 1997). The principle of systematicity is a structural expression of our tacit preference for coherence and deductive power from the mapping. Basically, there are four different types of similarity that support the mapping from the base concepts onto the target concepts and provide different levels of coherence to the resultant model (Gentner, 1983): First, literal similarity provides the highest coherence because it matches both relational structure and object-descriptions (e.g., The order processing system is like that of company X I analyzed last year.). Second, analogy provides the second highest coherence because it matches relational structure and disregards object descriptions (e.g., The order processing system is like the library system I analyzed two years ago.). Third, abstraction provides the third highest coherence because the base structure is an abstraction of the target structure (e.g., I want to use the base structure, A-structure "external entity, request, process", from Data Flow Diagram to match the target structure). Fourth and finally, surface similarity provides the lowest coherence because it matches some aspects of object descriptions and disregards relational structure (e.g., I want to model customer as external entity, and order as data store).

2. The design strategy based on systematicity

The principle of systematicity is not a principle for deriving self-asking questions, but rather a principle for selecting the best set of questions among several sets of answerable questions. When there are several sets of self-asking questions can be raised to induce answers for information requirements,

systematicity will select the set of self-asking questions that can induce answers to lead to the most coherent requirement specifications. On the basis of systematicity, the self-asking questions in table 3 will have the highest priority because they are derived from the principle of basic objects and provide a mapping on the basis of literal similarity. The self-asking questions in table 6 will receive the second highest priority because they are derived from the principle of (global) coherence and provide mappings on the basis of abstractions. The self-asking questions in table 5 also provide mappings on the basis of abstraction. However, they will have lower priority than those in table 6 because the questions in table 5 are derived from local coherence with lower order relations. Finally, the questions in table 1 can get only the lowest priority because they are generic questions and provide only the mapping on the basis of surface similarity.

The self-asking questions that provide a mapping on the basis of analogy, once identifies, will have higher priority than those on the basis of abstraction, or surface similarity. However, it is not easy to identify analogical relations between a base structure and a target structure at the first place (Maiden & Sutcliffe, 1992). Therefore, we do not actively pursue a set of self-asking questions for analogical mappings at the very early stage of requirement specification.

AN EXAMPLE OF SELF-QUESTIONING FOR DATA FLOW DIAGRAM MODELING

In this section, we will assume a requirement sentence, "The customer first sends an order to John, the order clerk", in a problem statement of an order processing system as an example to illustrate how self-questioning leads to correct requirement specifications. First, the novice information analysts will read the requirement sentence. Then guided by the principle of systematicity, they will raise different set of self-asking questions on the basis of their domain and modeling knowledge as discussed below:

Situation 1: Follow the Principle of Basic Concepts

1. The question-answering

When the information analyst has the prior experience on order processing system, he or she may raise the self-asking questions in table 3. The answers for the questions are in table 7 as below:

Table 7: Answer the Questions Derived from the Principle of Basic Concepts

Question	Answer
Are there people performing an activity of processing order?	John, the order clerk
Is there an order file?	Need further investigation
Is there a customer file?	Need further investigation
Are there customers outside of this system?	Yes
Are there orders flowing in the system from the customers?	Yes

2. The outcome

After confirming the existence of an order file and a customer file in further reading of the problem statement, the information analyst can achieve a validated data flow diagram like in figure 1.

Situation 2: Follow Analogical Reasoning

1. The question-answering

When the information analyst has the prior experience on information systems such as library information systems that are similar to the order processing system under investigation, he or she may raise the self-asking questions derived from analogy. The answers for the questions are in table 8 as below:

Table 8: Answer the Questions Derived from Analogical Similarities

Question	Answer
Is there an activity similar to checking out books done by librarian?	Order processing by John, the order clerk
Is there a file similar to a check-out books file?	Need further investigation
Is there a file similar to student file?	Need further investigation
Are there people similar to students outside of this system?	Yes, customers
Are there similar dataflows from the students?	Yes, orders

2. The outcome

The information analyst can also achieve a validated data flow diagram like in situation 1. However, the information analyst may not pursue the analogical mappings because the analogical similarities are difficult to identify (Maiden & Sutcliffe, 1992).

Situation 3: Follow the Principle of Coherence at the Global Level

1. The question-answering

When the information analyst has no prior experience on information systems similar to the order processing system under investigation, he or she may raise the self-asking questions derived from global coherence after fully understanding the intention of the whole problem statement. The answers for the questions are in table 9 as below:

Table 9: Answer the Questions Derived from Global Coherence	
Question	Answer
What is the event?	Order entry
What is the source?	Customer
What are the destinations?	None
What is the trigger?	Order
What are the responses?	None
What is the process?	Process order
What are the data stores?	Customer file, Order file
Does the process update any data stores?	Order file
Does the process use any data stores?	Customer file

2. The outcome

The information analyst can achieve a validated data flow diagram like in situation 1. However, the information analyst needs first to be able to understand the requirements of the problem statement at a global level.

Situation 4: Follow the Principle of Coherence at the Local Level

1. The question-answering

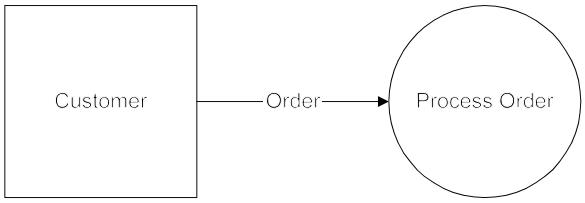
If the information analyst has no prior experience on information systems similar to the order processing system under investigation, and does not have an adequate understanding of the intention of the whole problem statement, he or she may pursue the answers for the self-asking questions from the principle of coherence at the local level. The information analyst first assumes that the incoherent concept "Customer" is an external entity. He or she then may try to validate his or her assumption by answering the questions in table 10 as below:

Table 10: Answer the Questions Derived from A-structure: (external entity---request---process)	
Question	Answer
What is the request of the external entity?	Order
To what process?	Order processing by John, the order clerk

2. The outcome

On the basis of local coherence, the information analyst is focused on the concepts in the requirement sentence without referring to the rest part of the requirement statement. The information analyst can thus achieve a validated partial data flow diagram like in Figure 3.

Figure 3: Partial requirement specification



Situation 5: Answer Generic Questions

1. The question-answering

If the information analyst has no domain knowledge on information systems similar to the order processing system under investigation, and does not have adequate modeling knowledge on Data Flow Diagram, he or she may try to specify the information requirements on the basis of the definitions of the constructs as in table 2. Therefore, the answers to the generic questions as in table 1 are like in table 11 as below:

Table 11: Answer Generic Questions for Data Flow Diagram	
What are the external entities in the problem statement?	Customer, Order Clerk
What are the data flows in the problem statement?	Order
What are the processes in the problem statement?	Process order
What are the data stores in the problem statement?	?

2. The outcome

On the basis of the definitions of the constructs of Data Flow Diagram, the information analyst is focused on the surface meaning of the concepts in the requirement sentence and tries to categorize the concepts into different constructs. Some errors may happen due to the lack of validation from the coherent structure of the requirement specification.

CONCLUSION

Traditionally, the research studies for improving the modeling performance of novice information analysts have been focused on providing access to domain and modeling knowledge through computer-aided software engineering (CASE) tools. However, this research is concerned with how to improve novices' cognitive abilities for requirement specification. On the basis of the principles of basic concepts, coherence, and systematicity, this research has proposed design strategies for effective self-asking questions. Guided by the effective self-asking questions, novice information analysts can specify more correct information requirements.

In sum, the design of self-questioning mechanism of requirement specification contributes to the research in information requirement specification in the following three aspects. First, this design can be used as the theoretical model for testing hypotheses about the impact of self-questioning mechanism on the performance of information requirement specification. Second, this design provides a theoretical basis for training novice information analysts to use self-questioning to improve their performance of requirement specification. In addition, the self-questioning mechanism can also facilitate the transition of information analysts from novice to expert. Third and finally, this design provides a theoretical basis for the development of computer-aided software engineering (CASE) tools. On the basis of this research, how to guide novice information analysts to ask right questions to themselves at local and global levels of problem statements is an important issue for the research in CASE tools.

REFERENCES

Adelson, B. & Soloway, E. (1985). The role of domain experience in software design," *IEEE Transactions On Software Engineering, se-11* (11), 1351- 1360.

Agarwal, R., & Tanniru, M. R. (1990). Knowledge acquisition using structured interviewing: an empirical investigation. *Journal of Management Information Systems, 7*(1), 123-140.

Allwood, C. M. (1986). Novices on the computer: a review of the literature. *International Journal of Man-Machine Studies, 25*, 633-658.

Batra, D. & Davis, J. G. (1992). Conceptual data modeling in database design: similarities and differences between expert and novice designers. *International Journal of Man-Machine Studies, 37*, 83-101.

Batra, D. & Sein, M. K. (1994). Improving conceptual database design through feedback. *International Journal of Human-Computer Studies, 40*, 653-676.

Chin, C. & Chia, L. G. (2004). Problem-based learning: Using students' questions to drive knowledge construction. *Science Education, 88*(5), 707-727.

Collins Cobuild English Dictionary (1995). Harper Collins Publishers.

Couger, J. D., Colter, M. A. & Knapp, R. W. (1982). *Advanced System Development/ Feasibility Techniques.* New York: Wiley.

Dardenne, A., Lamsweerde, A. V. & Fickas, S. (1993). Goal-directed requirements acquisition. *Science of Computer Programming, 20*, 3-50.

Davis, A. M. (1988). A comparison of techniques for the specification of external system behavior. *Communications of The ACM, 31* (9), 1098-1115.

Doerr, H.M. & Tripp, J.S. (1999). Understanding how students develop mathematical models. *Mathematical Thinking and Learning, 1*(3), 231-254.

Dorfman, M. (1990). System and software requirements engineering. In R. H. Thayer & M. Dorfman (Eds.), *System and Software Requirements Engineering* (pp. 4-16). Los Alamitos, CA: IEEE Computer Society Press.

Fraser, M. D., Kumar, K. & Vaishnavi, V. K. (1991). Informal and formal requirements specification languages: bridging the gap. *IEEE Transaction On Software Engineering, 17* (5), 454-465.

Gernsbacher, M. A. (1990). *Language Comprehension As Structure Building*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Gentner, D. (1983).Structure-Mapping: A Theoretical Framework for Analogy, *Cognitive Science*, (7), 155-170.

Gentner, D.& Markman, A. B.(1997). Structure Mapping in Analogy and Similarity, *American Psychologist, 52*(1), 45-56.

Glaubman, R. & Ofir L. (1997). Effects of self-directed learning, story comprehension, and self-questioning in Kindergarten, *The Journal of Educational Research, 90*(6), 361-374

Graesser, A. C. (1995). Inference generation and the construction of situation models. In Weaver, C. A., Mannes, S., and Hetcher , C. R. (eds.). *Discourse Comprehension*. Hillsdale, New Jersey: Lawrence Erbaum Associates, Publishers. 117-139.

Graesser, A. C. & Person, N.K. (1994). Question asking during tutoring. *American Educational Research Journal. 31*(1), 104-137.

Graesser, A,, Singer, M. & Trabasso, T. (1994). Constructing inferences during text comprehension. *Psychological Review, 101* (3), 371-395.

Greenspan, S. J. & Mylopoulos, J. (1982). Capturing More World Knowledge in the Requirements Specification, *Proceedings of International Conference in Software Engineering*, IEEE Computer Society Press, 225-234.

Guindon, R. & Curtis, B. (1988). Control of Cognitive Process During Software Design: What Tools Are Needed? *Proceedings of CHI'88*, 263-268.

Guindon, R., Krasner, H. & Curtis, B. (1987). Cognitive process in software design: activities In early, upstream design, in *Human-computer Interaction- INTERACT'97*, Bullinger, H. J., and Shackel, B. (eds), North-Holland, 383-388.

Hsia, P., Davis, A. M. & Kung, D. C. (1993). Status report: requirements engineering. *IEEE Software*, 75-79.

Huang, I. & Burns, J. R. (2000). A cognitive comparison of modeling behavior between novice and expert information analysts. *Proceedings of AMCIS Conference*, 1316-1322.

Huang, I. & Burns, J. R. (1998). A cognitive model of information requirement analysis on the basis of structure building theory of language. *Proceedings of AMCIS conference*, 659-661.

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures*, American Educational Research Journal, 29*(2), 303-323.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures, *Contemporary educational Psychology, 14*, 366-381.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95* (2), 163-182.

Kintsch, W. (1974). *The Representation of Meaning in Memory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Kletzien, S.B. & Bednar, M.R. (1988). A framework for reader autonomy: an integrated perspective. *Journal of Reading, 32*(1), 30-33.

Komatsu, L. K. (1992) Recent views of conceptual structure. *Psychological Bulletin, 112*(3), 500-526.

Koubek, R. J., Salvendy, G., Dunsmor, H. E. & Lebold, W. K. (1989). Cognitive issues in the process of software development: review and reappraisal. *International Journal of Man-Machine Studies, 30*, 171-191.

Long, D. L., Oppy, B. J. & Seely, M. R. (1997). A "global-coherence" view of event comprehension: Inferential processing as question answering. In van Broek, P. W., Bauer, P. J. & Bourg, T. (eds.) *Developmental Spans in Event Comprehension And Representation: Bridging fictional and actual events*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Maiden N.A.M. & Sutcliffe, A.G.(1992). Exploiting reusable specification through analogy. *Communications of the ACM, 35*(4), 55-64.

Martin, J. (1989). *Information Engineering, Book I: Introduction*. Prentice-Hall Corporation.

McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99* (3), 440-466.

McMenamin, S.M. & Palmer, J.F. (1984*). Essential Systems Analysis*. Yourdon Press.

Nolan, T.E. (1991). Self-questioning and prediction: Combining metacognitive strategies. *Journal of reading, 35*(2), 132-138.

Ortony, A. (1978). Remembering, understanding, and representation. *Cognitive Science, 2*, 53-69.

Osman, M.E. & Hannapin, M.J. (1994). Effects of advance questioning and prior knowledge on science learning. *Journal of Educational Research, 88*(1), 5-13.

Ozgungor, S. & Guthrie, J.T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text, *Journal of Educational Psychology, 96*(3),

Roman, G. (1985). A taxonomy of current issues in requirements engineering. *IEEE Computer*, 14-22.

Rosch, E., Mervis, C. B., Gary, W. D., Johnson, D. M. & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Satzinger, J.W., Jackson, R. B. & Burd, S. D. (2002*). Systems Analysis and Design.* Second Edition, Course Technology.

Scharer, L. (1981). Pinpointing Requirements, *Datamation,* April, 17-22.

Schenk, K. D., Vitalari, N. P. & Davis, K. S. (1998). Differences between novice and expert systems analysts: what do we know and what do we do*? Journal of Management Information Systems, 15* (1), 9-50.

Shemer, I. (1987). System analysis: a systematic analysis of a conceptual model. *Communications of the ACM*, 30 (6), 506-512.

Singer, H. &Donlan, D. (1982). Active comprehension: problem-solving schema with question generation for comprehension of complex short stories. *Reading Research Quarterly 27*, 166-186.

Sutcliffe, A. & Maiden, N. (1992). Analysing the novice analyst: cognitive model in software engineering. *International Journal of Man-Machine Studies, 36*, 719-740.

Swanson, P.N. (1998). Teaching effective comprehension strategies to students with learning and reading disabilities. *Intervention in School of Clinic, 33*(4), 209-219.

Trabasso, T. (1989). Causal representation of narratives. *Reading Psychology, 10*, 67-83.

Vessey, I. & Conger, S. A. (1993). Learning to specify information requirements: the relationship between application and methodology. *Journal of Management Information Systems, 10* (2), 177-201.

Vessey, I. & Conger, S. A. (1994). Requirement specification: learning object, process, and data methodologies. *Communications of The ACM, 37* (5), 102-113.

Vitalari, N. P. & Dickson, G. W. (1983). Problem solving for effective system analysis: an experimental exploration. *Communications of the ACM, 26* (11), 948-956.

Wieringa, R. (1998). A survey of structured and object-oriented software specification methods and techniques. *ACM Computing Survey, 30*(4), 459-527.

Wong, B.Y.L. (1985). Self-questioning instructional research: a review. *Review of Educational Research, 55*(2), 227-268.