ISSN: 1940-9893 (Print), 1940-9907(Online) Vol.3, No.1, March 2010

Scientific Research

Journal of Service Science and Management

www.scirp.org/journal/jssm

Journal Editorial Board

ISSN: 1940-9893 (Print), 1940-9907(Online)
http://www.scirp.org/journal/jssm

J. Service Science & Management, **2010, 3*:* 1-179**
Published Online March 2010 in SciRes (http://www.SciRP.org/journal/jssm/)

TABLE OF CONTENTS

Volume 3 Number 1 March 2010

J. Service Science & Management, **2010, 3*:* 1-179**
Published Online March 2010 in SciRes (http://www.SciRP.org/journal/jssm/)

J. Service Science & Management, **2010, 3*:* 1-179**
Published Online March 2010 in SciRes (http://www.SciRP.org/journal/jssm/)

Journal of Service Science and Management (JSSM)

Journal Information

SUBSCRIPTIONS

The *Journal of Service Science and Management* (Online at Scientific Research Publishing, www.SciRP.org) is published quarterly by Scientific Research Publishing, Inc., USA.

Subscription rates:

Print: $50 per issue.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jssm@scirp.org

J. Service Science & Management, **2010, 3: 1-15**
doi:10.4236/jssm.2010.31001 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Offering Appropriate Information Technologies at Different Stages in the Customer Service Life Cycle for Improved Service Delivery

Jennifer E. Gerow, Janis Miller

Clemson University, Clemson, USA.
Email: jgerow@clemson.edu, janism@clemson.edu

Received October 28th, 2009; revised December 4th, 2009; accepted January 9th, 2010.

ABSTRACT

The service sector is growing in importance in the US, particularly in B2B contexts. Despite this, research in these areas lags behind manufacturing and B2C studies. The purpose of this article is to begin addressing this issue by looking at how information technology can improve service delivery in B2B contexts by looking at the fit between the CSLC and a mix of customer contact technology modes. Specifically, we look at how Web 2.0 technologies facilitate interorganizational communication even for small and medium businesses by providing a wider variety of technologies at a lower cost. This allows businesses to use technology to reduce business customer uncertainty and equivocality, hence improving the quality of their service delivery. Due to the increasing use of the Internet as a source of communication and the evolution of customer expectations, this is becoming an increasingly important research topic. Propositions and implications are presented.

Keywords*: *Service Science, B2B, Customer Service Contact model, Customer Service Life Cycle, Media Richness Theory

1. Introduction

Although the service sector contributes 79.6% to the United States' GDP (Gross Domestic Product) [1], service research lags behind that of manufacturing. This is significant because the distinctive characteristics of managing services require a different approach than what is found in manufacturing [2]. IBM and others have recently urged participation in the service science movement (or Service Science, Management and Engineering – SSME). On their web page, IBM describes service science as "a growing multi-disciplinary research and academic effort that integrates aspects of established fields like computer science, operations research, engineering, management sciences, business strategy, social and cognitive sciences, and legal sciences" [3]. The goal of this movement is to improve quality and productivity, increase efficiency and scalability, and nourish learning and innovation rates in the service industry by combining interagency, interdisciplinary and innovative resources [3,4].

In addition to the shortage of research on the service industry, we also find that even though 85% of e-commerce is business-to-business (B2B) [5], most of the e-commerce research has been in business-to-consumer (B2C) [6]. Just like the service science movement, researchers are calling for more research in B2B contexts [7]. This paper resides in that rare intersection of B2B service research and examines how technology, specifically Web 2.0, can be used to facilitate efficient and effective service delivery.

To address this need, we bring together the customer service contact model [8] from the field of operations management, which describes various modes of customer contact, and the customer service life cycle model [9] from the field of management information systems, which describes the stages a customer must go through when contacting a vendor. Our intention is to bring insight to the question: "What form of information technology improves service delivery in specific situations?" To address the question, we created the matrix shown in Figure 1 (see Table 1 for construct definitions).

In this matrix, we cross Froehle and Roth's [8] technology-mediated customer service contact model with Ives and Mason's [9] customer service life cycle (CSLC) model. The technology-mediated customer service contact model indicates a range of service delivery modes ranging from technology-free, where two individuals conduct business without the aid of any technology, to

↓Contact Mode CSLC Stage→	Re-quire-ment	Acqui-sition	Owner-ship	Retire-ment
Technology-Free				
Technoogy-Assisted				
Technoogy-Facilitated				
Technoogy-Mediated				
*Peer-to-Peer				
Technoogy-Generated				
*Technology-Only				

*New contact modes added in this study to Froehle and Roth's framework [8]

Figure 1. Customer service contact mode / customer service life cycle matrix

technology-only, where business transactions are conducted without the presence of any individuals. For the second dimension of our matrix, we chose to use Ives and Mason's [9] CSLC model because it provides a framework for applying IT to external, customer-focused applications [14]. The framework of their model contains four stages: requirements, acquisition, ownership, and retirement [9] as illustrated in the matrix. Specifically, this matrix links the way services are delivered to a business customer depending on the business customer's stage in the customer service life cycle. Our model will aid in the selection of the appropriate technology level based on the business customer's position within the customer service life cycle.

This paper contributes to the existing literature in the following ways. First, we present the matrix which provides strategic insight into which customer contact modes

Table 1. Construct definitions

Construct	Definition
Customer Contact	physical presence of the buyer in the service system [10]
Full-Service	the buyer interacts directly with the provider in the front office (e.g. technology-free and technology-assisted modes of customer contact) [8]
Self-Service	the customer service representative does not actively participate in service performance (e.g. technology-generated mode of customer contact) [8]
Auto-Service	no human involvement is required for business transactions to occur (e.g. technology-only mode of customer contact)
Mode of Customer Contact	
Technology-Free	the buyer only makes contact with the service representative and technology is not involved at all in the process [8]
Technology-Assisted	the customer only makes contact with the service representative but the service representative has access to technology in order to process the transaction [8]
Technology-Facilitated	the technology system available to both its service representatives as well as the service customer [8]
Technology-Mediated	the service customer and service representative interact with technology but they do not interact with each other [8]
Peer-to-Peer	the customers discuss (with other customers) the supplier's service on a supplier-provided platform
Technology-Generated	the service customer interacts directly with the technology without the assistance of a service representative [8]
Technology-Only Mode	the computer systems at both companies interact with each other with no human involvement
Customer Service Life Cycle Opportunity	
Requirements	analyze and define purchase requirements, gather service information from suppliers, and identify service providers that meet requirements [9,11]
Acquisition	search for and identify potential suppliers, order the service, transfer funds, and evaluate the service [9,12,13]
Ownership	integrate the service into business processes, monitor the interaction [9,12]
Retirement	analyze expenses and switching costs [9,12]

should be utilized at the various stages in the customer service life cycle. Second, we add two additional customer contact modes, peer-to-peer and technology only, to the work of Froehle and Roth [8]. Third, we provide insight into how Web 2.0 can be used for both intra- and inter-company coordination. Finally, we present research propositions based on our framework, leading to future research that can test our conceptual framework.

This paper is organized as follows. In the next section we provide a background on Web 2.0, the characteristics of business-to-business transactions, and the concepts of equivocality and information richness. The succeeding section presents our model and propositions. The final section discusses implications for future research.

2. Background

In this section we provide information which will set the stage for the presentation of our model in Section 3. Our first background section covers Web 2.0, followed by a section on B2B, and finishes with a discussion of equivocality and information richness.

2.1 Web 2.0

Because Web 2.0 is a fairly new technology, we will define Web 2.0, describe four concepts of Web 2.0 that are suitable for B2B usage, and discuss how businesses can exploit these concepts for their B2B transactions. First, we define Web 2.0 as a collection of websites that use information technologies and applications to encourage user participation, information sharing, social interaction, and collaboration [5]. Unlike the more traditional websites (referred to as Web 1.0), Web 2.0 sites are typically used as services to accomplish tasks, typically with other people [5]. In a B2B context, companies can leverage Web 2.0 concepts such as Rich Internet Applications, Software-as-a-Service, collective intelligence, and mashups [15]. These are described in detail in the following paragraphs.

The first Web 2.0 concept that applies to B2B environments is Rich Internet Applications (RIA). These applications allow business customers to specify the type of information they want to receive from a supplier so that the supplier can proactively provide the information [15] and the buyer can then organize and view the information they receive more effectively [16]. For example, Web 2.0 can enable a business customer to automatically and securely access information about their account or perform complex data entry for highly customized purchase orders through an easy-to-use desktop-like experience such as drag and drop [17,18]. In short, Web 2.0 technology allows a business customer to get real-time information about their account or current business transactions in an information-rich environment with the added benefit of being able to do it from anywhere at any time [17]. This results in improved responsiveness and operational efficiency on the part of the supplier [15,19].

Second, Web 2.0 delivers functionality as a service rather than packaged software [5]. The B2B concept is Software-as-a-Service or SaaS. This means instead of purchasing software to install on company computers, a supply chain member can access the software, and all of the relevant data, through the internet by use of a web browser (and the correct password) from any location at any time. Since the software is not purchased and installed, payment to the software supplier is commonly pay-per-use. Also because the software and data are not installed on local computers, updates and maintenance are transparent to end-users, allowing software suppliers to continuously update the software and users to have access to the latest version of the software [15].

A third, applicable concept for B2B contexts is that Web 2.0 is based on collective intelligence (a group of people can both contribute to and use the information gathered). One of the best examples of collective intelligence, although not B2B or even business-related, is Wikipedia, as it supports more than 1.5 million articles written completely by user contributions. Another form of this collective intelligence is blogs where users post their solutions to problems they have personally encountered, or respond to requests to solve the problems of others [5]. In a B2B environment, collective intelligence can be used to improve forecasts or understand demand patterns. Additionally, customers can exchange their views on the supplier's service [15–20] or share ideas on how to improve the service [18]. Several of the top Enterprise Resource Planning (ERP) software vendors support blogs where experienced users can discuss issues encountered in using the ERP applications as well as the solutions they arrived at to address the issues. As these systems attract more user participation, they will become even more valuable, unique, and richer [18]. And when users help each other, the supplier doesn't have the expense of providing quite as much customer support.

The fourth and final Web 2.0 concept is the support of mashups. A mashup is the result of the transformation that occurs when local data is integrated with one of several powerful web services such as Google Maps [15]. For example, it is possible to analyze local firm data about the location of supply chain members and/or shipments by combining it with the Google Maps' application to get a visual display of the routes of all shipments to and from a plant. Viewing data in new ways may enable supply chain efficiencies, such as sharing transport or minimizing deadhead miles, which can benefit the entire supply chain.

So how can businesses exploit these concepts in their B2B transactions? Many small businesses have hesitated to invest in interorganizational systems, such as electronic data interchange (EDI), in the past [21–23] due to resource constraints and failure to understand the strate-

gic benefits of these types of systems [24,25]. In addition, the negative repercussions of investing in these systems can be major [22]. Typically, small and medium businesses are more likely to adopt these systems because of external pressures such as competitive pressure, dependence on trading partner, trading partner pressure, or industry pressure [21,23]. However, Web 2.0 technologies provide low-cost options for smaller businesses. In fact, small businesses may be able to integrate Web 2.0 technologies before many of the larger businesses, because the large businesses are tied to their legacy systems [26]. This likely explains why Forrester Research Inc. predicts investments in Web 2.0 technologies will increase an estimated 43% per year over the next 5 years with a total of 4.6 billion investment dollars being spent on the technology by 2013 [27].

2.2 Business-to-Business

Parasuraman and Zinkhan [7] wrote in a special issue of the Journal of the Academy of Marketing Science that most of the scholarly research has been focused on B2C even though the internet affects both B2C and B2B and the "economic magnitude of B2B transactions are estimated to be substantially higher than that of B2C transactions" (p292). They conclude there is "clearly a need for more research in B2B contexts" (p293). LaPlaca and Katrichis [6] give two reasons for the imbalance of B2B and B2C articles in the marketing literature. The first reason is B2C data is more available than B2B data. The second is that consumer marketing is more popular than business marketing among students.

When compared to B2C relationships, developing effective B2B relationships through the use of technology has the potential to significantly benefit all members of a supply chain. Within a long supply chain (as illustrated in Figure 2), all but one of the transactions occur between businesses. The only transaction that involves the customer is the final transaction in the chain. In lean supply chains, businesses actively strive to establish long-term relationships with a small number of suppliers [28,29]. This makes it worthwhile and economically feasible for both business suppliers and business customers to invest the required money, time, and effort needed to automate their supply chain relationships [30]. In addition, the long history of EDI has demonstrated that many B2B transactions are fairly standard, making them amenable to automation [31].

2.3 Equivocality and Information Richness

Daft and Lengel [32] discussed why organizations gather and process information and conclude that "information is processed to accomplish internal tasks, to coordinate diverse activities, and to interpret the external environment" (p555). They further state that, in order to accomplish these goals, organizations must overcome both uncertainty and equivocality. They defined uncertainty as the absence of information. To reduce uncertainty, the organization gathers information that is assumed to exist. When enough information is obtained, a clear decision can be made.

Equivocality is similar to uncertainty, but it presumes ambiguity. This means even if a plethora of data is obtained, it could still result in multiple, conflicting interpretations and an unclear solution [32]. Daft and MacIntosh [33] used the term "unanalyzability" to describe a similar concept. In order to reduce equivocality, managers cannot just collect more data; instead, they must ponder the issue at length, use judgment coming from experience, and elicit discussions with those of both similar and opposing views. So the solution to equivocality is what Daft and Lengel [32] call *rich information*. They define information richness as "the ability of information to change understanding within a time interval. Communication transactions that can overcome different frames of reference or clarify ambiguous issues to change understanding in a timely manner are considered rich. Communications that require a long time to enable understanding or that cannot overcome different perspectives are lower in richness" (p560).

Daft, Lengel and Trevino [34] further the discussion of information richness and identify four characteristics of rich media, which they define as media which has the capacity to process rich information. Their first characteristic of rich media is that it provides *immediate feedback* to both the message sender and the message receiver so they can check interpretations without delay. The second characteristic of rich media is it provides *multiple cues* which the message receiver is able to infer information not explicitly stated in the message (e.g. tone of voice and body language). Their third characteristic of rich media is it is able to convey material in a variety of formats such as pictures, symbols and verbal statements. They call this characteristic *language variety*. Their fourth and final characteristic of rich media is that rich messages can be tailored with a *personal focus* so that it fits personal circumstances and can be infused with emotions and personal feelings. They then go on to order media from most rich (face-to-face) to least rich (impersonal, written documents). Daft *et al.* [34] also found that managers who could recognize the equivocality of a message and then chose appropriate media to convey the message (media high in richness to convey high-equivocality messages and media low in richness to convey low-equivocality messages), had higher performance ratings than managers who were unable to match the media to the equivocality of the message.

In this paper, we intend to illustrate that the stages in the customer service life cycle [9] vary in equivocality, and that the modes of customer contact [8] vary in media

richness, allowing us to recommend appropriate customer contact modes for different stages in the customer life cycle.

3. Research Model

In this section we present the details of our model first introduced as Figure 1. We will first discuss the mode of customer contact (rows) followed by the stages in the customer service life cycle (columns).

3.1 The Mode of Customer Contact Dimension

To address the different customer contact technology modes, we will utilize Froehle and Roth's [8] model of conceptual archetypes of customer contact in relation to technology, and describe how they vary in the B2B context. We will first provide an overview of this model and describe each customer contact mode. Then, we will discuss how each type of contact (full-, self-, or auto-service) can be addressed by different levels of media richness according to Media Richness Theory. We will present propositions for each type of contact.

Figure 3 illustrates the broad range of conceptual archetypes (or technology modes) of customer contact. They can range from traditional, technology-free services to completely automated e-services ("comprised of all interactive services that are delivered on the Internet using advanced telecommunications, information, and multimedia technologies" [35 p175]). According to Froehle and Roth [8], the first type of contact, technology-free customer contact (Figure 3A), occurs when the business customer only makes contact with the service representative and technology is not involved at all in the process. A second form of contact, technology-assisted customer

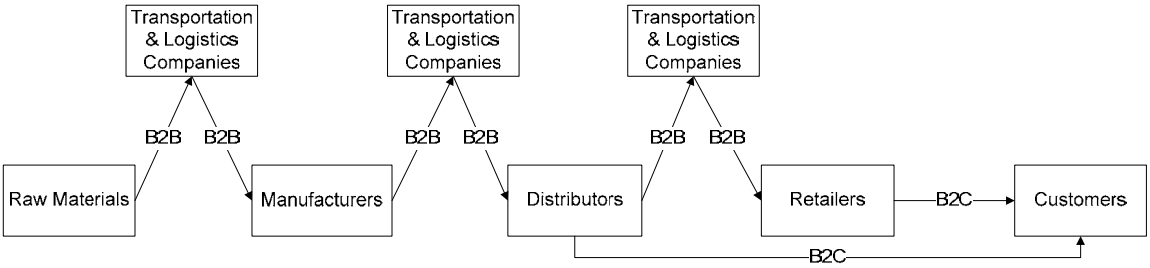

Figure 2. Example of a supply chain [Adapted from PowerPoint slides of 5]

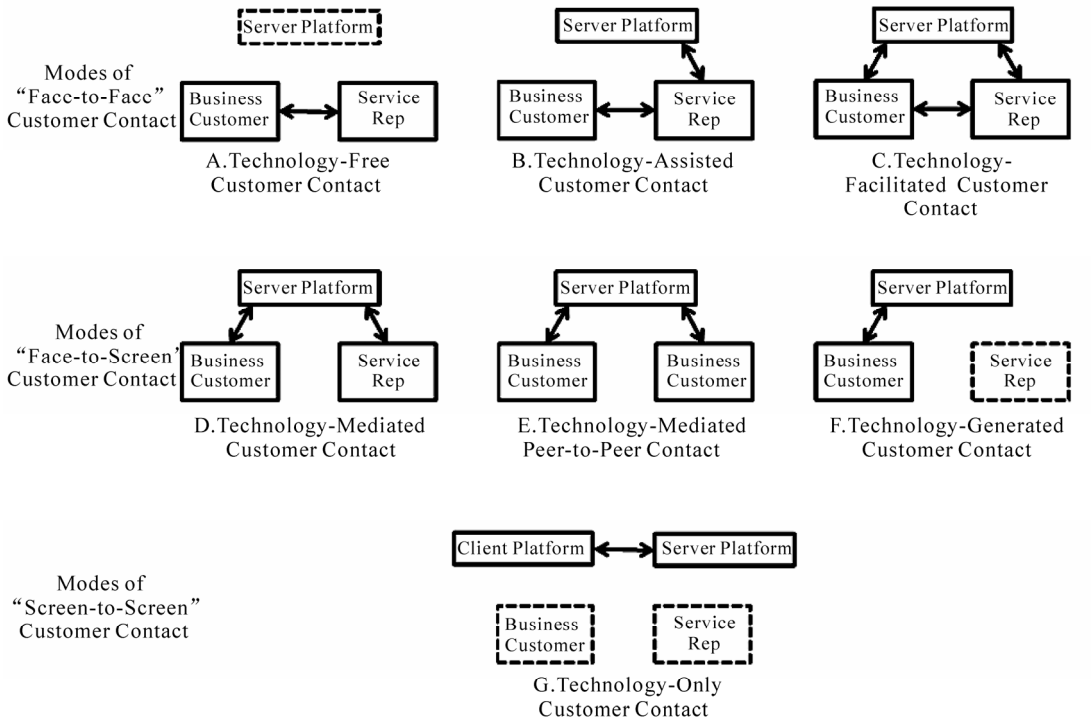

Figure 3. Froehle and Roth's [8] customer contact technology modes with B2B marketplace adaptations for S2S and P2P

contact (Figure 3B), occurs when the business customer makes contact with the service representative and the service representative makes use of technology (such as Web 2.0 technologies like wikis or blogs) in order to serve the business customer. These two types of customer contact are referred to as *full-service contacts*. A third form of contact, technology-facilitated customer contact (Figure 3C), occurs when a supplier makes the technology system available to both its service representative and its business customer. In addition to interacting with the system, the service representative and business customer also interact directly with each other. All three of these customer contact models (Figures 3A, 3B, and 3C) are considered "face-to-face" solutions to customer service.

"Face-to-screen" contact occurs, also according to Froehle and Roth [8], when the business customer interacts only with technology and not with the service representative. The first "face-to-screen" mode is technology-mediated customer contact (Figure 3D) where both the business customer and service representative interact with technology but they do not interact directly with each other. In technology-generated customer contact (Figure 3F), the business customer interacts directly with the technology without the assistance of a service representative; therefore, Froehle and Roth [8] considered this *self-service contact*.

One additional form of self-service interaction for business customers not mentioned by Froehle and Roth [8] is as a member of a virtual community where a buyer can interact with other buyers in lieu of a service representative (Figure 3E). An example of this is the "collective intelligence" Web 2.0 concept. In this case, business customers can interact directly with each other on a site provided by the supplier (e.g. http://blogs.oracle.com/) [36]. Businesses can install platforms to facilitate peer-to-peer interactions that can be used to solve and share solutions to mutual problems related to the company's service [17,18,20]. This interaction can be of two distinct types: viewing and posting. Customers can view (read) comments written by other customers or post (write) their own comments or questions that can be viewed by others. We call this technology-mediated peer-to-peer (P2P) contact.

These six technology-mediated customer contact modes can apply to both B2C and B2B marketplaces because a company's buyer can have the same type of relationship with a service supplier as an individual consumer can. However, B2B marketplaces, effectively, have an additional mode we call technology-only customer contact (Figure 3G)[1]. In this scenario, the technology systems at both the buyer's and supplier's companies interact with each other with no human involvement. For example, organizations have deployed inter-organizational B2B applications such as purchase order processing, invoice and payment processing, and procurement analysis [37]. Since this is not face-to-face or face-to-screen, so we call this *auto-service contact*.

[1]While it could be argued that personal, automatic payments could be an example of B2C technology-only customer contact, we argue this type of contact is actually B2B since the automatic payments are processed by either a bank or credit card company. Therefore, this type of interaction is ultimately B2B even though it seems B2C from the consumer's point of view.

3.1.1 Full-Service (Figures 3A and 3B)

There are three main benefits for the business customer in full-service customer contact modes. First, the buyer receives high human contact. This is important because, according to Vickery *et al.* [38], rich information "should be particularly relevant to customer contact in an industrial service environment" (p1108). The two media channels with the highest information richness include face-to-face and the telephone [32,34,38]. Only the full-service offers this type of human interaction because all the other modes involve the business customer interacting with a computer. A second benefit is most or all of the risk is the responsibility of the server. This assumption of risk can be critical for those customers who are inclined toward full-service [39]. Full-service gives these buyers the opportunity to avoid any risk in the performance of the service because the server is accountable for all of the activities after the intent to purchase is stated [39–41]. The final benefit of full-service contact is an avoidance of customer effort. Just like risk, the burden of activity resides with the server. This is particularly important for "non-standard" transactions [35] such as configuring specialized, industrial electric motors. Because the configuration can significantly influence the reliability of the motor [42], it requires a knowledgeable individual. If the buyer simply places the order, the burden of proper configuration is placed on the supplier [39].

Proposition 1: Business customers who need rich information, are risk-averse, and want to avoid extra effort will be more likely to opt for full-service contact with their suppliers, if given a choice.

3.1.2 Self-Service (Figure 3F)

In this type of service offering, the service representative does not actively participate in service performance. These types of interactions are called Self-Service Technologies (SSTs) [8]. SSTs are defined as the "technological interface that allows customers to produce and consume services without direct assistance from employees" [43 p283]. For example, employees at General Electric (GE) use a system called Trading Partner Network Register to electronically order office supplies from pre-approved vendors because it saves them money [44], wait times are shorter, no paperwork needs to be filled out, and authorizations are done electronically [41,45]. In addition to these benefits, SSTs offer an interactive, convenient, consistent, and personal way of communi-

cating with the seller [35,41,46]. While SSTs can provide immediate feedback and a personal focus, they do not offer the same number of cues (e.g. body language and tone are not conveyed) or language variety (e.g. verbal statements combined with pictures or graphs); therefore, they are considered less rich [34].

For the supplier, SSTs can also offer a number of benefits. First, they can reduce costs [47] by eliminating costly specialists and experts [35]. Second, they allow the organization to meet customer demand for alternative ways of purchasing services. Third, SSTs can allow a company to generate new markets by introducing a new channel that will reach worldwide markets [47].

It is important to note all of these benefits are not automatic or all-encompassing. First, the supplier must analyze its buyers' wants, needs, and capabilities to determine if self-service is the appropriate offering [48,49]. For example, the supplier needs to determine whether the customer willing to accept the lack of verbal communication inherent in less media rich channels [34]. Second, if SSTs are appropriate, the supplier has to try to customize the web interface to meet specific buyer needs. This typically results in a more complex website because of the level of personalization that is required. However, the website must still be simple, accessible, easy to use [35,41] and provide immediate feedback through pictures, on-screen displays, or text messages.

Proposition* 2: *Business customers who do not require verbal cues or statements* (*i.e. less rich information*) *will be more likely to opt for self-service contact with their suppliers, if given a choice.

3.1.3. Auto-service (Figure 3G)

Auto-service customer contact solutions are at the extreme of technology-mediated customer contact solutions. In this scenario, no human involvement is required from either supplier or customer for the business transactions to occur. In other words, the buyer and supplier work together to create a shared infrastructure utilizing the Internet and other technologies as a means for commercial exchange. These systems are referred to as interorganizational information systems [50,51]. A traditional example is Electronic Data Interchange (EDI) [49]. EDI is a standard protocol for electronically exchanging documents like purchase orders, invoices, and advanced ship notices. In the past, data was converted using expensive translation software and exchanged between corporations through a Value-Added Network (VAN) or a Virtual Private Network (VPN). Since it was so expensive and complicated, implementation of these types of technologies was usually restricted to large businesses [52]. If small businesses did choose to use interorganizational systems, they did so in response to competitive, trading partner, or industry pressure [21,23]. However, Web 2.0 technologies like enterprise application integration (EAI), Internet-based EDI, and point-to-point or system-to-system has made "auto-service" modes of customer contact available to even small and medium firms since these technologies are services rather than software packages that need to be purchased [15].

The main benefit of these auto-services is they allow firms to exchange information quickly [38] and more frequently because their systems are linked together [53]. Some other benefits of this technology include reducing transaction costs, reducing order cycles, improving partner relationships, improving the flow of data, and improving planning and forecasting [54].

Just like full- and self-service, auto-service does have disadvantages. For example, electronic media, like EDI, are considered "lean" or less rich [38,53]. Electronic media receive this label because the channel restricts visual communication [38]. Another disadvantage of auto-service is the substantial coordination costs necessary to connect the two systems [21]. However, Web 2.0 promises to reduce those costs because it can be used as a service (e.g. SaaS) or the software can be hosted either by the supplier or by a third-party host, so we propose:

Proposition* 3: *As Web* 2.0 *becomes more widely accepted, supplier companies will be able to provide richer information to their customers at a lower cost, so they are more likely to use Web* 2.0 *technologies than traditional interorganizational systems like EDI.

3.1.4 Mode Combination (Figures 3C and 3D)

As shown in Figure 3, there are middle-ground options available to sellers. First, the company can make their technology available to both the service representative and the service customer in a technology-facilitated customer contact. Second, the organization can support a technology-mediated customer contact environment where both the customer and representative mutually interact with the technology but not directly with each other [55]. These combined with full-service, self-service, and auto-service offer a wide variety of options that can be mixed-and-matched to best meet the customer's needs.

There are other ways that a company can utilize various modes of customer contact. One example of this combination is a company that provides full-service on the first transaction in order to demonstrate usage of the system while subsequent transactions are performed by the customer in self-service mode. A company could also split its high-volume, standardized services from its low-volume, customized services. The supplier could make its high-volume, standardized services available on their website while the customized services are handled by full-service customer representatives. Therefore, the business customer receives the benefit of both rich and lean media depending on the circumstance. In summary, the business customer can employ "media switching"

based on the task at hand [56].

Proposition* 4: *Business customers will use a variety of contact modes depending on the previous number of transactions or the type of transaction.

3.2 The Customer Service Life Cycle Dimension

To address the different customer life cycle dimensions, we will utilize Ives and Mason's [9] model of the customer service life cycle. As illustrated in Figure 4, there are four stages in this life cycle. The first stage of the customer life cycle is the *requirements* stage. This stage presents the supplier with a number of opportunities for serving the customer more effectively with technologies such as Web 2.0. For example, suppliers can develop websites specifically designed to provide customers with the information they need (e.g. inventory levels, tracking information). These websites can offer real-time information (i.e. immediate feedback). They can also specifically address services the business customer is interested in purchasing (e.g. personal focus) and present the information in multiple locations (e.g. a service could be viewed through the supplier's website or an offsite blog[2]) and multiple ways (e.g. language variety through demonstrations, trial packages, or written descriptions of the service). In other words, the supplier can use the Internet to provide a greater richness of information [8,57] that is unconfined by individual salespeople having to contact the customer during normal business hours [58,59].

In addition, the data can be more comprehensive, including both service details and specifications as well as potential application suggestions [58]. Since most organizations are utilizing their websites to provide more comprehensive information [13], the buyer can easily become overwhelmed by the sheer volume of options available to him and his company for purchasing services [60]. This is combined with the fact the buyer may already be addressing an ambiguous business need, particularly when it comes to the nature of services [61]. In order to deal with the complexity, uncertainty, equivocality, and information overload introduced by technology-mediated marketplaces [11,62,63], suppliers need to be proactive about providing service recommendations or substitution queries that help buyers filter unwanted services and alleviate some of the confusion and information-overload problems.

The second stage in Ives and Mason's [9] CSLC is the *acquisition* stage. In this stage, the buyer first searches for and identifies potential suppliers for the services identified in the requirements stage [64]. Through Web 2.0-supported peer-to-peer sites, the buyer can gather information from those who have provided comments about the service. Once a procurement source has been chosen, the buyer orders the service from the supplier [12]. The buyer then has the option to follow the progress of the order [13]. Once the order arrives and the buyer takes possession, he can transfer the appropriate funds [12,65]. Finally, the buyer can perform a service evaluation [13] to ensure the service meets expectations [12]. Unlike the first stage, this stage has a lower level of equivocality since there are a finite number of suppliers available.

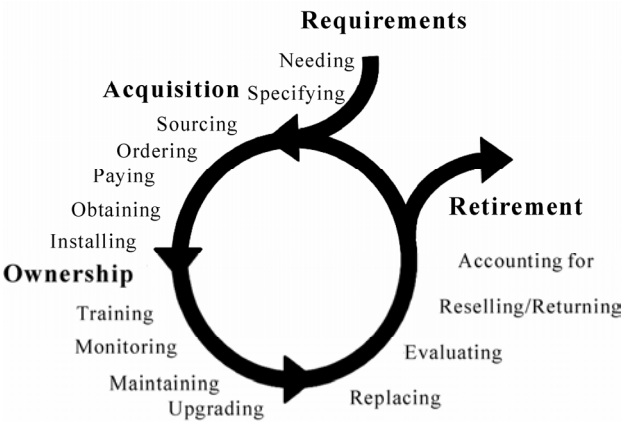

Figure 4. Ives and Mason's customer service life cycle [9]

The third stage in Ives and Mason's [9] CSLC is the *ownership* stage. This stage involves integration with other services, monitoring the access and use of the service, and upgrading to higher levels of service as needed to meet changing requirements [12,14]. As more companies refocus their efforts on core competencies, this stage can also include pay-for-usage models typical of outsourced activities [58]. This stage has an even lower level of ambiguity than the previous two stages because the company is likely going to have well-defined purposes for the purchased service. The only equivocality at this stage is the dyadic choice of upgrading to higher levels of service or not.

Finally, the buyer enters the fourth stage in Ives and Mason's [9] CSLC – *retirement*. This encompasses the analysis of expenses and the transfer or cancellation of the service [12]. At this stage, the buyer and supplier have the opportunity to establish a relationship instead of continuing to search for new partners that satisfy their service needs [66]. Business customers in this stage demonstrate the lowest level of equivocality because they only have to choose between continuing with the current supplier and moving to another supplier. If they choose to select a new supplier, they would move back into the requirements stage with its highest level of equivocality.

[2]SAP is a notable example of this. Not only does SAP provide an "SAP Community" for its business customers on its website (see http://sap.ittoolbox.com/), but they also have a separate blog site (http://sapro.blogspot.com/). Through these two sites, SAP customers can request weekly newsletters, e-mail alerts, and join groups. They have access to blogs, wikis, white papers, and other resources. In the blogs, SAP employees, SAP consultants, and SAP customers discuss their implementation challenges, the issues they experience using the SAP software, and how they overcome the problems they face with SAP.

While the model proposed by Ives and their model is B2C in nature, it can be modified to reflect the needs of B2B buyers. Unlike B2C marketplaces, B2B buying situations typically involve different decision-makers along the life cycle. This means the person who makes decisions in the requirements stage may be different from the person who actually purchases the service in the acquisition stage [67].

3.3 Mode of Customer Contact / CSLS matrix

With the two dimensions defined, the complete matrix shown in Figure 1 is considered. From the supplier's perspective, moving from the top to the bottom of the table reduces labor requirements. The top four rows (technology-free, -assisted, -facilitated, and -mediated) require service representative to interact with the business customer at some level. The bottom three rows, on the other hand, don't require interaction with any employee (peer-to-peer, technology-generated, and technology-only). From the business customer's perspective, the upper three rows of the table involve face-to-face interaction with a service representative, the next three rows are self-service in which the business customer is required to interact with the technology, and the final row does not require any interaction by either party.

By analyzing each stage in the customer service life cycle, a service-provider can identify opportunities for appropriate customer contact technology modes and avoid excessive IT spending that might result in unnecessary technology and potential buyer loss. Full-service customer contacts give customers human contact and risk-free service activities. Self-service technologies, on the other hand, offer customers control, convenience, and consistency across their service experience. By analyzing the customer's needs/wants and the service provider's technological ability, the company can offer the appropriate service for each circumstance.

We posit that levels of equivocality decrease when moving from left to right in the matrix. The most ambiguous phase is the requirements phase which has a higher level of equivocality than the acquisition phase which has higher levels than the ownership phase which has higher levels than the retirement phase. As the customer moves through the service process life cycle, from left to right in our matrix, from requirements, to acquisition, to ownership to retirement, the transactions become more routine and require fewer customer contact modes and less media richness and are more suitable for automating.

We also posit that media richness decreases when moving from top to bottom of the matrix. Technology-free, or full-service, is the richest mode of customer contact because the business customer and service represent interact directly via face-to-face or telephone conversations which are not constrained by any technology. Moving down through the rows of the matrix, contact with the service provider decreases and contact with technology increases. As the technology contact increases, this indicates a move toward less-rich media characterized by limited visual channels of communication [38]. The seventh row, auto-service, needs to be extremely structured so that the transaction can occur without intervention from the service representative.

As equivocality decreases right to left and media richness decreases top to bottom, we suggest the most appropriate matching of customer contact with stages of the customer life cycle will appear in a broad diagonal band going from the upper left to lower right. Because not every task in each stage of the customer life cycle has the same equivocality, it is possible to effectively use almost every mode of customer contact in every stage of the life cycle (see Table 2). For example, even if the two companies have an EDI-type relationship, it may be necessary to resort to technology-free face-to-face communication in order to solve a specific problem [53]. Additionally, it is possible a business customer in the requirements phase may want to peruse the comments in an online community discussion or download detailed specifications, resulting in effective off-diagonal positions. We now go on to discuss the diagonal positions and generate propositions for each stage.

3.3.1 Requirements

The requirements phase, the leftmost column in the matrix, is the time during which the new business customer moves from first realizing a need exists to specifying the attributes that will meet that need[3]. During this stage, the supplier company needs to educate potential business customers about the purpose of their service and help these business customers distinguish their service from the services of their competitors [68].

The decisions made in this stage have high levels of equivocality because gathering more data can easily result in more interpretations of the data, especially for services of which the buyer does not have experience. Because of the high levels of equivocality in this phase, a company should provide customers with as many contact modes as possible, including fact-to-face options, in order to make the communication with the potential customer as rich as possible. Therefore, we propose:

Proposition* 5: *In the CSLC requirements stage, customers will experience the highest levels of equivocality.

Proposition* 6: *In the CSLC requirements stage, rich media will be utilized to assist the customer to manage high levels of equivocality.

[3]Repeat customers move directly from the retirement phase to the acquisition phase without cycling through the requirements phase, so only new and inexperienced customers enter the customer service life cycle through the requirements stage.

Table 2. A matrix for linking CSLC opportunities and customer contact technology modes

		Customer-Service Life Cycle Stage			
		Requirements	**Acquisition**	**Ownership**	**Retirement**
Technology-Mediated Service	**Technology-Free**	· presales support and service specialists for high-demand customers [69] · understand the buyer's problem and seek solutions to it [58]	· "non-standard" transactions [35] · sales force to maintain public relations [10] · request for information (RFI) service - written information about the capabilities of various suppliers [64]		
	Technology-Assisted	· single-point sales contact - he utilizes technology to maintain the vast amount of service information [58]	· automate quote generation and order tracking for the sales force [78]		
	Technology-Facilitated	· marketing/advertising tool [13]	· customer service call center applications - routing, queue management [78] · transactional interaction (e.g. order entry, delivery, order tracking) [57,65,69,79]	· routine questions via electronic interaction; complex problems via live interactions [69]	
	Technology-Mediated	· contact sales staff through the website [13]	· purchase on-line, pick-up in person [59] · pay for orders online [13]	· management information and reporting [48] · online support [13]	
	Peer-to-Peer	· research/verify the experiences of existing customers [17]		· problem resolution through interactions with customers encountering similar problems [18]	
	Technology-Generated	· service catalogues [11,57, 65,71,79] · automated caller ID units and voice response units that suggest solutions to meet needs [58] · customer forums [58]	· service and pricing transparency real-time [59,88] · global marketing automation - segmentation, targeting, response tracking [59,78] · personalized purchasing website [69] · order tracking [13]		· receive automatic, real-time information about the supplier's inventory levels to make restock decisions or adjust forecasts [15]
	Technology-Only		· EDI · automated payment systems [71]	· automate current information flows [35,49,83]	

As illustrated in Table 2, research has shown that suppliers do employ a wide spectrum of customer contact technology modes to handle the complexity of the service being sought or the depth of information that must be gathered in order to suit many different buyer needs. For example, high-demand customers might need face-to-face presales support to handle all their specific service questions [69]. On the other hand, some buyers might have some generic questions that can be resolved by automated caller ID units that ascertain customer needs and then make suggestions without direct contact with a salesperson [58]. Another possibility is these issues could be resolved by the business customers themselves by accessing comments from an online users' community that can be used to determine attributes of actual service usage as well as company responsiveness issues [18]. Since different buyers will have different levels and timing of customer demand [70], a single technology-mediated service is not sufficient. Therefore, we propose the following:

Proposition* 7: *All modes of customer contact except auto-service will be commonly utilized in the requirements stage.

Proposition* 8: *Peer-to-peer contact will be dominated by viewing, with little or no posting.

3.3.2 Acquisition

Customers can enter the acquisition stage from the requirements phase or from the retirement phase of a previous order cycle[3]. Customers who are entering from the requirements phase experience a lower level of equivocality than they did in the first stage because now the supplier choice has been limited to those that offer the required services. On the other hand, customers entering from the retirement phase do not experience any equivocality because they've already done business with the supplier, so they know what to expect. Therefore, we propose:

Proposition* 9: *In the CSLC acquisition stage, equivocality will be lower than it was in the requirements phase.

Proposition* 9*a*: *In the CSCL acquisition stage, new customers will experience high levels of equivocality.

Proposition* 9*b*: *In the CSCL acquisition stage, customers retiring previous services with the same supplier will experience low levels of equivocality.

Just like the requirements stage, the acquisition stage presents a number of opportunities for the supplier to support the buyer through a mix of customer contact technology modes. First, this stage presents opportunities for automated payment systems [71]. For example, buyers and suppliers can exchange invoices via EDI or utilize automated clearinghouse (ACH) transactions. These systems allow the two parties to bundle transactions and avoid manual entry and reconciliation. Therefore, they can reduce or eliminate financial and human resources personnel for those processes [72]. Another opportunity is the supplier can make sure its website has clear and accurate pricing information [73,74]. This presents at least two benefits for both the buyer and the supplier. The first benefit is the buyer's uncertainty about the transaction process is reduced because the pricing information is readily available. This makes the buyer more likely to purchase from the supplier and it reduces the coordination costs of the transaction for both parties [75]. A second benefit is the price-sensitive buyer can accurately weigh the supplier against potential competitors [76]. Once the buyer chooses a supplier, he might be willing to establish a collaborative relationship with the supplier. This would mean the two parties could improve their efficiency as they synchronize their planning and scheduling activities [46,65]. Finally, the supplier can offer financial services or logistics arrangements in order to facilitate the order process [65,66]. Providing financial services, for instance, gives the buyer access to real-time payments or credit approvals [65] that can speed-up the order entry process for buyers. Since research indicates buyers are reducing their number of suppliers [77], any opportunity the supplier can find and meet will give that supplier an advantage.

As illustrated in Table 2, different customer contact technology modes provide a wide range of ways to fulfill these needs. For example, the supplier can empower its sales force by creating automatic systems that generate quotes and orders, provide contact management tools, and facilitate pipeline management [78]. Armed with this information, these salespeople can quickly and flexibly respond to the buyer just like the buyer wants [59]. For technology-facilitated services, the supplier can provide transactional interactions that allow the buyer to perform his own order entry or order tracking via the website [57,65,69,79]. This gives the buyer another valuable channel [80] for meeting his needs consistently and accurately [74]. In addition, interaction with an online users' community can assist with installation and early usage issues [18]. Because we can address the acquisition opportunities with a wide spectrum of customer contact technology modes, we propose:

Proposition* 10: *All modes of customer contact will be commonly utilized in the acquisition stage, but the use of self-service and auto-service will be higher than in the requirements stage.

Proposition* 11: *Peer-to-peer contact will be primarily viewing, with some posting.

3.3.3 Ownership

Typically, the supplier's role is to provide basic training [69] or education about the benefits of the service. Routine, unambiguous questions can be addressed electronically (e.g. FAQ sections on the supplier website)

while the more complex, equivocal problems can be handle via live interactions (e.g. demonstrations over the phone or through a salesperson visit) [69]. Customers can leverage online users' communities to solve unique problems and to find innovative uses for the service [18]. Since this stage doesn't require the interaction-intensity of the previous two stages [71], we propose the following:

Proposition* 12: *In the CSLC ownership stage, customers will experience lower levels of equivocality than in the acquisition stage.

Proposition* 13: *In the CSLC ownership stage, the use of self-service customer contact modes will increase and the use of full-service customer contact modes will decrease when compared to their use in the acquisition stage.

Proposition* 14: *Peer-to-peer contact will be both viewing and posting.

3.3.4 Retirement

Firms can establish their value-added partnership by automating their current, standardized/unambiguous, information flows through interorganizational systems like EDI [49]. This means firms can coordinate their simple processes (e.g. inventory management) [81,82] or automate their high-volume, repetitive restocks [35,49, 83,84]. As the services become more standardized, local knowledge and relationships become less crucial [85]. Therefore, the supplier and buyer can utilize different modes of technology services in order to reduce their coordination costs. In addition, as customers gain more experience with the service, they will become more active in participating in online users' communities because they will be able to offer suggestions of their own related to the service [15,20]. Therefore, we propose the following:

Proposition* 15: *In the CSLC retirement stage, customers will experience the lowest levels of equivocality.

Proposition* 16: *In the CSLC retirement stage, the use of self-service customer contact modes will increase and the use of full-service customer contact modes will decrease when compared to their use in the ownership stage.

Proposition* 17: *Peer-to-peer contact will be primarily posting, with some viewing.

4. Implications and Conclusions

In future research, scholars should explicitly examine the relationship between each CSLC opportunity and the mix of implemented customer contact technology modes. By using the customer as a unit of analysis, researchers can identify how different customers respond to different service offerings. In these studies, it will be important to control for different supplier sizes because they adopt B2B portals to different degrees. For example, smaller suppliers are less likely to create B2B websites because they are expensive and are likely to have fewer customers that will utilize these sites [86]. However, smaller suppliers are more likely to establish these types of sites in the future as Web 2.0 becomes more common. Larger suppliers, on the other hand, have access to more financial resources as well as in-house expertise so they find it easier to adopt B2B portals [52].

It would also be beneficial to study each technology-mediated service as a different variable. Future researchers could compare service experiences for each model to determine if one model is preferred over another in B2B marketplaces. Again, it is important to control for supplier size to determine if resource expenditures are the cause of higher acceptance ratings.

Future research needs to empirically examine the effectiveness of the modes of customer contact during the different CSLC stages. During each stage of the customer life cycle, the customer contact modes can be examined individually. In addition, the interactions between these modes should be examined to detect synergies obtained by using multiple modes of customer contact. Effectiveness can be measured in a number of ways including accurate information transmitted to the customer, the proportion of present customers that are retained and the percentage of new ones recruited, the percentage of customers that can be sold higher-levels of or complementary services, the number of "hits" on the online users' community as well as the number of ideas obtained from this platform and integrated into company operations.

In conclusion, the primary purpose of this article is to propose the fit between the CSLC and a mix of customer contact technology modes can improve service delivery. Due to the increasing use of the Internet as a source of communication and the evolution of customer expectations [77,87], this is becoming an increasingly important research topic. Therefore, this article proposes a theoretical lens for studying the buyer-supplier relationship in a B2B context.

REFERENCES

[1] Central Intelligence Agency, "GDP - Composition by sector," 10 August, 2009, Accessed https://www.cia.gov/library/publications/ the-world-factbook/fields/2012.html.

[2] J. A. Fitzsimmons and M. J. Fitzsimmons, "Service management: Operations, strategy," Information Technology, McGraw-Hill, Irwin, 2008.

[3] IBM, "Why SSME?" 8 October, 2009, Accessed http:// www.ibm.com/ developerworks/spaces/ssme.

[4] R. B. Chase and U. M. Apte, "A history of research in service operations: What's the big idea?" Journal of Operations Management, Vol. 25, pp. 375–386, 2007.

[5] R. K. Rainer and E. Turban, "Introduction to information systems: Supporting and transforming business," John Wiley and Sons, 2008.

[6] P. J. LaPlaca and J. M. Katrichis, "Relative presence of business-to-business research in the marketing literature," Journal of Business to Business Marketing, Vol. 16, pp. 1–22, 2009.

[7] A. Parasuraman and G. M. Zinkhan, "Web2.0 promises to enhance B2B activities," Journal of Academy of Marketing Science, Vol. 30, pp. 286–295, 2002.

[8] C. M. Froehle and A. V. Roth, "New measurement scales for evaluating perceptions of the technology-mediated customer service experience," Journal of Operations Management, Vol. 22, pp. 1–21, 2004.

[9] B. Ives and R. O. Mason, "Can information technology revitalize your customer service?" Academy of Management Executive, Vol. 4, pp. 52–69, 1990.

[10] R. B. Chase, "Where does the customer fit in a service operation?" Harvard Business Review, Vol. 56, pp. 137–142, 1978.

[11] T. Lee, J. Chun, J. Shim, and S. Lee, "An ontology-based product recommender system for B2B marketplaces," International Journal of Electronic Commerce, Vol. 11, pp. 125–155, 2006.

[12] B. Ives and T. Willinger, "To dell or be delled: A leading edge view of electronic commerce," University of Houston, Houston, 1999.

[13] B. H. Reber and S. Fosdick, "Building business relationships online: Relationship management in business-to-business e-commerce," Journal of Website Promotion, Vol. 1, pp. 13–29. 2005.

[14] J. Huang and C. Lin, "Empower internet services in hotel industry - A customer service life cycle concept," The Journal of American Academy of Business, Cambridge, Vol. 9, pp. 99–103, 2006.

[15] C. M. Goh , S. P. Lee, P. S. Tan, W. He, "Web 2.0 concepts and technologies for dynamic B2B integration," IEEE Conference on Emerging Technologies and Factory Automation, pp. 315–321, 2007.

[16] P. Anderson, "What is Web 2.0? ideas, technologies, and implications for education," 2007.

[17] J. Hammond, J. R. Rymer, R. Heffner, C. Schwaber, E. Driver, and J. Stone, "Developing enterprise Web 2.0 applications," Forrester Research, Inc., City, 2007.

[18] T. A. Yang, D. J. Kim, V. Dhalwani, and T. K. Vu, "The 8C framework as a reference model for collaborative value Webs in the context of Web 2.0," Proceedings of the 41st Hawaii International Conference on System Sciences, pp. 1–10, 2008.

[19] W. He, P. S. Tan, C. M. Goh, S. P. Lee, and E. W. Lee, "An investigative approach on improving B2B Interactions and communication capabilities for enterprise integration using Web 2.0 technologies," IEEE Conference on Emerging Technologies and Factory Automation, pp. 330–335, 2007.

[20] D. Iliyev, K. H. Choi, and K. J. Kim, "Dangers of applying Web 2.0 technologies in e-commerce solutions," International Conference on Information Science and Security, pp. 17–25, 2008.

[21] P. Chwelos, I. Benbasat, and A. S. Dexter, "Research report: Empirical test of an EDI adoption model," Information Systems Research, Vol. 12, pp. 304–321, 2001.

[22] M. Hughes, W. Golden, and P. Powell, "Inter-organisational ICT systems: The way to innovative practice for SMEs?" Journal of Small Business and Enterprise Development, Vol. 10, pp. 277–286, 2003.

[23] M. Quaddus, and G. Hofmeyer, "An investigation into the factors influencing the adoption of B2B trading exchanges in small businesses," European Journal of Information Systems, Vol. 16, pp. 202–215, 2007.

[24] W. H. DeLone, "Determinants of success for computer usage in small business," MIS Quarterly, Vol. 12, pp. 51–61, 1988.

[25] S. R. Magal, P. Kosalge, and N. M. Levenburg, "Using importance performance analysis to understand and guide e-business decision making in SMEs," Journal of Enterprise Information Management, Vol. 22, pp. 137–151, 2009.

[26] L. Rosencrance, "No smooth takeoff for US airways IT conversion. computerworld," Vol. 41, pp. 6, 2007.

[27] W. Kern, "Web 2.0 - end of accessibility? analysis of most common problems with Web 2.0 based applications regarding Web accessibility," International Journal of Public Information Systems, Vol. 2, pp. 131–154, 2008.

[28] W. Grey, T. Olavson, and D. Shi, "The role of e-marketplaces in relationship-based supply chains: A survey," IBM Systems Journal, Vol. 44, pp. 109–123, 2005.

[29] R. Shah, and P. T. Ward, "Defining and developing measures of lean production," Journal of Operations Management, Vol. 25, pp. 785–805, 2007.

[30] N. Thammakoranonta, A. Radhakrishnan, S. Davis, J. C. Peck, and J. L. Miller, "A protocol for the order commitment decision in a supply network," International Journal of Production Economics, Vol. 115, pp. 515–527, 2008.

[31] B. Massetti and R. W. Zmud, "Measuring the extent of EDI usage in complex organizations: Strategies and illustrative examples," MIS Quarterly, Vol. 20, pp. 331–345, 1996.

[32] R. Daft and R. Lengel, "Organizational information requirements, media richness and structural design," Management Science, Vol. 32, pp. 554–571, 1986.

[33] R. L. Daft and N. B. A Macintosh, "Tentative exploration into the amount and equivocality of information processing in organizational work units," Administrative Science Quarterly, Vol. 26, pp. 207–224, 1981.

[34] R. L. Daft, R. H. Lengel, and L. K. Trevino, "Message equivocality, media selection, and manager performance: Implications for information systems," MIS Quarterly, Vol. 11, pp. 354–366, 1987.

[35] K. K. Boyer, R. Hallowell, and A. V. Roth, "E-Services: Operating strategy - a case study and a method for analyzing operational benefits," Journal of Operations Management, Vol. 20, pp. 175–188, 2002.

[36] Oracle Welcome to Oracle Blogs, 2009, http://blogs. oracle.com/.

[37] T. S. H. Teo and C. Ranganathan, "Adopters and non-adopters of business-to-business electronic commerce in Singapore," Information and Management, Vol. 42, pp. 89–102, 2004.

[38] S. K. Vickery, C. Droge, T. P. Stank, T. J. Goldsby, and R. E. Markland, "The performance implications of media richness in a business-to-business service environment: Direct versus indirect effects," Management Science, Vol. 50, pp. 1106–1119, 2004.

[39] J. E. G. Bateson, "Self-Service consumer: An exploratory study. journal of retailing," Vol. 61, pp. 49–76, 1985.

[40] K. R. Jayasimha and R. Nargundkar, "Adoption of self service bill payment technologies (SSBPTS: A conceptual model," Journal of Services Research, Vol. 6, pp. 119–134, 2006.

[41] S. Globerson and M. J. Maggard, "A conceptual model of self-service," International Journal of Operations and Production Management, Vol. 11, pp. 33–43, 1991.

[42] D. Troyer, "Bearings for greased motors," MROToday, April/May, 2005.

[43] J. M. Curran and M. L. Meuter, "Encouraging existing customers to switch to self-service technologies: Put a little fun in their lives," Journal of Marketing Theory and Practice, Vol. 15, pp. 283–298, 2007.

[44] L. Cohn, D. Brady, and D. Welch, "B2B: The hottest net bet yet?" Business Week, Vol. 1, No. 17, pp. 36–37, 2000.

[45] D. Pujari, "Self-Service with a smile? self-service technology (SST. encounters among canadian business-to-business," International Journal of Service Industry Management, Vol. 15, pp. 200–219, 2004.

[46] L. M. Huete and A. V. Roth, "The industrialisation and span of retail banks' delivery systems," International Journal of Operations and Production Management, Vol. 8, pp. 46–66, 1988.

[47] M. J.Bitner, A. L. Ostrom, and M. L. Meuter, "Implementing successful self-service technologies," Academy of Management Executive, Vol. 16, pp. 96–108, 2002.

[48] J. L. Hopkins and D. F. Kehoe, "Testing an interrelationship matrix-based method designed for assisting in e-marketplace selection," Electronic Markets, Vol. 17, pp. 210–230, 2007.

[49] E. D. Rosenzweig and A. V. Roth, "B2B seller competence: Construct development and measurement using a supply chain strategy lens," Journal of Operations Management, Vol. 25, pp. 1311–1331, 2007.

[50] A. Zaheer and N. Venkatraman, "Determinants of electronic integration in the insurance industry: An empirical test," Management Science, Vol. 40, pp. 549–566, 1994.

[51] P. O'Reilly and P. Finnegan, "B2B marketplaces sharing is/it infrastructures: An exploration of strategic technology alliances," Journal of Enterprise Information Management, Vol. 20, pp. 304–318, 2007.

[52] V. Cho, "Factors in the adoption of third-party B2B portals in the textile industry," Journal of Computer Information Systems, Vol. 46, pp. 18–31, 2006.

[53] K. J. Petersen, G. L. Ragatz, and R. M. Monczka, "An examinatino of collaborative planning effectiveness and supply chain performance," The Journal of Supply Chain Management: A Global Review of Purchasing and Supply, Vol. 41, pp. 14–25, 2005.

[54] M. R. Lind and R. W. Zmud, "Improving interorganizational effectiveness through voice mail facilitation of peer-to-peer relationships," Organization Science, Vol. 6, pp. 445–461, 1995.

[55] A. V. Roth and L. J. Menor, "Insights into service operations management: A research agenda," Production and Operations Management, Vol. 12, pp. 145–164, 2003.

[56] A. R. Dennis and S. T. Kinney, "Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality," Information Systems Research, Vol. 9, pp. 256–274, 1998.

[57] K. Zhu and K. L. Kraemer, "e-Commerce metrics for net-enhanced organizations: Assessing the value of e-commerce to firm performance in the manufacturing sector," Information Systems Research, Vol. 13, pp. 275–295, 2002.

[58] J. N. Sheth and A. Sharma, "The impact of the product to service shift in industrial markets and the evolution of the sales organization," Industrial Marketing Management, Vol. 37, pp. 260–269, 2008.

[59] P. Oliveira, A. V. Roth, and W. Gilland, "Achieving competitive capabilities in e-services," Technological Forecasting and Social Change, Vol. 69, pp. 721–739, 2002.

[60] A. Barua, S. Ravindran, and A. B. Whinston, "Efficient selection of suppliers over the internet," Journal of Management Information Systems, Vol. 13, pp. 117–137, 1997.

[61] E. Ambrose, D. Marshall, B. Fynes, and D. Lynch, "Communication media selection in buyer-supplier relationships," International Journal of Operations and Production Management, Vol. 28, pp. 360–379, 2008.

[62] D. E. Leidner and T. Kayworth, "Review: A review of culture in information systems research: Toward a theory of information technology culture conflict," MIS Quarterly, Vol. 30, pp. 357–399, 2006.

[63] U. Schultze and W. J. Orlikowski, "A practice perspective on technology-mediated network relations: The use of internet-based self-serve technologies," Information Systems Research, Vol. 15, pp. 87–106, 2004.

[64] E. Bartezzaghi and S. Ronchi, "E-Sourcing in a buyer-operator-seller perspective: Benefits and criticalities," Production Planning and Control, Vol. 16, pp. 405–412, 2005.

[65] Q. Dai and R. J. Kauffman, "Business models for internet-based B2B electronic markets," International Journal of Electronic Commerce, Vol. 6, pp. 41–72, 2002.

[66] L. M. Ellram, W. L. Tate, and C. Billington, "Offshore outsourcing of professional services: A transaction cost economics perspective," Journal of Operations Management, Vol. 26, pp. 148–163, 2008.

[67] G. Chakraborty, V. Lala, and D. Warren, "What do cus-

tomers consider important in B2B Websites?" Journal of Advertising Research, Vol. 43, pp. 50–61, 2003.

[68] G. Piccoli, B. R. Spalding, and B. Ives, "A framework for improving customer service through information technology," Cornell Hotel and Restaurant Administration Quarterly, Vol. 42, pp. 39, 2001.

[69] T. Baumgartner, R. H. John, and T. Naucler, "Transforming sales and service," McKinsey Quarterly, pp. 80–91, 2005.

[70] J. A. Buzacott, "Service system structure," International Journal of Production Economics, Vol. 68, pp. 15–27, 2000.

[71] J. L. Hopkins and D. F. Kehoe, "The theory and development of a relationship matrix-based approach to evaluating e-marketplaces," Electronic Markets, Vol. 16, pp. 245–260, 2006.

[72] M. J. Cotteleer, C. A. Cotteleer, and A. Prochnow, "Cutting checks: Challenges and choices in B2B e-payments," Communications of the ACM, Vol. 50, pp. 56–61, 2007.

[73] T. Shifrin, "'Bad design' B2B Websites hits sales," Computer Weekly, June 13–14, 2006.

[74] M. Petre, S. Minocha, and D. Roberts, "Usability beyond the website: An empirically-grounded e-commerce evaluation instrument for the total customer experience," Behaviour and Information Technology, Vol. 25, pp. 189–203, 2006.

[75] T. W. Malone, J. Yates, and R. I. Benjamin, "Electronic markets and electronic hierarchies," Communications of the Association for Computing Machinery, Vol. 30, pp. 484–497, 1987.

[76] R. Brennan, L. Canning, and R. McDowell, "Price-setting in business-to-business markets," Marketing Review, Vol. 7, pp. 207–234, 2007.

[77] E. K. Clemons, S. P. Reddi, and M. C. Row, "The impact of information technology on the organization of economic activity: The move to the middle hypothesis," Journal of Management Information Systems, Vol. 10, pp. 9–35, 1993.

[78] A. V. Roth and C. H. Mason, "The right path to scm-crm integration," City, 2008.

[79] T. M. Laseter and S. E. Bodily, "Strategic indicators of B2B e-marketplace financial performance," Electronic Markets, Vol. 14, pp. 322–332, 2004.

[80] R. Gulati, and J. Garino, "Get the right mix of bricks and clicks," Harvard Business Review, Vol. 78, pp. 107–114, 2000.

[81] M. G. Jacobides, "Industry change through vertical disintegration: How and why markets emerged in mortgage banking," Academy of Management Journal, Vol. 48, pp. 465–498, 2005.

[82] N. Rackoff and C. Wiseman, "Information systems for competitive advantage: Implementation of a planning process," MIS Quarterly, Vol. 9, pp. 285–294, 1985.

[83] P. Walley and V. Amin, "Automation in a customer contact environment," International Journal of Operations and Production Management, Vol. 14, pp. 86–100, 1994.

[84] D. Brandon, "The future of automated forecasts and ordering," Chain Store Age, Vol. 83, pp. 134, 2007.

[85] D. Tan and J. T. Mahoney, "Why a multinational firm chooses expatriates: Integrating resource-based, agency and transaction cost perspectives," Journal of Management Studies, Vol. 43, pp. 457–484, 2006.

[86] M. Beckinsale, M. Levy, and P. Powell, "Exploring internet adoption drivers in SMEs," Electronic Markets, Vol. 16, pp. 361–370, 2006.

[87] S. Leek, P. Naude, and P. W. Turnbull, "Interactions, relationships and networks in a changing world," Industrial Marketing Management, Vol. 32, pp. 87–90, 2003.

[88] S. Bandyopadhyay, J. M. Barron, and A. R. Chaturvedi, "Competition among sellers in online exchanges," Information Systems Research, Vol. 16, pp. 47–60, 2005.

J. Service Science & Management, **2010, 3: 16-22**
doi:10.4236/jssm.2010.31002 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Heuristics for Production Allocation and Ordering Policies in Multi-Plants with Capacity

Jie Zang[1,2], Jiafu Tang[1]

[1]School of Information Science and Engineering, Northeastern University, Shenyang, China; [2]School of Information, Liaoning University, Shenyang, China.
Email: jiez00509@sina.com, jftang@mail.neu.edu.cn

Received October 19th, 2009; revised November 27th, 2009; accepted January 5th, 2010.

ABSTRACT

Joint decisions in production allocation and ordering policies for single and multiple products in a production-distribution network system consisting of multiple plants are discussed, production capacity constraints of multi-plants and unit production capacity for producing a product are considered. Based on the average total cost in unit time, the decisive model is established. It tries to determine the production cycle length, delivery frequency in a cycle from the warehouse to the retailer and the economic production allocation. The approach hinges on providing an optimized solution to the joint decision model through the heuristics methods. The heuristic algorithms are proposed to solve the single-product joint decision model and the multi-products decision problem. Simulations on different sizes of problems have shown that the heuristics is effective, and in general more effective than Quasi-Newton method (*QNM*).

Keywords*: *Joint Decisions, Production-Distribution; Multiple Plants, Capacitated, Heuristic Algorithm

1. Introduction

In the past, logistic decision among material procurement management, production and distribution were made in isolation. Previous studies have examined production, transportation and inventory separately. These major activities are closely related with each other and should be coordinated effectively to enhance its profit in today's competitive market. Uncoordinated and isolated decision-making among functional related activities in supply chain system may weaken its system-wide competitiveness. Hence, more efforts are now being made to integrate coordinate production and distribution, production and transportation, production and inventory, as well as transportation and inventory in the form of supply chain management.

King [1] described the implementation of a coordinated production-distribution system, a major tire manufacturer with four factories and nine major distribution centers. Williams [2] considered the problem of joint scheduling of production and distribution in a complex network, the objective of the problem was to minimize average production and distribution cost per period. Hill [3] discussed production-delivery policies in a single manufacturer and a single retailer. David [4] attempted to identify lot sizing and delivery scheduling in a single manufacturer and a single retailer system. Kim [5] discussed the production and ordering policies in a supply chain consisting of a single manufacturer and a single retailer. He proposes an efficient heuristic algorithm to determine the near optimal production allocation ratios. Kim [6] extended their paper and develops joint economic production allocation, lot-sizing, and shipment policies in a supply chain where a manufacturer produces multiple items in multiple production lines and ships the items to the respective retailers. Their formulations are often based on economic order quantity (EOQ) and mathe- matical programming. Accordingly, the corresponding solution methods are EOQ [7,8], heuristics [5,6,9] and decomposition [10,11].

In recent studies, model for coordinating production-distribution network systems have tended to focus on joint decisions on all activities. More complicated integrated decisions on production, transportation, and inventory have received relatively little attention, as in [12] and [13]. Tang [12] discussed an integrated decision on production assignment, lot-sizing, transportation, and order quantity for a multiple-supplier/multiple-destinations logistics network in a global manufacturing system and proposed a heuristics to solve medium and large-scale integrated decision problems. Yung [13] attempted

to tackle joint decisions in assigning production, lot-size, transportation, and order quantity for sing and multiple products in a production-distribution network system with multiple suppliers and multiple destinations. He provided an optimized solution to solve the joint decision model through a two-layer decomposition method that combines several heuristics.

This paper addresses the issue of how to effectively allocate production requirement to multiple plants in supply chain system. Kim [5,6] discussed the production and ordering policies in a supply chain consisting of a single manufacturer with multiple plants and a single retailer or multiple retailers. The retailers place orders based on the EOQ-like policy, and the multiple plants produce demand requirement from the retailers. Each of multiple plants has its production and transfer rates. In real life, all the plants in the manufacturer have production capacity constraints. All the plants should produce within its capacity to meet the demands of the retailers. The problem discussed in this paper extends the model proposed by [5,6], and production capacity constraints of multi-plants and unit production capacity for producing a product are considered in the model. The heuristics methods have been developed to solve the problem with single product and multiple products, respectively.

In this paper, the model for a single product will be discussed in Section 2, followed by detailed discussion to solve multiple products in Section 3. One illustrated example with several testing problems and their respective simulation results and analyses are presented in Section 4.

2. Formulations and Heuristics with Single Product

2.1 Problem Formulations

In a global manufacturing enterprise, there are plants each producing multiple parts and multiple assemblies that serve multiple assembly plants in a year, or alternatively, each assembly plant demands multiple parts from many different suppliers. Hence, such a global manufacturing enterprise can be formulated as a combined production-distribution network consisting of multiple suppliers and multiple destinations. In this paper, we consider a production-distribution network composed of a single manufacturer with multiple plants and multiple retailers. The retailers are given annual demand of the product. To meet the annual demands of the product, the manufacturer procures the materials and multi-plants produce within their capacity in the manufacturer. The multi-plants of the manufacturer have their production rate. The finished products are transferred to the common warehouse at the plants' transfer rate. Finally, the warehouse delivers the ordered lots of a fixed size to the retailer periodically. The network is shown in **Figure 1**. The cost components considered include two parts, the first part is the ordering cost from raw materials, the production setup cost, the ordering cost at the warehouse, and the ordering cost of the retailer; the second part is the holding costs for raw materials, work-in-process inventories, finished items at the warehouse and the retailer.

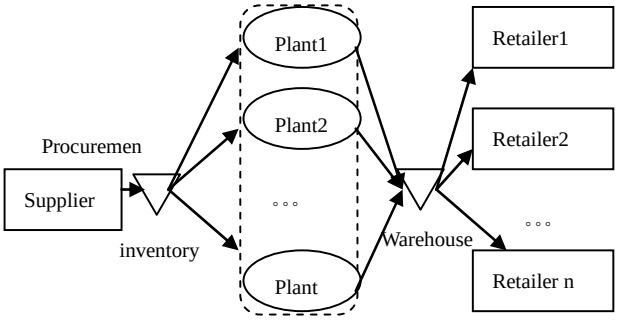

Figure 1. Production-distribution network

Assume that there are m plants in a manufacturer, where each of the plants is indicated by the subscripts j. The following notations and decision variables are applied.

P_j = annual production rate at plant j (unit/year)

Q_j = annual production capacity at plant j (year)

d_j = annual transfer rate from plant j to the warehouse (unit)

u_j = production capacity needed to produce unit product at plant j (year)

h_j = holding cost for work-in-processes at plant j ($)

S_p = production setup cost at the manufacturer ($)

A_m =ordering cost for raw materials at the manufacturer ($)

A_w = order handling cost for finished products at the warehouse ($)

A_r = ordering cost at the retailer ($)

H_m = holding cost for raw materials at the manufacturer ($)

H_w = holding cost for finished products at the warehouse ($)

H_r = holding cost for finished products at the retailer ($)

D =demand rate in units at the retailer (unit/year)

T = decision variable, production cycle length at the manufacturer (year)

m = decision variable, delivery frequency in a production cycle from the warehouse to the retailer

$\lambda = (\lambda_1, ..., \lambda_j)$ decision variable, production allocation for multiple plants

These notations will be extended in Section 3 to include multiple products. Accordingly, from the above parameters and decision variables, $\sum_j d_j \geq D$ and $P_j \geq d_j$ should be satisfied for the relevance of the pro-

posed model.

2.2 Joint Decision Model for a Single Product

The average cost components considered in this problem include two parts, the first part is the ordering cost; the second part is the holding costs these two parts of the costs are denoted by $F_1(T, m)$ and $F_2(m, \lambda)$ respectively. In a production cycle has m delivery from the warehouse to the retailer, so the ordering cost $F_1(T, m)$ are given as

$$F_1(T,m) = [(A_m + S_p) + m(A_w + A_r)]/T \tag{1}$$

For the second part of the costs, the average inventory levels for raw materials, work-in-process in plant (j), and finished products at the warehouse and the retailer over the production cycle are denoted by I_m, I_j, I_w and I_r, respectively. I_m and I_w can be derived by the appendix of Reference [5]. From the decision variables, we can derived the production lot size is DT, and the apportioned production lot size for plant i is $\lambda_j DT$. During a production cycle, the production time is $\lambda_i DT / P_i$, the delivery time is $\lambda_j DT / d_j$, as illustrated in **Figure 2**. It can be shown that, the average inventory for work-in-process I_j is

$$I_j = \frac{1}{T}[\frac{1}{2}(\lambda_j DT / d_j)\lambda_j DT - \frac{1}{2}(\lambda_j DT / P_j)\lambda_j DT]$$
$$= (D^2T/2)[(\lambda_j^2 / d_j)(1 - d_j / P_j)]$$

Hence, I_m, I_j, I_w and I_r [5] are given as

$$I_m = (D^2T/2)\sum_{i=1}^{n} \lambda_j^2 / P_j \tag{2}$$

$$I_j = (D^2T/2)[(\lambda_j^2 / d_j)(1 - d_j / P_j)] \tag{3}$$

$$I_w = (DT/2)(1 + 1/m) - (D^2T/2)\sum_{i=1}^{n} \lambda_j^2 / d_j \tag{4}$$

$$I_r = DT / 2m \tag{5}$$

Hence, the holding cost $F_2(m, \lambda)$ are given as

$$F_2(m,\lambda) = H_m I_m + \sum_{j=1}^{n} h_j I_j + H_w I_w + H_r I_r \tag{6}$$

Substituting (2)–(5) into (6), we can obtain

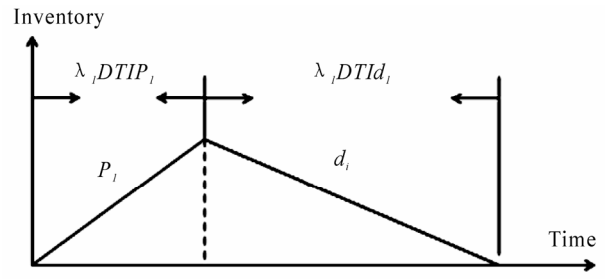

Figure 2. Inventroy trajectory for work-in-process in plant *j*

$$F_2(m,\lambda) = (DT/2)[H_w + (H_w + H_r)/m - D\sum_{j=1}^{n} H_j \lambda_j^2] \tag{7}$$

where $\mathrm{H}_j = \dfrac{h_j - H_m}{P_j} + \dfrac{H_w - h_j}{d_j}$

The integrated decisions of the economic production allocation and delivery policies are expressed as the following model:

$$\begin{aligned} Min\ W &= F_1(T, m) + F_2(m, \lambda) \\ &= [(A_m + S_p) + m(A_w + A_r)]/T \\ &+ (DT/2)[H_w + (H_w + H_r)/m - D\sum_{j=1}^{n} H_j \lambda_j^2] \end{aligned} \tag{8}$$

$$\text{s.t. } \sum_j \lambda_j = 1 \tag{9}$$

$$0 \le \lambda_j \le d_j / D \qquad \forall j = 1, 2, ..., n \tag{10}$$

$$0 \le \lambda_j \le Q_j / Du_j \qquad \forall j = 1, 2, ..., n \tag{11}$$

In this model, (8) is the objective of minimizing the average ordering and holding cost for raw materials, work-in-process, finished products at the warehouse and the retailer. The constraint (9) is the allocation vector for multiple plants. The constraints (10) and (11) should be satisfied by definition, respectively.

2.3 Heuristics Solution Procedures

The model is a fractional nonlinear programming model that is neither convex nor concave and is difficult to be solved. So we transform this model with the decision variables (T, m, λ) into a more simplified and equivalent problem with a decision variable λ, the last transformed problem is computed using a heuristic procedure.

First, the problem is strictly convex with respect to T, thus the optimal cycle length $T^*(m, \lambda)$ for a fixed pair of m and λ can be uniquely derived by solving $dW/dT=0$:

$$\mathrm{T}^* = \left\{ \frac{2[(A_m + S_p) + m(A_w + A_r)]}{[H_w + (H_w + H_r)/m - \sum_{j=1}^{n} (DH_j \lambda_j^2)]D} \right\}^{1/2} \tag{12}$$

Substituting T^* into (8), we can derive E(m, λ):

$$\begin{aligned} E(m,\lambda) &= W(T^*, m, \lambda) \\ &= \{2[(A_m + S_p) + m(A_w + A_r)] \\ &[H_w + (H_w + H_r)/m - \sum_{j=1}^{n} (DH_j \lambda_j^2)]D\}^{1/2} \end{aligned} \tag{13}$$

For (13), we can derive:

$$S(m/\lambda) = [(A_m + S_p) + m(A_w + A_r)]$$
$$[H_w + (H_w + H_r)/m - \sum_{j=1}^{n} (DH_j \lambda_j^2)]D \tag{14}$$

We can obtain (15) for fixed λ:

$$\frac{dS(m/\lambda)}{dm} = (A_w + A_r)[H_w - \sum_{i=1}^{n} DH_j\lambda_j^2] - \frac{(A_m + S_p)(H_w + H_r)}{m^2}$$
$$\frac{d^2S(m/\lambda)}{dm^2} = \frac{2(A_m + S_p)(H_w + H_r)}{m^3} \quad (15)$$

Since $d^2S(m/\lambda)/dm^2 > 0$, we can obtain m from $dS/dm=0$ and is given by

$$m(\lambda) = \{(A_m + S_p)(H_w + H_r)/(A_w + A_r)[H_w - \sum_{j=1}^{n}(DH_j\lambda_j^2)]\}^{1/2} \quad (16)$$

Since other terms in (17) are constant regardless of λ except $\sum_{j=1}^{n} DH_j\lambda_j^2$, we reformulate the next problem equivalently as follows:

$$MaxG(\lambda) = \sum_{j=1}^{n} H_j\lambda_j^2$$
s.t. (9),(10),(11)

This problem belongs to the class of quadratic maximization problems subject to linear constraints with a positive definite quadratic term. Reference [14] has proved it is an NP-hard problem. Since this problem aims to assign production allocation λ_j, a heuristic procedure is proposed as follows to solve it.

The heuristic algorithm steps

*Step*1. Resequence H_i in the descending order, such that $H_1 \geq H_2 \geq H_3 \geq \cdots \geq H_m$;

*Step*2. Let *t* be the current index number of the plant to be assigned, and $R_t = \sum_{i=1}^{t} \lambda_i$ be the total amount of the production allocation *t=0*, R_t *=0*;

*Step*3. *t=t+1* assignment to production to the *t*th plant point:

If $R_{t-1}<1$ set

$$\lambda_t = \min\{1 - R_{t-1}, d_t / D, Q_t / Du_t\}$$
$$R_t = R_{t-1} + \lambda_t$$

Else $\lambda_t = 0, R_t = R_{t-1} + \lambda_t$

End if

*Step*4. If *t*<*m*, go to *Step 3*; else, go to *Step*5;

*Step*5. Calculate the *MaxG*(λ), then stop.

After deriving λ *, we can obtain *m** and *T** from (16) and (12).

3. Joint Decisions for Multiple Products

3.1 Formulation with Multiple Products

In many real cases, the manufacture often produces multiple products to meet the need of the retailers. In this production-distribution network of multiple products, the main issue is how joint decisions can be made annually on production cycle length, delivery frequency and production allocation at a minimal average cost to the network. To derive the solution, the notations are defined as follows:

P_{ij} = annual production rate for product *i* at plant *j* (unit/year)

Q_{ij} = annual production capacity for product *i* at plant *j* (year)

d_{ij} = annual transfer rate for product *i* from plant *j* to the warehouse (unit)

u_{ij} = production capacity needed to produce unit product *i* at plant *j* (year)

h_{ij} = holding cost for product *i* at plant *j*($)

S_i = production setup cost for product *i* at the manufacturer ($)

A_i^R = ordering cost of raw materials for product *i*($)

A_i^W = order handling cost for finished product *i* at the warehouse ($)

A_i^C = ordering cost for product *i* at the retailer ($)

H_i^R = holding cost of raw materials for product *i*($)

H_i^W = holding cost for finished product *i* at the warehouse ($)

H_i^C = holding cost for finished product *i* at the retailer ($)

D_i = demand rate for product *i* (unit/year)

T = decision variable, production cycle length at the manufacturer (year)

m_i = decision variable, delivery frequency for product *i* in a production cycle from the warehouse to the retailer

λ_{ij} = decision variable, production allocation for product *i* in plant *j*

Similar to the average cost structure of a single product, the ordering costs and the holding costs are represented as follows, respectively:

$$F_1(T, m_i) = \sum_i [(A_i^R + S_i) + m_i(A_i^W + A_i^C)]/T \quad (18)$$

The second part is the holding costs for raw materials, work-in-process inventories, finished items at the warehouse and the retailer. They are denoted by I_i^R, I_{ij}, I_i^W, I_i^C respectively

$$I_i^R = (D_i^2T/2)\sum_{j=1}^{n} \lambda_{ij}^2 / P_{ij} \quad (19)$$

$$I_{ij} = (D_i^2T/2)\sum_j [(\lambda_{ij}^2 / d_{ij})(1 - d_{ij}/P_{ij})] \quad (20)$$

$$I_i^W = (D_iT/2)(1 + 1/m_i) - (D_i^2T/2)\sum_{j=1}^{n} \lambda_{ij}^2 / d_{ij} \quad (21)$$

$$I_i^C = D_iT/2m_i \quad (22)$$

Hence, the holding cost $F_2(m_i, \lambda_{ij})$ are given as

$$F_2(m_i,\lambda_{ij}) = \sum_i H_i^R I_i^R + \sum_{ij} h_{ij} I_{ij} + \sum_i H_i^W I_i^W + \sum_i H_i^C I_i^C \tag{23}$$

Substituting (19)–(22) into (23), we can obtain

$$F_2(m_i,\lambda_{ij}) = (T/2)\sum_i D_i[H_i^W + (H_i^W + H_i^C)/m_i - D_i\sum_j H_{ij}\lambda_{ij}^2] \tag{24}$$

$$H_{ij} = (h_{ij} - H_i^R)/P_{ij} + (H_i^W - h_{ij})/d_{ij} \tag{25}$$

The integrated decisions of the economic production allocation and delivery policies are expressed as the following model:

$$\begin{aligned}\min F &= F_1(T,m_i) + F_2(m_i,\lambda_{ij}) \\ &= \sum_i[(A_i^R + S_i) + m_i(A_i^W + A_i^C)]/T \\ &+ (T/2)\sum_i D_i[H_i^W + (H_i^W + H_i^C)/m_i - D_i\sum_j H_{ij}\lambda_{ij}^2]\end{aligned} \tag{26}$$

$$s.t \quad \sum_j \lambda_{ij} = 1 \qquad \forall i \tag{27}$$

$$0 \le \lambda_{ij} \le d_{ij}/D_i \quad \forall i, j \tag{28}$$

$$0 \le \lambda_{ij} \le Q_{ij}/D_i u_{ij} \quad \forall j \tag{29}$$

3.2 Heuristics Method for Multiple Products

The model is a fractional nonlinear programming model that is the same as the model with the single product. It can be solved by traditional nonlinear programming techniques, such as GINO, gradient search methods, where only the local optimal solution may be found. A heuristics is proposed to solve this problem.

First, the problem is strictly convex with respect to *T*, thus the optimal cycle length $T^*(m_i, \lambda_{ij})$ for a fixed pair of m_i and λ_{ij} can be uniquely derived by solving $dF/dT=0$:

$$T^*(m_i,\lambda_{ij}) = \left\{\frac{2\sum_i[(A_i^R + S_i) + m_i(A_i^W + A_i^C)]}{\sum_i D_i[H_i^W + (H_i^W + H_i^C)/m_i - D_i\sum_j H_{ij}\lambda_{ij}^2]}\right\}^{1/2} \tag{30}$$

Substituting *T** into (26), we can derive:

$$\begin{aligned}\min F' = \{&2[\sum_i((A_i^R + S_i) + m_i(A_i^W + A_i^C))] \\ &[\sum_i D_i(H_i^W + (H_i^W + H_i^C)/m_i - D_i\sum_j H_{ij}\lambda_{ij}^2)]\}^{1/2}\end{aligned} \tag{31}$$

For (31), we can derive:

$$\begin{aligned}S_E(m_i/\lambda_{ij}) = &\{\sum_i[(A_i^R + S_i) + m_i(A_i^W + A_i^C)]\} \\ &\{\sum_i D_i[H_i^W + (H_i^W + H_i^C)/m_i - D_i\sum_j H_{ij}\lambda_{ij}^2]\}\end{aligned} \tag{32}$$

We can obtain (33) for fixed λ_{ij} :

$$\begin{aligned}\frac{dS_E(m_i/\lambda_{ij})}{dm_i} = &(A_i^W + A_i^C)\sum_i D_i(H_i^W - D_i\sum_j H_{ij}\lambda_{ij}^2) \\ &-\frac{(H_i^W + H_i^C)D_i\sum_i(A_i^R + S_i)}{\sum_i m_i^2}\end{aligned} \tag{33}$$

$$\frac{d^2S_E(m_i/\lambda_{ij})}{d^2m_i} = \frac{2(H_i^W + H_i^C)D_i\sum_i(A_i^R + S_i)}{\sum_i m_i^3} \tag{34}$$

Since $d^2S_E(m_i/\lambda_{ij})/dm_i^2 > 0$, we can obtain *m* from $dS_E/dm_i=0$ and is given by

$$m^0{}_i(\lambda_{ij}) = \left[\frac{(H_i^W + H_i^C)D_i\sum_i(A_i^R + S_i)}{(A_i^W + A_i^C)\sum_i D_i(H_i^W - D_i\sum_j H_{ij}\lambda_{ij}^2)}\right]^{1/2} \tag{35}$$

Substituting (35) into (31), we get $F(\lambda_{ij})$:

$$\begin{aligned}F(\lambda_{ij}) = &\{2\sum_i(A_i^R + S_i)[\sum_i D_i(H_i^W - D_i\sum_j H_{ij}\lambda_{ij}^2)]\}^{1/2} \\ &+\sum_i[2(A_i^W + A_i^C)(H_i^W + H_i^C)D_i]^{1/2}\end{aligned} \tag{36}$$

Since other terms in (36) are constant regardless of λ_{ij} except $D_i\sum_j H_{ij}\lambda_{ij}^2$, we reformulate the next problem equivalently as follows:

$$Maximize \quad G_i(\lambda_i) = \sum_j H_{ij}\lambda_{ij}^2$$

s.t (27), (28), (29)

This problem belongs to the class of quadratic maximization problems subject to linear constraints with a positive definite quadratic term. Since this problem aims to assign production allocation λi , a heuristic procedure is proposed as follows to solve the model. The heuristics is

*Step*1. Resequence H_{ij} in the descending order for product i, such that $H_{i1} \ge H_{i2} \ge H_{i3} \ge \cdots \ge H_{in}$;

*Step*2. *i=i+1,* Let *t* be the current index number of the plant to be assigned, and

$$R_{it} = \sum_{j=1}^{t}\lambda_{ij}, t = 0, R_{it} = 0$$

*Step*3. *t=t+1* assignment to production to the *t*th plant point:

If R_{it}<1 set

$$\lambda_{it} = \min\{1 - R_{it}, d_{it}/D_i, Q_{it}/D_i u_{it}\}$$

$$R_{it} = R_{i,t-1} + \lambda_{it} \text{ go } Step\ 4$$

Else $\lambda_{it} = 0, R_{it} = R_{i,t-1} + \lambda_{it}$

End if

*Stetp*4. if *t*<*n*, go to *Step 3*; else, go to *Step5*;

*Step*5. Calculate $G_i(\lambda_i)$;

*Step*6. if *i*<*m*, go *Step 5*

*Step*7. Calculate $\sum_i D_i^2 \sum_j H_{ij}\lambda_{ij}^2$

*Step*8. Calculate m_i^*, if m_i^* isn't interger, then

$$m^* = \arg_m \min\{F'/m = (m_1,...,m_I),$$
$$m_i \in \{\lfloor m_i^0(\wedge^*)\rfloor, \lceil m_i^0(\wedge^*)\rceil\}\}$$

Else go *Step* 2

*Step*9. Calculate T*

*Step*10. Calculate F, stop

4. Simulations and Performance Analysis

To test the performance of the heuristics, some computation experimentations are conducted and their simulation results, as well as analysis, are presented in this section. The comparison of the heuristics and traditional quasi-Newton method are reported and analyzed. To shorten the length of the paper and without loss of generality, this section presents an example with multiple products to illustrate the application of the model and the heuristics. For simplicity, multiple plants at the manufacturer are denoted by 1, 2, 3, they can all produce three products A, B, C. The production rate, annual production capacity, Transfer rate and the holding cost are presented in **Table 1**. The production setup cost, the ordering cost and the holding cost at the manufacturer, the warehouse and the retailer and the demand rate are presented in **Table 2**.

Table 1. The parameters used in simulation tests

Plants	Production rate(unit)			Annual Capacity(hours)		
	A	B	C	A	B	C
1	6000	7400	5000	0.2	0.35	0.45
2	5700	11000	5800	0.2	0.35	0.45
3	8000	10000	5700	0.2	0.35	0.45
Plants	Transfer rate(unit/year)			Holding cost($/unit)		
	A	B	C	A	B	C
1	3400	2800	3100	6	5	2
2	3000	3200	3000	5	8	2
3	4100	5200	3100	7	8	6

Table 2. Basic data of the test

costs	A	B	C
Setup cost ($)	600	500	200
OR of raw material ($)	100	150	90
Shipping cost($)	25	20	20
OR at the retailer ($)	50	60	40
HC of raw material ($)	2	3	2
HC at warehouse ($)	8	8	8
HC at the retailer ($)	8	8	10
Demand rate(unit)	6000	7200	4300

From **Table 3**, the production cycle length, delivery frequency in a cycle from the warehouse to the retailer, the production allocation of the two solutions with heuristics and QNM are compared. A near-optimal solution with relative deviation of 0.928% is obtained from the best solutions by QNM with feasible initial solution. Hence, one can conclude that the heuristics method is more effective than QNM in the case of the above example. In particular, the results have pointed out the significance of assigning production among the plants. It reveals that business operations, including production and distribution among the plants, should be considered in an integrative manner so as to reduce costs and enhance the enterprise's competitiveness.

To illustrate the effectiveness of the heuristics, four randomly generated examples with 3*5, 5*5, 5*10, 10*10 (plants*products) are cited to make the comparison between the heuristics and the QNM. From **Table 4**, one can see that the heuristic algorithm is better than the QNM in these examples in terms of quality of the solution.

5. Conclusions

One of the core problems of supply chain management is the coordination of production and distribution. This paper considers joint decisions in production cycle length, delivery frequency and production allocation for a single

Table 3. Comparison of the heuristics and QNM

Delivery frequency			Production allocation		
	heuristics	QNM		heuristics	QNM
A	6	6	1-A	0.500	0.500
B	5	5	1-B	0.500	0.500
C	5	5	1-C	0.000	0.000
Production cycle	heuristics	0.209	2-A	0.389	0.301
	QNM	0.207	2-B	0.000	0.000
The total costs			2-C	0.611	0.699
heuristics		26705	3-A	0.721	0.721
QNM		26953	3-B	0.279	0.140
difference/%		0.928	3-C	0.000	0.139

Table 4. Comparison of the heuristics and QNM for different size of examples

Problem size	Costs		
	heuristic	QNM	Diff (e/%)
3*5	26539	26916	1.42
5*5	45339	46413	2.37
5*10	45209	46520	2.89
10*10	82336	83223	1.07

product and for multiple products in a production- distribution network system with multiple plants and multiple retailers. All plants are all capacitated. Based on the production capacity and the unit production capacity for producing a product, the mathematical programming model is presented to distribute the demand of the retailer to multi-plants to achieve an objective of minimizing the average costs. Two effective heuristic methods are developed to solve the joint decision problem with single product and multiple products. The simulation results have shown that the heuristics is easily implemented and effective for the decision problems.

Future work includes: the economic allocation of the complex product in multiple plants.

6. Acknowledgment

The paper is financially supported by the Natural Science Foundation of China (NSFC 70625001, 70721001), National 973 program (2009CB320601) of China, and 111 project of Ministry of Education (MOE) in China with number B08015.

REFERENCES

[1] R. H. King and R. R. Love, "Coordinating decisions for increased profits," Interfaces, Vol. 10, pp. 4–19, 1980.

[2] J. F. Williams, "Heuristic techniques for simultaneous scheduling of production and distribution multi-echelons structures: Theory and empirical comparisons," Management Science, Vol. 27, pp. 336–352, 1981.

[3] R. M. Hill, "The single-manufacturer single-retailer integrated production-inventory model with a generalized policy," European Journal of Operational Research, Vol. 97, pp. 493–499, 1997.

[4] I. David and M. Eben-Chaime, "How far should JIT vendor– buyer relationships go?" International Journal of Production Economic, Vol. 81, pp. 361–368, 2003.

[5] T. Kim, Y. Hong, and J. Lee, "Joint economic production allocation and ordering policies in a supply chain consisting of multiple plants and a single retailer," International Journal of Production Research, Vol. 43, pp. 3619–3632, 2005.

[6] T. Kim and Y. Hong, "Technical note: Production allocation, lot-sizing, and shipment policies for multiple items in multiple production lines," International Journal of Production Research, Vol. 46, pp. 289–294, 2008.

[7] R. W. Hall, "On the integration of production and distribution: Economic order and production quantity implications," Transportation Research, Part B, Vol. 30, pp. 387–403, 1996.

[8] J. Hahm and C. A. Yano, "The economic lot and delivery scheduling problem: The single item case," International Journal Production Econom, Vol. 28, pp. 238–252, 1992.

[9] J. Benjamin, "An analysis of inventory and transportation costs in a constrained network," Transportation Science, Vol. 23, pp. 177–183, 1989.

[10] D. E. Blumenfeld, L. D. Burns, and C. F. Daganzo, "Synchronizing production and transportation schedules," Transportation Research, Vol. 25, pp. 23–37, 1991.

[11] F. Fumero and C. Vercellis, "Synchronized development of production, inventory and distribution schedule," Transportation Science, Vol. 33, pp. 330–340, 1999.

[12] J. F. Tang, K. L. Yung, and A. W. H. Ip., "Heuristics based integrated decisions for logistics network systems," Journal Manufacturing Systems, Vol. 23 , pp. 1–13, 2004.

[13] K. L. Yung, J. F. Tang, A. W. H. Ip., and D. W. Wang, "Heuristics for joint decisions in production, transportation, and order quantity," Transportation Science, Vol. 40, pp. 99–116, 2006.

[14] S. Sahni, "Computationally related problems," SIAM Journal on Computing, Vol. 3, pp. 262–279, 1974.

J. Service Science & Management, **2010, 3: 23-32**
doi:10.4236/jssm.2010.31003 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Scientific Research

Risk Averse Members Coordination with Extended Buy-Back Contract

Zheng Qin, Xiaofeng Xue*

School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai, China.
Email: qinzheng@mail.shufe.edu.cn, *xf_xue@126.com

Received December 20th, 2009; revised January 31st, 2010; accepted February 23rd, 2010.

ABSTRACT

This paper considers how to coordinate a supply chain (*SC*) *consisted of one supplier and one retailer who possess different risk aversion preference with a contract. Based on the classical buy-back contract, this paper presents an extended buy-back contract. In addition to the member's objective of maximizing his expected profit, downside risk constraint is used to represent the SC member's risk aversion preference. Under different risk aversion preference combination, the SC perfect solution existence conditions are identified and the specific contract is provided accordingly. This research finds out that, with a low risk aversion supplier and a high risk aversion retailer, the supplier as the SC coordinator can give the retailer incentive to increase the order quantity so as to reach SC perfect coordination. Finally a numerical analysis verifies the effectiveness of the extended buy-back contract.*

Keywords*: *Downside Risk, Buy-Back Contract, Supply Chain Coordination, Risk Aversion Preference

1. Introduction

Supply chain (SC) performance is affected by many uncertainty factors. In the supply chain contract (SCC) literatures two major uncertainty factors receives most focus. These factors include market demand uncertainty and manufacture's production capacity uncertainty. These uncertainty factors make direct impacts on the SC performance and incur SC members' profit risks. The SC members' profit risks are different from the SC members' risk preferences. The former are objective while the latter are subjective. Many SCC literatures consider the objective profit risks. In these literatures SCCs are designed to share profit risks through transfer payment between SC members. The SC members are assumed to be risk neutral, though they may have different risk preferences in reality. The risk preferences have great impacts on the members' decisions. Therefore some papers that consider the members' risk preferences are presented in recent years.

Eeckhoudt *et al*. research the effect of risk aversion in the single period newsboy problem. They examine what changes in price and cost parameters relate to the member's risk aversion [1]. Agrawal and Seshadri study the impacts of demand uncertainty and retailer's risk aversion on the retailer's pricing and ordering decision. It is proved that with different models that depict the relationship between price and demand distribution, the risk aversion retailer will price higher or price lower than the risk neutral retailer [2]. Lau and Lau model a two echelon SC consisted of different risk attitude members in a single period. They use mean–variance utility function as the member's objective. It is found that the optimal return policy may change from "no return allowed" to "unlimited return with full credit" based on the member's risk attitude [3]. Gan *et al*. consider the channel coordination issue that the SC is consisted of a risk–neutral supplier and a downside-risk retailer. They put forward a downside risk formulation and design a risk-sharing contract that can achieve channel coordination [4]. Choi *et al*. study the channel coordination and risk control issues of a SC. They make comparison of the SC expected performance between the model with risk control and one without risk control. A mean-variance framework is used in [5]. Gan *et al*. consider a two echelon SC consisted of risk aversion members. They use mean-variance trade-off and expected utility as the member's objective. The set of Pareto optimal solutions are identified [6]. Wei and Choi use the mean-variance framework to analyze the supply chain coordination issue with wholesale pricing and profit sharing scheme (WPPS). In the decentralized case, they find that there exists a unique equilibrium of the Stackelberg game with WPPS [7]. Mastui research a SC consisted of one manufacture and one retailer. The two players are risk averse. He presents a new contract that is a trade-off between outright sales contract and

full-credit return policy and investigates the economic outcome differences [8]. Xiao and Choi use dynamic game theory to analyze a more complicated structure SC consisted of two manufactures and two retailers who are all risk averse. Their research focuses on the impacts of the retailer's risk sensitivity and other properties on the channel structure strategies and wholesale prices of the manufactures [9].

The above mentioned papers all involve risk aversion SC members. Only [4–7] consider the issue of SC coordination. The impacts of the risk aversion of the retailer on the SC coordination contract are considered in [4]. The impacts of the risk aversion of both the retailer and the supplier are researched and the existence conditions of Pareto optimal solution are identified in [6]. The numerical analysis of the channel coordination contract properties is presented for a SC consisted of different risk aversion members in [5]. SC coordination contract WPPS in [7] is different from the contract researched in this paper.

Unlike the previous literatures, this paper is devoted to coordination of a supply chain consisting of a supplier and a retailer with different risk aversion preference. The SC member wishes to maximize his expected profit and has to satisfy the risk aversion constraint. The major purposes of this paper include: a) Identify the conditions that may achieve SC perfect coordination; b) Design SCC for the decentralized SC to achieve perfect coordination; c) Analysis the impacts of member's risk aversion on the SC efficiency.

The remainder of this paper is organized as follows. In Section 2, we describe the basic models of the problems facing with the supplier, the retailer and the supply chain as a whole. In Section 3, we give the detail of supply chain contract that can achieve perfect coordination under SC members' different risk aversion preference combination. In Section 4, we provide a numerical analysis to verify the contract. Finally in Section 5, we summarize the main findings of this paper and possible future research of this study.

2. Assumptions, Notations and Models

2.1 Assumptions, Notations

As stated earlier, we consider a SC consisted of one supplier (she) and one retailer (he). The SC members have different risk aversion preferences. The supplier provides the retailer with a kind of short life cycle products. There is only one selling season. Prior to the selling season, the retailer should decide the order quantity and submit it to the supplier. The supply capacity of the supplier is unlimited. The supplier will ship the ordered products before the selling season. Market demand is stochastic and with a specific known distribution. Both the supplier and the retailer have the complete knowledge about the demand, profits, costs, risk-aversion preferences and prices. We analyze this problem as a Stackelberg game in which the supplier acts as the leader while the retailer acts as the follower. The supplier provides the contract. Then the retailer decides order quantity that maximizes his expected profit as well as satisfies his risk aversion constraint according to the contract. The supplier as the leader may give the retailer incentive to increase his order quantity.

Notations used in this paper are defined as follows. Let q denote the retailer's order quantity and q_i^* denote the optimal order/supply quantity. The subscript i might be r(retailer), s(supplier), sc(supply chain). The following notations' subscripts are defined in the same way. Let X denote the stochastic market demand with distribution F(·) and density f(·). Let c denote the supplier's unit supplying cost; ω denote the unit wholesale price charging the retailer; b denote the unit buy back price paying the retailer for all returned products at the end of the selling season; p denote the unit retail price selling to the end customers. Let $\Pi_i(\cdot)$ be the members' profit. Let $\pi_i(\cdot)$ denote the members' expected profit, therefore $\pi_i(\cdot) = E[\Pi_i(\cdot)]$. Let α_i be the target profit; β_i ($0 \le \beta_i \le 1$) be the maximum risk level that the member will accept. We define γ_i ($0 \le \gamma_i \le 1$) as the member's downside risk. We assume there is a relationship among the parameters of c, b, ω, p that is $0 < c \prec b \prec \omega \prec p$.

2.2 Definition of the Member Risk Aversion Preferences

We use downside risk to measure SC member's risk level. The downside risk is used in [4,10]. Downside risk of the member is the probability that his realized profit is less than or equal to his specified target profit. The member's risk aversion preference is defined as the downside risk must be less than or equal to his or her maximum risk level.

The member's downside risk definition formula is shown as below:

$$\gamma_i \equiv P\{\Pi_i \le \alpha_i\} \tag{1}$$

The member's risk aversion preference definition formula is shown as below:

$$\gamma_i \le \beta_i \tag{2}$$

2.3 Modeling on Supply Chain Consisted of Members with Risk Aversion Preferences

Gan *et al*. provide the definition of the supply chain coordination of one risk neutral supplier and one downside risk aversion retailer [4]. We reference their definition and make adjustment in line with our conditions.

Definition 1. The supply chain is perfectly coordinated if the following conditions are satisfied:

1) the retailer and the supplier get payoffs not less than their respective reservation payoffs,

2) both the retailer's and the supplier's downside risk constraint are satisfied,

3) the supply chain's expected profit is maximized.

The model of the supply chain consisted of risk aversion members is shown as below.

For the supplier:

$$\max_{q\geq 0} \pi_s(q) \quad s.t.\ \ \gamma_s \leq \beta_s \tag{3}$$

For the retailer:

$$\max_{q\geq 0} \pi_r(q) \quad s.t.\ \ \gamma_r \leq \beta_r \tag{4}$$

For the supply chain:

$$\max_{q\geq 0} \pi_{sc}(q) \tag{5}$$

2.4 The Benchmark Model

We select the model of the buy-back contract (also called return policy) with risk neutral members as the benchmark. To differentiate the benchmark contract from the contract presented in this paper, we call the former initial buy-back contract while the latter extended buy-back contract. Pasternack has proved that in the risk neutral case, buy-back contract will coordinate the supply chain if the wholesale price ω and buy-back price b satisfy the Formula (6). The retailer's optimal order quantity to maximize the supply chain profit is shown as the Formula (7) [11].

$$\frac{p-b}{p} = \frac{w-b}{c} \tag{6}$$

$$\hat{q} = F^{-1}\left(\frac{p-c}{p}\right) \tag{7}$$

2.5 The relationship between order/supply quantity and member's downside risk under initial buy-back contract.

SC member's minimum supply or order quantity equals to the least amount of products to be sold so as to reach her or his target profit α_i.

Lemma 1. Assumed the member's target profit is $\alpha_i\ (i:s,r)$, the supplier' minimum supply quantity q_s^0 and the retailer's minimum order quantity q_r^0 are shown as the formulas below.

$$q_s^0 = \frac{\alpha_s}{\omega - c} \tag{8}$$

$$q_r^0 = \frac{\alpha_r}{p-\omega} \tag{9}$$

Proof: see the appendix A.1.

Theorem 1. If the supply quantity q_s is less than or equal to the minimum supply quantity q_s^0 then the supplier's downside risk γ_s equals to 1; if q_s is larger than q_s^0 then γ_s equals to $F\left(\frac{\alpha_s-(\omega-c-b)q_s}{b}\right)$, where $\alpha_s \geq (\omega-c-b)q_s$.

Further description about Theorem 1: Let $\Delta\omega = \omega - c$, $\Delta\omega$ approximately equals to the net profit of the supplier. b is unit buy-back price. $\Delta\omega$ and b represent the relative bargaining power between the supplier and the retailer. We have assumed that the supplier is the SC leader and the retailer is the follower. Therefore it is reasonable to assume the supplier is dominant and thus $\Delta\omega \succ b$. $\Delta\omega \leq b$ contradicts our assumptions. We only consider the case of $\Delta\omega \succ b$ in the rest of this paper.

Proof: see the appendix A.2.

Theorem 2. If the order quantity q_r is less than or equal to the minimum order quantity q_r^0 then the retailer's downside risk γ_r equals to 1; if q_r is larger than q_r^0 then γ_r equals to $F\left(\frac{\alpha_r+(\omega-b)q_r}{p-b}\right)$.

Proof: see the appendix A.3.

2.6 Threshold of the Member's Downside Risk

Definition 2. The threshold of the SC member's downside risk is defined as the maximum value of the member's downside risk when $q_i = \hat{q}$ and

$$\alpha_i = \pi_i(\hat{q}) \quad i:s,r.$$

Theorem 3. The downside risk threshold of the supplier equals to that of the retailer and the value equals to $F(\hat{q} - \int_0^{\hat{q}} F(x)dx)$.

For notation convenient we represent this value as BN. That is $BN = F(\hat{q} - \int_0^{\hat{q}} F(x)dx)$.

Proof: see the appendix A.4.

2.7 The Relationship between the Member Risk Aversion Preference and the Order/Supply Quantity

Corollary 1 describes the relationship between the supplier's risk aversion preference and her optimal supply quantity. It can be inferred from Theorem 1 and Theorem 3.

Corollary 1. If $\beta_s \prec F\left[\frac{\pi_s(\hat{q})-(\omega-c-b)q_s^0}{b}\right]$ then the

supplier's optimal supply quantity that maximize her expected profit while satisfy her risk aversion constraint is defined by Formula (10).

$$q_s^* = \begin{cases} \hat{q} & BN \le \beta_s \prec F\left[\dfrac{\pi_s(\hat{q}) - (\omega - c - b)q_s^0}{b}\right] \\ \dfrac{\pi_s(\hat{q}) - bF^{-1}(\beta_s)}{\omega - c - b} & \beta_s \prec BN \end{cases} \quad (10)$$

Similar to Corollary 1, Corollary 2 describes the relationship between the retailer's risk aversion preference and his optimal order quantity.

Corollary 2. If $\beta_r \succ F\left[\dfrac{\pi_r(\hat{q}) - (\omega - b)q_r^0}{p - b}\right]$ then the retailer's optimal order quantity that maximize his expected profit while satisfy his risk aversion constraint is defined by Formula (11).

$$q_r^* = \begin{cases} \hat{q} & \beta_r \ge BN \\ \dfrac{(p-b)F^{-1}(\beta_r) - \pi_r(\hat{q})}{\omega - b} & F\left[\dfrac{\pi_r(\hat{q}) - (\omega - b)q_r^0}{p - b}\right] \prec \beta_r \prec BN \end{cases} \quad (11)$$

3. Supply Chain Contracts with Different Risk Aversion Preference Members

In this section we will give specific supply chain contracts for three scenarios that have different risk aversion preferences combination. The first scenario depicts that both the supplier and retailer show low risk aversion preferences. The second scenario has a low risk aversion supplier and a high risk aversion retailer. In the third scenario the supplier is high risk aversion while the retailer's risk aversion might be high or low.

3.1 Both the Supplier and the Retailer are Low Risk Aversion ($\beta_s \ge BN$ and $\beta_r \ge BN$)

Step 1. Compute the optimal order/supply quantity of the member with risk aversion preference.

It can be inferred from Corollary 1 and Corollary 2 that both the retailer's optimal order quantity q_r^* and the supplier's optimal supply quantity q_s^* equal to the supply chain optimal order quantity $\hat{q}$ that can achieve the maximum expected profit of the supply chain.

Step 2. Design supply chain coordination contract.

It is obviously that the SC member's risk aversion preference constraint doesn't impact on the optimal solution when her (his) risk aversion preference is low. The supply chain contract in this scenario is the same as that with risk neutral SC members. The supply chain contract can be buy-back contract, revenue sharing contract or other contracts. For the purpose of comparison we select buy- back contract here.

The initial buy-back contract that the wholesale price ω and buy-back price b satisfy the Formula (6) will coordinate the supply chain with low risk aversion preference SC members. In that the supply chain is perfectly coordinated.

3.2 The Supplier is Low Risk Aversion while the Retailer is High Risk Aversion ($\beta_s \ge BN$ and $\beta_r \prec BN$)

Step 1. Compute the optimal order/supply quantity of the SC member with risk aversion preferences.

It can be inferred from Corollary 1 and Corollary 2 that the retailer's optimal order quantity q_r^* is less than the supply chain optimal order quantity $\hat{q}$ while the supplier's optimal supply quantity q_s^* equal to $\hat{q}$.

Step 2. Design supply chain coordination contract.

For the sake of high risk aversion preference, the retailer's optimal order quantity q_r^* is less than the supply chain optimal order quantity $\hat{q}$. In order to improve the SC efficiency the supplier as the SC coordinator has to give incentive to the retailer to increase the order quantity.

We present an extended buy-back contract to coordinate the SC in this scenario. The main contents of the contract are shown as follows.

1) If the retailer's order quantity q is less than or equal to q_r^* then the supplier provides the retailer with the initial buy-back contract;

2) If the retailer's order quantity q is larger than q_r^* but no larger than $\hat{q}$ then the supplier charges the retailer with the wholesale price ω' and gives full refund of ω' for all unsold for the part of order quantity that is equal to $q - q_r^*$ and the supplier provides the retailer with initial buy-back contract for the part of order quantity that is equal to q_r^*;

3) If the retailer's order quantity q is larger than $\hat{q}$ then the supplier treat the retailer the same as that in (ii) but with a maximum full refund return quantity being equal to $\hat{q} - q_r^*$.

Proof: the extended buy-back contract will satisfy all requirements of Definition 1.

Let $\Pi_r^*(q, X)$, $\Pi_s^*(q, X)$ and $\pi_r^*(q)$, $\pi_s^*(q)$ denote the profit and expected profit of the retailer and the supplier respectively with extended buy-back contract.

1) Firstly, we'll prove that the SC member's expected profit is no less than her (his) original expected profit with the extended buy-back contract.

a) If $q = \hat{q}$ then the retailer's profit is

$$\Pi_r^*(\hat{q},X)=\begin{cases}\Pi_r(q_r^*,X) & X\le q_r^*\\ \Pi_r(q_r^*,X)+(p-\omega')\ (X-q_r^*) & q_r^*\prec X\le\hat{q}\\ \Pi_r(q_r^*,X)+(p-\omega')\ (\hat{q}-q_r^*) & X\succ\hat{q}\end{cases} \tag{12}$$

If $q=\hat{q}$ then the retailer's expected profit is:

$$\begin{aligned}\pi_r^*(\hat{q})&=\int_0^{q_r^*}\Pi_r(q_r^*,x)f(x)dx\\&+\int_{q_r^*}^{\hat{q}}\left[\Pi_r(q_r^*,x)+(p-\omega')\ (X-q_r^*)\right]f(x)dx\\&+\int_{\hat{q}}^{\infty}\left[\Pi_r(q_r^*,x)+(p-\omega')\ (\hat{q}-q_r^*)\right]f(x)dx\end{aligned}$$

$$\begin{aligned}\pi_r^*(\hat{q})-\pi_r(q_r^*)&=(p-\omega')\left[\int_{q_r^*}^{\hat{q}}(x-q_r^*)f(x)dx\right.\\&\left.+(1-F(\hat{q}))(\hat{q}-q_r^*)\right]\end{aligned}$$

If $\omega'\le p$ then $\pi_r^*(\hat{q})-\pi_r(q_r^*)\ge 0$.

A sample comparison between the retailer's profit with initial buy-back contract Π_r and the profit with extended buy-back contract Π_r^* is made with the parameters that are declared in Section 4. The comparison is illustrated in Figure 1.

If $q=\hat{q}$ then the supplier's profit is:

$$\Pi_s^*(\hat{q},X)=\begin{cases}\Pi_s(q_r^*,X)-c(\hat{q}-q_r^*) & X\le q_r^*\\ \Pi_s(q_r^*,X)+\omega'(X-q_r^*)-c(\hat{q}-q_r^*) & q_r^*\prec X\le\hat{q}\\ \Pi_s(q_r^*,X)+(\omega'-c)\ (\hat{q}-q_r^*) & X\succ\hat{q}\end{cases} \tag{13}$$

If $q=\hat{q}$ then the supplier's expected profit is:

$$\begin{aligned}\pi_s^*(\hat{q})&=\int_0^{q_r^*}[\Pi_s(q_r^*,x)-c(\hat{q}-q_r^*)]f(x)dx\\&+\int_{q_r^*}^{\hat{q}}\left[\Pi_s(q_r^*,x)+\omega'(x-q_r^*)-c(\hat{q}-q_r^*)\right]f(x)dx\\&+\int_{\hat{q}}^{\infty}\left[\Pi_s(q_r^*,x)+(\omega'-c)\ (\hat{q}-q_r^*)\right]f(x)dx\end{aligned}$$

$$\begin{aligned}\pi_s^*(\hat{q})-\pi_s(q_r^*)&=\omega'[\int_{q_r^*}^{\hat{q}}(x-q_r^*)f(x)dx\\&+(1-F(\hat{q}))(\hat{q}-q_r^*)]-c(\hat{q}-q_r^*)\end{aligned}$$

If $\omega'\ge\dfrac{c(\hat{q}-q_r^*)}{\int_{q_r^*}^{\hat{q}}(x-q_r^*)f(x)dx+(1-F(q))(\hat{q}-q_r^*)}$ then $\pi_s^*(\hat{q})-\pi_s(q_r^*)\ge 0$.

If $\dfrac{c(\hat{q}-q_r^*)}{\int_{q_r^*}^{\hat{q}}(x-q_r^*)f(x)dx+(1-F(q))(\hat{q}-q_r^*)}\le\omega'\le p$ then the SC member's payoff will be no less than his or her reservation payoff.

A sample comparison between the supplier's profit with initial buy-back contract Π_s and the profit with extended buy-back contract Π_s^* is illustrated in Figure 2.

2) Secondly it is proved that with the extended buy-back contract, the SC member's risk aversion constraint is satisfied.

a) The retailer's risk aversion constraint is satisfied.

According to Formula (12):

If $X\le q_r^*$, $\Pi_r^*(\hat{q},X)=\Pi_r(q_r^*,X)$

then $P\{\Pi_r(q_r^*,x)\le\pi_r(\hat{q})\}=\beta_r$;

If $X\succ q_r^*$, $\Pi_r^*(\hat{q},X)\succ\Pi_r(q_r^*,X)$

then $P\{\Pi_r^*(\hat{q},x)\le\pi_r(\hat{q})\}\prec\beta_r$.

Therefore the retailer's risk aversion constraint is satisfied under the extended buy-back contract.

b) The supplier's risk aversion constraint is satisfied.

It can be proved that $\Pi_s^*(\hat{q},\hat{q})\succ\pi_s(\hat{q})$. The proving is simple and is omitted here. Then $\exists\, q_B$, $q_B\in(0,\hat{q})$, q_B meets the requirement that $\Pi_s^*(\hat{q},q_B)=\pi_s(\hat{q})$. If $\beta_s\ge F(q_B)$ then the supplier's risk aversion constraint is satisfied.

3) The supply chain order quantity is consistent with the system perfect solution i.e. $q=\hat{q}$ with the extended buy-back contract.

Summarize 1 to 3, we have the result that the supplier can give the retailer incentive to increase the order quantity and achieve the system perfect solution by providing the extended buy-back contract when the supplier is low risk aversion e and the retailer is high risk aversion.

3.3 The Supplier is High Risk Aversion ($\beta_s\prec BN$)

According to Corollary 1, we have $q_s^*<\hat{q}$. In this case no matter what the retailer's risk aversion is the supply chain can't achieve system perfect solution.

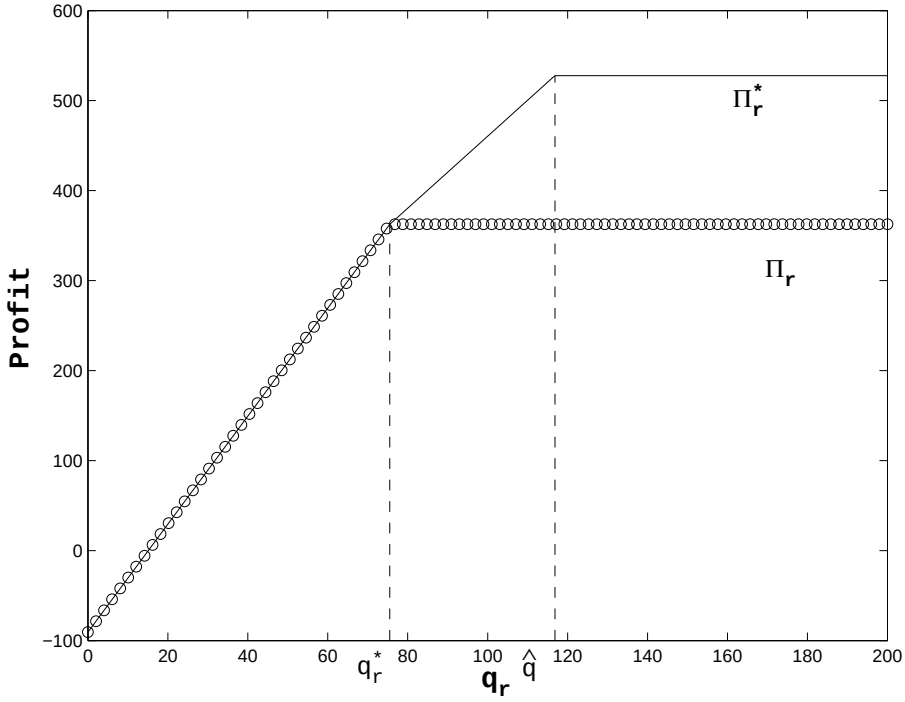

Figure 1. Retailer's profits comparison between Initial and extended buy-back contract

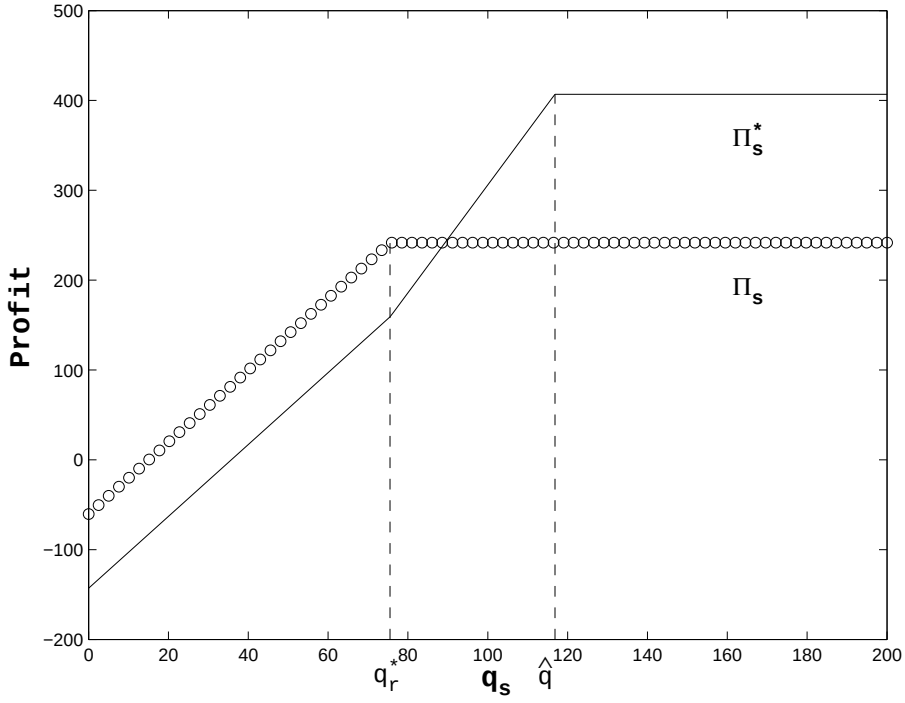

Figure 2. Supplier's profits comparison between Initial and extended buy-back contract

4. Numerical Analysis

In this section we'll give a numerical analysis for the case of low risk aversion supplier and high risk aversion retailer. It is shown that with the extended buy-back contract, with the retailer's risk aversion increasing, the optimal order quantity decreases and the two SC member's expected profits get better off comparing with that of the initial buy-back contract.

Table 1. Impact of the retailer's risk aversion on the members' expected profit with extended buy-back contract

β_r	q_r^*	$\Delta\pi_r$	$\Delta\pi_s$
0.35	90.74	7.19	0.59
0.3	81.89	13.05	25.82
0.25	72.34	20.74	63.15
0.2	61.71	30.3	111.93
0.15	49.31	41.98	172.85
0.1	33.71	56.92	251.16

The parameters used in this section are setup as follows. Let the stochastic market demand $X \sim N(100, 20^2)$. Set the supplier's unit supplying cost c = 2, the retailer's unit retail price $p = 10$ .With the initial buy-back contract assume the unit wholesale price $\omega = 5.2$, the unit buy back price b=4. With the extended buy-back contract the supplier provides the retailer with the unit wholesale price (buy-back price) $\omega' = 8.8$ for the amount of order quantity that is greater than his optimal quantity. Assume the retailer's risk aversion threshold $\beta_r = 0.15$, the supplier's risk aversion threshold $\beta_s = 0.7$ and BN = 4.6. Based on these parameters it can be obtained that $q_r^* = 49.31$, $\pi_r(q_r^*) = 236.46$, $\pi_s(q_r^*) = 157.65$ with the initial buy-back contract; $q_r^* = \hat{q} = 49.31$, $\pi_r(\hat{q}) = 278.44$, $\pi_s(\hat{q}) = 330.5$ with the extended buy-back contract.

The supplier gives the retailer incentive to increase the order quantity and finally increase the whole channel efficiency via the extended buy-back contract. How the role of the extended buy-back contract changes with different retailer's risk aversion levels needs to be analyzed in detail. With the retailer's risk aversion threshold changing from 0.35 to 0.1, the corresponding values of q_r^*, $\Delta\pi_r$, $\Delta\pi_s$ are presented in Table 1. q_r^* is the retailer's optimal order quantity with the initial buy-back contract. $\Delta\pi_r = \pi_r^*(\hat{q}) - \pi_r(q_r^*)$ is the retailer's expected profit increment with the extended buy-back contract. $\Delta\pi_s = \pi_s^*(\hat{q}) - \pi_s(q_r^*)$ is the supplier's expected profit increment with the extended buy-back contract. By analyzing the data in Table 1 we will find that the member's expected profit increment is increasing in the retailer's risk aversion threshold. This confirms that the extended buy-back contract definitely increase the whole channel's efficiency. Moreover it is shown that the extended buy-back contract is more attractive for the supplier than the retailer.

5. Conclusions

In this study, we consider a supply chain that is consisted of a supplier and a retailer who posses different risk aversion preferences in a single period. Based on the initial buy-back contract we provide the extended buy-back contract model. We analyze the condition under which the supply chain can achieve system perfect solution and provide the corresponding contract terms. We report in this paper that: (1) if both the supplier and the retailer are low risk aversion, then the supply chain can achieve system perfect solution with the initial buy-back contract. (2) if the supplier is low risk aversion while the retailer is high risk aversion, then the supplier can give the retailer incentive to achieve system perfect solution with the extended buy-back contract. (3) if the supplier is high risk aversion (no matter what risk aversion the retailer is then the supply chain can not achieve system perfect solution. At the end of this paper we present a numerical analysis to verify the extended buy-back contract.

There are many related problems that need to be further explored. First, the model we considered is in single period. A good extension to this research may be to consider the model in multiple periods. Secondly, the structure of the SC in this paper is simple. The structure of the supply chain is complicated in reality. It is more appreciate to investigate in a more complicated structure.

6. Acknowledgement

The research is supported by China National Nature Science Foundation under Grant 70971083 and SHUFE Postgraduate Innovation Foundation under Grant CXJJ-2009-336.

REFERENCES

[1] L. Eeckhoudt, C. Gollier, and H. Schlesinger, "The risk averse (prudent) newsboy," Management Science, Vol. 41, No. 5, pp. 786–794, 1995.

[2] V. Agrawal and S. Seshadri, "Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem," Manufacturing & Service Operations

Management, Vol. 2, No. 4, pp. 410–423, 2000.

[3] H. S. Lau and A. H. L. Lau, "Manufacturer's pricing strategy and returns policy for a single-period commodity," European Journal of Operational Research, Vol. 116, No. 2, pp. 291–304, 1999.

[4] X. Gan, S. Sethi, and H. Yan, "Channel coordination with a risk neutral supplier and a downside risk-averse retailer," Production and Operations Management, Vol. 14, No. 1, pp. 80–89, 2005.

[5] T. M. Choi, D. Li, and H. Yan, "Mean-variance analysis of a single supplier and retailer supply chain under a returns policy," European Journal of Operational Research, Vol. 184, No. 1, pp. 356–376, 2008.

[6] X. Gan, S. Sethi, and H. Yan, "Coordination of supply chain with risk-averse agents," Production and Operations Management, Vol. 13, No. 2, pp. 135–149, 2004.

[7] Y. Wei and T. M. Choi, "Mean-variance analysis of supply chains under wholesale pricing and profit sharing schemes," European Journal of Operational Research, Vol. 204, No. 2, pp. 255–262, 2010.

[8] K. Mastui, "Returns policy, new model introduction, and consumer welfare," International Journal of Production Economics, in press 2009.

[9] T. Xiao and T. M. Choi, "Purchasing choices and channel structure strategies for a two-echelon system with risk-averse players," International Journal of Production Economics, Vol. 120, No. 1, pp. 54–65, 2009.

[10] A. Charnes and W. W. Cooper, "Chance-constrained programming," Management Science, Vol. 6, No. 1, pp. 73–79, 1959.

[11] B. Pasternack, "Optimal pricing and returns policies for perishable commodities," Marketing Science. Vol. 4, No. 2, pp. 166–176, 1985.

Appendix

A.1 Proof of Lemma 1

(1) Proof of Formula 8

The supplier's profit function is

$$\prod_s (q, X) = (\omega - c)q - b(q - X)^+$$

The least amount products means all products sold in the retail market i.e. $X \geq q$.

The above formula transforms to

$$\prod_s (q, X) = (\omega - c)q$$

Substitute $\prod_s$, q with α_s, q_s^0 into above formula we get

$$\alpha_s = (\omega - c)q_s^0$$

$$q_s^0 = \frac{\alpha_s}{\omega - c}.$$

(2) Proof of Formula 9

The retailer's profit function is

$$\prod_r (q, X) = p \min\{q, X\} - \omega q + b(q - X)^+$$

All products are sold out i.e. $X \geq q$.

The above formula transforms to

$$\prod_r (q, X) = (p - \omega)q$$

Substitute $\prod_r$, q with α_r, q_r^0 into above formula we get

$$\alpha_r = (p - \omega)q_r^0$$

$$q_r^0 = \frac{\alpha_r}{p - \omega}$$

A.2 Proof of Theorem 1

1) In the case of $q_s \leq q_s^0$

$$\begin{aligned}&\prod_s (q_s, X)\\ &= (\omega - c)q_s - b(q_s - X)^+\\ &\leq (\omega - c)q_s\\ &\prec (\omega - c)q_s^0 = \alpha_s\end{aligned}$$

Therefore $P\{\prod_s (q_s, X) \leq \alpha_s\} = 1$

ii) In the case of $q_s \succ q_s^0$

If $X \succ q_s$

$$\begin{aligned}&P[\{\prod_s (q_s, X) \leq \alpha_s\} \cap \{X \succ q_s\}]\\ &= P\{(\omega - c)q_s \leq \alpha_s\}\\ &\because (\omega - c)q_s \succ (\omega - c)q_s^0 = \alpha_s\\ &\therefore P\{(\omega - c)q_s \leq \alpha_s\} = 0\end{aligned}$$

If $X \leq q_s$

$$\begin{aligned}&P[\{\prod_s (q_s, X) \leq \alpha_s\} \cap \{X \leq q_s\}]\\ &= P\{(\omega - c)q_s - b(q_s - X) \leq \alpha_s\}\\ &= P\{(\omega - c - b)q_s + bX \leq \alpha_s\}\\ &= P\left\{X \leq \frac{\alpha_s - (\omega - c - b)q_s}{b}\right\}\\ &= F\left(\frac{\alpha_s - (\omega - c - b)q_s}{b}\right)\end{aligned}$$

Therefore if $q_s \succ q_s^0$ then the supplier's downside risk is $F\left(\frac{\alpha_s - (\omega - c - b)q_s}{b}\right)$.

A.3 Proof of Theorem 2

1) In the case of $q_r \leq q_r^0$

$$\begin{aligned}&\prod_r (q_r, X)\\ &= p \min\{q_r, X\} - \omega q_r + b(q_r - X)^+\\ &\leq (p - \omega)q_r\\ &\prec (p - \omega)q_r^0 = \alpha_r\end{aligned}$$

Thus $P\{\prod_r (q_r, X) \leq \alpha_r\} = 1$

Therefore if $q_r \leq q_r^0$ then γ_r equals to 1.

2) In the case of $q_r \succ q_r^0$

If $X \succ q_r$

$$\begin{aligned}&P[\{\prod_r (q_r, X) \leq \alpha_r\} \cap \{X \succ q_r\}]\\ &= P\{(p - \omega)q_r \leq \alpha_r\}\\ &\because (p - \omega)q_r \succ (p - \omega)q_r^0 = \alpha_r\\ &\therefore P\{(p - \omega)q_r \leq \alpha_r\} = 0\end{aligned}$$

If $X \leq q_r$

$$\begin{aligned}&P[\{\prod_r (q_r, X) \leq \alpha_r\} \cap \{X \leq q_r\}]\\ &= P\{(p - b)X - (\omega - b)q_r \leq \alpha_r\}\\ &= P\left\{X \leq \frac{\alpha_r + (\omega - b)q_r}{p - b}\right\}\\ &= F\left(\frac{\alpha_r + (\omega - b)q_r}{p - b}\right)\end{aligned}$$

Therefore if $q_r \succ q_r^0$ then γ_r equals to

$$F\left(\frac{\alpha_r + (\omega - b)q_r}{p - b}\right).$$

A.4 Proof of Theorem 3

Firstly we'll show how to get the downside risk threshold of the supplier.

$$\pi_s(\hat{q})$$
$$= (\omega - b - c)\hat{q} + bE[\min\{\hat{q}, X\}]$$
$$= (\omega - b - c)\hat{q} + b(\hat{q} - \int_0^{\hat{q}} F(x)dx)$$

$$P\{\prod_s(\hat{q}, X) \le \pi_s(\hat{q})\}$$
$$= P\left\{\min\{\hat{q}, X\} \le \hat{q} - \int_0^{\hat{q}} F(x)dx\right\}$$

If $\hat{q} \prec X$

$$P\{\prod_s(\hat{q}, X) \le \pi_s(\hat{q})\}$$
$$= P\left\{\hat{q} \le \hat{q} - \int_0^{\hat{q}} F(x)dx\right\}$$
$$= 0$$

If $\hat{q} \ge X$

$$P\{\prod_s(\hat{q}, X) \le \pi_s(\hat{q})\}$$
$$= P\left\{X \le \hat{q} - \int_0^{\hat{q}} F(x)dx\right\}$$
$$= F(\hat{q} - \int_0^{\hat{q}} F(x)dx)$$

Therefore the downside risk threshold of the supplier equals to $F(\hat{q} - \int_0^{\hat{q}} F(x)dx)$.

In the similar way we can prove that the downside risk threshold of the retailer also equals to $F(\hat{q} - \int_0^{\hat{q}} F(x)dx)$.

J. Service Science & Management, **2010, 3: 33-44**
doi:10.4236/jssm.2010.31004 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Evaluating a Student MIS in the Tertiary Education Sector: Addressing Functional-Operational Misalignment through Continuous Participative Evaluation

Marian Carcary

Kemmy Business School, University of Limerick, Limerick, Ireland.
Email: marian.carcary@ul.ie

Received September 16th, 2009; revised November 2nd, 2009; accepted December 7th, 2009.

ABSTRACT

The Information and Communications Technology (ICT) evaluation literature spans several decades. ICT evaluation approaches range from objective and positivistic to subjective and interpretive. While positivistic approaches have dominated the past, there is increasing recognition of the value of interpretivist methods and the need for ongoing project evaluation. Formative continuous participative evaluation (CPE) offers several benefits in terms of project control, enhanced stakeholder relationships and benefit realisation; nonetheless this is often ignored in practice. There is a paucity of ICT evaluation within the Higher Education sector. The 14 Irish Institutes of Technology (IoTs) recently underwent an extensive transformation of their ICT systems, through a nationwide implementation of a suite of integrated IS. This research study, centred on the evaluation of the Student MIS implementation was interpretive in nature; case studies were conducted in five IoTs. This paper focuses specifically on one issue uncovered through the research i.e. the misalignment between the Student MIS and the IoTs requirements. The paper proposes a set of guidelines for addressing this issue through focusing on the theoretical underpinnings of CPE and its importance for organisational learning and benefit realisation.

Keywords*: *ICT Investment Management, ICT Evaluation, Formative Evaluation, Continuous Participative Evaluation, MIS, ICT in Tertiary Education

1. Introduction

ICT evaluation research has attracted the interests of academics and practitioners for several decades. Evaluation approaches can be viewed along a spectrum that ranges from objective, rational, positivist approaches to subjective, interpretive approaches [1]. The former have historically dominated ICT evaluation research [2], at the expense of contextual issues. However, in recent years, interpretive approaches have gained a stronger foothold. There is growing consensus that evaluation needs to be formative in nature i.e. ongoing throughout the project lifecycle [3–6]. Gemmell and Pagano (2003) [7] suggest that organisations would benefit from moving towards a continuous evaluation approach that was integrated into the project management process. Arguments for continuous evaluation include project control, adaptability, Value realization and improved relationships [8]. The formative approach seeks to explore all project issues, identify unexpected impacts, assess the degree to which the project proceeds as planned and provide feedback to improve project performance. It focuses on complex interactions between people and technology, and considers the system's short-term and intermediate effects and its influence on stakeholders.

Many authors have argued that ICT evaluation needs to be dynamic and continuous across the project lifecycle (see for example [9–13]). This is necessary in order for what Willcocks (1992) [14] terms the "*islands of evaluation*" to become integrated and for continuous improvement opportunities to be identified. A particular type of formative evaluation is Continuous Participative Evaluation (CPE). Continuous participative approaches for example Remenyi *et al*'s (1997) [15] Active Benefit Realisation (ABR[1]), promote stakeholder involvement, motivation and commitment and improve ICT decision-making

through establishing dialogue with all affected parties. Cordoba's (2009) [16] development of a methodological framework to support the process of IS planning emphasises continuous identification of concerns from stakeholders; it promotes continuous dialogue, listening and mutual collaboration between participants and facilitates critical reflection in the exploration of possibilities for improvement. CPE is integrated into organisational daily activities with a continuous focus on stakeholders, ICT quality and evolving business objectives. Stakeholder participation helps improve functionality, helps identify and manage intangible impacts, stimulates organisational learning, focuses attention on issues that impair ICT success, helps maximise ICT benefit realisation and reduces the culture gap between business and ICT domains. Throughout the 1980s and 1990s the concept of an "*evaluation party*" [17] emerged to emphasise the importance of stakeholder views.

Despite its usefulness, formative continuous participative evaluation is often not conducted in practice. Willcocks *et al.* (1999) [18] suggested that few companies evaluate strategic, business, end-user and technical performance in an integrated manner throughout the system's lifecycle. Evaluation declines following project feasibility assessment; it is poorly linked across lifecycle stages; and little learning for future ICT investments is derived. Further, Hillam and Edwards (2001) [19] stated that user perception is rarely considered. Lack of involvement can result in decreased commitment or negative attitudes. For example, Markus (2004) [20] claimed that approximately 75% of organisational change efforts driven by technology fail because of negative stakeholder reactions to work practice, business process and technology changes.

This paper discusses the need for formative continuous participative evaluation mechanisms in addressing some of the problems resultant from a large-scale standard Student MIS implementation in the Irish Institutes of Technology (IoTs). To date, there is a paucity of evaluation research in the tertiary education sector [21]. Further research is required in this area as the ICT systems that support student administrative operations represent the means for competitive parity with or advantage over other educational establishments. This paper discusses one of the key findings that were distilled from the Student MIS evaluation processes undertaken – the issue of functional-operational misalignment. It proposes a set of guidelines to help the IoT sector address this issue through focusing on the importance of formative CPE approaches. The evidence suggests that the approaches adopted in the implementation and operationalisation of the Student MIS did not emphasise the importance of evaluation exercises; however, this paper suggests that future benefit potential realisation can be enhanced through adopting formative CPE methods.

2. The MIS Project in the IoT Sector

A proposal to investigate a collaborative acquisition of a MIS for the Irish IoT sector was initially raised in 1991/1992. This system sought to support new modes of education delivery, support IoT administrative operations, improve services to all stakeholders, streamline workflow and improve organisational communications and competitiveness. The project involved representatives from the Department of Education and Science (DoES) and the Council of Directors of the Irish IoTs. The system was selected by a team of 12 representatives from 8 IoTs. However, the mechanism used in tender evaluation was not sufficiently in-depth or not applied with the necessary rigour; hence it was felt by many stakeholders that the system selection decision was a misguided one. The initiative resulted in the implementation of a suite of integrated Information Systems (IS) for library, human resources, finance and student management functions in 15 Institutions. These systems were rolled out to the IoT sector in a series of implementation waves between 2000 and 2006. This paper focuses primarily on the impact of the Student MIS.

In an attempt to maintain a common national standard system design, a central project team was responsible for all system development. Any required system changes were managed through a central system change request process. Hence, the majority of IoTs did not have system development autonomy. While the central team offered the benefits of resource efficiency, development of specialisations, access to scarce resources etc, it was felt that the change request process to the central project team for expost system development was unnecessarily bureaucratic. Changes refused by the central team, on the basis that they were not required by the majority of IoTs, resulted in some sites using functionality they found to be unsuitable. Because any tailoring was designed to meet common requirements, the changes made were often compromises on specific needs. The bureaucracy of system change requests resulted in IoTs deviating from the common system standard through in-house developments that were adhoc and nonstandard across the sector.

[1]ABR shifts the focus away from technical issues towards stakeholder involvement. Each stakeholder level plays a co-creation and coevolutionary role in systems development and in achieving results, and is responsible for ICT management. In the process, the primary issues and evaluation information are summarised in documents called pictures-a business picture and a supporting financial picture and project picture. Evaluation consists of three phases-setting the course, formative evaluation, and moving forward after feedback; and seven key activities-initialisation, production of pictures, agreement to proceed, systems development, evidence collection, participative evaluation and development of updated pictures. This evaluation process continues in an iterative manner.

Commissioning of the Student MIS across the various implementation waves resulted in considerable work for both the central team and individual IoTs. There was a perceived lack of preparation for system commissioning which gave rise to negative stakeholder perceptions and detracted from staff willingness to embrace the system. System operationalisation resulted in multiple problems; these were linked to the system's complexities, lack of system configuration, lack of IoT preparation, and loss of IoT control to the central structure. Further, in many IoTs, the project teams established for system implementation were disbanded/scaled back too early, hence there were little resources to support problems or exploit system capabilities. However, following a period of system use, staff's increased familiarity with the system led to the realisation of a not insignificant flow of benefits. Benefits included for example improved structure, data standards and data access; job and management related benefits; improved quality procedures; and staff empowerment; and the Student MIS also served as a platform for future IoT ICT developments.

3. The Student MIS Explained

This Student MIS was oriented towards the US market and was anticipated to cater for all stages of a student's interaction with an IoT. The system had comprehensive functionality for course and subject management, student data, admissions application processing, student registration, maintenance grants payments, accounts receivable and fees assessment, examinations and academic history, student progression and graduation. Details of its modules are discussed in **Table 1**:

Figure 1 provides an overview of how these Student MIS modules were integrated. Firstly, biodemographic data is captured to facilitate admissions applications processing. These applications may come from the CAO or be made directly to the IoT. The applications are verified against the catalogue of existing courses. Following acceptance of a course place, the individual's data is used to create a student record. All students are registered at the individual subject level; this is made possible by the breakdown of subjects into their respective delivery modes based on the ACS. This registration is required for student fee assessment, maintenance grants payments and for examination results processing. All results are transferred to an academic history archive; they become the basis for student graduation from a completed programme or progression to the following academic year, where a new student record is created.

4. Research Methodology

The research methodology used in evaluating the impact of the Student MIS was interpretive in nature. The interpretive paradigm offered the opportunity to develop an in-depth understanding of the ICT system's impact; it facilitated the capture of contextual depth and detailed, nuanced descriptions; and avoided the unproblematic, value-free view of organisations associated with positivist approaches. The study's research methodology is outlined in **Figure 2**. The case study was the selected research method and was based on data collected from five sources-organisational websites, project documentation, newspaper articles, independent reports and semi-structured interviews.

Table 1. Student MIS modules

Catalogue	The database of courses and subjects offered by an Institution
Approved Course Schedule (ACS)	The listing of subjects linked to a course; it outlines teaching hours, examination and continuous assessment components
Sections and Blocks	Breaks down subjects into specific delivery modes, such as lecture, practical and tutorial hours. Students are registered on these sections. A block is a grouping of sections by stage
CAO Interface	Enables download of student data from the Central Applications Office website. The CAO is the body responsible for recording all student applications to third level education institutions in Ireland
Admissions	Stores applicant data and processes applications. It supports both CAO and direct admission applicants
General Person	Captures biographical and demographic data
General Student	Records current and historical student data. The record is originally created when an applicant accepts a course place offered
Registration	Records registration information for a given term
Maintenance Grants	Facilitates grant payments to a student's bank account
Accounts Receivable	Tracks financial transactions such as student fee assessment, contract assignment, payments, and receipt, invoice and journal production
Examinations	Facilitates student exam results and production of relevant reports
Academic History	Enables historical academic results to be recorded in the Student MIS
Web for Faculty	Empowers academics to enter student results, and view class and student details. Grades are broken into their component parts through electronic gradebook functionality
Letter Generation	Enables selected data to be merged with predefined letters
Graduation	Produces graduate lists and records data relating to graduation ceremonies

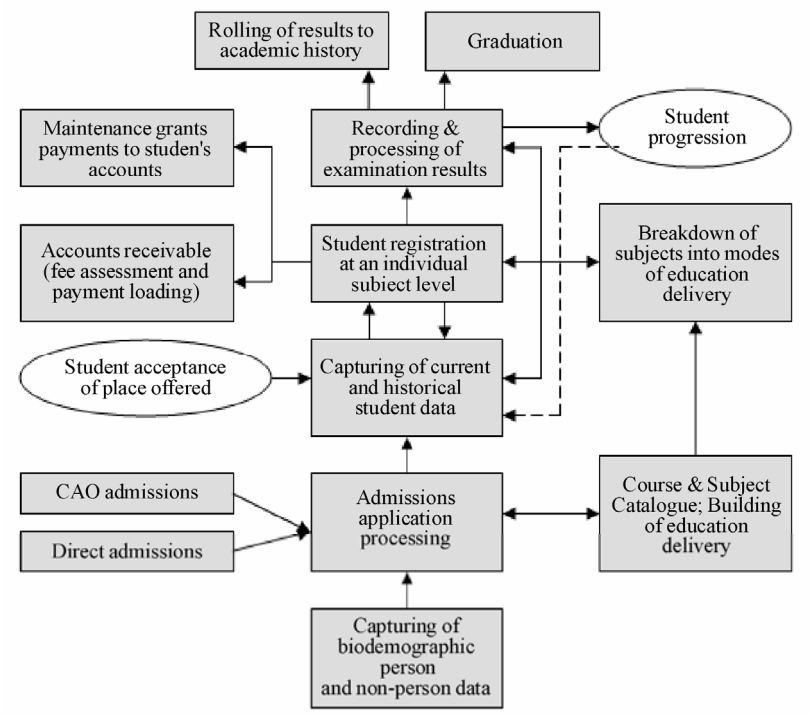

Figure 1. Functionality and operations of the student MIS

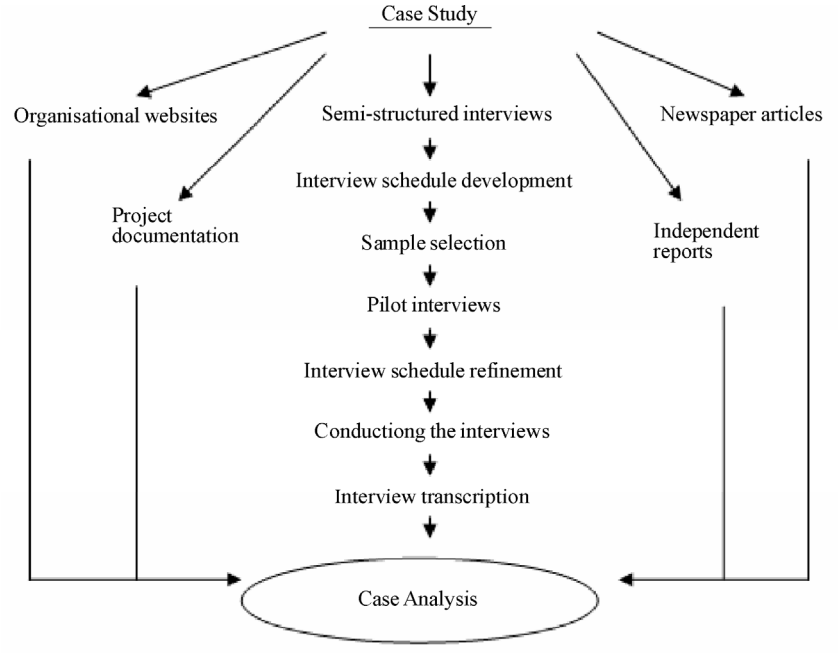

Figure 2. Research methodology

Case studies were conducted within five IoTs. Purposive sampling was used in case site selection as this sampling strategy ensures that key research themes are addressed and that diversity in each category is explored. The five case sites were selected due to their diversity in a number of respects. They participated in different implementation waves, were geographically dispersed and differed in their student population sizes and academic programme offerings. The following points give a brief synopsis of the five case sites:

- Site One was one of the first IoTs to implement the student MIS. This site had significant in-house MIS team capabilities, who were responsible for system implementation, support and had considerable autonomy in system development.
- Site Two was a member of the first implementation wave. Due to difficulties experienced with the central change request process, its in-house MIS unit also developed additional functionality to meet end-user needs.
- Site Three was a member of the second implementation wave. This IoT was smaller than the previous two sites and had more limited technical resources.
- Site Four was the smallest IoT examined in this study and was also a member of the second implementation wave. Its project team experienced a number of personnel changes during the implementation effort.
- Site Five was a member of the final implementation wave. It experienced difficulties in resourcing a dedicated project team and its initial system start-up date was delayed. At the time of research, the system was used to a limited extent and work was ongoing in implementing core functionality.

Within those IoTs, 49 semi-structured interviews were carried out between 30 November 2005 and 24 May 2006 with senior management personnel, MIS team personnel and system end users. The selected informants were closely involved in the ICT project and had in-depth knowledge of the subject area. Each interview lasted between 60 and 90 minutes, was recorded with the informants' permission and was later transcribed. The informants were given the opportunity to verify the transcripts prior to analysis. Further, the supporting documentation was valuable in corroborating the evidence collected in the semi-structured interviews. It provided a means of triangulation in that it supplied specific details, and helped to augment and substantiate the interview data. The data analysis process is outlined in **Figure 3**.

Data analysis was carried out using a variant of Glaser and Strauss's (1967) [22] grounded theory method. GT is one of the most widely used qualitative frameworks in business and management studies [23]. This GT analysis was supported by a Computer Aided Qualitative Data Analysis Software (CAQDAS) package called N-vivo. The N-vivo package facilitates efficient data indexing and management, and supports analysis through for example relationship and model exploration. As outlined in **Figure 3**, the interview transcripts were initially imported into this software. Examination of these transcripts led to key words/ideas being identified and these were coded using N-vivo. Groupings of these codes that contain similar content are referred to as concepts in GT. As coding progressed, it became apparent that many concepts were related and these were reclassified into a se-

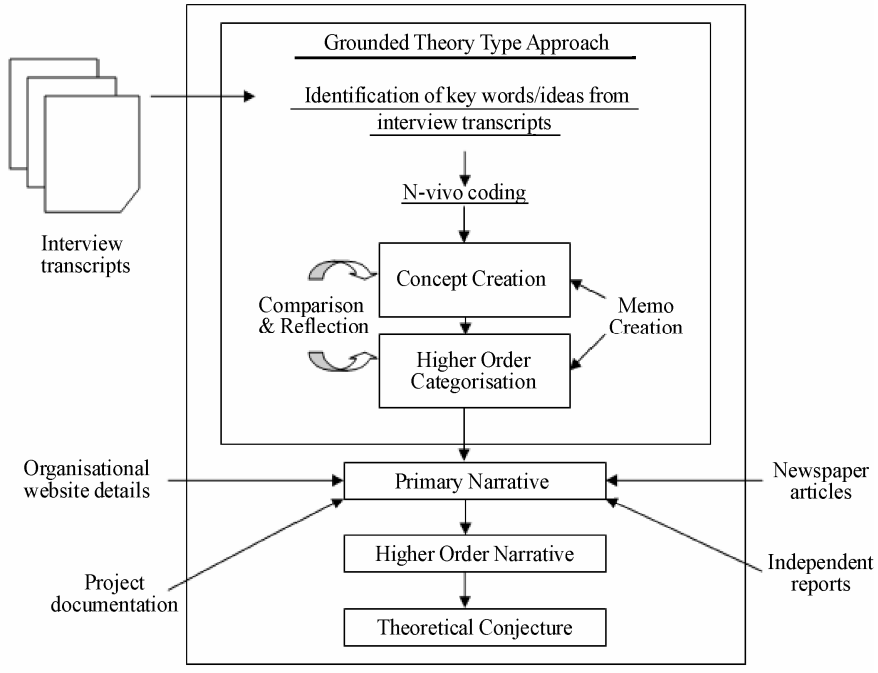

Figure 3. Qualitative data analysis

Table 2. Degree of system exploitation

	IoT's Degree of System Exploitation				
System Capabilities	Site One	Site Two	Site Three	Site Four	Site Five
Catalogue	Partly Exploited	Partly Exploited	Exploited	Partly Exploited	Exploited
ACS	Unexploited	Partly Exploited	Partly Exploited	Partly Exploited	Partly Exploited
Sections & Blocks	Partly Exploited	Exploited	Partly Exploited	Partly Exploited	Exploited
CAO Interface	Exploited	Partly Exploited	Partly Exploited	Partly Exploited	(Was being implemented)
Admissions	Partly Exploited	Partly Exploited	Partly Exploited	Partly Exploited	Partly Exploited
General Person	Exploited	Exploited	Partly Exploited	Partly Exploited	Partly Exploited
General Student	Partly Exploited	Exploited	Partly Exploited	Partly Exploited	Partly Exploited
Registration	Exploited	Partly Exploited	Partly Exploited	Partly Exploited	Partly Exploited
Maintenance Grants	Exploited	Exploited	Exploited	Exploited	Unexploited
Accounts Receivable	Unexploited	Partly Exploited	Partly Exploited	Partly Exploited	Unexploited
Examinations	(Was being implemented)	Partly Exploited	Partly Exploited	Partly Exploited	(Was being implemented)
Academic History	(Was being implemented)	Exploited	Exploited	Exploited	Unexploited
Web for Faculty	(Was being implemented)	Unexploited	Partly Exploited	Unexploited	Unexploited
Letter Generation	Unexploited	Unexploited	Unexploited	Unexploited	Unexploited
Graduation	Unexploited	Unexploited	Unexploited	Unexploited	Unexploited

ries of categories (i.e. broad groups of similar content that are later used to generate theory) and related sub-categories using N-vivo's hierarchical tree structure. This organised related concepts in relation to the overall research and facilitated greater understanding of the body of evidence through examining the key themes. Memo creation to clarify ideas and identify relationships between categories, constant concept comparison and iterative reflection on what was already coded were important steps in this coding process.

The key concepts and categories identified through N-vivo coding, as well as important details from the other four sources of case study evidence were synthesised into a detailed cross-case primary narrative of the Student MIS project. Narratives play an important role in the social world; they are a form of knowledge and communication [24], as complex situations can be better understood in story format. Hence, they enable a researcher to shape various interview stories into a coherent account of the key themes. Through significant reflection on the primary narrative, it was reduced to the principal findings or themes. Reflection on the primary narrative involved considering three questions: "*what does the text say*?" "*why does the text say what it does*?", and "*what is my understanding of what is taking place*?". This approach was useful in providing a conceptual separation of three ways of examining the primary narrative and in expanding my interpretation over a series of stages. The processes involved in distillation of the key findings involved both creativity and flexibility. Diagrammatic representation was important in understanding the phenomenon's diversity and in exploring relationships and complex processes.

5. Findings-the Issue of Functional Operational Misalignment

Findings on several different aspects of the project were uncovered including system selection; system development for the Irish IoTs; system commissioning; ex-post performance in the early years; and ex-post performance at the time of research. These findings are discussed by Carcary (2009) [21]. This paper focuses on one of those findings i.e. the misalignment between system capabilities and IoT requirements. The main challenge facing the IoTs was the misalignment between what the system was used for at the time of research and what it was capable of delivering, and between IoT requirements and the extent to which they were met. As explained in the following sections, IoTs use of the student MIS fell short of system potential; hence, IoTs were not leveraging the system's potential advantages. Sub-optimal system usage was partly due to the system's inability to meet some IoT requirements.

5.1 An Evaluation of System Functional Capability Exploitation

Benefit realisation from the student MIS was sub-optimal within the IoTs. The system offered functionality for all aspects of student administration, yet many of its capabilities remained unused. For example, several informants within Site Three suggested that system usage at the time of research was "a fire fighting" exercise and was "just touching the tip of the iceberg". One end-user estimated functionality exploitation to be 35%-40%,

while a project team member suggested 50%-60%. Within Site Two, only the baseline functionality was implemented; work was ongoing in streamlining operational processes. Site Five's exploitation was limited to course set-up and registrations functionality, and significant work was required to advance the knowledge base. Site Four's system usage was equated to "scratching the surface"; estimates of exploitation ranged between 5%-10% according to the ICT Manager and 70%-75% according to a project team member. Despite being the first to implement the system, Site One had according to the system's development team leader used only 10% of the baseline functionality. For example, examination-related modules were only being introduced at the time of informant interviewing. Hence, in all IoTs, informants suggested that the system was not fully exploited. **Table 2** provides a synopsis of the degree of IoT's system exploitation. This suggests that certain modules were exploited, partially exploited or completely untouched.

The variances experienced in system functional capability exploitation can be linked to a number of factors:

- *Implementation Wave:* From Site Five's perspective, limited system exploitation was partly due to its postponed implementation deadline. However in the other case study sites, the system was stabilised between three and five years. Those IoTs had a longer timeframe to exploit the system but had not taken full advantage of this.
- *Staff Knowledge and Resources:* Across all IoTs, reasons for sub-optimal system exploitation were related to limited knowledge of further system capabilities and the time required to achieve system familiarity. This was linked to the scaling back of project teams and the associated training problem. Site Two explicitly cited lack of resources as key in failing to explore the system's fuller potential and implement further modules.
- *Ability to Meet Requirements:* The inability of some modules to meet requirements and lack of tailoring for those sites under the central structure impacted on the degree of system exploitation. The central structure had a phased approach for system development, which was restrictive for urgent IoT requirements.

5.2 An Evaluation of the Degree of IoT Operational Requirements Met

Across case study sites, the efficiency with which requirements were met was questionable. Many baseline modules did not meet Site One's requirements and, as stated, resulted in numerous in-house customisations and development projects. The system involved significant work for Site One's administrative staff; however in comparison to legacy applications, all informants agreed that it had substantially improved student record management. The degree of requirements met by the baseline system within Site Two was estimated by a managerial figure at 20%. The functionality was too generic and deficient in reporting capabilities; hence the system was used primarily for data storage while all reporting, manipulating and controlling was done in-house. Within Site Three, the required further development of the Student MIS had not taken place. Hence, the system remained bespoke towards full-time students and was inflexible in dealing with any other applicant types. A project team member of Site Four regarded it as meeting 85%-90% of requirements, but suggested a complete suite of integrated products would be more beneficial. Site Five also encountered some limitations. Approximately 30% was regarded by one team member as unsuited to the Irish tertiary educational system. **Table 3** outlines the degree of requirements met by the various modules; it highlights that modules either met all requirements, some requirements or were unsuited.

5.3 Functional-Operational Misalignment

The functional-operational misalignment uncovered impacted the extent to which administrative staff could effectively use the system. IoTs ability to address this issue depended on a number of factors:

- *Site autonomy over the Student MIS:* Site One had greater control over its system development. Much functionality used on a day-to-day basis was developed in-house. This was in contrast to the other case study sites that had little local autonomy. They were restricted in their ability to customise the system due to the control exercised by the central structure.
- *Degree of adherence to the common national system standard versus in-house development:* As stated, IoTs under the central structure's directive were tied to a common national standard. However, the central structure's development time for mission critical reports and for applications outside the full-time student population was slow and based on priority demand. Further, development was generic rather than customised to IoT specifications. Hence, all case study sites diverged to some extent from the common standard. This was primarily in developing external reporting capabilities.
- *Degree of in-house development resource capability:* A lack of resources was apparent in Site Two, Three and Four due to the scaling back of project teams. This impacted on:
  - IoT's ability to advance the system for pro jects such as modularisation and semesterisation, and tailor it for different student co-

Table 3. Degree of IoT requirements addressed

	Degree of IoT Requirements Met				
System Capabilities	**Site One**	**Site Two**	**Site Three**	**Site Four**	**Site Five**
Catalogue	Unsuited	Met Past Requirements	Partly Met Requirements	Met Requirements	Met Requirements
ACS	Unsuited	Met Past Requirements	Met Requirements	Met Past Requirements	Met Requirements
Sections & Blocks	Partly Met Requirements	Met Requirements	Met Requirements	Met Requirements	Met Requirements
CAO Interface	Met Requirements	Met Requirements	Met Requirements	Met Requirements	(Unknown[2])
Admissions	Met Requirements	Met Requirements	Partly Met Requirements	Met Requirements	Met Requirements
General Person	Met Requirements	Partly Met Requirements	Met Requirements	Partly Met Requirements	Met Requirements
General Student	Unsuited	Partly Met Requirements	Met Requirements	Partly Met Requirements	Met Requirements
Registration	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements
Maintenance Grants	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements	(Unknown)
Accounts Receivable	Unsuited	Unsuited	Partly Met Requirements	Unsuited	Unsuited
Examinations	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements	Partly Met Requirements	(Unknown)
Academic History	(Unknown)	Met Requirements	Met Requirements	Met Requirements	(Unknown)
Web for Faculty	Met Requirements	(Unknown)	Partly Met Requirements	(Unknown)	Met Requirements
Letter Generation	Unsuited	Unsuited	Unsuited	Unsuited	Unsuited
Graduation	(Unknown)	Unsuited	Unsuited	Unsuited	(Unknown)

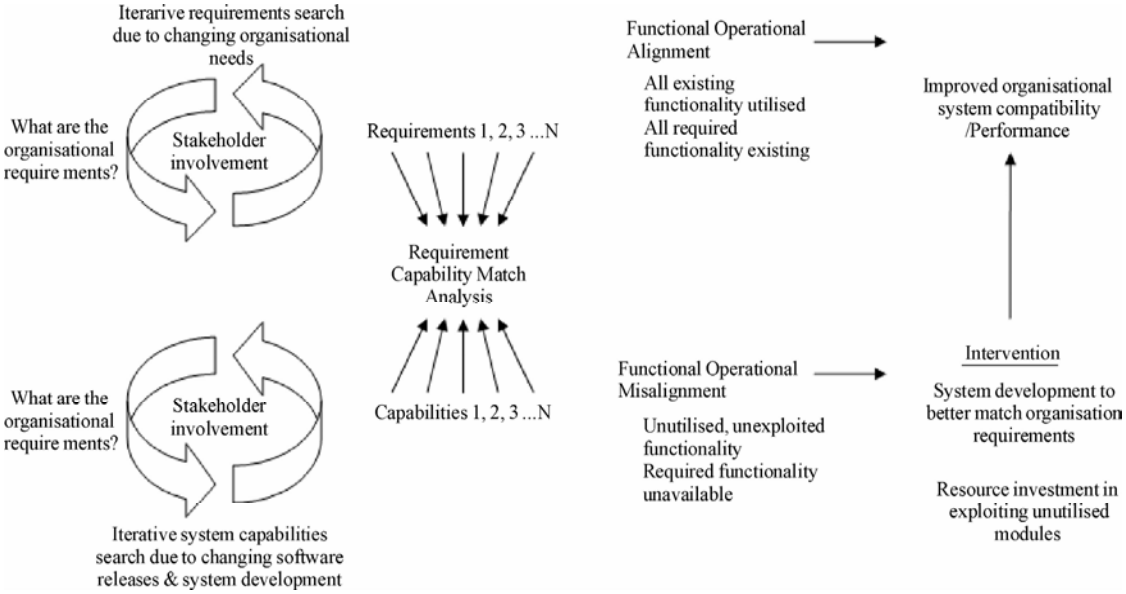

Figure 4. Addressing functional-operational misalignment

[2]The compatibility of a system module with IoT requirements was unknown when the module had not been exploited or examined by MIS personnel

horts;
- Time lags for report development;
- The degree to which unutilised modules could be further exploited;
- The degree to which staff could be trained in further system capabilities and to which more efficient work practices could be found.

Those IoTs with greater resources were better positioned to address the above four points.

5.4 Towards CPE-Addressing Functional-Operational Misalignment

The following model provides a visual depiction of functional-operational misalignment. It puts forward a number of guidelines/steps that IoTs should consider in addressing this issue. These guidelines draw on the theoretical underpinnings of formative CPE.

The following five steps provide an explanation of **Figure 4**.

Step One. Organisational requirements constantly evolve. Examples of such changes in the Irish IoTs include:

- Changes in education delivery modes. These include the introduction of semesterisation, modularisation, e-learning and blended learning, and a move towards distance learning through exploiting the virtual campus concept. Each IoT had a specific timeframe for embracing these changes.
- Changes in student profiles. Examples of non-standard students include students with subject exemptions, Socrates students, repeat students, ACCS students, Fáilte Ireland and FÁS apprentice students, adult learners and students undertaking more than one course.

Determining functional-operational alignment requires IoTs to iteratively establish what their requirements are. This process is outlined in **Figure 5**. Key stakeholder groups need to be involved. In the Student MIS project these stakeholders include senior management who make academic policy decisions and administrative staff who use the system.

Step Two. System capabilities also evolve. Changes to the Student MIS capabilities in the Irish IoTs are due to:

- Improvements and upgrades from the system vendor;
- Developments by the central structure or in- house MIS personnel.

Determining functional-operational alignment requires IoTs to iteratively review what the system capabilities are. This process is outlined in **Figure 6**. Key stakeholders need to be involved. In the Student MIS project, these stakeholders include MIS personnel who understand the system's changing functionality. Such changes also need to be promoted to administrative staff to increase their awareness of the functionality available.

Step Three. The output of the previous two steps is examined in a requirements-capability match analysis (**Figure 7**). It investigates the degree to which organisational-operational requirements and system functional capabilities converge. Each operational requirement is examined against relevant system components to identify how well it is supported. In totality, this exercise outlines the extent to which the system meets IoT requirements. Each system module is also examined to determine its degree of utilisation. In totality, this outlines the extent to which the IoT exploits system capabilities. This process should be documented in the format of detailed descriptions, which would better help in identifying steps for remedial action.

Step Four. The outcome of the requirements- capability match analysis (**Figure 8**) will be a state of functional-operational alignment and/or a state of functional-operational misalignment. Functional-operational

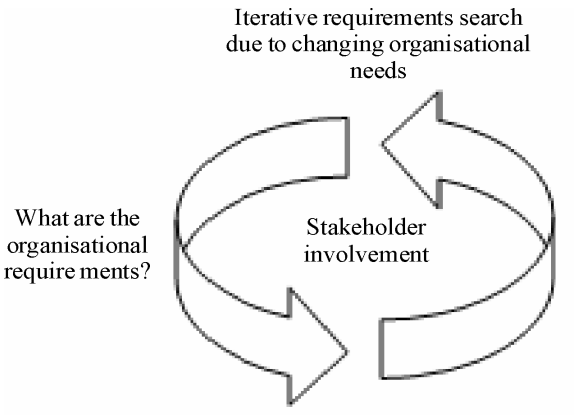

Figure 5. Determining organisational requirements

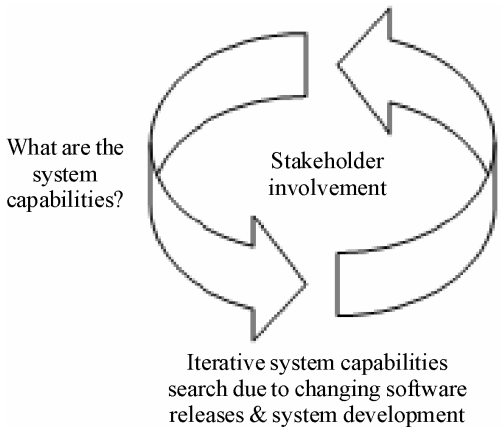

Figure 6. Determining system capabilities

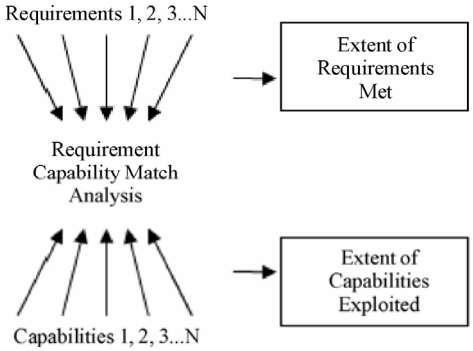

Figure 7. Requirements-capability match analysis

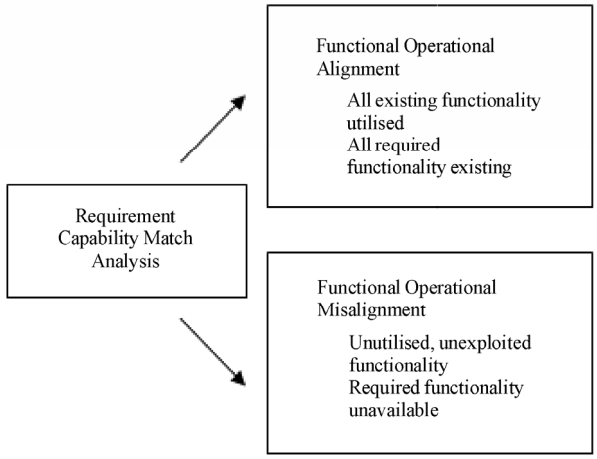

Figure 8. Functional-operational alignment or misalignment

alignment arises when all existing functionality is used and all organisational requirements are met by available functionality. Functional-operational misalignment occurs when existing functionality is unused and organisational requirements are not met by available functionality. Achieving an ideal match between requirements and capabilities is complex. The misalignment experienced across all IoT case study sites was unsurprising given the system's US orientation, the common national standard, the phased approach for system development and changing IoT requirements.

Step Five. Areas of misalignment uncovered are addressed. As outlined in **Figure 9**, this includes:

- System development to better match organisational requirements. This involves development by the central project team for general IoT sector needs. It also requires a degree of in-house tailoring to meet site specific demands.
- Resource investment to exploit necessary unutilised functionality. Resource investment for further system exploitation is important only where unused functionality would be of benefit to the organisation. This includes reinstating functional team members to identify unexploited capabilities and to train and educate users in the broader system functionality.

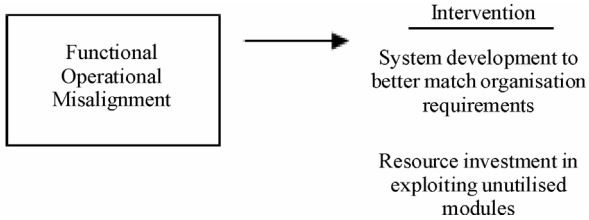

Figure 9. Addressing functional-operational misalignment

The goal of this activity is to achieve greater convergence between organisational requirements and system capabilities, with a view to improving organisational-system compatibility/performance and realisation of greater system benefit potential.

The above guidelines reflect in five steps how the functional-operational misalignment issue may be improved. This model may be viewed as an alternative continuous participative ICT investment evaluation tech nique that may be iteratively applied throughout the system's lifecycle. Through examining the alignment between requirements and capabilities, the IoTs can highlight problematic areas, identify steps for corrective action, set future performance targets and track progress by involving key stakeholder groups.

6. Discussion and Conclusions

ICT investment evaluation is important in identifying and addressing the misalignment that exists between the standard student system's functional capabilities and the diverse operational requirements of the multiple autonomous IoTs. Functional-operational misalignment results from the inability of organisations to exploit system capabilities. There is also an issue regarding whether such a system meets organisational needs. Prior to this research, the impact of functional-operational misalignment in leveraging system benefit potential was not clearly addressed in the literature. Some researchers discussed the match between system capabilities and organisational requirements. For example Avram (2001) [25] highlighted "*definitional mismatch*", i.e. the degree of risk faced by an organisation in terms of how well functional requirements meet operational needs; and Saastamoinen (2005) [26] suggested that the most difficult factor in ICT evaluation is the match between system and processes. However, previous research had not explored this issue in-depth or suggested how it may be addressed.

In this respect, this paper expands the body of existing research. It examined how CPE offers the opportunity to understand and exploit the unused function

ality in a standard system through involving key stakeholder groups. This not only increases stakeholder awareness of evolving technological capabilities and organisational requirements but also promotes a more favourable acceptance to organisational changes through fostering their commitment and motivation. It can lead to enhanced stakeholder relationships through recognising the impact of organisational change and the interactions that exist between people and technology. Hence, through a continuous participative evaluation process that emphasises dialogues and stakeholder feedback, there is potential for improving functional-operational alignment and this is important for enhanced system benefit realisation.

The author's development of this CPE approach is in line with a growing volume of ICT evaluation research that recognises the value of more formative evaluation methods over rational, objective approaches that do not capture contextual depth or address important human complexities. For example, Gemmell and Pagano (2003) [7], who conducted ICT evaluation research in the Higher Education sector argued for the importance of continuous evaluation throughout the project lifecycle. Similarly, several researchers in other sectors discuss the importance of formative methods in identifying the unexpected impacts of ICT investment; monitoring progress, increasing control and improving project performance; identifying continuous improvement opportunities; improving relationships with stakeholders; and increasing benefit realisation (see for example [3–6,8–13]). As evident in the author's discussion in Section 5.4, the most important aspect in addressing the functional-operational misalignment issue is the promotion of stakeholder involvement; enabling more effective ICT decisions on how to develop the system and staff capabilities through ongoing dialogue with affected parties. This is the underlying fundamental principle of widely cited CPE methods such as Remenyi *et al*'s (1997) [15] ABR and Cordoba's (2009) [16] IS Planning Framework. The approach developed in this paper is a further useful step in enhancing this CPE body of research in that it addresses the misalignment issue, which was not previously explored in-depth. It further adds to the limited body of evaluation research in the Higher Education Sector.

7. Avenues of Further Research

This study has taken a small step towards addressing the paucity of ICT evaluation research in HEIs and in providing greater understanding of the challenges in operationalising a standard ICT system in organisations with diverse requirements. However, there are many opportunities for further research in this sector.

- The research findings may be examined in a wider context. Interviews may be conducted with informants in the ten other Institutions. This would provide a basis for comparing issues across the IoT sector and for confirming or extending this study's findings. It would also determine the finding's inferential transferability to a broader setting.
- Secondly, the study's scope could be expanded to include the viewpoints of the central project team and the DoES. Due to the nature of their Student MIS project involvement; it is probable that their project perceptions differ from those directly impacted by system introduction. If this were found to be true, it would make for an interesting examination of the diversity in social groups' perceptions regarding large-scale MIS implementations.
- Thirdly, the model for resolving functional-operational misalignment could be iteratively applied across the IoT sector. This CPE approach would provide for greater stakeholder involvement in determining evolving IoT requirements and system capabilities, and in evaluating the degree of functional-operational alignment. Further, it would focus IoT activities on problematic areas in order to improve system utilisation and benefit realisation. This project would be longitudinal; however, through setting various benchmarks, it would determine the model's usefulness in optimising ICT performance.
- Finally, the research was confined to the Irish education market. Comparative research in other countries would be of benefit. This would determine whether issues, similar to those experienced in the Irish IoTs, were encountered in implementing and operationalising large-scale standard ICT systems, and how such issues were overcome. In this respect it might be interesting to look to the UK or even the USA for comparisons.

REFERENCES

[1] M. Wilson and D. Howcroft, "Power, politics and persuasion in IS evaluation: A focus on 'relevant social groups'," Journal of Strategic Information Systems, Vol. 14, pp. 17–43, 2005.

[2] W. Chen and R. Hirschheim, "A paradigmatic and methodological examination of information systems research from 1991 to 2001," Information Systems Journal, Vol. 14, pp. 197–235, 2004.

[3] S. Cronholm and G. Goldkuhl, "Strategies for information systems evaluation-six generic types," Electronic Journal of Information Systems Evaluation, Vol. 6, No. 2, pp. 65–74, 2003.

[4] C. Miskelly, A. Moggridge, C. Stephenson, and M. Williams, "Formative and interpretive approaches to the evaluation of community focused web-fronted projects," Proceedings of the 11th European Conference on Information Technology Evaluation, Amsterdam, The Netherlands, Academic Conferences, Reading, pp. 281–288, 11–12 November, 2004.

[5] A. Roudsari, E. J. Berridge, M. Vodel, and E. Carson, "Design and evaluation of ICT-based patient information systems to support the management of chronic diseases," Proceedings of the 10th European Conference on Information Technology Evaluation, Madrid, Spain, MCIL, Reading, pp. 547–556, 25–26 September, 2003.

[6] N. Shiratuddin and M. Landoni, "Evaluation of content activities in children's educational software," Proceedings of the 8th European Conference on Information Technology Evaluation, Oxford, United Kingdom, MCIL, Reading, pp. 79–87, 17–18 September, 2001.

[7] M. Gemmell and R. Pagano, "A post-implementation evaluation of a student information system in the UK higher education sector," Electronic Journal of Information Systems Evaluation, Vol. 6, No. 2, 2003.

[8] N. Doherty and L. McAulay, "The relationship between the ex ante and ex post information systems evaluation: Reflections from the literature and the practice of evaluating e-commerce investments," Proceedings of the 9th European Conference on Information Technology Evaluation, Paris, France, MCIL, Reading, pp. 129–136, 15–16 July, 2002.

[9] E. Berghout and M. Nijland, "Full lifecycle management and the IT management paradox," In Remenyi, D. and Brown, A. (eds.), Make or Break Issues in IT Management, Butterworth Heinemann, Oxford, pp. 77–107, 2002.

[10] N. Doherty and M. King, "The treatment of organisational issues in systems development projects: The implications for the evaluation of information technology investments," Electronic Journal of Information Systems Evaluation, Vol. 4, No. 1, 2001.

[11] B. Farbey, F. Land, and D. Targett, "Evaluating investments in IT: Findings and a framework," In Willcocks, L.P. and Lester, S. (eds.), Beyond the IT Productivity Paradox, Wiley, Chichester, pp. 183–215, 1999.

[12] B. Farbey, F. Land, and D. Targett, "Moving IS evaluation forward: Learning themes and research issues," Journal of Strategic Information Systems, Vol. 8, No. 2, pp. 189–207, 1999.

[13] Z. Irani, A. M. Sharif, and P. E. D. Love, "Linking knowledge transformation to information systems evaluation," European Journal of Information Systems, Vol. 14, pp. 213–228, 2005.

[14] L. P. Willcocks, "Evaluating information technology investments: Research findings and reappraisal," Journal of Information Systems, Vol. 2, No. 3, pp. 243–268, 1992.

[15] D. Remenyi, M. Sherwood-Smith, and T. White, "Achieving maximum value from information systems," Wiley, Chichester, 1997.

[16] J. R. Cordoba, "Critical reflection in planning information systems: A contribution from critical systems thinking," Information Systems Journal, Vol. 19, No. 2, pp. 123–147, 2009.

[17] V. Serafeimidis and S. Smithson, "Information systems evaluation as an organisational institution experiences from a case study," Information Systems Journal, Vol. 13, pp. 251–274, 2003.

[18] L. P. Willcocks, G. Fitzgerald, and M. Lacity, "To outsource IT or not? Research on economics and evaluation practice," In Willcocks, L. P. and Lester, S. (eds.), Beyond the IT Productivity Paradox. Wiley, Chichester, pp. 293–333, 1999.

[19] C. E. Hillam and H. M. Edwards, "A case study approach to evaluation of Information Technology/Information Systems (IT/IS) investment evaluation processes within SMEs," Electronic Journal of Information Systems Evaluation, Vol. 4, No. 1, 2001.

[20] M. L. Markus, "Technochange management: Using IT to drive organisational change," Journal of Information Technology, Vol. 19, No. 1, pp. 4–20, 2004.

[21] M. Carcary, "ICT evaluation in the Irish higher education sector," Electronic Journal of Information Systems Evaluation, accepted for publication in Vol. 12, No. 2, 2009.

[22] B. G. Glaser and A. L. Strauss, "The discovery of grounded theory: Strategies for qualitative research," Aldine De Gruyter, New York, 1967.

[23] M. Alvesson and K. Sköldberg, "Reflexive methodology-new vistas for qualitative research," Sage Publications, London, 2000.

[24] B. Czarniawska, "Narratives in social science research-introducing qualitative methods," Sage Publications, Thousand Oaks, London, 2004.

[25] G. Avram, "Evaluation of investments in ecommerce in the Romanian business environment," Electronic Journal of Information Systems Evaluation, Vol. 4, No. 1, 2001.

[26] H. Saastamoinen, "Exception-based approach for information systems evaluation: The method and its benefits to information systems management," Electronic Journal of Information Systems Evaluation, Vol. 8, No. 1, pp. 51–60, 2005.

J. Service Science & Management, **2010, 3: 45-50**
doi:10.4236/jssm.2010.31005 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

End Stage Renal Disease Economics and the Balance of Treatment Modalities

John D. Sullivan

Boston University, Boston, USA.
Email: jsulliva@bu.edu

Received June 25th, 2009; revised August 19th, 2009; accepted October 1st, 2009.

ABSTRACT

Periodically, research articles emerge arguing the economic benefits of peritoneal dialysis, or PD, over the traditional in-center hemo dialysis for patients suffering from End Stage Renal Disease. Resulting conclusions indentify PD as the ideal therapy to reduce Medicare expenditures for this expensive treatment. However, despite this possible economic benefit to the United States taxpayer, the number of PD patients remains relatively flat with an increasing amount of patients being prescribed in-center hemo dialysis. A simplistic view of controlling the rising costs, on a per treatment basis, associated with the treatment of this disease would be to increase the number of patients from in-center hemo dialysis to peritoneal dialysis. This paper will argue why this shift is both unlikely and unrealistic, and why the resulting potential cost savings to this segment of the Medicare program is a myth.

Keywords*: *End Stage Renal Disease, Dialysis, Medicare, Reimbursement, Economics, Mergers, Acquisitions

1. Introduction

The primary function for our kidneys is to remove excess fluids and toxins from the body through urine. The production of urine is a complex process of excretion and reabsorption. Because of this process, kidneys are also known as the body's chemists, balancing and stabilizing salt, potassium, and acid content. Kidneys also produce a potent Vitamin D and hormones that impact the function of other organs. Erythropoietin, one of these significant hormones, acts as stimulation for the production of red blood cells.

Chronic kidney disease, or CKD, impacts the lives of more than 26 million patients in the United States with millions of others at an increased risk of developing the disease [1]. There are five stages of kidney disease tracking renal function with stage 5 representing near or complete renal failure. Progression of the disease is measured through several tests such as the Glomerular filtration rate, urine protein tests, and blood pressure.

As of today, there is no cure for end stage renal disease. Patients with stage 5 renal CKD require either a kidney transplantation or dialysis for continued survival. Without immediate treatment, patients will suffer from uremia, a build up of toxins and fluid in the blood. This condition will lead to death in a matter of days or weeks depending on existing renal function as each kidney deteriorates. Symptoms of uremia include, but are not limited to, nausea, vomiting, loss of appetite, weight loss and ultimately changes in mental status, confusion, reduced awareness, psychosis, seizures, agitation, and coma.

The most common cause of kidney disease is diabetes and hypertension. Incidence of renal failure is increasing as a result of these disorders as well as the aging population. Race can also be a factor in the incidence of ESRD. African Americans, American Indians, Hispanics, Asians, and Pacific Islanders have a higher incidence of renal failure due to higher rates of diabetes and high blood pressure [2,3].

The most ideal treatment for patients with End Stage Renal Disease is transplantation. Transplantation is purely a treatment and not a cure. Patients, even with ideal tissue matches, must take immunosuppressive drugs to ensure that the organ is not rejected by the body's natural immune system. The number of transplants in the United States has increased significantly from the late 1980's as technology and drugs have improved. But since 2003, the total number of transplantations have leveled off at approximately 16,500 per year with the waiting list for a kidney growing to 85,000 [4]. (Figure 1). The transplant rate remains flat due to the lack of available organs. Studies have concluded that transplantation is not only the best option for a patient's quality of life, but also is the most cost effective therapy [5,6].

2. Dialysis Treatment Options

For patients not fortunate enough to receive a kidney transplant, there are two other treatment options: hemodialysis and peritoneal dialysis. Hemodialysis is primarily performed in a dialysis outpatient center although some hospital based programs continue to operate. Patients that utilize this therapy visit the dialysis center three times each week for treatments that will last between 3 1/2 and 4 1/2 hours. Due to the time commitment for hemodialysis and many of the other medical problems that are associated with patients that require this treatment, most patients cannot maintain full-time employment and drop out of the workforce. By far, hemodialysis is the majority of the treatment modality chosen by nephrologists. (Figure 2)

In the past few years, some research has suggested that the treatment time for hemodialysis should be longer if the three treatments per week model are followed [7]. Current literature has moved from this theory to one that prescribes shorter dialysis times, but on a daily basis. But since the present reimbursement structure is based on a thrice a week structure, it is unlikely that this change in treatment time and frequency will occur on a national wide-scale despite the probable benefits to the patient.

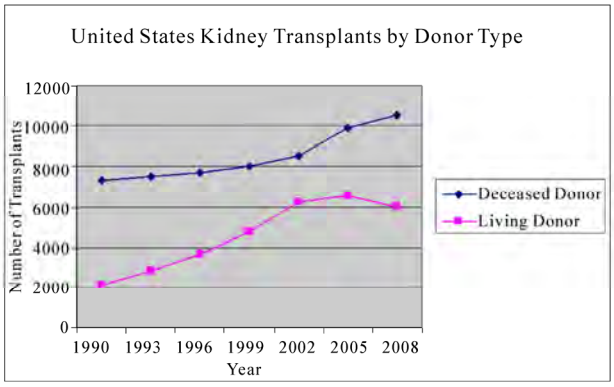

Figure 1. Source: United network for organ sharing

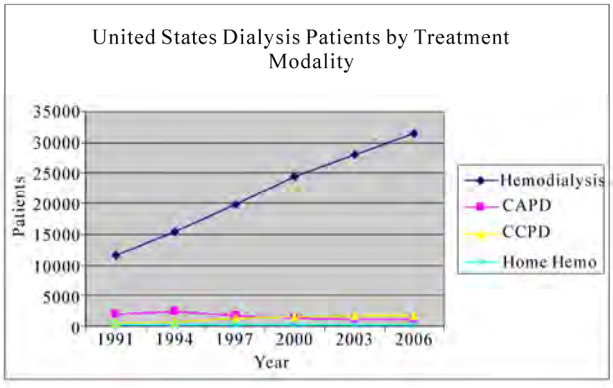

Figure 2. Source: USRDS 2008 annual report

Not all hemodialysis is performed in outpatient or hospital based clinics. A small portion of dialysis patients do receive treatment in their home. Many of these patients utilize machines built by NxStage and undergo extensive training in their treatment. Despite the logical reasoning behind this treatment modality, home hemodialysis doesn't seem to be getting any traction and the number of patients undergoing this treatment is insignificant [8].

No matter in a clinic or in the home setting, hemodialysis requires pumping the blood from the patient to the dialysis machine. The machine pushes the blood along through arterial lines to an artificial kidney where the blood flows against dialysate, a fluid that removes toxins through osmosis. The blood then returns to the body by way of venus lines. To accommodate the flow of blood, patients must undergo surgery to improve vascular access. For all hemodialysis patients, vascular access is both required and presents an ongoing challenge for both nephrologists as well as vascular surgeons. The site often clots and requires hospitalizations and repeated medical procedures to ensure proper blood flow for dialysis.

One of the positive aspects of hemodialysis therapy is the social interaction between patients and the center's staff. Patients in this setting tend to bond and exchange information about their illness and treatment progress. Many patients also prefer the clinic over treatment at home due to the stress and potential medical problems that can occur outside a qualified medical facility.

Peritoneal dialysis, as a therapy, is very different than its hemodialysis counterpart. PD dialyzes a patient through the use of the peritoneal cavity. Each patient has a catheter implanted in their lower abdomen and infuses dialysis fluid into the peritoneal cavity. As the fluid is removed from the abdomen, toxins flow out along with the dialysate. The peritoneal membrane essentially acts as the filter for the dialysis process. There are two types of PD available for patients: Continuous Ambulatory Peritoneal Dialysis (CAPD) and Continuous Cycling Peritoneal Dialysis (CCPD). Not all patients are suited for peritoneal dialysis [9]. Patients must be compliant and careful in their treatment. Both therapies require that patients undergo extensive training, usually at an approved dialysis center. Patients must visit that center frequently to pick up their supplies and to receive any infusion drugs that may be prescribed.

CAPD requires multiple exchanges of dialysate during the day and has no equipment. CCPD is performed at night while the patient sleeps and requires the use of a "cycling" machine. In the past, most patients that were suitable for PD were prescribed CAPD. Presently, mostly as a result of technologically advanced cyclers, the amount of patients on PD has flipped with CCPD being the preferred prescription [10]. Each PD treatment, unlike hemodialysis' thrice a week regime, is performed each day. Reimbursement for this therapy is identical to

hemodialysis whereby the same amount is paid for a three treatment week as the PD seven day a week therapy.

A major advantage for patients that choose PD as their treatment modality is flexibility since trice a week travel is not required for treatment. This would also enforce the shift from CAPD to CCPD given that the therapy is done while the patient sleeps. However, use of the peritoneal membrane as a filter does have its drawbacks for patients. Most patients, at one time or another, can develop peritonitis. This inflammation and infection of the peritoneal cavity is painful and requires patients to convert to hemodialysis while being treated for their infection and many patients cannot convert back to PD.

For both therapies, reimbursement is currently based on a basic treatment with additional payments for ancillary drugs and testing. PD does appear less expensive than hemodialysis. From the payment perspective, PD prescriptions for the basic treatment are identical to hemodialysis treatment although individual non-government payers may have separate payment structures for each therapy. Although the payment structures for hemodialysis and PD are the same for the basic treatment, the difference in cost seems to be ancillary drugs. Patients that are prescribed PD are typically healthier than their hemodialysis counterparts and often don't require additional drugs to maintain a better quality of life [11]. Peritoneal dialysis patients also tend to possess a greater kidney function that would require less external treatments for such disorders as anemia. Research studies conclude that only 20% of PD patients are using EPO for anemia as compared with 80% of hemodialysis patients [12].

2.1 The End Stage Renal Disease Program

In July of 1973, under the direction of the Federal Government, health care in the United States changed dramatically for patients with renal failure. Through Medicare Part B, virtually every citizen, after an initial waiting period, would be entitled to coverage, paid by tax dollars, for this disease regardless of age. The ESRD program has, at times, been hailed as a success story and at other moments in its history, been criticized for the programs growing economic burden and high mortality rates.

Initial results were difficult to assess because of the slow build up of physicians and clinics to meet the needs of ESRD patients across the country. Initially, the argument behind the program was that treatment options, such as transplant, would keep the cost of providing care low. The first year of the life saving treatment cost the government $229 million for approximately 16,000 patients. The current Medicare payment is $20.3 billion for 437, 334 patients [13]. (Figure 3)

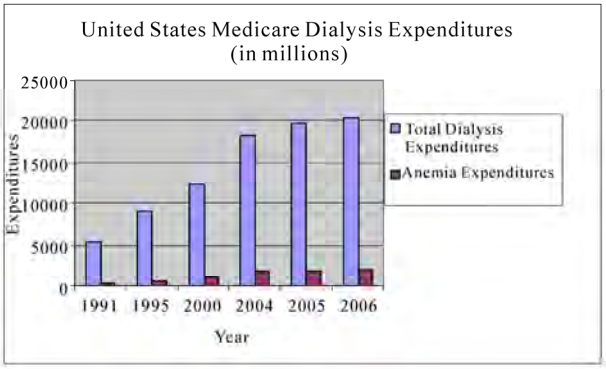

Figure 3. Source: USRDS 2008 annual report

The idea behind dialysis was that it would serve as a model for a universal health care system in the United States. The problem with this theory is that the renal patient community is not an accurate reflection of the mainstream health care needs of the general population. As reflected by the very nature of the constant therapy combined with the numerous hospitalizations, the ESRD program cannot be an example of how universal health care may work. The exploding costs have placed more and more of an economic burden on Medicare consuming 6% of the budget for less than 1% of the Medicare population [14].

2.2 The Dialysis Business Model

The dialysis market, from the providers' side, is basically controlled by two entities: Fresenius Medical Care and DaVita. The combination of these two companies control 82.1% of the top ten providers in the dialysis market [15]. (Figure 4). Other smaller either niche or "mom and pop" clinics control the remaining share along with hospital dialysis based units. For profit clinics dominate the market and have significantly increased their size over the last fifteen years. (Figure 5)

Fresenius and DaVita, by their very business model, are two very distinct companies. Fresenius, or FMC, is a fully integrated company that both provides dialysis care to patients and manufactures the equipment and supplies for that care. FMC is a global health care company with the largest patient base of any provider.

By increasing the production of artificial kidneys, or dialyzers, FMC has virtually ended the process of reprocessing dialyzers to reduce costs through economies of scale production. Reprocessing dialyzers was a common practice in the 1970's through the 1990'sat first due to the lower ability to manufacture enough dialyzers. As production increased, reuse was purely a means to reduce medical supply costs. Studies are inconclusive as to whether or not reuse impacts the quality of treatment or effects mortality rates.

DaVita on the other hand, is a pure dialysis provider

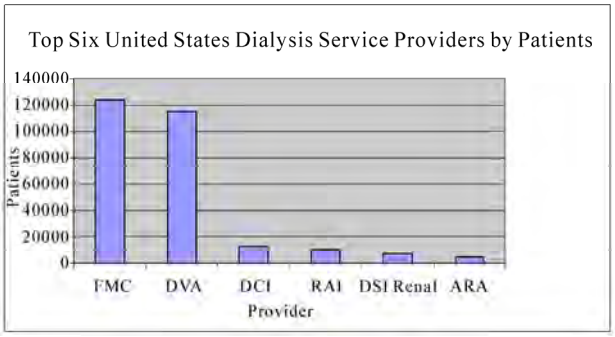

Figure 4. Source: Nephrology news and issues july 2009

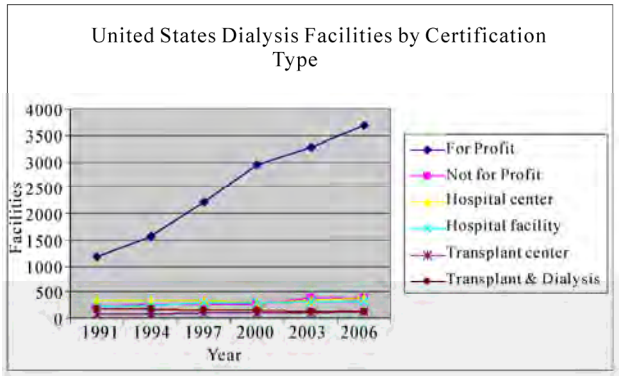

Figure 5. Source: USRDS 2008 annual report

only treating patients in the United States. While not at all of their clinics, many of DaVita's clinics still reuse patient dialyzers. For patients worried about this practice, DaVita maintains information regarding reuse on their website [16].

Consolidation over the past thirty years has been driven by the incentive to cut costs rather than to increase market share and expand. Dialysis, by its very nature, is a business that requires an infrastructure to support multiple clinics with minimum administrative or corporate staff. Mergers and acquisitions add the patient volume and base revenue while administration costs from the target company can be eliminated.

The driving force behind this strategy has been Medicare's inability or unwillingness to raise the composite rate. With a level reimbursement rate from the majority of the patient population combined with inflationary pressures from the medical labor market, providers were forced to look to this "roll up" strategy to maintain profitability. Larger dialysis companies operate at a significant advantage to their smaller competitors. Expenses can be leveraged through economies of scale, but with stagnant reimbursement from government sources, this advantage can be limited. Thus, leverage in revenue is also critical to their advantage over smaller providers. By controlling larger shares of the market, dialysis chains can leverage payments from commercial insurance providers making up more than the difference from the Medicare and Medicaid shortfall.

Outpatient dialysis hemodialysis units are expensive investments with leasehold improvements and equipment ranging from $1.5 to $2.5 million depending on size and location. From a finance perspective, these are large fixed cost businesses [17]. To minimize or leverage these fixed costs, volume of patients is critical to maximize profitability. For profit dialysis centers tend to have higher numbers of patients per unit than non-profit and hospital based units. (Figure 6)

Because of the increase in fixed asset investment by dialysis companies, there may be an incentive to push towards flowing the volume of patients from a PD therapy to hemodialysis. Adding to this pressure of sending renal patients to clinics is the emerging structure within the small to medium size dialysis companies. To motivate younger nephrologists or perhaps to give older physicians the ability to divest a portion of their investment in a dialysis business, these smaller dialysis companies, many supported by private equity investment dollars, is to offer a joint venture arrangement with the nephrologist. Whether a de novo unit or an existing clinic, the physician will have an incentive to prescribe hemodialysis over PD to drive volume and fully leverage their large fixed asset investment.

PD, by its very nature, is a variable cost treatment without the overhead required in a sophisticated outpatient clinic. Labor costs do exist, but are far less utilized as patients usually only visit the clinic once a month to pick up supplies and drop off blood samples. Variable costs for PD patients are a different story. Because PD patients dialyze seven days a week, medical supply costs are significantly higher than in-center hemodialysis patients. The difference in cost experienced by the payer is attributable to the lower amount of prescription drugs.

If PD is indeed a lower cost with similar revenue streams as in-center hemodialysis patients, why is the industry so reluctant to embrace this therapy? The overwhelming majority of the market is controlled by

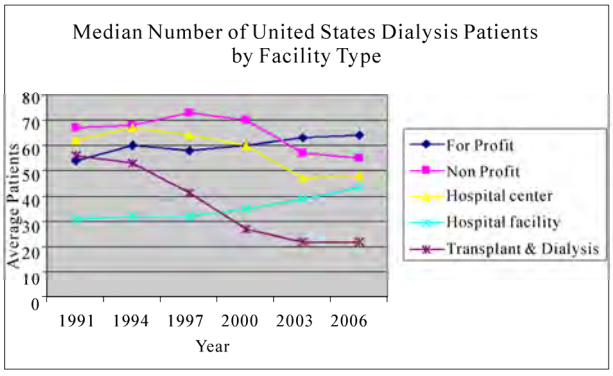

Figure 6. Source: USRDS 2008 annual report

for-profit entities with the top two being publicly traded. Small to mid size other players are backed by private equity firms responsible to their investors. Peritoneal dialysis may be the lowest cost from the payers' perspective, but is it the most profitable to the providers? With the immense pressure from Wall Street and Private Equity Firms for high returns, it at least appears that the significant factor to profitability for dialysis companies is ancillary drugs and not the basic dialysis treatment. Otherwise, capital resources would be directed to invest in areas with the highest potential return. This theory may be complimented by the number of nephrologists that believe in the effectiveness of peritoneal dialysis [18]. Regardless of motive, the simple fact is that the level of PD patients continues to be flat while the in-center hemodialysis patient population is growing [19].

3. Current and Future Reimbursement Structure for Dialysis

The renal community is on the edge of a monumental change in reimbursement by the federal government. The present system is based on a fee for service with additional charges for drugs and medical tests. The new payment structure would entail a "bundled" rate. This rate would be inclusive of both the dialysis treatment and any additional drugs the patient may require. The Medicare Improvements for Patients and Providers Act of 2008, passed on July 15 and survived a veto from President Bush, requires this new bundled system to be in place for ESRD by 2011.

Bundled rate systems are not new to renal providers. Prior to the invention of Amgen's anemia drug, Epogen in 1989, the Medicare composite rate was essentially a "bundled payment" covering nearly all the costs of providing care to patients. Some commercial insurance companies have already moved to a bundled reimbursement as a means to control costs and simplify the billing process.

The reasoning behind this change is likely to control costs associated with additional necessary drugs while creating a high quality environment thorough quality measures in outcomes to ensure that care isn't being reduced under bundling to secure higher profitability [20]. The problem with this approach is that it does not take into account adverse consequences such as providers pushing patients that are non-compliant into competing clinics. Nor does the plan specify how small and medium providers will afford to invest in the systems necessary to gather and analyze this data.

Implementation for this program will start regionally and as it is tested, be rolled out on a nationwide basis. Obviously, geographical areas will play a key role in the reimbursement rate based on the cost of living. It may be likely that adjustments to the bundle would occur as data is collected similar to Medicare's first implementation of the ESRD program in 1973. A report will be issued during the summer of 2009 with more details although much hasn't been disclosed by way of bundled reimbursement by modality.

The question that remains regarding this new structure is what the goals of Medicare will be beyond maintaining a quality program? If it is to ensure profitability for providers, even slightly beyond break-even, Medicare will be forced to increase what they're currently paying for these services [21]. Thus, the bundled rate will likely be very similar to the current reimbursement.

4. Conclusions

The majority of dialysis that is performed in the United States is conducted under the high fixed-cost outpatient dialysis units. Despite significant improvements in the delivery of PD therapy, the numbers of patients utilizing either CAPD or CCPD remains flat as the total renal failure population experiences consistent growth. This is primarily due to the high fixed cost infrastructure already in place, which many physicians don't believe that their patients are suitable for PD, and many nephrologists unwillingness to prescribe PD for treatment. If profitability is the goal of most providers, it would appear that in-center patients, as long as the center is at a high utilization, are more profitable than at home hemodialysis or PD patients. Otherwise peritoneal dialysis would be emphasized as an alternative therapy and the rate of utilization would at least grow at the rate of the incidence of renal failure.

From the payers' perspective, peritoneal dialysis treatment is more cost effective than traditional in-center and home hemodialysis. However, in the current clinical and economic environment, any shift in treatment modality seems unlikely. The pressure from investors, including physicians, emphasizes a volume in-center treatment model that maximizes financial returns. With the aging population combined with higher incidence rates of diabetes and hypertension, the overall cost of the ESRD program will continue to consume a greater portion of the Medicare budget.

REFERENCES

[1] National Kidney Foundation. Chronic kidney Disease (CKD), The facts about chronic kidney disease (CKD), May 26, 2009, http://www.kidney.org/kidneydisease/ckd/index.cfm.

[2] N. Tareen, *et al*., "Chronic kidney disease in African American and Mexican American populations," Kidney International, No. 68, pp. S137–S140, 2005.

[3] N. Burrows, Y. Li, and D. Williams, "Racial and ethnic differences in trends of end stage renal disease: United States 1995 to 2005," Advances in Chronic Kidney Dis-

ease, Vol. 15, No. 2, pp. 147–152.

[4] United Network for Organ Sharing, View Data Reports, National Data, Transplants by Donor Type, Kidney, May 25, 2009, http://optn.transplant.hrsa.gov/latestData/rptData.asp.

[5] K. Howard, *et al.*, "The cost effectiveness of increasing kidney transplantation and home based dialysis," Nephrology, Vol. 14, No. 1, pp. 123–132, 2009.

[6] E. Yen, *et al.*, "Cost effectiveness of extending Medicare coverage of immunosuppressive medications to the life of a kidney transplant," American Journal of Transplantation, Vol. 4, No. 10, pp. 1703–1708, 2004.

[7] R. Saran, *et al.*, "Longer treatment time and slower ultrafiltration in hemodialysis: Associations with reduced mortality in the DOPPS," Kidney International, Vol. 69 No. 7, pp. 1222–1228, April 2006.

[8] United States Renal Data System 2008 Annual Report.

[9] J. D. Sullivan, "Making peritoneal dialysis and home hemodialysis more economically viable," Nephrology News and Issues and Issues, pp. 54–58, July 2006.

[10] United States Renal Data System Annual Data Report, Chapter 2, Incidence and Prevalence, 2008, http://www.USRDS.org.

[11] F. Coronel, *et al.*, "Erythropoietin requirements: a comparative multicenter study between peritoneal dialysis and hemodialysis," Journal of Nephrology, Vol. 16, No. 5, pp. 697–702, September–October, 2006.

[12] J. J. Snyder, R. N. Foley, D. T. Gilberston, E. F. Vonesh, and A. J. Collins, "Hemoglobin levels and erythropoietin doses in hemodialysis and peritoneal dialysis patients in the United States," Journal of the American Society of Nephrology, Vol. 15, No. 1, pp. 174–179, 2004.

[13] United States Renal Data System Annual Data Report, Section K, The Economic Costs of ESRD, pp. 292, 2008, http://www.USRDS.org.

[14] 1999 HCFA statistics. Baltimore, Health Care Financing Administration, April 2000, (HCFA publication no. 03421).

[15] Nephrology News and Issues, July 2009.

[16] http://www.davita.com/dialysis/treatment/a/1867.

[17] J. D. Sullivan, "Making peritoneal dialysis and home hemodialysis more economically viable," Nephrology News and Issues, pp. 54–58, July 2006.

[18] http://www.aakp.org/aakp-library/Fewer-Patients-Peritoneal-Dialysis/.

[19] United States Renal Data System, 2008 Annual Report

[20] J. Sullivan, "Bundling and its potential impact on dialysis service providers," Nephrology News and Issues, October 2008.

[21] J. D. Sullivan, "Bundling and its potential impact on dialysis service providers," Nephrology News and Issues, October 2008.

J. Service Science & Management, **2010, 3: 51-66**
doi:10.4236/jssm.2010.31006 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Scientific Research

Evolution or Revolution of Organizational Information Technology – Modeling Decision Makers' Perspective

Niv Ahituv[1*], Gil Greenstein[2]

[1]Faculty of Management, Tel Aviv University, Tel Aviv, Israel; [2]Faculty of Technology Management, Holon Institute of Technology, Holon, Israel; *Corresponding Author.
Email: ahituv@post.tau.ac.il, gilgr@hit.ac.il

Received December 28th, 2009; revised January 28th, 2010; accepted February 21st, 2010.

ABSTRACT

This paper suggests a new normative model that attempts to analyze why improvement of versions of existing decision support systems do not necessarily increase the effectiveness and the productivity of decision making processes. Moreover, the paper suggests some constructive ideas, formulated through a normative analytic model, how to select a strategy for the design and switching to a new version of a decision support system, without having to immediately run through a mega conversion and training process while temporarily losing productivity. The analysis employs the information structure model prevailing in Information Economics. The study analytically defines and examines a systematic informativeness ratio between two information structures. The analysis leads to a better understanding of the performances of decision support information systems during their life-cycle. Moreover, this approach explains normatively the phenomenon of "leaks of productivity", namely, the decrease in productivity of information systems, after they have been upgraded or replaced with new ones. Such an explanation may partially illuminate findings regarding the phenomenon known as the Productivity Paradox. It can be assumed that the usage of the methodology that is presented in this paper to improve or replace information structure with systematically more informative versions of information structures over time may facilitate the achievement of the following major targets: increase the expected payoffs over time, reduce the risk of failure of new versions of information systems, and reduce the need to cope with complicated and expensive training processes.

***Keywords*:** *Decision Analysis, Decision Support Systems, Productivity and Competitiveness, Information Technology Productivity, the Productivity Paradox*

1. Introduction

A major and continuing problem in the information technology (IT) profession is the high rate of failure of new information systems (IS) or upgraded versions of them. From a rational point of view it may be assumed that IS professionals usually analyze and design IS "properly". But is it really so? Are they aware of the possibility of limits in perception among IS users, especially decision makers? Do they realize that "improvement" of decision support information systems might lead sometimes to a result opposite to what has been expected, namely degradation in the level of the productivity of the firms, since new and unfamiliar decision rules have not been fully implemented and adopted by the decision-makers?

This article suggests a new normative model that attempts to explain that improvement of versions of existing information systems do not necessarily increase the effectiveness and productivity of decision making processes. It also suggests some constructive ideas, formulated through a normative analytic model, how to select a strategy of switching to new version of a system, without having to immediately run through a mega training program, and to take a risk of losing productivity.

The methodological and theoretical foundations for the analysis presented here anchor in the literature on information economics. The earliest mathematical model presenting the relaying of information in a quantitative form was that of Shannon [1]. The model distinguished between two situations:

1) A noise-free system—a univalent fit between the transmitted input data and the received signals;

2) A noisy system—the transmitted input data (denot-

ing a state of nature) are translated into signals probabilistically.

In assigning an expected normative economic value to information, some researchers made use of Microeconomics and Decision Theory tools [2]. The combination of utility theory and the perception that information systems can be noisy led to the construction of a probabilistic statistical model that accords to an information system the property of transferring input data (states of nature) to output (signals) in a certain statistical probability [3–5]. This model, which delineates a noisy information system, is called the ***information structure*** model. It is based on the assumption that a system is noisy but it does not examine the nature of the noise. This paper expands the analysis by examining some patterns of noise. The consequences of that analysis are then demonstrated.

Over the years significant research was conducted to explore aspects of the phenomenon termed by Simon [6] as "***bounded rationality***"[1] and its main derivative—***satisficing*** behavior. Some of its aspects were presented comprehensively by Rubinstein [7]. Ahituv and Wand [8] showed that when satisficing is incorporated into the information structure model, there might be a case where none of the optimal decision rules will be pure anymore (unlike the results of optimizing behavior).

Ahituv [9] incorporated one of the aspects of bounded rationality into the information structure model: the inability of decision-makers to adapt instantaneously to a new decision rule when the technological characteristics of the information system, as expressed by the probabilities of the signals, are suddenly changed. Moreover, Ahituv [10] portrayed a methodology in which decision support systems are designed to act consistently during their lifecycle (in accordance with a constant decision rule). He suggested that this decision rule (that was an optimal decision rule in a previous version of the information system) guarantees improvement of expected outcomes, although it is not necessarily the optimal decision rule for later versions of this information system.

This study presents a conceptual methodology that combines aspects of bounded rationality [9,10] dealing with a rigid decision rule and the life-cycle of information systems, with elements of rational behavior presented in the information structure model [5].

The article raises some questions: Is it possible to improve an existing information system without adopting a new decision rule? What are the analytical conditions that enable a "smooth" (without much disturbance) upgrading or replacement of an information system? In a decision situation where two information structures are activated probabilistically, and one of them is generally more informative than the other, are there analytical conditions encouraging to enhance the percentage of usage of the superior system?

A normative framework is suggested to cope with essential processes (e.g.: implementation processes, correction of bugs, or upgrading of versions) during the life cycle of a decision support system [11]. By defining and analyzing a new informativeness relationship - "the systematic informativeness ratio", this paper demonstrates situations where decision-makers are equipped with partial information. Through these cases, it is explained how to assure a "smooth" implementation of new or upgraded information systems, as well as how to reduce the investment in implementation activities.

Moreover, It is shown that the existence of this new relationship (ratio) between two information structures enables to improve the level of informativeness without the awareness and the involvement of the users (the decision makers).

"The systematic informativeness ratio approach", which is presented and analyzed for the first time in this paper, contributes to better understanding of various aspects of the "productivity paradox" [12–14]. Furthermore, it portrays a methodology that suggests how to deal with some aspects of the "productivity paradox" which were explored in earlier studies [15,16].

The next section summarizes the information structure model and the Blackwell Theorem [5]. It describes the motivation to use convex combinations in order to describe processes during the life cycle of a decision support system. Section 3 describes, analyzes, and demonstrates a new informativeness relationship between two information systems—"***the systematic informativeness ratio***". Section 4 explores the existence of systematic informativeness ratio between un-noisy information structures. Section 5 presents some implications that could be extrapolated to noisy information structures. The last section provides a summary and conclusions, and presents the contribution of the study and the directions it opens for further research. Proofs of the theorems and lemmas appear in the appendix.

2. The Basic Models

2.1 The Information Structure Model and Blackwell Theorem

The source model employed in the forthcoming analysis is the information structure model [5]. This is a general model for comparing and rank ordering information systems based on the rules of rational behavior.[2]

The information structure model enables a comparison of information systems using a quantitative measurement reflecting their economic value. An information structure

[1]Simon termed the human decision-making process, which is affected by bounded rationality as "satisficing", and the decision-maker in accordance as a "satisficer" (aims to be satisfied with his or her decision). This is in contrast to the perception of the decision-maker under rational behavior assumptions in "classical" Utility theory who is an "optimizer" (aims to achieve the best out of his or her decision)

Q_1 is said to be more informative than an information structure Q_2 if the expected payoff of using Q_1 is not lower than the expected payoff of using Q_2. The expected payoff is trace $(\Pi^*Q^*D^*U)$[3], where trace is an operator that sums the diagonal elements of a square matrix. The objective function for maximizing the expected compensation is $\underset{D}{Max}(trace(\Pi * Q * D * U))$[4].

Let us examine a numerical example. Assume that an investment company serves its customers by using a web based information system. Let Q_1 be an information structure that predicts the attractiveness of investing in various alternative channels. The IS supports the decision-making of the investors. For simplicity, suppose there are three categories of states of nature: S_1 - accelerated growth (probability: 0.2), S_2 - stability (probability: 0.6), and S_3 - recession (probability: 0.2). Assume also that there are three possible decisions: A_1 - Invest in bank deposits; A_2 - Invest in stocks; A_3- Invest in foreign currency; Q_1 - The information system provides the following signals: Y_1- Accelerated growth is expected; Y_2 - Stability is expected; Y_3 - Recession is expected;

$$Q_1 = \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.4 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The compensations matrix U, which represents the expected percentage of profit or loss, is described as follows: $U = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 5 & 0 \\ -1 & -1 & 3 \end{pmatrix}$

$$\underset{D}{Max}\left(trace(\Pi * Q_1 * D * U)\right) = \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}$$

$$\begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.4 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} D_{1,1} & D_{1,2} & D_{1,3} \\ D_{2,1} & D_{2,2} & D_{2,3} \\ D_{3,1} & D_{3,2} & D_{3,3} \end{pmatrix}\begin{pmatrix} 3 & 0 & 1 \\ 0 & 5 & 0 \\ -1 & -1 & 3 \end{pmatrix} = 2.4$$

where $D^* = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$[5]. Invest "$A_1$", while the signal is Y_1 or Y_2. Invest "A_3", while the signal is Y_3.

Given two information systems that deal with the same state of nature and are represented by the information structures Q_1 and Q_2, Q_1 is defined as generally more informative[6] Given two information systems that deal with the same state of nature and are represented by the information structures Q_1 and Q_2, Q_1 is defined as generally more informative.[7] The rank ordering is transitive.[8]

Over the years, a number of researchers developed analytical models to implement the concept of the *information structure model* in order to evaluate the value of information technology. Ahituv [10], demonstrated the life cycle of decision support information system with the model. Ahituv and Elovici [17] evaluated the value of performances of distributed information systems. Elovici

[2]According to the information structure model, four factors determine the expected value of information.

The a priori probabilities of pertinent states of nature. Let S be a finite set of n states of nature: $S=\{S_1,..,S_n\}$. Let P be the vector of a priori probabilities for each of the states of nature: $P=(p_1,..,p_n)$.

The information structure – a stochastic (Markovian) matrix that transmits signals out of states of nature. Let Y be a finite set of n signals, $Y=\{Y_1,..,Y_m\}$. An ***information structure*** Q is defined such that its elements obtain values between 0 and 1, Q: $SxY\rightarrow[0,1]$. $Q_{i,j}$ is the probability that a state of nature S_{ii} displays a signal Y_j $\sum_{j=1}^{m} Q_{i,j}=1$

The decision matrix – a stochastic matrix that links signals with the decision set of the decision-maker. Let A be a finite set of k possible decisions, $A=\{A_1,..,A_k\}$. Let D be the decision function. Similar to Q, D is a stochastic (Markovian) matrix, namely, it is assumed that the decision selected for a given signal is not necessarily always the same. $D:Y\ x\ A\rightarrow[0,1]$

The payoff matrix – a matrix that presents the quantitative compensation to the decision-maker resulting from the combination of a decision chosen and a given state of nature. Let U be the payoff function: $U: A\ x\ S \rightarrow \Re$ (a combination of a state of nature and a decision provides a fixed compensation that is a real number). $U_{i,j}$ – is the compensation yields when decision maker decides –"A_i", while state of nature "S_j" occurs.

[3]Sometimes Q represents an un-noisy (noise free) information structure. In these cases Q represents an information function f, $f: S \rightarrow Y$ [4]. Q is a stochastic matrix that contains elements of 0 or 1 only. This means that for each state of nature the information structure will always act identically (will produce the same signal), although it is not guaranteed that the state of nature will be exclusively recognized.

[4]When the utility function is linear, that is, the decision-maker is of the type EMV [2], a linear programming algorithm may be applied to solve the problem, where the variables being the elements of the decision matrix D. It can be proved that at least one of the optimal solutions is in a form of a decision matrix whose elements are 0 or 1 (a pure decision rule), [5]. For numerical illustrations of the model, see [8–10].

[5] D^* is a decision matrix which represent the optimal decision rules in this decision situation.

[6]It should be noted that when we deal with two information functions rather than structures we use the term "fineness" to describe the general informativeness ratio [4].

[7]In terms of the information structure model, if for every possible payoff matrix U, and for every a priori probability matrix Π $\underset{D}{Max}\left(trace(\Pi^*Q_1^*D^*U)\right) \geq \underset{D}{Max}\left(trace(\Pi^*Q_2^*D^*U)\right)$, then Q_1 is generally more informative than Q_2, Denoted: $Q_1 \geq Q_2$. Blackwell Theorem states that Q_1 is generally more informative than Q_2 if and only if there is a Markovian (stochastic) matrix R such that $Q_1^*R = Q_2$. R is termed the ***garbling*** matrix

[8]It should be noted that the general informativeness ratio is a partial rank ordering of information structures. There is not necessarily rank order between any two information structures.

et al [18] used this method to compare performances of Information Filtering Systems. Ahituv and Greenstein [15] used this model to assess issues of centralization vs. decentralization. Aronovich and Spiegler [19] use this model in order to assess the effectiveness of data mining processes.

The model was expanded to evaluate the value of information in several aspects: the value of a second opinion [20]; the value of information in non-linear models of the Utility Theory [21]; analyzing the situation of case dependent signals (the set of signal is dependent on the state of nature, [22]); a situation of a two-criteria utility function [23]. The model was also implemented to evaluate empirically the value of information in postal services [24], and in analysis of Quality Control methods [25,26].

2.2 The Use of *Convex Combinations*[9] of Information Structures to Represent Evolution during Their Life-Cycle

A possible reason why we should consider probabilistic combination of information systems is the existence of decision support systems that use Internet (or intranet) based search engines. These engines can retrieve information from several information sources, and produce signals accordingly. The various sources are not always available.

Information sources are essential for the proper survivability of competitive organizations. As a result, the importance of proper functioning of information systems is increasing. When a certain source in unavailable it is possible to acknowledge the users about it by alarming them with a special no-information signal [15]. Another option, which is presented in this paper, is to consider implementation of a "mixture" of information systems. For example: suppose there is "a state of the art" organizational information center that can serve, during peak times only 90% of the queries. How will the rest 10% are served? One alternative is to reject them.[10] Another one is to direct those queries to a simpler (perhaps cheaper) information system whose responses are less informative. This leads to consider probabilistic usage of information systems that can be delineated by a convex combination.

The analysis focuses on the convex combination of information structures reflecting a probabilistic employment of a variety of information systems (structures), where the activation of each one of them is set by a given probability. The various systems react to the same states of nature and produce the same set of signals.[11]

The mechanism of convex combinations[12] of information systems is employed in an earlier research by Ahituv and Greenstein [15] which analyses the effect of probabilistic availability of information systems on productivity, and illuminates some aspects of the phenomenon that are termed as "the productivity paradox" [12–14].

3. The Systematic Informativeness Ratio

3.1 Definition of the Systematic Informativeness Ratio

As mentioned in Section 2, when an information structure Q_1 is more informative than an information structure Q_2 irrespective of compensations and a priori probabilities, a general informativeness ratio exists between the two of them [5].

If an information structure Q_1 is more informative than Q_2 when the optimal decision rule of Q_2 is employed, and given some certain a priori probabilities of the states of nature, then under some assumptions on the payoffs, an informativeness ratio under a rigid decision rule is defined between them [9].[13]

[9]The convex combination of information structures was discussed in earlier studies. Marschak [4] notices that the level of informativeness of convex combination of two information structures (denoted Q_1 and Q_2) which produce the same set of signals is not equivalent to the level of informativeness of using Q_1 with a probability p and Q_2 with the complementary probability (1-p).

[10]Sulganik [27] indicates that a convex combination of information structure could be used to describe experimental processes (with a probability p of success and (1-p) of failure). For example, he investigates the convex combination of two information structures: one presents perfect information and the other one no-information (its rows are identical).

[11]It should be noted that in case that the two information structures do not produce the same set of signals, the non-identical signals can be represented in by columns of zeroes respectively [9].

[12]A convex combination of two information systems is defined as follows: Let Q_1 and Q_2 be two information structures describing information systems. Let $S=\{S_1,\ldots,S_n\}$ be their set of the states of nature. Let $Y=\{Y_1,\ldots,Y_m\}$ be their set of signals. When a decision situation is given let p the probability that Q_1 will be activated, and (1-p) that Q_2 will be activated. Since, decision makers do not aware which information structure is activated, Q_3, the weighted information structure, is represented by a convex combination of Q_1 and Q_2.

$Q_3 = p^* Q_1 + (1-p) * Q_2$

[13]Given two information systems that deal with the same states of nature, produce the same set of signals, and are represented by the information structures Q_1 and Q_2 respectively, Q_1 will be considered more informative than Q_2 under a rigid decision rule if its expected payoff is not lower than that of Q_2 for the following conditions:

$\forall i, i = 1,..,n$. Let $U(k(i),i)$ the single maximum payoff when the state of nature S_i occurs.

Denote:

$\bar{u}^* = \max_i(U(k(i),i)); \ \underline{u}^* = \min_i(U(k(i),i)); \ \bar{u}0 = \max_{i,k\neq k(i)}(Uk,i); \ \underline{u}0 = \min_{i,k\neq k(i)}(Uk,i);$

$q1^* = (q1(1,(k(1)),\ldots,q1(n,(k(n)))^t \ q2^* = (q2(1,(k(1)),\ldots,q2(n,(k(n)))^t$, respectively.

$\bar{\delta}^* = \bar{u}^* - \bar{u}0; \quad \underline{\delta}^* = \bar{u}^* - \underline{u}0; \quad \underline{\delta}0 = \bar{u}0 - \underline{u}0;$

The theorem which is proved by Ahituv [9], states that:

If $\Pi^t(\underline{\delta}^* q1^* - \bar{\delta}^* q2^*) \geq \delta 0$ then Q_1 more informative than Q_2 with regard to U *and* Π.

The ratio will be denoted: $Q_1 \underset{R}{\geq} Q_2$

The informativeness ratio under a rigid decision rule is a partial rank ordering of information structures.

Assume those two informativeness ratios can be conceptually combined to a new informativeness ratio: Let Q_1 and Q_2 be two information structures that deal with the same state of nature and produce the same set of signals. Q_1 will be considered systematically more informative than Q_2 if for any decision situation (for any a priori probabilities vector-Π and any payoff matrix-U), its expected payoff is not lower than that of Q_2 while Q_1 operates under an optimal decision rule of Q_2. In terms of the information structure model, this is presented hereinafter by Definition 1.

Definition 1: Let Q_1 and Q_2 be two information structures representing two information systems operating on the same set of states of nature $S = \{S_1,\ldots,S_n\}$ and producing the same set of signals $Y = \{Y_1,\ldots,Y_m\}$. Q_1 is defined ***systematically more informative*** than Q_2, denote $Q_1 \underset{S}{\geq} Q_2$, if for any decision situation (irrespective of payoffs and a priori probabilities) Q_1 is more informative than Q_2 under an optimal decision rule of Q_2.

It means that if Q_1 is systematically more informative than Q_2, then for every decision situation[14] there exists an optimal decision rule of the inferior information structure Q_2, that can be used with the superior information structure Q_1, and guarantees at least the optimal outcomes of using Q_2.

Mathematically it looks this way:

$$\exists D_{Q2} \in \{D\max(Q_2)\}, trace(\Pi * Q_1 * D_{Q2} * U) \geq$$

$$\underset{D}{Max}(trace(\Pi * Q_2 * D * U)) = trace(\Pi * Q_2 * D_{Q_2} * U))$$

where $\{D\max(Q_2)\}$ - denotes the set of optimal decision rules, when $Q2$ is activated in this specific decision situation.

In contrast to the general informativeness ratio, in the systematic informativeness ratio the information structure Q_2 can be replaced with the superior systematically information structure Q_1, without an immediate awareness of the decision makers (the users), since the decision rule does not necessarily have to be changed instantaneously. It means that when the systematic informativeness ratio exists between two information structures, at least the same level of expected payoffs is guaranteed when the superior[15] information structure is activated. Hence the decision maker does not have to adopt a new optimal decision urgently.

Let us now examine the informativeness ratio between two information systems from the point of view of "smooth" implementation. This is presented in Definition 2.

Definition 2: Let Q_1 and Q_2 be two information structures representing two information systems operating on the same set of states of nature $S=\{S_1,\ldots,S_n\}$ and producing the same set of signals $Y=\{Y_1,\ldots,Y_m\}$. Assume Q_1 is generally more informative than Q_2. A ***smooth implementation*** of Q_1 instead of Q_2 is defined if for any level of usage p $(0 \leq p \leq 1)$ $p * Q_1 + (1-p) * Q_2 \geq Q_2$.

The importance of this ratio is that in any probabilistic level of usage of the superior information system Q_1, the mean of the expected payoffs (compensation) that the decision-makers gain is not less than that achieved by using only the inferior information system. It contributes to a smooth implementation of the superior information structure Q_1.

In our study we argue that those definitions (1 & 2) are equivalent. Theorem 1 proves analytically the equivalence of Definition 1 and 2.

Theorem 1[16]

Let Q_1 and Q_2 be two information structures operating on the same set of states of nature $S = \{S_1,\ldots,Sn\}$, and producing the same set of signals $Y = \{Y_1,\ldots,Ym\}$. Then

$$Q_1 \underset{S}{\geq} Q_2 \quad \Leftrightarrow \quad \forall p, 0 \leq p \leq 1, \quad p * Q_1 + (1-p) * Q_2 \geq Q_2$$

This theorem shows that the two ratios which have been defined above are identical. Replacement (or improvement, or upgrade) of an information structure with a more systematically informative, information structure than it, guarantees smooth implementation, and vice versa.

Moreover, from the aforementioned equivalence it is understood that during a smooth implementation of the superior information structure Q_1[17], we do not have to adopt a new decision rule, and we can stick to an optimal decision rule we used in the past with the inferior information structure Q_2. In fact this theorem sets a new normative perspective that defines the necessary and sufficient conditions for the ability to implement a superior information structure smoothly without immediate interference in the routine work of decision makers.

Using this method facilitates information systems professionals to plan systems under the assumption that during a certain transition period the decision-makers may act identically and stick to the same decision-rule [10]. The existence of this informativeness ratio reduces the criticality of an urgent implementation process.

3.2 A Framework to Examine the Existence of the Systematic Informativeness Ratio

In order to identify the existence of a systematic informativeness ratio between two information structures when one of them is generally more informative than the other, we would analyze a special case in which the number of signal and the number of states of nature are identical. In this case, the identity square matrix I is a complete and perfect information structure. We will try to find out whether I is systematically more informative than any other square stochastic matrix of similar

[14]A given set of a-priori probabilities - Π, and a given utility matrix - U.

[15]Systematically more informative than the other.

[16]The proof is provided in the appendix.

[17]Since $\forall p, 0 \leq p \leq 1, \quad p * Q_1 + (1-p) * Q_2 \geq Q_2$, $Q_1 \underset{S}{\geq} Q_2$.

dimensions.

The motivation to do this is provided by Lemma 1. Assume two information structures Q_1 and Q_2 act on the same set of states of nature, and respond with the same set of signals, and $Q_2 = Q_1 * R$, where R is a stochastic matrix (Blackwell Theorem's condition). In Lemma 1 it is shown that the existence of systematic informativeness ratio between I and R sets a pre-condition (sufficient condition) to the existence of the general informativeness ratio between Q_1 and Q_2.

Lemma 1[18]

Let Q_1 and Q_2 be two information structures operating on the same set of states of nature $S = \{S_1,\ldots,S_n\}$, and producing the same set of signals $Y = \{Y_1,\ldots,Y_m\}$**.** Assume that Q_1 is generally more informative than Q_2, implying that $Q_2 = Q_1 * R$, where R is a stochastic matrix [5]**.**

If $\forall p, 0 \le p \le 1, \quad p * I + (1-p) * R \ge R$,

Then $\forall p, 0 \le p \le 1, \quad p * Q_1 + (1-p) * Q_2 \ge Q_2$

From Lemma 1 it can be shown that if Q_1 is generally more informative than Q_2, namely $Q_2 = Q_1 * R$ (R *is a stochastic matrix) and* $I \underset{S}{\ge} R$ (I is systematically more informative than R) then $Q_1 \underset{S}{\ge} Q_2$ (Q_1 is systematically more informative than Q_2).

3.3 The Monotony of the Systematic Informativeness Ratio

The following lemma deals with the improvement of the accuracy level of information systems by enhancing the probability to receive perfect information.

Lemma 2[19]

Let I be an information structure that provides perfect information. Let Q be any information operating on the same set of states of nature $S = \{S_1,..,S_n\}$ and producing the same set of signals $Y = \{Y_1,\ldots,Y_n\}$**.**

If: for $0 \le p \le 1, p*I + (1-p)*Q \ge Q$ (every convex combination of I and Q is generally more informative than Q)

Then:

$$\forall q, 0 \le q \le p \le 1, p*I + (1-p)*Q \ge q*I + (1-q)*Q$$

Conclusion:

If: $\forall p, \ 0 \le p < 1, p*I + (1-p)*Q \ge Q$

Then:

$$\forall p, \forall q, \ 0 \le q \le p < 1, p*I + (1-p)*Q \ge q*I + (1-q)*Q$$

This lemma proves the monotony of the systematic informativeness ratio. Actually it is shown that an improvement in the accuracy level of information (expressed by increasing the probability of perfect information) is positively correlated with the general informativeness ratio of a convex combination.

3.4 The Systematic Informativeness Ratio – An Example

We continue the example of Q_1, an Information system for choosing an investment option, which was first demonstrated in Section 2. $Q_1 = \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.4 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Suppose it is intended to replace the information system with an improved one, Q_2: $Q_2 = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.9 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Due to technological and organizational limitations, e.g.: inability to implement the system simultaneously all across the organization and the need to monitor carefully the system's performances, the system is implemented step by step.

By using some of the lemmas and theorems presented above, it can be demonstrated that the information structure Q_2 is systematically more informative than Q_1.

Let Q_0 be an information system: $Q_0 = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Let Q_3 be an information structure, which represents perfect information. $Q_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Since,

$\forall p, \ 0 \le p \le 1, \quad (p * Q_3 + (1-p) * Q_0) * Q_0 = Q_0$ [20]

$$\left(p * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (1-p) * \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) * \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then, from Theorem 1 it is clear that Q_3 is systematically more informative than Q_0.

Let us present Q_1 and Q_2 as convex combination of Q_3 and Q_0.

$$Q_2 = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.9 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 0.8 * Q_3 + 0.2 * Q_0$$

[18]The proof and an example are provided in the Appendix.

[19]The proof is provided in the Appendix.

[20]We use the information structure Q_0, as a garbling (stochastic) matrix, either.

$$= 0.8*\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + 0.2*\begin{pmatrix}0.5&0.5&0\\0.5&0.5&0\\0&0&1\end{pmatrix}$$

$$Q_1 = \begin{pmatrix}0.6&0.4&0\\0.4&0.6&0\\0&0&1\end{pmatrix} = 0.2*Q_3 + 0.8*Q_0$$

$$= 0.2*\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} + 0.8*\begin{pmatrix}0.5&0.5&0\\0.5&0.5&0\\0&0&1\end{pmatrix}$$

According to Lemma 2 Q_2 is systematically more informative than Q_1. **Table 1** demonstrates the implication of the existence of the systematic informativeness ratio between Q_2 and Q_1.

This example illustrates that if an upgrading of an information system is based on the implementation of later versions of it which are systematically more informative than the earlier versions, then sticking to the old and familiar decision rule will not harm productivity.

The principle of developing information systems to be systematically more informative provides the luxury of training and on-site implementation which is "life-cycle independence". It facilitates the implementation of a new version of information system or insertion of minor changes, without the immediate awareness of the decision makers. Therefore, organizations can schedule the optimal timing of wide training processes. That is in contrary to the usual situation, when the scheduling of training and on-site implementation might interfere with other organizational considerations and requirement (e.g.: periodically tasks). Moreover, development of new decision support systems without adopting this principle may explain, normatively, "leaks of productivity". In other words it may explain the decrease in user performance of information systems, although they have been improved. This degradation in the expected outcomes while using improved information systems can be attributed to the inability of users to adapt immediately to new decision rules.

4. The Systematic Informativeness Ratio – A Noise Free Scenario.

4.1 Conditions for Existence of the Systematic Informativeness Ratio

Historically, the starting point for analyzing the value of information in noisy information structures was the analysis of the value of information in noise free information structures. These are also termed **information functions** [4,5]. Following this approach, we will start with a simple presentation of the informativeness ratio between in formation functions, which could be classified as unnoisy information structures.

In order to identify the existence of a systematic informativeness ratio between two information functions, a new aggregation ratio (a fineness ratio that keeps orders of signals) between information functions is defined, hereinafter.

Definition 3

Let f_I be the identity information function. Let $S = \{S_1,..,S_n\}$ be its set of states of nature and $Y = \{Y_1,..,Y_n\}$ the set of signals f_i produces. Hence, $f_I(S_i) = Y_i$.

Assume for simplicity that $S = \{1,..,n\}$, $Y = \{1,..,n\}$ then $f_I(i) = i$

Table 1. Comparison of two information structures, one of them is systematically more informative than the other

	The current information structure	The improved information structure	
Information structure	$Q_1 = \begin{pmatrix}0.6&0.4&0\\0.4&0.6&0\\0&0&1\end{pmatrix}$	$Q_2 = \begin{pmatrix}0.9&0.1&0\\0.1&0.9&0\\0&0&1\end{pmatrix}$	
The matrix of a priori probabilities for the states of nature + The matrix of compensation (percentage):	$\Pi = \begin{pmatrix}0.2&0&0\\0&0.6&0\\0&0&0.2\end{pmatrix}$ $U = \begin{pmatrix}3&0&1\\0&5&0\\-1&-1&3\end{pmatrix}$		
The Decision rules: *A1* – Invest in Bank Deposits *A2* – Invest in stocks *A3* – Invest in foreign currency	$D = \begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix}$	$D = \begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix}$	$D = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$
Expected compensation (percentage)	2.4	2.4	3.12

Let F be a **set** of information functions that acts on the same decision environment of f_I .

$$g \in F \Leftrightarrow \left(g(i)=i \quad or \quad if \quad g(i)=k\neq i, \ then \ \ g(k)=k\right)$$

Since there is an isomorphism between the representation of information functions and a set of information structures representing them, F can be defined analogically as a set of un-noisy information structures.

$$F_1 \in F \Leftrightarrow (F1_{i,i}=1 \quad or \quad if \quad F1_{i,k}=1, k\neq i, \ then \ \ F1_{k,k}=1)$$

Following that f_I is equivalent (for example) to I, an information structure that produces perfect information I is represented by the identity Matrix of the order nxn.

In fact F is a complete set of information functions that could be termed as aggregations of f_I . If $g \in F$ its way of transforming of states of nature into signals does not contradict the way f_I transforms states of nature into signals. In other words it can be said that f_I is a higher resolution of every information function belonging to the set F .

Theorem 2 sets the necessary and sufficient conditions for the existence of systematic informativeness ratio between I (the identity information structure), and any non-noisy information structure:

Theorem 2[21]

Let *F1* be an information function. Let $S=\{S_1,..,S_n\}$ be its set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be its set of signals. Let I be the identity information function, which represent perfect information, then $I \underset{S}{\geq} F_1$ **if and only if** $F_1 \in F$.

It is shown in Theorem 2 that for every information function (non-noisy information structure) $F1$, the necessary and sufficient condition that I is systematically more informative than $F1$, is that $F1^2 = F1$. In fact, four equivalent conditions are found as we will show below.

Let $F1$ be an information function. Let $S=\{S_1,..,S_n\}$ be its set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be its set of signals. The following conditions are equivalent.[22]

1) $I \underset{S}{\geq} F_1$

2) $\forall p, \ 0 \leq p \leq 1 \quad p*I+(1\text{-}p)*F_1 \geq F_1$

3) $F_1 \in F$

4) $F_1^2 = F_1$

The equivalence of the first and second conditions (which was demonstrated earlier by Definition 1 and 2 respectively) was proven by Theorem 1. Conditions 1 and 2 are not specific to un-noisy information structures, and can hold for any type of structure. In contrast to Conditions 1 and 2, the third and fourth conditions are relevant only to un-noisy information structures. By using those two latter conditions we can explicitly classify un-noisy and diagonal information structures into two separate classes:

1) Structures that the identity information structure is systematically more informative than them,

2) Structures that the identity information structure is not systematically more informative than them.

4.2 The Implications of the Systematic Informativeness Ratio – An Example

In the example that follows, two scenarios are presented, analyzed, and compared. The first scenario: upgrading an un-noisy information structure F_1 to *I*—the identity information structure, while *I* is systematically more informative than F_1.

The second scenario: upgrading an un-noisy information structure F_2 to *I*—the identity information structure, while *I* is **not** systematically more informative than F_2.

We use the situation of choosing an investment alternative, that was shown earlier in Section 3, except the fact that the current information structures are F_1 or F_2 respectively.

Suppose an un-noisy information structure that predicts the attractiveness of an investment in various channels is installed. This information structure does not distinguish between S_1 – accelerated growth, and S_2 – stability. It is intended to replace the information system with I - an information function that provides perfect information: $I=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$

Due to technological and organizational limitations, e.g.: inability to implement the system simultaneously all across the organization and the need to monitor carefully the system's performance, the system is implemented step by step.

First Scenario: The existing information structure is $F1$.

$$F_1=\begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix} \Rightarrow F_1^2=\begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix}*\begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0\\1&0&0\\0&0&1\end{pmatrix}=F_1 \Rightarrow F_1 \in F$$

Since $F_1 \in F$, it can be shown from Theorem 2 that $I \underset{S}{\geq} F_1$. **Table 2** demonstrates that while the probability

[21] The proof is provided in the Appendix.

[22] While, Theorem 1 proves the equivalence between expressions 1&2, Theorem 2 proves the equivalence between expressions 3&4, and then proves the equivalence between expressions 1&3.

for perfect information increases, the expected compensation increases too.

Second Scenario: Suppose the decision situation is identical to the previous one, but instead of F_1 the existing information structure is F_2 where:

$$F_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow F_2^2$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \neq F_2 \Rightarrow F_2 \notin F$$

Since $F_2 \notin F$, according to Theorem 2 I is not systematically more informative than F_2.

Table 3 demonstrates that although the probability of perfect information increases the level of informativeness declines.

The comparison between those two scenarios is demonstrated in **Figure 1**:

By observing the aforementioned example it can be concluded that, when a new (improved) information system is systematically more informative than the current information system two important goals are achieved:

1) "Decision situation independence" -The ability to implement the information system step by step and to improve the level of informativeness is guaranteed.

2) "Life-cycle independence" -The ability to implement the information system without interfering the users (the decision makers) and while existing expected outcomes are guaranteed (without the necessity to start training and testing processes).

5. Towards Assessing the Systematic Informativeness Ratio between Noisy Information Structures – the Dominancy of Trace

A characteristic of F is that its diagonal elements are (weakly) dominant (in accordance with Definition 4). From Theorem 3 it can be shown that this characteristic is a necessary condition for existence of the systematic informativeness ratio between I and Q:

Theorem 3[34]

Let I (the identity matrix), and Q be two information structures. $S = \{S_1,..,S_n\}$ is the common set of states of nature, and $Y = \{Y_1,..,Y_n\}$ is their same set of signals they produce.

$I \underset{S}{\geq} Q \Rightarrow \forall i, i = 1,..,n,\ Q_{i,i} \geq Q_{j,i}\ i \neq j$ (The diagonal elements are weakly dominant in each and every column).

This theorem implies that the dominancy of the diagonal elements in each and every column of an information structure is a necessary condition for the existence of a systematic informativeness ratio between the identity information structure (which represents complete information) and the non-identity information structure. This casts a preliminary condition for the existence of the informativeness ratio.

The following example demonstrates, by using Theorem 3, that the systematic informativeness ratio is not always transitive.

$$F_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, F_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Table 2. Expected compensation in various levels of prob. for perfect information (1st scenario)

Characteristics of the Decision situation	The probability to receive I.	The probability to receive F_1	Expected compensation
A-priory probabilities: (0.2,0.6,0.2)	0	1	2.6
	0.2	0.8	2.76
Perfect information $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	0.4	0.6	2.92
	0.6	0.4	3.08
	0.8	0.2	3.24
Partial information $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	1	0	3.4

Table 3. Expected compensation in various levels of probability of perfect inf. (2nd scenario)

Characteristics of the Decision situation	The probability to receive I	The probability to receive F_2	Expected compensation
A-priory probabilities: (0.2,0.6, 0.2)	0	1	2.6
	0.2	0.8	2.52
Perfect information $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	0.333 (1/3)	0.666 (2/3)	2.4666
	0.4	0.6	2.56
	0.6	0.4	2.84
	0.8	0.2	3.12
Partial information $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	1	0	3.4

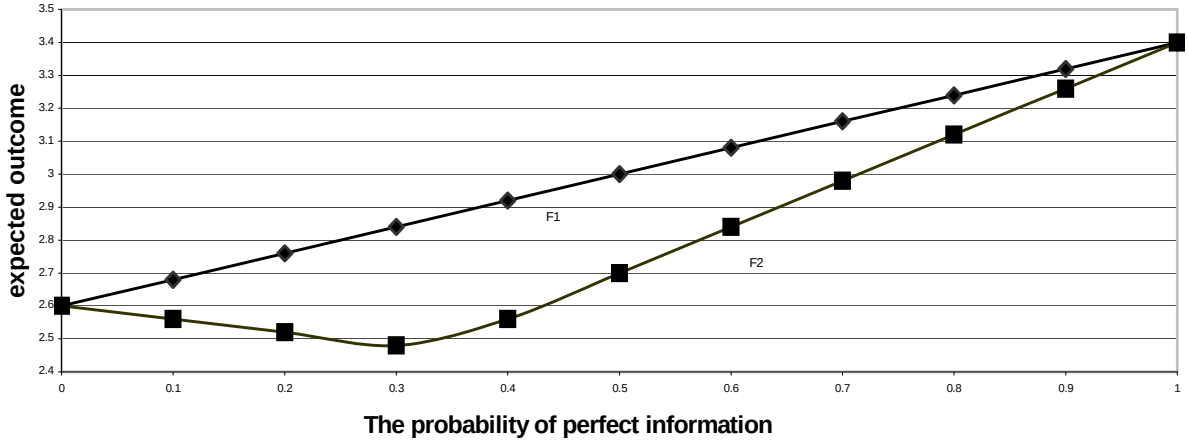

Figure 1. A comparison between the two scenarios

$$Q_1 = 0.25 * I + 0.75 * F_1 = 0.25 * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$+0.75 * \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.75 & 0.25 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$Q_2 = Q_1 * \left(0.4 * I + 0.6 * F_2\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0.75 & 0.25 & 0 \\ 0 & 0 & 1 \end{pmatrix} *$$

$$\left(0.4 * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 0.6 * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \right) =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0.75 & 0.25 & 0 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.75 & 0.25 & 0 \\ 0 & 0.6 & 0.4 \end{pmatrix}$$

Since $F_1 \in F$, from Theorem 2 it is concluded that $I \underset{S}{\geq} Q_1$. Moreover, since $F_2 \in F$, from Lemma 1 it is concluded that $Q_1 \underset{S}{\geq} Q_2$. However, from Theorem 3, it is concluded that since $Q_{2_{3,2}} > Q_{2_{2,2}}$, I is not systematically more informative than Q_2.

Since the systematic informativeness ratio is not always transitive, when there is a multi stage implementation and improvement program during the life-cycle of a decision-support information system and the informativeness ratio of this information system can be improved systematically, the preservation of systematically informativeness ratio is not automatically guaranteed during the whole life-cycle of information system. Hence, the importance of a long-range perspective arises. This can be achieved in one of two ways, depending on the ability to guarantee whether the last version of information structure can be systematically more informative than any previous version, or only superior to its predecessor version:

When the systematic informativeness ratio can be obtained between each and every two sequential versions of an information system during its lifecycle, then a long-range plan of the versioning mechanism is required. This could guarantee that the latest version of an information system will be systematically more informative than any of the previous versions. Moreover, it will guarantee a growth (or at least stability) in expected outcomes during the lifecycle of the decision support system, without alerting the decision makers. Hence, implementation and training processes between versions of the information system become less critical.

When the systematic informativeness ratio can be achieved only between the last version of an information system and its predecessor version, then a training and implementation plan is required. However, the existence of systematic informativeness ratio between consequent versions reduces the costs and lowers the criticality of the implementation and training processes.

6. A Summary and Conclusions

This paper analytically examines and identifies the systematic informativeness ratio between two information structures. The methodological approach presented here may lead to a better understanding of the performances of decision support information systems during their life-cycle.

This approach may explain, normatively, the phenomenon of "leaks of productivity". In other words it may explain the decrease in productivity of information systems, after they have been improved or upgraded. This degradation in the expected outcomes can be explained by the inability of the users to adapt immediately to new decision rules.

It can be assumed that the usage of the methodology that was presented in this paper to improve or replace information structure with systematically more informative versions of information structures over time may facilitate the achievement of the following major targets:

1) Increase the expected payoffs over time.

2) Reduce the risk of failure of new information systems as well as new versions of information systems.

3) Reduce the need to cope with complicated and expensive training processes during the implementation stages of information systems (as well as the implementation of new versions of the systems). Moreover, sometimes this process can be completely skipped during the installation of a new version of an information system.

The paper analyzes the conditions for the existence of a systematic informativeness ratio between I -the identity information structure which represents complete information, and another information structure. In the case of non-noisy information structures the necessarily and sufficient conditions for existence of the systematic informativeness ratio between I and a second information structure are set and proved comprehensively. As a result, some necessary and sufficient conditions are set, proved and demonstrated for the noisy information environment as well.

Further research can be carried out in some directions:

1) Exploration of additional analytical conditions for the existence of the systematic informativeness ratio between I, the identity information structure and noisy information structures.

2) Classification of cases where the systematic informativeness ratio inheritably exists by using the conditions those are set so far.

3) Devising empirical methods to examine the impact of using the principle of developing decision support information systems is systematically more informative over time, on the performance of decision-makers, as well as on their perceived satisfaction from using those systems.

4) Designing empirical studies (experiments, case studies and surveys) to validate the theoretical analysis provided here.

REFERENCES

[1] C. E. Shannon and W. Weaver, "The mathematical theory of communication," University of Illinois Press, Urbana, Illinois, 1949.

[2] H. Raiffa, "Decision analysis," Reading, Addison-Wesley, Massachusetts, 1968.

[3] G. A. Feltham, "The value of information," The Accounting Review, Vol. 43, No. 4, pp. 684–696, 1968.

[4] J. Marschak, "Economic of information systems," Journal of the American Statistician Association, Vol. 66, pp. 192–219, 1971.

[5] C. B. McGuire and R. Radner, (editors), "Decision and organization (ch. 5)," University of Minnesota Press, 2nd edition, Minneapolis, Minnesota, 1986.

[6] H. A. Simon, "The new science of management decisions," Harper and Row, New-York, 1960.

[7] A. Rubinstein, "Modeling bounded rationality," MIT. Press, Cambridge, Massachusetts, 1998.

[8] N. Ahituv and Y. Wand, "Comparative evaluation of information under two business objectives," Decision Sciences, Vol. 15, No. 1, pp. 31–51, 1984.

[9] N. Ahituv, "A comparison of information structure for a 'Rigid Decision Rule' case," Decision Science, Vol. 12, No. 3, pp. 399–416, 1981.

[10] N. Ahituv, "Describing the information system life cycle as an adjustment process between information and decisions," International Journal of Policy Analysis and Information Systems, Vol. 6, No. 2, pp. 133–145, 1982.

[11] G. Ariav and M. J. Ginzberg, "DSS design: A systemic view of decision support," Communications of the ACM, Vol. 28, No. 10, pp. 1045–1052, 1985.

[12] E. Brynjolfsson, "The productivity paradox of information technology," Communications of the ACM, Vol. 36, No. 12, pp. 67–77, 1993.

[13] E. Brynjolfsson, "Paradox lost," CIO, Vol. 7, No. 14, pp. 26–28, 1994.

[14] E. Brynjolfsson and L. M. Hitt, "Beyond the productivity paradox," Communications of the ACM, Vol. 41, No. 8, pp. 49–55, 1998.

[15] N. Ahituv and G. Greenstein, "Systems inaccessibility and the productivity paradox," European Journal of Operational Research, Vol. 161, pp. 505–524, 2005.

[16] M. C. Anderson, R. D. Banker, and S. Ravindran, "The new productivity paradox," Communications of the ACM, Vol. 46, No. 3, pp. 91–94, 2003.

[17] N. Ahituv and Y. Elovici, "Evaluating the performance of an application running on a distributed system," Journal of the Operational Research Society, Vol. 52, pp. 916-927, 2001.

[18] Y. Elovici, B. Shapira, and P. B. Kantor, "Using the information structure model to compare profile-based information filtering," Information Retrieval, Vol. 6, No. 1,

pp. 75–97, 2003.

[19] L. Aronovich and I. Spiegler, "CM-tree: A dynamic clustered index for similarity search in metric databases," Data & Knowledge Engineering, Vol. 3, pp. 919–946, 2007.

[20] N. Ahituv and B. Ronen, "Orthogonal information structure—a model to evaluate the information provided by a second opinion," Decision Sciences, Vol. 19, No. 2, pp. 255–268, 1988.

[21] Z. Safra and E. Sulganik, "On the nonexistence of blackwell's theorem-type, results with general preference relations," Journal of Risk and Uncertainty, Vol. 10, pp. 187–201, 1995.

[22] E. Sulganik and I. Zilcha, "The value of information the case of signal dependent opportunity sets," Journal of Economic Dynamics and Control, Vol. 21, pp. 1615–1625, 1997.

[23] B. Ronen and Y. Spector, "Theory and methodology evaluating sampling strategy under two criteria," European Journal of Operational Research, Vol. 80, pp. 59–67, 1995.

[24] N. Carmi and B. Ronen, "An empirical application of the information-structures model: The postal authority case," European Journal of Operational Research, Vol. 92, pp. 615–627, 1996.

[25] N. Margaliot, "Selecting a quality control attribute sample: An information-economics method," Annals of Operations Research, Vol. 91, pp. 83–104, 1999.

[26] B. Ronen, "An information-economics approach to quality control attribute sampling," European Journal of Operational Research, Vol. 73, pp. 430–442, 1994.

[27] E. Sulganik, "The value of information under diverse economic environments," Ph. D. Dissertation, Tel Aviv University, Faculty of Management (Hebrew), Tel Aviv, 1993.

Appendix

Theorem 1:

Let Q_1 and Q_2 be two information structures operating on the same set of states of nature $S = \{S_1,\ldots,S_n\}$, and producing the same set of signals $Y = \{Y_1,\ldots,Y_m\}$. Then

$$Q_1 \underset{S}{\geq} Q_2 \iff \forall p, 0 \leq p \leq 1,\ p*Q_1 + (1-p)*Q_2 \geq Q_2$$

First, Lemma 1.1 is proven.

Lemma 1.1:

Let Q_1 and Q_2 be two information structures describing information systems. Let $S = \{S_1,\ldots,S_n\}$ be their set of the states of nature of Q_1 and Q_2. Let $Y=\{Y_1,\ldots,Y_m\}$ be their set of signals. Then for any given decision situation described by Π (a matrix of a-priori probabilities of states of nature), U (a matrix of utilities or compensations), A (a set of decisions), where $\{D_{pure\text{-}\max}(Q_2)\}$ is the set of optimal decision rules when Q_2 is used, there exists $\varepsilon>0$, such that if $0<p\leq\varepsilon$, and Dp is an optimal decision rule of the Information structure $p*Q_1+(1-p)*Q_2$

Then $D_p \in \{D_{pure\text{-}\max}(Q_2)\}$.

Proof (of Lemma 1.1):

1) It can be assumed that every optimal decision rule is a convex combination of pure decision rules [10]. So we try to find the optimal decision rule of $p*Q_1+(1-p)*Q_2$ in the set of the optimal pure decision rules of Q_2, $D_p \in \{D_{pure\text{-}\max}(Q_2)\}$.

2) Let k be the number of possible decisions in this given decision situation. This means that there are k^m pure decision rules, denoted $D_1,..,D_{k^m}$.

Let $\{D_{pure}\}$ the full set of the possible pure decision rules for this given decision situation

3) If $\{D_{pure\text{-}\max}(Q_2)\} = \{D_{pure}\}$, that means that every pure decision rule is an optimal decision rule it is obvious that $D_p \in \{D_{pure\text{-}\max}(Q_2)\}$.

4) So, assume that $\{D_{pure\text{-}\max}(Q_2)\} \underset{\neq}{\subset} \{D_{pure}\}$.

5) Hence: $\{D_{pure}\} \setminus \{D_{pure\text{-}\max}(Q_2)\} \neq \phi$

6) Let's calculate for every pure strategy D_i the following values: $V_{1i} = trace(\Pi * Q_1 * D_i * U)$,

$V_{2i} = trace(\Pi * Q_2 * D_i * U)$

7) $trace(\Pi*(p*Q_1+(1-p)*Q_2)*D_i*U)) =$

8) $= p*trace(\Pi*Q_1*D_i*U)) + (1-p)*(trace(\Pi*Q_2*D_i*U)) =$

9) $= p*V_{1i}+(1-p)*V_{2i}$

10) Let's define in this specific decision situation:

- $V_1 \max = \underset{D}{Max}\{V_{1i}\}$ - The (optimal) expected value when using the information structure Q_1.
- $V_1 = \underset{Di\in Dpure\text{-}\max(Q2)}{Max}\{V_{1i}\}$ - The (optimal) expected value when using the information structure Q_1, when the set of decision rule is limited to the optimal set of pure decision rules when using the information structure Q_2.
- $V_{2\max} = \underset{Di\in Dpure\text{-}\max(Q2)}{Max}\{V_{2i}\}$ - The (optimal) expected value when using the information structure Q_2.
- $V_2 = \underset{Di\notin Dpure\text{-}\max(Q2)}{Max}\{V_{2i}\}$ - The (optimal) expected value when using the information structure Q_2, when the set of decision rule is limited to the non-optimal set of pure decision rules when using the information structure Q_2.

11) According to expression (4) $\{D_{pure\text{-}\max}(Q_2)\} \underset{\neq}{\subset} \{D_{pure}\}$

Hence: $\Delta V_2 = (V_{2\max} - V_2) > 0$

12) Moreover: $\Delta V_1 = (V_{1\max} - V_1) \geq 0$

13) Let's examine for $Di \in \{D_{pure}\} \setminus \{D_{pure\text{-}\max}(Q_2)\}$ when it is not an optimal decision rule of $p*Q_1+(1-p)*Q_2$. We try to identify a small value of p that will always give $p*V_{1i}+(1-p)*V_{2i} < p*V_1+(1-p)*V_{2\max}$. In fact, the purpose is to find an "environment" of Q_2 where an optimal decision rule of Q_2 is also an optimal decision rule of $p*Q_1+(1-p)*Q_2$.

14) From (9) it is concludes that

$$p*V_{1i}+(1-p)*V_{2i} \leq p*V_{1\max}+(1-p)*V_2$$

15) Let's examine weather exists:

$$p*V_{1\max}+(1-p)*V_2 < p*V_1+(1-p)*V_{2\max}$$

16)
$$\Leftrightarrow p*(V_{1\max}-V_1)+p*(V_{2\max}-V_2) < V_{2\max}-V_2$$
$$\Leftrightarrow p*(\Delta V_1+\Delta V_2) < \Delta V_2$$
$$\Leftrightarrow 0 < p < \frac{\Delta V_2}{\Delta V_1+\Delta V_2}$$

That's according to (10), (11)

$$\Delta V_1 \geq 0,\quad \Delta V_2 > 0,\quad \Delta V_1+\Delta V_2 > 0$$

17) Let's pick: $0 < \varepsilon < \frac{\Delta V_2}{\Delta V_1+\Delta V_2} \leq 1$

18) And in this environment (for every $0 \leq p \leq \varepsilon$) at least one optimal decision rule of $Q2$ is an optimal decision rule of $p*Q_1+(1-p)*Q_2$

Q.E.D (Lemma 1.1)

Proof (of the theorem itself):

First direction: assume that for every decision situation:

1)
$$\exists D_{Q2} \in \{D_{\max}(Q_2)\}, \text{trace}(\Pi*Q_1*D_{Q2}*U) \geq \underset{D}{\text{Max}}(\text{trace}(\Pi*Q_2*D*U))$$

where $\{D_{\max}(Q_2)\}$ is the set of optimal decision rules

when Q_2 is used in this specific decision situation.

Then implies:

2) $\underset{D}{Max}(trace(\Pi*(p*Q_1+(1-p)*Q_2)*D*U))=$

3) $=\underset{D}{Max}(p*trace(\Pi*Q_1*D*U)$
$+(1-p)*(trace(\Pi*Q_2*D*U))$

4) $\geq p*trace(\Pi*Q_1*D_{Q2}*U)$
$+(1-p)*trace(\Pi*Q_2*D_{Q2}*U)$

5) $\geq p*trace(\Pi*Q_2*D_{Q2}*U)$
$+(1-p)*trace(\Pi*Q_2*D_{Q2}*U)=$

6) $=\underset{D}{Max}(trace(\Pi*Q_2*D*U))$ (first direction is proven)

Second Direction:

7) $\forall p,0\leq p\leq 1, p*Q_1+(1-p)*Q_2\geq Q_2$

8) According to the lemma there exists ε>0, such that if 0<p1≤ε<1, then exists Dp_1 *an* optimal decision rule of the IS: $p_1*Q_1+(1-p_1)*Q_2$ that implies

$Dp_1\in\{D_{pure-\max}(Q_2)\}$.

9) Let D_{Q_2} this optimal decision rule

$D_{Q2}\in\{D_{pure-\max}(Q_2)\}$

10) $\underset{D}{Max}(trace(\Pi*(p_1*Q_1+(1-p_1)*Q_2)*D*U))=$

11) $=(trace(\Pi*(p_1*Q_1+(1-p_1)*Q_2)*D_{Q2}*U))=$

12) $p_1*trace(\Pi*Q_1*D_{Q_2}*U)$
$+(1-p_1)trace(\Pi*Q_2*D_{Q_2}*U)\geq$

13) $\geq trace(\Pi*Q_2*D_{Q_2}*U)$ (According to (7))

14) $=p_1*trace(\Pi*Q_2*D_{Q_2}*U)$
$+(1-p_1)trace(\Pi*Q_2*D_{Q_2}*U)$

15) From (12), (13), (14) $\Rightarrow$ $trace(\Pi*Q_1*D_{Q_2}*U)$
$\geq trace(\Pi*Q_2*D_{Q_2}*U)$

That is correct for every decision situation (any given Π and U)

Q.E.D

Lemma 1

Let Q_1 and Q_2 be two information structures operating on the same set of states of nature $S=\{S_1,\ldots,S_n\}$, and producing the same set of signals $Y=\{Y_1,\ldots,Y_m\}$. Assume that Q_1 is generally more informative than Q_2, implying that $Q_2=Q_1*R$, where R is a stochastic matrix [5].

If $\forall p,0\leq p\leq 1,\ p*I+(1-p)*R\geq R$,

then $\forall p,0\leq p\leq 1,\ p*Q_1+(1-p)*Q_2\geq Q_2$

Proof:

1) According to the second condition of Blackwell's theorem [5] for every p there exists R_p, where R_p is a stochastic matrix of the order nxn.

$(p*I+(1-p)*R)*R_p=R$

2) Therefore: $Q_1*(p*I+(1-p)*R)*R_p=Q_1*R$

3) Hence: $(p*Q_1+(1-p)*Q_2)*R_p=Q_2$

Q.E.D.

Lemma 2:

Let I be an information structure that provides perfect information. Let Q be any information operating on the same set of states of nature $S=\{S_1,..,S_n\}$ and producing the same set of signals $Y=\{Y_1,..,Y_n\}$.

If: for $0\leq p\leq 1, p*I+(1-p)*Q\geq Q$ (every convex combination of I and Q is generally more informative than Q)

Then:

$\forall q,0\leq q\leq p\leq 1, p*I+(1-p)*Q\geq q*I+(1-q)*Q$

Proof:

1) According to the 2nd condition of Blackwell's theorem, $\exists R_p$, R_p is a stochastic matrix, such that: $(p*I+(1-p)*Q)*R_p=Q$

2) Since, $p\geq q$, $\left[\frac{q}{p}*I+\frac{p-q}{p}*R_p\right]$ is a stochastic matrix.

3) Let's examine:

4) $[p*I+(1-p)*Q]*[\frac{q}{p}*I+\frac{p-q}{p}*R_p]=$
$=[p*\frac{q}{p}*I+(1-p)*Q*\frac{q}{p}*I]$
$+\frac{p-q}{p}*[p*I+(1-p)*Q]*R_p=$

5) $=q*I+\frac{q-p*q}{p}*Q+\frac{p-q}{p}*Q=q*I$
$+\frac{q-p*q+p-q}{p}*Q=q*I+(1-q)*Q$

6) According to the 2nd condition of Blackwell's theorem $(5\Rightarrow)$

$\forall q,0\leq q\leq p\leq 1, p*I+(1-p)*Q\geq q*I+(1-q)*Q$

Q.E.D

Theorem 2:

Let F_1 be an information function. Let $S=\{S_1,..,S_n\}$ be its set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be its set of signals. Let I be the identity information function, which represent perfect information, then $I\underset{S}{\geq}F_1$ if and only if $F_1\in F$

First 3 lemmas are demonstrated and proven:

Lemma 2.1

Let fI be the identity information function. Let $S=\{S_1,..,S_n\}$ be its set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be its sets of signals. $f_I(S_i)=Y_i$.

Without loosing generality (for the sake of simplicity)

Assume $S=\{1,..,n\}$, $Y=\{1,..,n\}$.

Let F be the set of information functions (without garbling of signals) that f_I is systematically more informative than each one of them:

$g \in F \Leftrightarrow (g(i)=i \quad or \quad (if \quad g(i)=k\neq i, \ then \ \ g(k)=k)$

Let $g_1, g_2 \in F$ and g_1 is finer than g_2 then:

$\forall i, i=1,..,n, \ g_2(g_1(i))=g_2(i)$

Proof (of Lemma 2.1):

1) Let's check all the possible situations, given:

$g \in F \Leftrightarrow (g(i)=i \quad or \quad (if \quad g(i)=k\neq i, \ then \ \ g(k)=k)$

2) 1st Case: $g_1(i)=g_2(i)=i \Rightarrow g_2(g_1(i))=g_2(i)=i$

3) 2nd Case: $g_1(i)=g_2(i)=k\neq i \Rightarrow g_2(k)$
$=k=g_2(g_1(i))=g_2(k)=k$

4) 3rd Case: $g_1(i)=i, \ g_2(i)=k\neq i \Rightarrow g_2(g_1(i))=g_2(i)=k$

5) 4th Case: $g_1(i)=k\neq i, g_2(i)=i \Rightarrow g_1(k)=k$

6) Since g_1 is finer than g_2: $g_2(k)=i$

7) Hence: $(5),(6) \Rightarrow g_2(g_1(i))=g_2(k)=i=g_2(i)$

8) 5th Case:

$g_1(i)=k\neq i, g_2(i)=j\neq i \Rightarrow g_1(k)=k, \ g_2(j)=j$

9) Moreover, since g_1 is finer than g_2: $g_2(k)=j$

10) Hence: $(8),(9) \Rightarrow g_2(g_1(i))=g_2(k)=j=g_2(i)$

11) It is proved for any possible situation that:
$\forall i, i=1,..,n, \ g_2(g_1(i))=g_2(i)$

Q.E.D (Lemma 2.1)

Following that an Adaptation to the information structure model is concluded straight forward: Let f_I be the identity information function. Let $S=\{S_1,..,S_n\}$ be its set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be its sets of signals. $f_I(S_i)=Y_i$.

Without loosing generality (for the sake of simplicity)

Assume $S=\{1,..,n\}$, $Y=\{1,..,n\}$. Let F be the set of information functions (without garbling of signals) that f_I is systematically more informative than each one of them:

Let $g_1, g_2 \in F$ and g_1 is finer than g_2. g_1 is equivalent to G_1, and g_2 is equivalent to G_2.

Then $G_1 * G_2 = G_2$

Lemma 2.2:

Let F_1, F_2 and F_3 be information structures. F_1 represents information functions accordingly. Let $S=\{S_1,..,S_n\}$ be their set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be their set of signals.

$$F_1=F_2=F_3 \Leftrightarrow \forall p, 0<p<1,$$
$$F_1=p*F_2+(1-p)*F_3$$

Proof (of Lemma 2.2):

1) First direction - Assume: $F_1=F_2=F_3$ then necessarily: $\forall p, 0<p<1, \ F_1=p*F_2+(1-p)*F_3$.

2) Second direction - Assume:

$\forall p, 0<p<1, \ F_1=p*F_2+(1-p)*F_3$. without loosing generality suppose (on the negative form) there exists an index $i, F_{1i,j}<F_{2i,j}$.

3) Then (by calculating):

$F_{3i,j}=\frac{1}{1-p}*(F_{1i,j}-p*F_{2i,j})<0$. It is a contradiction.

Q.E.D (Lemma 2.2)

Lemma 2.3

Let F_1 be an information structure, which represents information function. Let $S=\{S_1,..,S_n\}$ be its set of states of nature. Let $Y=\{Y_1,..,Y_n\}$ be its set of signals.

Then: $F_1 \in F \Leftrightarrow F_1^2=F_1$

Proof (of Lemma 2.3):

1) 1st direction: From the definition of F, and Lemma 2.1 it is obvious that $F_1 \in F \Leftrightarrow F_1=F_1^2$

2) 2nd direction: Assume (on the negative form) $F_1^2=F_1$ and $F_1 \notin F$

3) Let's examine f_1, which is described by F_1. $f_1:\{1,..,n\}\rightarrow\{1,..,n\}$

4) $F_1 \notin F$, Hence there exist an index i, $f_1(i)=j$, $f_1(j)\neq j$

5) $f_1(i)=j=f_1(f_1(i))\neq j$, a contradiction.

6) Hence, $F_1 \in F$

Q.E.D (Lemma 2.3)

Proof of the theorem:

1) 1st direction: it is clear from Lemma 2.3 that
$F_1 \in F \Rightarrow (q*I+(1-q)*F_1)*F_1$
$=q*F_1+(1-q)*F_1*F_1=F_1$

2) Hence: $\forall q, 0\leq q\leq 1 \quad q*I+(1-q)*F_1 \geq F_1$

3) From Theorem 1 it is proven that,

$$\forall q, 0\leq q\leq 1 \quad q*I+(1-q)*F_1 \geq F_1 \Leftrightarrow I \underset{S}{\geq} F_1$$

4) 2nd direction: From Theorem 1 it is proven that,

$$\forall q, 0\leq q\leq 1 \quad q*I+(1-q)*F_1 \geq F_1 \Leftrightarrow I \underset{S}{\geq} F_1 \quad .$$

5) According to the 2nd condition of Blackwell's theorem [5], there exists a stochastic matrix R,

$$F_1=(q*I+(1-q)*F_1)*R.$$

6) $F_1=q*R+(1-q)*F_1*R$

7) From Lemma 2.2 it is obvious that: $R=F_1$

8) Moreover, from Lemma 2.2 it is understood that $F_1*R=F_1$.

9) Hence: $F_1=F_1^2$.

Q.E.D

Theorem 3

Let I (the identity matrix), and Q be two information structures. $S=\{S_1,..,S_n\}$ is the common set of states of nature, and $Y=\{Y_1,..,Y_n\}$ is their same set of signals they respond with.

$I \underset{S}{\geq} Q \Rightarrow \forall i, i=1,..,n, \ Q_{i,i} \geq Q_{j,i} \ i\neq j$ (The diagonal

elements are weakly dominant in each and every column).

Proof:

1) Suppose (on the negative way): There exists an index j , $Q_{i,i} < Q_{j,i}\ i \neq j$ then $1 \geq Q_{j,i} - Q_{i,i} = \Delta > 0$ where (without loosing generality) $Q_{j,i}$ is the maximal element in the i column.

2) Let's examine a specific decision situation: suppose there are n possible decisions, and

$$\forall k,k = 1,...,n,\ \ \Pi_{k,k} = \frac{1}{n}.$$

Let's define U (the utility matrix) as follows:

$$\forall r,r = 1,...,n,\ \ \forall s,s = 1,...,n$$

$$U_{r,s} = \begin{cases} A, & (A > \Delta),\ r = j, s = i \\ A - \Delta, & r = i, s = i \\ \Delta, & r = i, s = j \\ 0, & else \end{cases}$$

3)

$$\underset{D}{Max}\left(trace\left(\Pi * Q * D * U\right)\right) = \frac{1}{n} * \underset{D}{Max}\left(trace\left(U * Q * D\right)\right)$$

$$\forall r,r = 1,...,n,\ \ \forall s,s = 1,...,n\ \ \left(Q*U\right)_{r,s} =$$

4)
$$= \sum_{m=1}^{n} Q_{r,m} * U_{m,s} = \begin{cases} A * Q_{i,s}, & r = j \\ \Delta * Q_{j,s} + \left(A - \Delta\right) * Q_{i,s}, & r = i \\ 0, & else \end{cases}$$

5)

$$\left(Q*U\right)_{j,i} = A * Q_{i,i} < \left(Q*U\right)_{i,i} = \Delta * Q_{j,i} + \left(A - \Delta\right) * Q_{i,i} =$$

$$= \Delta * \left(Q_{i,i} + \Delta\right) + \left(A - \Delta\right) * Q_{i,i} = A * Q_{i,i} + \Delta^2$$

6) Suppose D_1 represents the optimal decision rule. $(5) \Rightarrow D_{1_{i,i}} = 1$.

7) Moreover, $\sum_{\substack{m=1 \\ m \neq i}}^{n} \left(Q*U\right)_{j,m} = A * \left(1 - Q_{i,i}\right)$

.

8) From (5), (7) it is derived that:

$$trace\left(\Pi * Q * D_1 * U\right) \geq \frac{1}{n} * \left(A * \left(1 - Q_{i,i}\right) + A * Q_{i,i} + \Delta^2\right)$$

$$= \frac{1}{n} * \left(A + \Delta^2\right)$$

9) Moreover, $D_{1_{i,i}} = 1 \Rightarrow D_{1_{i,j}} = 0$

10) Hence,
$$trace\left(\Pi * I * D_1 * U\right) \leq \frac{1}{n} * \left(U_{i,j} + U_{i,i}\right)$$
$$= \frac{1}{n} * \left(\Delta + A - \Delta\right) = \frac{1}{n} * A$$

11) It means that:

$$trace\left(\Pi * I * D_1 * U\right) < trace\left(\Pi * Q * D_1 * U\right)$$

12) Under this decision situation Q is precisely more informative than *I*. Hence I is not systematically more informative than *Q*.

Q.E.D

J. Service Science & Management, **2010, 3: 67-71**
doi:10.4236/jssm.2010.31007 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Transformation of Semantic Analysis to Com+ Business Requirements Using MDA Approach

Mohammad Yamin[1], George Tsaramirsis[2], Manos Nistazakis[3], Naibai Zhang[2]

[1]King's Abdelaziz University, Jeddah, Saudi Arabia; [2]King's College London, London, UK; [3]Middlesex University, London, UK.
Email: myamin@kau.edu.sa, {george.tsaramirsis, naibai.zhang}@kcl.ac.uk, E.Nistazakis@mdx.ac.uk

Received September 11th, 2009; revised October 26th, 2009; accepted December 2rd, 2009.

ABSTRACT

Model Driven Architecture (*MDA*) *is one of many competing techniques that have the potential to contribute towards the development of better software systems. However the business support at Computation-Independent Model* (*CIM*) *level does not necessarily allow the development teams to utilize its power and it might become difficult to bridge the gap between analysis, design and implementation phases. Consequently, the clients usually do not get what they want. Semantic Analysis expressed as ontology charts offers a very sound business analytical method that can be used to overcome these problems. The aim of this paper is, with the help of Meta Object Facility* (*MOF*) *meta-model, to advocate the use of Semantic Analysis and to present a method of transforming it to the Component Object Model* (*COM+*) *architecture, which is the Microsoft standard for developing application programs.*

Keywords*: *Organizational Semiotics, MEASUR's Semantic Analysis, Ontology Chart, Component Diagram, COM+

1. Introduction

Poor analysis and design are the main causes of faulty systems. Software development, without proper business analysis and design, results in a poor software product having many problems and huge ongoing maintenance costs. Some of these problems become visible during the system implementation whereas some others become visible after the software system goes into production. Due to these and other problems discussed by some, e.g. [1], many software projects fail, at least partially, and most of them do not meet their requirements, quality and cost benchmarks. One of the most prominent reasons for failing to meet the expected objectives is the lack of appropriate requirement analysis, see [1]. Even if the proper requirement analysis is carried out, the end product would still have many problems if the developers ignore the dictates of requirement analysis and adopt an improper transformation methodology and architecture design.

Traditionally, the designers have been carrying the system analysis on black/white boards and/or papers. The resulting sketches are often discarded once their objective of core-system development was achieved. This practice still continues in many organizations. Unfortunately, implementation of the systems, that is the act of writing code for the system, is considered productive, whereas writing models and properly documenting the changes is viewed, by many, as unnecessary and overhead. In the absence of the sketches, models and charts, maintenance and up-gradation of systems becomes a hugely complex task which often results in a premature demise of many systems. A universally agreed method of making changes to an existing system specifies that the designers must refer to the manuals of analysis and design of the original system and understand the chronology of changes that may have taken place earlier. Failures to adhere to these objectives for making changes in the existing systems would result in system failure, at least partially. These objectives can be easily achieved if we follow the Model Driven Architecture (MDA) in which models themselves are primary artifacts in the development of software. The MDA not only helps to provide a smooth core-system development but is equally efficient for managing ongoing changes in organizational requirements.

Semantic Analysis Method (SAM), according to Stamper "leads to robust software constructed from reusable design component and also reduces development, support and maintenance costs". Once the Semantic

Analysis is completed, for demonstrating the behavior of the systems, it is expressed in the form, usually a number, of Ontology Charts. There are certain rules for drawing the Ontology Charts which are very helpful to describe the relationship amongst the constructs. To improve the communication between the business analyst and IT domains, a special communicational model is constructed in which communication acts are displayed. These communication models help the SAM phase of system development life cycle.

The main aim of this paper is to present a method of transforming Ontology Charts resulting from the Semantic Analysis to the Component Object Model (COM+) architecture while maintaining the benefits of semantic analysis such as temporal data and immunity of systems to malignant changes from requirements to design. The COM+ architecture is the Microsoft's standard for developing application programs. Thus the business analysis can be carried out with the Semantic Analysis resulting in the Ontology Charts which can then be transformed into COM+ design architecture to implement the systems. Before going into the details of our transformation methodology, let us explore various concepts and ideas which are crucial for developing our transformation method.

2. Model Driven Architecture (MDA)

2.1 MDA Overview

Model Driven Architecture (MDA) is a technique to produce a system that is based on models [2]. A model is a description of a system written in a well-defined language. There are certain rules to describe a model these rules are defined using a modeling language. The language which is used to describe a model and which includes the complete syntax of model is known as modeling language [2]. To define and describe a modeling language we shall use Meta model. MDA uses two modeling constructs known as Platform Independent Model (PIM) and Platform Specific Model (PSM). PIM is the primary model and from PIM we could construct PSM. We can have more than one PSM constructed from same PIM. An MDA tool may support transforming a model in several steps, from initial analysis to executable code. For a system to be successful the PIM must be accurate. The correctness of PIM depends upon the information collection, flow of the real world system and the satisfaction of the business analyst. Incorrect PIM design may result in complete system failure. Figure 1 is an example of a MDA workflow. In our case, PIM will be an Ontology Chart and PSM the COM+ design.

2.2 Models and Meta Models

A model is an abstraction of what it represents. Models are simply sketches of design ideas, often discarded once

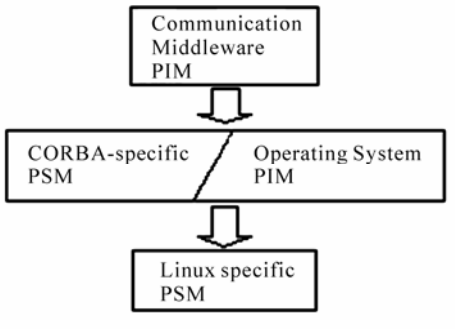

Figure 1. MDA workflow

the code is written [3]. This kind of modeling is done by software developers before writing the code; this is known as the code-centric development. Once the organization changes, the changes must be made to the newly developed or existing system as well. But in the case of code-centric development all the sketches which were developed during the SDLC are normally discarded. So, it becomes very difficult for the maintenance team to modify, extend and enhance the system. In MDA, model is a primary artifact in the development of software [3]. "Model" refers to a pattern, plan, representation or description designed to clearly demonstrate the understanding of the structure or working of an object, system, or concept. MDA uses models to represent key ideas in both problem and solution domains, and provides a conceptual framework for using models and applying transformations between them as part of a controlled, efficient software development process. The basic assumptions in and parameters governing MDA usage today, as described in [4,5] are:

- Model help people understand and communicate complex ideas.
- Many different kinds of elements can be modeled, depending on the context. These offer different views of the world that must ultimately be reconciled.
- We see commonality at all levels of these models – in both the problems being analyzed and the proposed solutions.
- Applying the ideas of different kinds of models and transforming them between representations provides a well-defined style of development, enabling the identification and reuse of common approaches.
- In what it calls "Model Driven Architecture," the OMG has provided a conceptual framework and a set of standards to express models, model relationships, and model-to-model transformations.
- Tools and technologies can help to realize this approach, and make it practical and efficient to apply.

During the development phase of a software/system, the system can be described in terms of models. In this way, the system can be developed at component level, and later be integrated as one.

2.3 Ontology Chart Meta Model

Even if UML is well supported, the scope of MDA doesn't diminish. MDA uses the Meta Object Facility (MOF) to allow any model to have its semantics represented in MOF participating in the MDA process [6]. The OMG [2] suggests four levels of abstraction for the models. Beginning with the actual system M0, moving to a model M1 that describes the specific system, the Meta model of this model is M2 which describes the semantics of the model and can be used for more models and finally, the Meta model of the Meta model that explains the semantics of the Meta model. "A metamodel makes statements about what can be expressed in the valid models of a certain modeling language", see [3]. Also in [3] "Given its metamodel representation, we can determine whether a model is valid". MOF is a model that can be used to define Meta models. This is illustrated by the following example: Given a system that consists of a number of java classes which is used to process loan applications, the actual system is the M0. The class diagram that represents the system for loan applications is the M1. The Meta model that describes class diagrams is the M2. MOF is used to describe the semantics of the Meta model that describes class diagrams. Obviously, there can be more layers between the M2 and MOF but in the real world it does not make much sense because it will increase the complexity at the cost of little or no benefit.

In the Meta model, as shown in Figure 2 which is reproduced from [7], the surrogate table holds all the information that a node of the Ontology Chart needs to hold. Ades and others [8] have proposed an ontology chart which is claimed to be the most complete. The sort attribute can be either universal or particular, and the type can be agent, affordance, role or communication act.

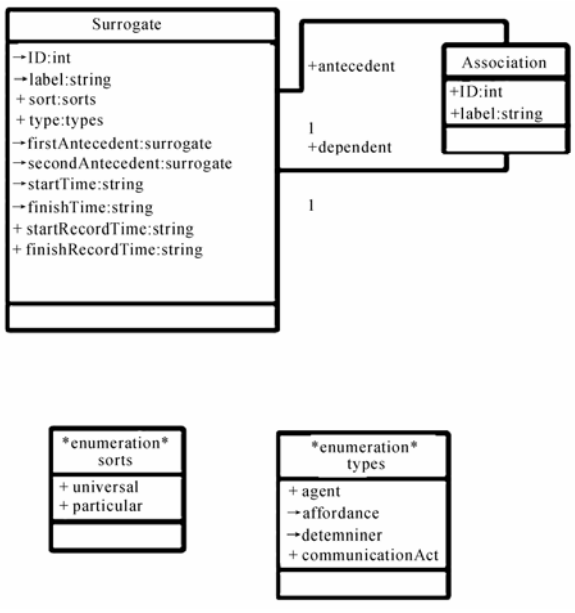

Figure 2. M2 Meta model for ontology chart

3. Semantic Analysis and Com+

SAM is a business analysis method used to capture organisational knowledge and is represented in the form of a number of diagrams known as Ontology Charts. Each of these diagrams consists of nodes with different shapes and their associations. Each node of the ontology chart can be agent, affordances and determiners [9]. Agents are entities from the real world capable to taking legal responsibility. Usually they are physical or legal persons (organisations). Determiners are attributes of agents or affordances, such as name, address and so on. Affordance stands for everything else. Ontology Charts are limited to binary associations. Each node is ontologically depended on one or two ancestor node(s) with the "root" node being the only exception. The "root" does not have any ancestor nodes and it is used to set the scope of the system. The "root" would disappear during transformation. Ontologically dependency requires that if the ancestor node seizes to exist then the child node would also seize to exist. Another property of the Ontology Charts is that each node has a start and finish time, giving a temporal dimension to the chart. Ontology Charts are capable of reflecting the complete structure of the organization describing the relationship between different entities.

COM+ is an approach for developing application programs and is an architectural design standard introduced by Microsoft [10]. The COM+ architecture supports both an object oriented programming and a set of operating system services. It adds a new set of services for application components while they are running, such as notifying them of significant events or ensuring they are authorized to run. COM+ is intended to provide a model that makes it relatively easy to create business applications that work well with the Microsoft technologies. Here we produce transformation rules for converting Ontology Chart to COM+ format. These transformations are intended to provide developers with a direct mapping to COM+ application which is a Microsoft standard.

4. Case Study: An Ontology Chart for A Law Firm and Its Transformation

4.1 Law Firm Ontology Chart

Consider the following case from a law firm. The "Law Firm" is connected to its dependant "Contacts" and it is again connected to "Client". That means that if "Law Firm" ceases to exist, "Contacts" would be meaningless and also cease to exist. Also if there is no "Client" the norm "Law Firm" still exists. A "Client" is a "person" that has a "Contact" with "Law Firm" regarding his/her case. Finally an "Employee" has a function to calculate the "hourly rate". Also the "Person" has role "Lawyer"

who is "Employed By" the "Law Firm". Figure 3 shows the ontology char for the above case.

In the above Ontology Chart, a physical person can be employed by a law firm as a lawyer. This employee has an hourly rate based on his employment contract. A client contacts the law firm and the law firm assigns a lawyer for this client.

4.2 The Transformation

The Key component in MDA is the transformations between models. According to [11], "A transformation language must provide for complete automation and must be expressive, unambiguous, and Turing complete". The transformation can be endogenous if the source and the target models have the same Meta model or exogenous if they have a different Meta model. For example the transformation between a class diagram and a more detailed class diagram is endogenous where the transformation from use case diagrams to a class diagram is exogenous. The transformation does not necessary have to be between a PIM and a PSM. It can also be from a more abstract PIM to a more detailed PIM. For example, use case diagrams can be transformed to sequence diagrams free of technical details, or from a PSM to another PSM (e.g. from EJBs to COM+). The official transformation language of OMG is QVT; however, other transformation languages such as Kermeta or even Java are also available.

This mapping will define the transformation rules for Meta models. In order to write the transformation rules, fist of all, we have to identify the key elements in the source model which in our situation are Ontology Charts. So, the key elements on which the transformation rules need to be applied are,

a) Agents
b) Affordances
c) Roles
d) Determiners

Now, after identifying the key elements of the source model now we have to identify what are the potential key elements of the target model which in our situation is COM+. The potential key elements in COM+ are,

a) Components
b) Interfaces
c) Connections (Either Delegate or Assembly)

Once the key elements are identified in the source and target model than the most important step is the mapping of each source element to the relevant target element. Before writing the transformations using any transformation language such as Kermeta or QVT, it is important to identify the transformation rules. We present the following rules for mapping Ontology Chart to COM+:

1) For each surrogate that is of type agent, a COM+ component is generated. For each surrogate determiner, it will become an attribute of that component.

2) For each affordance, an interface will be generated depending on the antecedent or dependent with required or provided interface respectively, depending on whether the generated component requires or provides data.

3) For a type role, a special component that is sub unit of Agent will be created.

4) Roles connected with affordances in Ontology Chart will become a sub-COM+ component

Once the key rules are written using normal English Language we can use any transformation language or even plain English to covert the source model in to target model. The result of COM+ is modelled in Figure 4 (Note that the root in Ontology Chart is dropped).

The output generated after the transformation is in the form of xmi which is easy to manipulate by transformation languages like Kermeta or even java or C#.

5. Conclusions and Future Work

The work in this paper demonstrates as to how an Ontology Chart can be mapped to COM+ based component

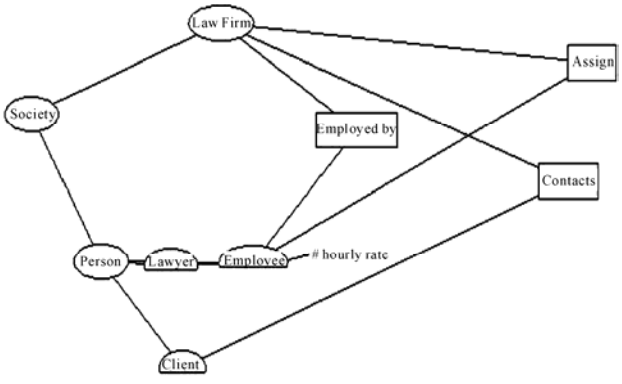

Figure 3. An ontology chart from a law firm

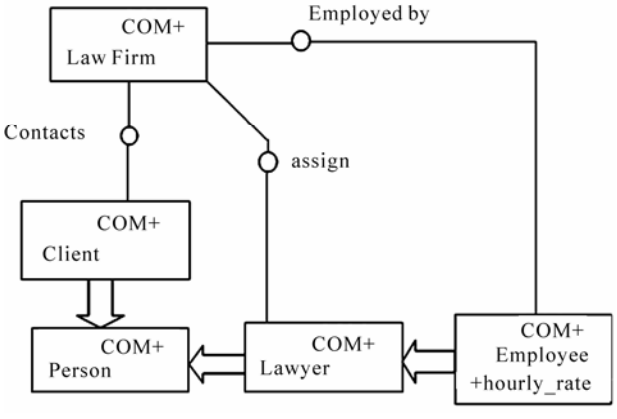

Figure 4. The result COM+ model

architecture. Our research has proposed a MOF Meta model for Ontology Chart and its transformation to COM+ diagram. This development would allow business analysts to design their Ontology Charts which can directly be coded without worrying about the loss of quality during the systems development. It should be emphasized that the key to correct transformation is to define a set of transformation rules accurately. Although the transformation from Ontology Chart to COM+ model has been achieved but still more work needs to be done regarding the services which are implemented in COM+.

REFERENCES

[1] BCS, "The challenges of complex it projects," The report of a working group from The Royal Academy of Engineering and The British Computer Society, April 2004.

[2] OMG, Meta ObjectFacility (MOF) Specification, Object Management Group, 2003, http://doc.omg.org/formal/02-04-03.

[3] S. Sendall and W. Kozaczynski, "Model transformation: The heart and soul of model-driven software development," IEEE Software, IEEE Computer Society, October 2003.

[4] E. Seidewitz, "What models mean," IEEE Softw, Vol. 20, No. 5, pp. 26–32, September 2003, http://dx.doi.org/10.1109/MS.2003. 1231147.

[5] A. Kleppe, J Warmer, and W. Bast, MDA Explained: The Model Driven Architecture(TM): Practice and Promise The Addison-Wesley Object Technology Series, 2000.

[6] I. Poernomo, A type theoretic framework for formal metamodelling, in R. Reussner and J. Stafford and C. Szyperski (Eds.), Architecting Systems with Trustworthy Components, Springer, LNCS, Vol. 3930, pp. 262–298, 2006.

[7] Y. Ades, N. Karimi Sani, M. Nistazakis, I. Poernomo, G. Tsaramirsis, and M. Yamin "Implementing SNF compliant software," Proceedings of ICISO'09, ISBN: 978-0-9806057-2-3, pp. 71, 2009.

[8] Y. Ades, F. Ben-Oman, I. Poernomo, and G. Tsaramirsis, "Mapping ontology charts to class diagrams," ICOS'07, 2007.

[9] K. Liu, "Semiotics in information systems engineering," Cambridge University Press, ISBN 0521 593352, Microsoft, COM+, 2000, http://msdn.microsoft.com/en-us/library/ms685978(VS.85).aspx.

[10] OMG, MDA guide version 1.0.1, 2000, www.omg.org/docs/omg/03-06-01.pdf.

[11] R. Stamper, "MEASUR—Methods of theory and analysis of information systems," Proceedings of IWRA, pp. 135–160, 2008.

J. Service Science & Management, **2010, 3: 72-77**
doi:10.4236/jssm.2010.31008 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Sustainable Logistics Networks in Sparsely Populated Areas

Wei Deng Solvang*, Moulay Hicham Hakam

Department of Industrial Engineering, Narvik University College, Narvik, Norway.
Email: *wds@hin.no, hakam@hin.no

Received August 17th, 2009; revised November 17th, 2009; accepted December 24th, 2009.

ABSTRACT

A logistics network is the 'connective tissue' in a comprehensive supply chain and often the decisive factor for the chain's ultimate success. Logistics networks in sparsely populated areas are facing different challenges than those that operating in relatively densely inhabited ones. This paper discusses, firstly, various critical success factors of a logistics network in order for it to sustain in sparsely populated areas. The challenges of establishing and operating a logistics network in sparsely populated areas are then identified and discussed. A case study has been conducted in the context of a region in Northern Norway where the average population rate is 4.3 *persons each square kilometer. The paper suggests that, by combining reverse logistics networks with forward ones, social, economic and environmental sustainable logistics networks can be achieved. The challenges of integrating forward and reverse logistics networks in Northern Norway are discussed in the final part.*

Keywords**:** *Logistics Network, Sparsely Populated Areas, Sustainability*

1. Introduction

A logistics network is traditionally defined as a set of resources which interactively connected with each other purposing of timely and cost-effective transfer of material and products to specified places. The definition itself doesn't limit the objects being transferred along the network, may it be material, products, services or even the end-of-life products which are also termed as waste.

Currently, challenges to logistics networks can often be seen in the context of globalization, individualization/mass customization and needs for economic development. The increasing globalization which is motivated by taking geographical advantages i.e., low-cost labor, easier access to necessary technology and local markets [1] often places huge challenges to logistics network due to, for instance, large distance among geographically scattered resources. Moreover, the increasing individualization [2] leads to the needs to mass-customized products. This requires that logistics networks which associated to these products are also able to provide mass-customized services. Further, the increasing public concerns on climate changes, environmental degradation and energy crisis as well as continuously strengthened international environmental and energy regulations lead to that a logistics network can no longer survive by taking only economic and operational performances into consideration [3–5]. That leads to the fact that economic sustainability needs to be offset by environmental sustainability in both shorter and longer terms.

Harris and Startup [6] state that all organizations attempting to provide universal geographical coverage face difficulties in supplying the populations of sparsely populated areas. The enlargement of the European Union has also accentuated differences between peripheral /sparsely populated areas and centrally/heavily urbanized areas in Europe. EEA EFTA states "that EU internal market legislation and other policy instruments should to a larger extent reflect such regional differences such that actions and legislation at EU level do not contribute to reduced competitiveness of industries in peripheral and remote areas" [7].

Currently, literatures that related to population sparseness and sustainability of its logistics networks are limited. Sandow [8] emphasizes the critical importance of labor mobility for regional competitiveness and sustainability in sparsely populated areas and further pointed that "a well functioning transport (logistics) system" is vital for securing this mobility of competent labor. Nilsson *et al.* [9] present another Swedish study which focuses on challenges of current low heat density areas to competitiveness of district heating in sparsely populated areas. They stated that, in order to improve the economic

sustainability of district heating in less densely populated areas, the rate of expansion in the detached-house segment needed to be increased. Muilu and Rusanen [10] document an increasing urbanization in Finland from 1970 to 2000. Same trend can be confirmed in Norway and Sweden. According to Muilu and Rusanen [10], as Finnish population continuous to increase, the number and proportion of young people have declined in rural municipalities. This resettlement of young people from peripheral areas to the built-up ones resulted in a dichotomy between the sparsely and densely populated areas which further accentuated by the distortion of population structure and gender ratio.

According to Neto *et al.* [11], the increasing focus on climate changes and pollution reduction from both publics and governments in recent years has aroused great interests of as well as large pressure on how logistics networks should operate. They proclaim that the objective in design a logistic network has changed, from cost minimization only to cost and environmental impact minimization. Tsoulfas and Pappis [12] propose some environmental principles for companies to be able to obtain a sustainable supply chain. Their argument is that by obtaining sustainable functions inside, outside and among companies along the chain, a competitive supply chain can be achieved in a long-term perspective. The logistics function is defined as one of five focuses in their research. Unfortunately, the study is still on conceptual level and hasn't provided any practical measures that can be used for logistics networks to achieve their sustainability.

This paper aims to shed light on challenges to economic and environmental sustainability of logistics networks in sparsely populated areas. The economic and environmental sustainability are interrelated and need to be detailed into some manageable concepts. Therefore three critical success factors are explored and discussed in the section that followed. A review of forward and reverse logistics networks is conducted purposing to provide the theoretical background for a case study of current logistics networks in Northern Norway. Based on the analysis on current logistics networks, the authors give several suggestions on how economic as well as environmental sustainability might be improved in these networks.

2. Critical Success Factors for the Success of Logistics Networks that Operating in Sparsely Populated Areas

A logistics network is a critical element for national, regional and community development, especially in a fragmented landscape with high demographic sparseness. Providing necessary level of logistics excellence in such an area prevents depopulation from the region and is vital for the region's economic development. Comparing with logistics networks in densely populated areas, those in sparsely populated areas are facing different strategic and operational challenges.

Generally, an organization may sustain when competing with any of three advantages, namely, overall cost leadership, differentiation and focus [13]. The first factor is concerned with cost-effectiveness; the second represents the uniqueness of the product as perceived by the customer. As for the third factor, the key idea is to serve a particular target very well.

One of the most significant differences between densely and sparsely populated areas is the size of population a logistics network aims to serve. A densely populated area has usually large enough concourse which provides necessary basis for the network to take the advantages of both economy-of-scale and economy-of-scope. Logistics networks in sparsely populated area, however, usually do not possess necessary prerequisites for taking these advantages.

Another difference is the distance. A logistics network in a densely populated area is less distant between two resources (i.e., between a distribution centre and a retailer). This may easily enable, for example, the network to achieve more mass-customized logistics service with diverse focuses.

The third difference is that a sparsely populated area is usually situated in a peripheral region of a country. Developing logistics networks that connecting sparsely and densely populated areas needs to be viewed in conjunction with offering universal geographical coverage for social and economic development as well as sustainable regional development in whole territory. By doing so, risks such as social degradation and depopulation can be prevented or minimized.

Different strategies needs to be applied to handle the challenges mentioned above. The critical success factors for logistics networks operating in sparsely populated areas are therefore:

- Cost-effectiveness: refers to the extent to which customer requirements are met given a limited level of cost.
- Cost-efficiency: a measure of how economically a network's resources are utilized when providing a given level of customer satisfaction.
- Eco-efficiency: the efficiency with which resources are converted into product.

3. Forward and Reverse Logistics Network

As mentioned earlier, a logistics network links difference resources in the way that tangible or intangible objects (i.e., material, parts, products, service and waste) being transferred within the network and towards a clearly defined destination. While a forward logistics network brings raw material or new products from origin (e.g., a

raw material supplier) to end-customers, a reverse network refers to the network structure dealing with the end-of-life products from end-customers (Figure1, made based on [14]).

Reverse logistics is the process of moving goods from their typical final destination to another point, for the purpose of capturing value otherwise unavailable, or for the proper disposal of the products [14].

In recent years, the term 'closed-loop supply chain' has becoming a topical issue as increasing attention has been attached to pollution prevention and minimization rather than end-of-pipe pollution control [4]. A Closed-loop supply chain can be regarded as a system without waste. The primary material flow of forward logistics network has been linked with material flow of reverse logistics network through waste dis charge and management system (direct treatment, indirect treatment and purification). The closed-loop supply chain system can also achieve high level of logistics excellence as capacity of forward and reverse logistics network can be utilized optimally.

4. Case study

Norway has especially fragmented territory and Norwegian territorial structure can be characterized as [15]:

- Long distance to main markets
- Extensive mountain areas
- Insular and coastal regions
- Isolated border regions
- Arctic and sub-arctic climate
- Demographic sparseness
- An extremely diffused urban system with sharp differences regarding settlement structures, functional profiles and centralities

Norway is ranked 214 out of 241 countries according to average population density based on data for 2005 [16].

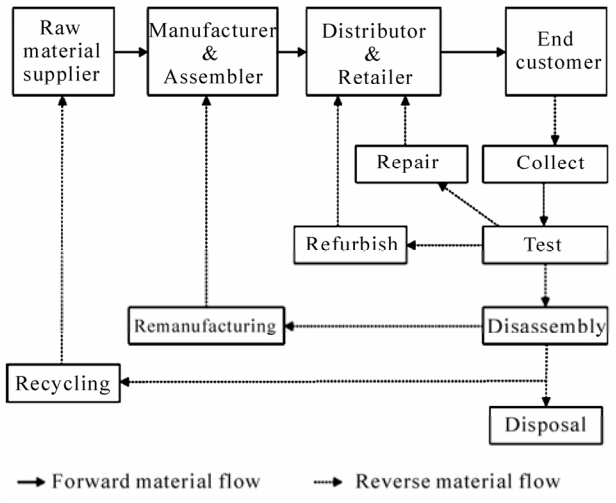

Figure 1. Forward and reverse logistics networks

The population density is 15.4 persons per each square kilometer (Table 1). Areas with less than 12.5 persons per square kilometer cover 87% of the land area. 286 of 430 municipalities has population density lower than 12.5 persons per square kilometer.

Norway has nineteen counties. Among which, three of them are in Northern part of Norway (Nordland, Troms and Finnmark, as shown in Table 1). The area of this part counts up to 106,715 km^2 and is over 35% of total Norwegian territory. In 2007, less than 10% of Norwegian population is resident in this part and population density is 4.3 persons per square kilometer. This extremely low population density places huge challenge to the development and sustainability of the logistics networks within this region. The long distance between the Northern and Southern parts of Norway also places fierce challenge to sustainability of their connecting logistics networks.

4.1 Current Logistics Networks that Link North with South

4.1.1 Staple Goods and Fishes

Currently, the main stream of staple goods (i.e., food, mail) from south to north is carried by Arctic Rail Express (ARE) which is an established service offered by CargoNet [17]. Yearly, up to 23,500 containers (either twenty feet or twenty-five feet container) are carried by ARE train from Oslo Alnabru to Narvik through Swedish territory (**Figure 2**). The whole journey takes 27 hours

Table 1. Population density in Norway (2007 data)[A]

Unit	Population	Area (km^2)	Density
Norway	4681134	304280	15,4
Northern Norway	462237	106715	4,3
- Nordland	235436	36074	6,5
- Troms	154136	24884	6,2
- Finnmark	72665	45757	1,6

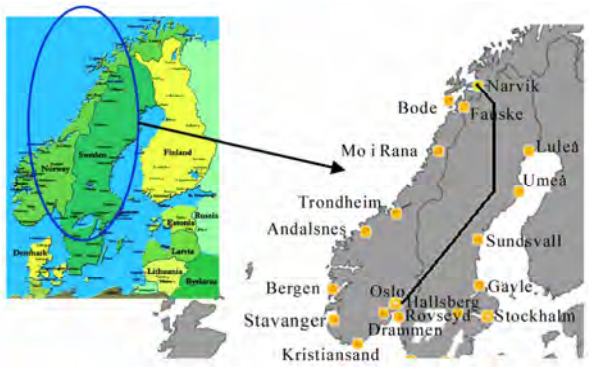

Figure 2. ARE train

and total length of railway is 250 km. These goods are further distributed to three main regions (Tromsø, Harstad and Narvik) by third party logistics providers (3PL) (i.e., Schenker and Nor Cargo). These further transportations are conducted on road. 55% (approx. 13,000 containers) of the staple goods travel around 250 km to Tromsø, 10% being transported 135 km to Harstad, and 35% being distributed in local Narvik area (**Figure 3**).

According to CargoNet, the back loading rate of ARE is barely 60%. Of which the majority is raw fish and fish products. This means that nearly 9,500 containers return empty from Narvik back to Oslo. There is no doubt a considerable potential in cost-efficiency improvement.

The ARE runs twenty-two trains each week. That is eleven trains in each direction. Together with other 3PLs, ARE satisfies basic demands (food, communication) of Northern Norway with a population of nearly 470,000. Eco-efficiency is satisfactory considering the transportation is mainly on railroad.

4.1.2 Wastes

It is said that the waste amount is increasing as the economy is accelerating in a region. Since 1995, the entire amount of wastes in Norway has been increased by nearly 30% and has already reached 9,6 million tons in 2006 [18]. Recognizing this, the Norwegian government has, since 1970s, been working systematically in establishing a sustainable waste management system.

In Norway, the enactment of Product Control Act in 1976 and Pollution Control Act in 1981 as well as Product Responsibility Act in 1988 established the fundament for crystallizing the 'polluter pays' principle to entire nation. Various environmental legislations are enacted since 1980s with purpose of promoting pollution prevention and pollution minimization (through reuse and recycling) and avoiding waste being deposed in landfills.

Currently, sixteen waste management companies are established in Northern Norway handling up to 900,000 ton wastes per year (the volume is calculated based on data from Statistics Norway). Take one of the waste management companies in Northern Norway as an example, Hålogaland Ressursselskap (HRS) is an intra-municipal companies which owned by twelve municipalities. Its activities consist of collecting, sorting, transportation and treatment of household, industrial and special wastes (**Figure 4**).

Its current waste is categorized as:

- glass and bottles
- timber
- iron
- food waste
- paper

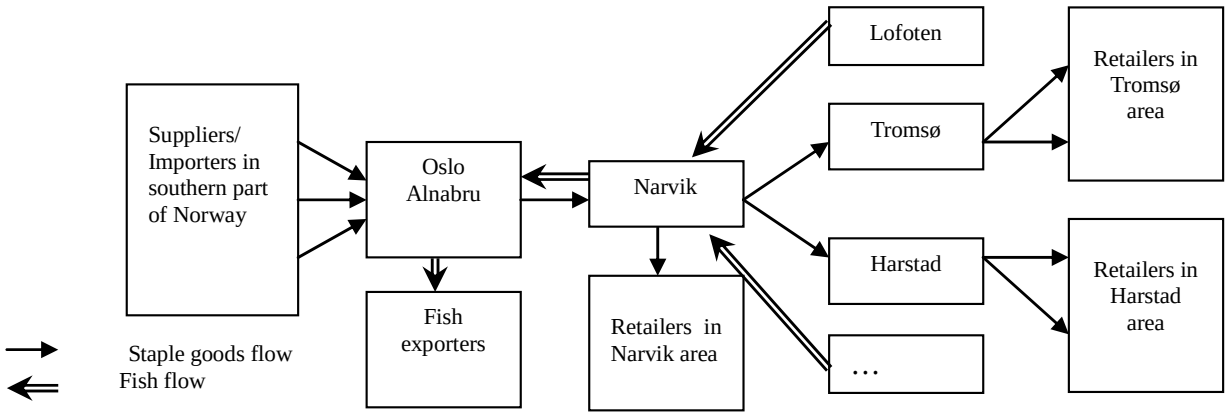

Figure 3. Main material stream south-to-north and north-to-south with Narvik as intersection

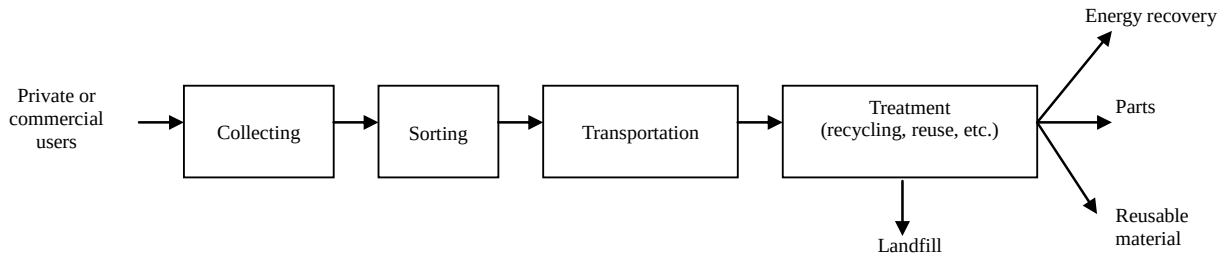

Figure 4. Reverse logistic processes

- corrugated cardboard
- plastics from industries
- beverage paperboard box
- waste of electrical and electronic equipment (WEEE)
- hazardous waste

Most of the waste goes to recycling, nearly 90% in 2005. Taking this as key figure, the recycling volume can be estimated up to around 800,000 tons in Northern Norway.

Currently, large portion of the recyclable wastes is transported to the Southern part of Norway where the majority of recycling facilities are. The transportation is mainly on road with trucks or semi-trailers.

4.2 The Potential and Challenges of Integrating Forward and Reverse Logistics Networks

It is undoubtedly that there is a huge potential of integrating forward logistic network (of staple goods and fish) with reverse logistics network (of wastes) with Narvik as the junction point. By doing so, the biggest gain is the reduction of empty container back to Oslo. This will, to a large extent, increasing cost-efficiency of both logistics networks. This will also lead to the reduction of transportation price. Together with regular as well as frequent transportation between south and north, a cost-effective logistics network can be achieved. Further, as the integration will 'lift' goods from road to track, significant environmental profit (i.e., less pollution and noise) and social gains (i.e., less traffic accidents) can be obtained.

However, challenges for this integration are many. First of all, current waste management system has obtained itself an exclusive logistics network. Take Elretur AS for example, as one of four take-back companies approved by Norwegian Pollution Control Authority (SFT), Elretur AS collects over 76,000 ton WEEE in 2007. The company cooperates with seven 3PLs for collecting from entire Norway [19]. One of these companies is Østbø AS which has full authority and responsibility for collection of WEEE in Northern Norway. Østbø AS has its own closed system for transporting WEEE to treatment plants in Southern part of Norway. The transportation from Northern Norway to Østbø AS and from Østbø AS to Southern Norway is mainly on road.

Secondly, the hygienic requirements of transporting foods and other staple goods from south to north may place the challenge to return wagons loaded by wastes. The traditional perception of waste may also create scepticism at staple goods owners and therefore may result in reduced logistics flow from south to north.

5. Summary and Discussion

The paper presents a conceptual study on logistic networks in sparsely populated area. As opposite to densely populated areas, sparsely populated ones place larger challenge to their logistic networks. The limitations of such areas include that the demand volume does not justify economy-of-scale or economy-of-scope. Larger distance between two resources within a network is also a significant challenge. That a sparsely populated area usually situated in a peripheral to densely populated ones place also the challenge to the logistics connection of these areas.

This paper proposed three critical success factors for logistics networks operating in sparsely populated area. By conducting the case study, the authors argued that, in order to achieve cost-efficiency, cost-effectiveness and eco-efficiency of main logistics network that connecting Northern Norway with the Southern part, the potential of utilizing existing south-north logistics network (for staple goods and fish) for north-south transportation of wastes is considerably large. However, currently exclusive waste management systems and traditional perception of waste may place the largest challenges to this possible solution.

Benchmarking of performances in terms of cost- effectiveness, cost-efficiency, and eco-efficiency of current logistics networks (separated networks for staple goods and waste) with that of suggested logistics network (transport waste with logistics network of staple goods) need to be conducted in order to provide the quantified justification for the suggested solution. The authors also suggest that a study that mapping all goods flows at a more detail level should be conducted in order to provide an overall optimal logistics solution in Northern Norway.

6. Acknowledgment

The authors have received great support from different organizations and companies during this study. Here we would thank especially Jan Viggo Fredheim at CargoNet, Line Dalhaug at HRS and Ole Viggo Svendsen at Elretur AS for enthusiasm and information.

REFERENCES

[1] S. Chopra and P. Meindl, "Supply chain management - strategy, planning, and operation," Pearson Educational International, 2004.

[2] E. Alfnes, "Enterprise reengineering – a strategic framework and methodology," Department of Production and Quality Engineering, Faculty of Engineering Science and Technology, Trondheim, Norwegian University of Science and Technology, 2005.

[3] F. Du and W. Evans, "A bi-objective reverse logistics network analysis for post-sale service," Computers and Operations Research, pp. 2617–2634, 2008.

[4] W. D. Solvang, "Architecture for supply chain analysis and methodology for quantitative measurement of supply chain flexibility," Department of Production and Quality

Engineering. Trondheim, Norwegian University of Science and Technology, 2001.

[5] W. D. Solvang, Z. Deng, and B. Solvang, "A closed-loop supply chain model for managing overall optimization of eco-efficiency," IN POM (Ed.) POMS 18th Annual Conference. Dallas Texas, USA, 2007.

[6] C. C. Harris and R. Startup, "Researching a traditional territorial organization: Analogy, and the sociological imagination," International Journal of Sociology and Social Policy, Vol. 18, 1998.

[7] European Economic Area Standing Committee of the EFTA States, EEA EFTA Comment on the EU Freight Transport Agenda, 27 February 2008.

[8] E. Sandow, "Commuting behaviour in sparsely populated areas: Evidence from northern Sweden," Journal of Transport Geography, No. 16, pp. 14–27, 2008.

[9] S. F. Nilsson, C. Reidhav, K. Lygnerud, and S. Werner, "Sparse district-heating in Sweden," Applied Energy, No. 85, pp. 555–564, 2008.

[10] T. Muilu and J. Rusanen, "Rural young people in regional development - the case of Finland in 1970–2000," Journal of Rural Studies, No. 19, pp. 295–307, 2003.

[11] J. Q. F. Neto, J. M. Bloemhof-Ruwaard, J. A. E. E. van Nunen, and E. van Heck. "Designing and evaluating sustainable logistics networks," International Journal of Production Economics, No. 111, pp. 195–208, 2008.

[12] G. T. Tsoulfas and C. P. Pappis, "Environmental principles applicable to supply chains design and operation," Journal of Cleaner Production, No. 14, pp. 1593–1602, 2006.

[13] M. E. Porter, "Competitive strategy – techniques for analyzing industries and competitors," The Free Press, 1980.

[14] S. K. Scrivastava, "Network design for reverse logistics," The International Journal of Management Science, pp. 535–548, 2008.

[15] J. Sandal and O. Godal, "Norwegian input to the EU green paper on territorial cohesion," In Royal Ministry of Local Government and Regional Development, N. (Ed.), 2008.

[16] List_of_countries_by_population_density available at: http://en.wikipedia.org/wiki [Last checked: November 2009].

[17] Cargonet webpage available at: http:// cargonet.no /Wide Page.aspx?id=3729&lang=NO [Last checked: November 2009].

[18] Norwegian statistical data available at Statistics Norway: www.ssb.no [Last checked: November 2009].

[19] Elretur webpage available at: http://www.elretur.no/nor/Elretur/Retursystemet/Transportoerer, Last checked: November 2009.

J. Service Science & Management, **2010, 3: 78-83**
doi:10.4236/jssm.2010.31009 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Design on the Incentive Contract of University Achievements Commercialization Offices Based on Principal-Agent Theory

Meifang Li[1], Yongxiang Zhao[2], Feng Shi[3]

[1]School of Management, Wuhan University of Technology, Wuhan, China; [2]Wuhan University of Technology, School of Computer Science and Technology, Wuhan, China; [3]Wuhan Academy of Social Sciences, Wuhan, China.
Email: poplimeif@126.com, zhaosanhe@263.net, sf196293@163.com

Received October 13th, 2009; revised November 21st, 2009; accepted December 30th, 2009.

ABSTRACT

Based on principal-agent theory, an incentive contract model of university achievements commercialization offices (*UACO*) *was constructed in this paper, and an optimal incentive contract between university and UACO was researched into. The conclusion indicates that many factors, such as working ability, working willingness, risk aversion degree of UACO, as well as the outside uncertain factors and so on, have important influences on the contract design. The efficiency of commercialization of university inventions has a squared forward growth relation with working ability, a direct proportion with working willingness, and has an inverse proportion with risk aversion degree of UACO and with outside uncertainty. Additionally, the level of hard working of UACO under the condition of information asymmetry is strictly less than that of information symmetry.*

Keywords**:** *University Achievements Commercialization Offices* (*UACO*), *Achievements Commercialization, Principal-agent, Incentive Contract*

1. Introduction

In order to promote the commercialization of research findings, and to raise technology transfer rate, almost all the universities in China have set up a department for connecting among industry, university and research institute. The department usually conducts the work of patents applications and achievements transformation for teachers and scientists of the university. At present, there is still not uniform name for this department in Chinese universities, some of them is called university industry management office, some is named S&T achievements transfer office, and still others is known as technology transfer office and so on. In this paper it is named university achievements commercialization office (UACO) for the sake of convenience.

In the recent years, the transformation of achievements of universities has been attached great importance in all levels of the governments in China, and universities also offers a series of incentives for the commercialization of research findings. However, the result of transformation is still not desirable. The reasons would be various but UACO has to bear the blame, because it plays an important role of bridge and link between university and businesses. But the cause from UACO hasn't been given enough attention. At the present time there is less research focusing on this problem domestically, and the research has been existed abroad but numerically small. For example, it is fully affirmed the important roles of UACO in the papers of Leitch, *et al* [1] and Colm, *et al* [2]. They hold that UACO should and could play an important and unique role in the process of the commercialization of university achievements. Donald, *et al* [3] and Swamidass, *et al* [4] proposed that inappropriate administrative staff and irrational reward system are the major hindrances to the achievements transformation of university. Chapple *et al*. pointed out that it is very important for university technology transfer officers and managers in England to upgrade business skills and capabilities in order to increase university technology transfer efficiently [5]. Markman *et al*. showed that commercialization outcomes would be enhanced when UACO employ diverse licensing strategies, enjoy greater autonomy, and be compensated well [6]. In fact, there is a principal-agent relationship between university and UACO, and commercialization efficiency could be en-

hanced through the design of incentive contract of these two parts, but scholars in China and abroad pay little attention to it. In this paper, principal-agent theory is deployed to construct a principal-agent model between university and UACO, so as to design an incentive contract thereby increasing the efficiency of commercialization, and hardening society-serving function of universities.

2. The Principal-Agent Relations of University Achievements Commercialization and the Necessity of Incentive Contract Designing

2.1 The Principal-Agent Relations of University Achievements Commercialization

Any interrelation involving asymmetric information could be called principal-agent relation from the sense of economics. One party with information superiority is agent, and the other at information disadvantage is principal. Furthermore, personal information of agent has some influence on the interests of principal who is uninitiated [7].

From the perspective that UACO is commissioned to commercialize achievements and patents of university scientists, there is a principal-agent relation between university and UACO. In case of information asymmetry, UACO has private information, and this information such as how hard it works, has a great impact on the interests of university, for example impacting on the commercialization efficiency of the university.

According to the theory of information economics, information asymmetry is easy to result in moral hazard problem. In the principal-agent relation between university and UACO, university (as principal) were not accessible to direct observation on which level of hard working UACO had selected, and what university could observe was another variables, such as technology transfer rate and so on. But the rate of technology transfer is determined by the level of hardworking and other outside random factors together. Consequently, there is an incomplete information game with uncertainty but couldn't be supervisory [8]. In this game, the task confronting university (principal) is how to design an incentive contract to motivate UACO (agent) select behavior that fit for the interests of the university.

2.2 The Necessity of Incentive Contract Designing

To some extent, UACO performs practically no function in the universities in China, and it is lack of influence and ability in the work of technology transfer. Many university scientists are reluctant to commercialize their achievements through UACO, but to establish communication with corporations by themselves or just declare their achievements for prize and acknowledgment and then put them on the shelf.

There are many causes giving rise to this situation. First, staff structure is unfit for requirements of commercialization. The officers in UACO have, in many cases, neither technical advantages nor marketing skills, and have not the competence to be engaged in the work of achievements commercialization. Second, the compensation system is not rational. What UACO basically executes is fixed salary in China for a long time, and there is no difference for officers between good business and bad one, no difference between hardworking and slacking. As a result, there is no enthusiasm for officers of UACO to involve in the work of commercialization. Third, it lacks the necessary autonomy in UACO. A great majority of UACO in China are attached to departments of Science & Technology of universities. Officers work in UACO just as a matter of routine, and it is absence of mental stress for their jobs. Fourth, it is lack of active action of hunting for market demands. The outdated way of working that doing nothing but waiting for buyer cannot meet the needs of a market economy.

Consequently, in order to make UACO play a greater role for universities in serving local economy and in pressing ahead with combination of industry and university, the urgent task is to introduce market mechanism into UACO, to reform the existing compensation system and grant more autonomy, and to enhance enthusiasm for the work of commercialization. Under these circumstances, it is urgently necessary for universities to design an optimal contract, so as to encourage UACO to work doubly hard on enhancing the efficiency of commercialization of research findings in universities.

3. Design on Incentive Contract of University Achievements Commercialization Offices

On the basis of the preceding analysis on the principal-agent relation between university and UACO, and supposing that principal and agent are based on rational-economic man hypothesis, whether agent choose to work hard to enhance the efficiency of achievements transformation or not is totally dependent upon the incentive contract that university provided. According to the parameter method of distribution function brought forward by Mirrlees [9] and Holmstrom [10], the principal-agent model between university and UACO could be constructed.

3.1 Construction of Incentive Model for UACO

For the convenience of research working, and in a situation of no influence on conclusion, the following assumptions are proposed, supposing e is a one-dimension variable of hardworking level of agent (UACO), and it is related to initiative and resources injection of UACO. The parameter k is representative the coefficient of

working capability of UACO, and it is relative to management ability, scientific and technical level, marketing skills of UACO. Parameter θ denotes uncertain factors that influences achievements transformation of universities, and it is a normal approximation distribution and let it be a random variable with mean of 0 and variance of σ^2. Then the output of UACO could be expressed with the following linear function:

$$\pi = ke + \theta \tag{1}$$

Here, $E\pi = ke$, $\text{var}(\pi) = \text{var}(\theta) = \sigma^2$. That is to say, the level of hardworking and the level of working capability determine the mean of output of UACO, and the variance of output is only connected with that of endogenous random variable. Here, E represents mathematic expectation operator, and var means variance.

Supposing that principal (university) is risk-neutral and agent (UACO) is risk-avoidance, and optimal contract is linear under the hypothesis of rational-economic agent [11], then linear contract is considered as following

$$s(\pi) = \alpha + \beta\pi \tag{2}$$

$s(\pi)$ refers to income of UACO (or its employees), and α is fixed salary. β is the share of UACO in its outputs, it could be also called incentive intensity that university provides to UACO. A further formula could be gotten if expression (2) is substituted with (1), that is

$$s(\pi) = \alpha + \beta(ke + \theta) \tag{3}$$

then the revenue of university could be expressed by $v(\pi, s(\pi)) = \pi - s(\pi) = -\alpha + (1-\beta)(ke+\theta)$, and expected utility equals to expected revenue according to the assumption of risk-neutral for principal (university), that is

$$\begin{aligned} Ev(\pi, s(\pi)) &= E[-\alpha + (1-\beta)(ke+\theta)] \\ &= -\alpha + (1-\beta)ke \end{aligned} \tag{4}$$

Let ρ be the level of utter risk avoidance, and $\rho > 0$. To express risk cost of agent with $r(\rho)$, then $r(\rho) = \rho\beta^2\sigma^2/2$. If $c(e)$ refers to the cost of working hard of UACO, and $c(e) = be^2/2$ is supposed for simplifying the problem. Here, b refers to cost coefficient and $b > 0$. Therefore, the deterministic equivalence income of UACO is

$$\begin{aligned} W &= E[s(\pi) - c(e) - \rho\beta^2\sigma^2/2] \\ &= \alpha + \beta ke - be^2/2 - \rho\beta^2\sigma^2/2 \end{aligned} \tag{5}$$

Let $\overline{w}$ be conservative income of agent, then individual rationality (IR) of agent could be expressed by

$$\alpha + \beta ke - be^2/2 - \rho\beta^2\sigma^2/2 \geq \overline{w} \tag{6}$$

and the incentive compatibility (IC) of agent is

$$e \in \arg\max\{\alpha + \beta ke - be^2/2 - \rho\beta^2\sigma^2/2\} \tag{7}$$

3.2 Solution of the Incentive Model

In order to compare expected income of university and UACO under the condition of asymmetric with that of symmetric information, the solution of incentive model under the condition of symmetric information is necessary to be discussed in brief.

1) Optimal contract under the condition of information symmetry

The level of hardworking of agent could be observed by principal under the condition of symmetric information. The incentive compatibility (IC) of agent at the moment does not work because any level of e could be gotten and realized through an enforced contract meeting the requirement of individual rationality (IR). Therefore, the optimal contract under the condition of information symmetry could be obtained by solving the following optimal problem

$$\begin{cases} \max\limits_{\alpha,\beta,e} Ev = -\alpha + (1-\beta)ke \\ s.t.(IR)\ \ \alpha + \beta ke - be^2/2 - \rho\beta^2\sigma^2/2 \geq \overline{w} \end{cases} \tag{8}$$

The equal-sign in constraint condition of IR is ture in the above Formula (8) under the optimal circumstances of information symmetry, because it is not necessary for principal to pay agent more when information is symmetric. Then the optimal problem of Formula (8) could have another expression as following

$$\max_{\beta,e}(ke - be^2/2 - \rho\beta^2\sigma^2/2 - \overline{w}) \tag{9}$$

To solve first order condition of the optimal problem in Equation (9), the expressions of Pareto optimal hardworking level (e^*) and Pareto optimal incentive intensity (β^*) could be gotten as following

$$e^* = k/b, \quad \beta^* = 0, \tag{10}$$

Substituting the equation of IR in the Formula (8) with the above results in the Formula (10), the expression of optimal fixed salary (α^*) under the condition of information symmetry could be obtained

$$\alpha^* = \overline{w} + k^2/2b \tag{11}$$

2) Optimal contract under the condition of information asymmetry

The level of hardworking of agent could not be observed by principal under the condition of asymmetric information. The incentive compatibility (IC) of agent at this moment does work. The maximization model of expected utility function of agent could be established as following

$$\max_e Eu = \max_e W = \alpha + \beta ke - be^2/2 - \rho\beta^2\sigma^2/2 \quad (12)$$

For any given incentive contract, agent always selects an optimal hardworking level to maximize expected utility function. According to Mirrlees (1974) [12] and Holmstrom (1979) [10], the constraint of incentive compatibility (IC) could be replaced by first order condition of maximization model of expected utility function of agent. For the Formula (12), first order condition of maximization is $\frac{\partial Eu}{\partial e} = \beta k - be = 0$. Solving the equation, the expression of hardworking level under the condition of information asymmetry could be gained, and the constraint of incentive compatibility (IC) could also be described as the expression

$$\text{(IC):} \quad e = \beta k / b \quad (13)$$

Therefore, the optimal contract under asymmetric information should satisfy the solution of the following conditions

$$\begin{cases} \max\limits_{\alpha,\beta} Ev = -\alpha + (1-\beta)ke \\ s.t.\ (IR)\ \alpha + \beta ke - be^2/2 - \rho\beta^2\sigma^2/2 \geq \overline{w} \\ \quad (IC)\ e = \beta k/b \end{cases} \quad (14)$$

Supposing that principal and agent are rational-economic, the equal-sign in constraint condition of IR is true in the Formula (14) when designing the optimal contract. Substituting the target function of the Formula (14) with the expressions of IR and IC, the optimal condition of the Formula (14) could be reformulated by

$$\max_\beta \beta k^2/b - \beta^2 k^2/2b - \rho\beta^2\sigma^2/2 - \overline{w}$$

and first order condition of it could be expressed as $k^2/b - \beta k^2/b - \rho\beta\sigma^2 = 0$, then the share of output that university provides to UACO (namely incentive intensity) is

$$\beta = \frac{k^2}{k^2 + b\rho\sigma^2} \quad (15)$$

Putting expressions of Formulas (15) and (13) into the equal equation of IR in Formula (14), the expression of fixed salary (α) under the condition of information asymmetry could be gained as

$$\alpha = \overline{w} - (k^2 - b\rho\sigma^2)\Big/2b(1+b\rho\sigma^2/k^2)^2 \quad (16)$$

Consequently, the optimal contracts under information symmetry and asymmetry are listed in Table 1. fol- lowing the synthetic results of the above analysis.

4. Analysis and Results

The results will be analyzed from two sides of incentive contract and university revenue. The efficiency of achievements transformation in universities is of brutally vital importance for the outputs of UACO and the revenue of university, so these latter two parts will be equivalent to the efficiency of achievements transformation in the following analysis.

Table 1. The design of incentive contracts of UACO under the conditions of symmetric and asymmetric information

	Information symmetry	Information asymmetry
Hard-working level	$e^* = k/b$	$e = \beta k/b$
Share of output	$\beta^* = 0$	$\beta = k^2/(k^2 + b\rho\sigma^2)$
Fixed salary	$\alpha^* = \overline{w} + k^2/2b$	$\alpha = \overline{w} - (k^2 - b\rho\sigma^2)\Big/2b(1+b\rho\sigma^2/k^2)^2$

4.1 Analysis of Incentive Contracts

Compared the incentive contracts under the conditions of information symmetry and information asymmetry, the following results could be brought to:

Result 1: Incentive mechanism should be introduced in the salaries of UACO to enhance the efficiency of achievements commercialization in universities, for example implementing the wage system comprised fixed salaries and efficiency-related wages.

According to the incentive contract under the condition of asymmetric information, the share of outputs that UACO should be provided is $\beta = k^2/(k^2 + b\rho\sigma^2)$. Here $0 < \beta < 1$ because $b, \rho, \sigma^2 > 0$, that is, to get the best incentive effectiveness, $k^2/(k^2 + b\rho\sigma^2)$ unit should be provided to UACO if 1 unit output value of achievements commercialization is produced by UACO. From this point of view, the unitary fixed wage system implemented for a long time has seriously hindered the advancement of achievements transformation, and it is urgent for universities to reform the wage system of UACO.

Result 2: When the work ability of UACO is higher (larger k), or the willingness of working hard is stronger (smaller b), or the degree of risk aversion of UACO is smaller (smaller ρ), or outside uncertain factors are smaller (smaller σ^2), the wage system of lower fixed salary plus higher proportion of efficiency-related wage could be more effective for UACO to work hard and to boost the efficiency of commercialization. In certain circumstances, the wage mechanism only for efficiency-related wage, even that state of it which we term zero fixed salary, could be adopted.

The reason is that, according to Formula (15), first-order derivatives of β respectively satisfy the conditions of $\partial\beta/\partial k = 2kb\rho\sigma^2/(k^2 + b\rho\sigma^2)^2 > 0$, $\partial\beta/\partial b < 0$, $\partial\beta/\partial\rho < 0$, $\partial\beta/\partial\sigma^2 < 0$, (the above calcula-

tions are omitted here). Those first-order derivatives show that, the share of outputs (β) is a monotone increasing function to the work ability of UACO (k), and is a monotone minus function to the willingness of working hard (b), a monotone minus function to the degree of risk aversion (ρ), a monotone minus function to outside uncertain factors (σ^2). According to Formula (16), $\alpha < \overline{w}$ when $k^2 > b\rho\sigma^2$, that is, the fixed salary could be below to conservative income for agent when certain condition is satisfied. Furthermore, $\alpha = 0$ when $k^2 = \overline{w} + b\rho\sigma^2$, it means that, the condition of zero fixed salary is happened. Certainly, the proportion of outputs required to share would be very high.

Remarkably, zero fixed salary is a state of fully- marketing operation that UACO are in a high degree of autonomy, or even become an independent legal entity detached from universities. It is an operating pattern of UACO worth of being probed into.

Result 3: The level of hardworking of UACO under the condition of asymmetric information is strictly less than that of symmetric information

It is because $0<\beta<1$, so $e=\beta k/b<k/b$, that is $e < e^*$.

The result provides a rational explanation for the current states that it is lack of working enthusiasm and lack of active action of hunting for market demands in UACO. It is further verified that incentive measures should be drawn out to encourage UACO work hard to promote the efficiency of achievements commercialization, especially under the circumstances of information asymmetry.

Result 4: Under the condition of information symmetry, UACO could be only paid for fixed salary because of $\beta^* = 0$. When fixed salary exactly equals to conservative income plus the cost of hardworking (that is $\alpha^* = \overline{w} + k^2/2b$), a desired optimal hardworking level of universities could be reached, and Pareto optimal risk apportion and optimal hardworking level could be realized simultaneously.

4.2 Analysis of University Revenue (Efficiency of Commercialization)

Connecting equations of (13), (14) and (15), the expression Formula (18) of university revenue ($E^a v$) under the condition of asymmetric information could be obtained, and substituting Formula (8) with Formula (10), then we can get the expression Formula (17) of university revenue ($E^s v$) under the condition of symmetric information. The results after collecting the forms could be seen in Table 2.

The following results could be gotten from Formula (17) and (18).

Result 5: The working ability of UACO is of great importance to the efficiency of commercialization for universities. The stronger the working ability, the more revenue is produced, then the higher efficiency of commercialization is taken place. Furthermore, a vital important result is that, the efficiency of achievements commercialization has squared forward growth relation with working ability of UACO.

Table 2. University revenues under the conditions of symmetric and asymmetric information

	University revenue	
Information symmetry	$E^s v = \frac{k^2}{2b} - \overline{w}$	(17)
Information asymmetry	$E^a v = \frac{k^4}{2b(k^2 + b\rho\sigma^2)} - \overline{w}$	(18)

It is because of

$$\frac{\partial E^a v}{\partial k^2} = \frac{k^4 + 2k^2 b\rho\sigma^2}{2b(k^2 + b\rho\sigma^2)^2} > 0,$$

that is, the efficiency of achievements commercialization ($E^a v$) is a monotone increasing function to square of working ability (k^2).

What we learn from result 5 is that, it is need to adjust staff structure of UACO, to stress importance of technical ability, marketing skills and administrative ability for officers of UACO in China, and UACO managers should have all these abilities and skills.

Result 6: The efficiency of achievements commercialization is directly proportional to the willingness of hardworking of UACO, and is inversely proportional to the degree of risk aversion of UACO and to outside uncertain factors.

It is because the efficiency of achievements commercialization of universities ($E^a v$) is a monotone minus function to willingness of working hard (b), to risk aversion degree (ρ), and to outside uncertain factors (σ^2).

The revelation of this result is that, for the enhancement of commercialization efficiency of universities, the breakthrough points of it could also be from correcting the working attitudes and reducing the mood of risk aversion of UACO officers. Governments at all levels and universities should create a favorable atmosphere for achievements commercialization of universities, and reduce the adverse impact of outside uncertain factors on the commercialization of research findings in universities.

5. Conclusions

The value of this paper is that, a principal-agent relation between university and UACO is analyzed, and the optimal incentive contract designing model is proposed on the question of enhancing the efficiency of universities

achievements commercialization. The results show that, the efficiency of universities achievements commercialization has a squared forward growth relation with UACO working ability, and it is directly proportional to willingness of hardworking of UACO, and is inversely proportional to risk aversion degree of UACO and to outside uncertainty. The level of hardworking of UACO under the condition of information asymmetry is strictly less than that of information symmetry. For the enhancement of commercialization efficiency, universities need to set up incentive mechanism in the wage system of UACO, and to determinate optimal fixed salary and incentive intensity on the basis of UACO working ability, willingness of hardworking, degree of risk aversion and outside uncertain factors.

In addition, the tentative idea of fully-marketing operation pattern for UACO that produced from zero fixed salary, and a series of problem initiated from it, for instance, the pattern is feasible under what kind of circumstance and how does the patter work, are all worth of being discussed further.

6. Acknowledgements

This research is supported by the National Natural Science Foundation of China (No. 70772074) and National Social Science Foundation of China (No. BIA090049). The authors want to give their acknowledgement to Professor Liu Guoxin and Doctor Yan Junzhou who gave valuable comments on the paper.

REFERENCES

[1] C. M. Leitch and R. T. Harrison, "Maximising the potential of university spin-outs: The development of second-order commercialization activities," R&D Management, Vol. 35, No. 3, pp. 257–272, 2005.

[2] C. O'Gorman, O. Byrne, and D. Pandya, "How scientists commercialize new knowledge via entrepreneurship," Journal of Technology Transfer, Vol. 1, pp. 23–43, 2008.

[3] D. S. Siegel, D. A. Waldman, L. E. Atwater and A. N. Link, "Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: Qualitative evidence from the commercialization of university technologies," Journal of Engineering and Technology Management, Vol. 21, No. 2, pp. 115–142, 2004.

[4] P. M. Swamidass and V. Vulasa, "Why university inventions rarely produce income? Bottlenecks in university technology transfer," Journal of Technology Transfer, Vol. 34, No. 4, pp. 343–363, 2009.

[5] W. Chapple, A. Lockett, D. Siegel, and M. Wright, "Assessing the relative performance of UK university technology transfer offices: Parametric and non-parametric evidence," Research Policy, Vol. 34, No. 3, pp. 369–384, 2005.

[6] G. D. Markman, P. T. Gianiodis, and P. H. Phan, "Supply-side innovation and technology commercialization," Journal of Management Studies, Vol. 46, No. 4, pp. 625–649, 2009.

[7] W. Zhang, "Game theory and information economics," Shanghai, Shanghai People's Publishing House (in Chinese), pp. 238–239, 2004.

[8] S. Xie, "Economic game theory," Shanghai, Fudan University Press (in Chinese), pp. 137–139, 2002.

[9] J. Mirrlees, "The optimal structure of authority and incentives within an organization," Bell Journal of Economics, Vol. 7, No. 8, pp. 105–131, 1976.

[10] B. Holmstrom, "Moral hazard and observability," Bell Journal of Economics, Vol. 10, No. 12, pp. 74–91, 1979.

[11] X. Huang, X. Tao, and W. Xing, "Design of energy saving contract based on principal-agent theory," Industrial Engineering Journal (in Chinese), Vol. 12, No. 4, pp. 19–22, 2009.

[12] J. Mirrlees, "Notes on welfare economics, information and uncertainty," In Essays on Economic Behavior under Uncertainty, edited by M. Balch, D. McFadden and S. Y. Wu, Amsterdam: North-Holland, 1974.

J. Service Science & Management, **2010, 3: 84-90**
doi:10.4236/jssm.2010.31010 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Does Learning Process Mediate the Relationship between Social Control and Production Innovation of International Joint Ventures in China?

Cheng Wang, Hailin Lan, Hongming Xie

School of Business Administration, South China University of Technology, Guangzhou, China.
Email: fly2018@yahoo.com.cn

Received October 23rd, 2009; revised December 2nd, 2009; accepted January 16th, 2010.

ABSTRACT

A comprehensive model that delineates the interrelationships among social control, learning process and production innovation in international joint ventures is absent. This study aims to fill this void. Unlike previous research, this study investigates the role of social control, in facilitating learning process and production innovation in China. In our framework, we argued that learning process mediated the impact of social control on production innovation. In order to test the feasibility of this framework, we conducted an empirical study. This study employed a survey instrument, which contained data collected from 300 *organizations international joint ventures in South China. A total of* 96 *usable responses were analyzed. The results indicate that social control has a positive impact on learning processes and production innovation. Learning process partial mediate the influence of social control on production innovation. The implications of the study are provided, and future research is suggested.*

Keywords**:** *Learning Process, Production Innovation, Social Control, International Joint Ventures*

1. Introduction

International joint ventures (IJV) are now a common organizational form. Increased international joint venture activity has been seen as one of the major changes in international business environment in the past decade. The exercise of managerial control has been one of the most important subjects in the IJV literature. Control has long been recognized as an important aspect of an organization's management of exchange relationships. Although control considerations are important for any organization, they are especially so when managing organizational exchanges across diverse national environments. The major function of control involves bringing about conformance to organizational requirements and achievement of ultimate purpose of the organization.

While the literature reflects a fair amount of effort directed toward an understanding of control exercised as well as its antecedents and consequences [1–3]. It provides very little, if any, guidance about the effectiveness of various control mechanisms across different innovation modes used in structuring international principal agent exchange relationships.

Although firms form IJVs for a variety of reasons, interorganizational learning has become an important rationale for their creation [4,5]. While an IJV's parents often seek to learn from one another, IJV learning from its foreign parents is considered to be essential for its survival, hence, the realization of the parents' strategic goals [6,7].

Consider first technology-related capabilities. A highly technologically turbulent environment is characterized by a short cycle of technological innovation and obsolescence. In high turbulence, technology-related capabilities (such as innovation) should enable an IJV to shape or react to these environmental conditions [8]. For example, the timely introduction of new products to replace obsolete products may become crucial to firm success [9].

The purposes of this study are: a) to examine the role of social control in facilitating learning process; b) to realize the impact of learning process on the effect of adopting social control to facilitate production innovation.

This article proceeds as follows. First, the theory of social control is examined, focusing on the role of distinctive capabilities in innovation-based competitive strategy. Second, the conceptual framework describing the focal constructs and theoretical relationships intended to be tested is discussed. Third, the method used to test the hypotheses is discussed. Next, the results, based on analysis of data collected from 96 firms of international

joint ventures in China, are presented. The paper concludes by discussing implications for innovation theory and practice, identifying limitations of the study and providing directions for future research.

2. Theoretical Backgrounds and Hypothesis

2.1 Social Control

For many decades the topic of control has been a source of considerable discussion and thus it is not surprising to find a variety of approaches to conceptualizing and operational sing control.

Social control represents an informal control mechanism based on prevailing social perspectives and patterns of interpersonal interactions within a firm. In socially controlled exchange relationships, behavior is specified by the organizational culture, and performance is viewed as a social obligation maintained via mechanisms of social pressure [10]. Social control relies on implicit organization wide culture to monitor agents' behavior and to orient them toward organizational rather than parochial goals.

Similarly, socially controlled relationships are characterized by shared values and a common sense of well-being among members. Reliance on social control in external relationships assures that each partner is being treated fairly which in turn facilitates community interests fostering greater relationship flexibility. However, in internal exchanges, social controls are expected to be especially instrumental in making agents feel an integral part of the team, fostering shared values and aligning interest thereby leading to much greater relationship flexibility and adaptability. As a result, principals and agents can adopt a longer term perspective by balancing immediate results with long term nurturing relationships with each other. This is necessary because existence and continuity of social control requires stability of the exchange partners which in turn necessitate compromise and adaptability.

Through the process of socialization and indoctrination, the agents can be accounted wide latitude enabling them to respond to conditions quickly and in a manner consistent with long-term organizational goals and objectives. The socialization process can be used to create shared values and common interest among independent parties to encourage flexible and negotiated adjustment as well as to improve economic performance in terms of both sales growth and market share increases. Furthermore, by using social controls that sanction rather than punish risk taking, agents can be encouraged to try innovative approaches to achieve organizational performance objectives [11].

2.2 Learning Process

Acquiring knowledge related to those capabilities from foreign parents critical to the transitional economy IJVs' survival. As noted earlier, the survival of the IJV is a prerequisite for the domestic and foreign parents to achieve their goals. It is also a prerequisite for the domestic government to achieve its goals [12,13]. The common interest of all stakeholders in the survival of a transitional economy IJV creates strong incentives for the IJV and its parents to focus on its learning from the foreign parent. Hence, the analysis of organizational learning has become an increasingly important study area in international joint ventures. Organization learning involves the basic elements and processes of organizational development and growth. From a learning perspective, organizations grow when there is an increase in shared understanding involving the organization, its environment and the relationship between the two. A parent adopting a learning-oriented cooperative strategy usually possesses clear learning intent, which, in this study, refers to the level of desire and will of the parent with respect to learning from the international joint venturing experience.

Discussions of individual learning generally refer to the product of the learning process as changes in beliefs (or cognition) and changes in behavior. The term skill is often used to describe behaviors for which an individual has developed a level of expertise or proficiency. In a reliable learning process, an organization develops common understandings of its experience and makes its interpretations public, stable and shared. In a valid learning process, an organization understands, predicts and controls its environment. As organizations engage in learning efforts, neither reliability nor validity is assured because different people and groups in an organization approach historical experience with different expectations and beliefs. As we found, shared understanding about the value of IJV learning experiences was often obstructed by the variety and differences in managerial beliefs.

2.3 Production Innovation

Resource-based theory views a firm as a unique bundle of tangible and intangible resources and emphasizes the protection of firm core competencies comprising these resources. Staying on top of innovation and creativity in business is necessary for a business to stay ahead of the competition. New product introduction indicates the potential significance of a firm's innovation activities. New products are a central mechanism by which organizations diversify, adapt, and reinvent themselves in changing market and technical conditions [14].

In the past 20 years, many potentially useful typologies have been proposed, each providing insight to our understanding of the innovation process. Three of the most popular typologies are based on the distinctions between administrative and technical innovations, prod-

uct and process innovations, and radical and incremental innovations [15,16]. Technical innovations include products, processes and technologies used to produce products or render services related to the basic work activity of an organization. Product and process innovations are distinguished based on the different areas and activities that each of them affect within the firm [17]. Product innovations are outputs or services that are introduced for the benefit of customers or clients.

Accordingly, the following hypothesis is advanced:

H1: The higher the level of social control, the higher the level of learning process

H2: social control is a function of its production innovation

H3: learning process will significantly impact on production innovation

In order to explore the impact of social control and learning process on production innovation, we developed the research framework in **Figure 1**. There are two research questions. a) Does social control play a role in facilitating learning process? b) Does learning process mediate the effect of social control on facilitating production innovation?

3. Methods

3.1 Sample

Data were collected from international joint ventures of China through a survey instrument. An initial version of the survey instrument was developed based on the theory-grounded operationalization of the various constructs. This version was subsequently revised through pretesting with academic and experts. The multiple phases of instrument testing and development resulted in a significant degree of refinement and restructuring of the survey instrument as well as establishing the initial content validity [17]. We had three sampling criteria: the firms had to be 1) at least 1years; 2) at least 50 employees; 3) manufacturing and service firms in South China. We received responses from 118 of 300 IJV firms, a response rate of 39.3 percent. Of the 118 returned questionnaires, 16 were excluded because they did not meet all sampling criteria and 6 were excluded because of incomplete answers, leaving 96 usable responses. Hence, our usable response rate was approximately 32 percent.

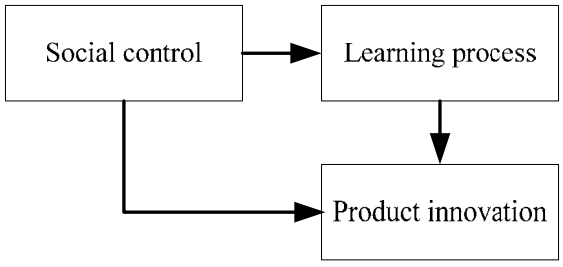

Figure 1. Research framework

Location and age data on no respondents indicate no significant differences between respondents and no respondents. Because those responding late are argued to be similar in composition to no respondents, we also tested for response bias by comparing early against late respondents. We found no significant differences in terms of sales, employees, or international sales. The results of statistical tests for response bias suggest that our sample is representative of the firm satisfying our criteria.

3.2 Reliability and Validity

Factor analysis using principal components factor analysis with factor extraction and VARIMAX rotation was conducted to examine the convergent and discriminant validity [18]. The four commonly employed decision rules were applied to identify the factors [19]: 1) minimum Eigen value of 1; 2) minimum factor loading of 0.4 for each indicator item; 3) simplicity of factor structure; and 4) exclusion of single item factors. Reliability was evaluated by assessing the internal consistency of the indicator items of each construct by using Cronbach's α [20].

The results of factor analysis relating to convergent validity are shown in Appendix.

3.3 Measures

Social control represents an informal control mechanism where monitoring occurs through interpersonal interactions. As Jaeger and Baliga (1985) [21] notes, in socially controlled exchanges, one would expect the interactions between the principal and foreign agents to be weighted in favor of personal contacts such as visits, social meetings, and use of expatriates. Based on this conceptualization, a five-item scale was developed to capture the informal interactions and socialization opportunities in principal-agent relationships. Alpha coefficient for this five item scale is 0.80.

Following Chou (2003) [22], learning process consist of decision to contribute and decision to adopt. For an individual to contribute knowledge one has to formulate the tacit and explicit knowledge about what has been learned, what the problems were, what kind of know-how I used to solve the problems, and what the context for a solution was. The other activity for 'contribute' is to delivery such knowledge to the person who needs it. In order to do so, knowledge has to be represented in a way that is meaningful and easy to understand to others. In order to adopt knowledge, an individual has to search for the possible solutions and match the problems to the appropriate solutions. Alpha coefficient for this seven-item scale is 0.85.

Product innovation was adapted from the scale developed by XIE Hongming and Wang Cheng *et al.* (2007) [23]. Product innovation is measured by a five-item scale. Production innovation Alpha coefficient for this five-item scale is 0.85.

4. Analysis and Results

The means, standard deviations and correlations of all the variables are reported in **Table 1**. The bivariate correlations show that similarity is significantly related to social control, learning process and production innovation. Although the bivariate relationships support our hypothesis, we use structural equation modeling approach to actually test the hypothesis, thereby taking into account any relationships between the latent variables and observable indicators.

A structural equation modeling approach was applied to the data using the AMOS version 7.0 software packing. **Figure 2** and **Table 2** shows the results of the competing model analysis. Although the χ^2 statistic for each model is significant, other relevant fit indices indicate good overall fit. The results show that learning process partial mediates the relationship between social control and production innovation. First, positive relationships exist between social control and learning process (H1: β=0.27, t=2.099) and learning process and production innovation (H3: β=0.36, t=2.913). Second, the significant relation ship between social control and production innovation

Table 1. Means, standard deviations, correlations

Multi-item scales	Means	S.D.	1	2	3
(1) Social control	3.316	0.663	1.000		
(2)Learning process	3.826	0.498	0.276***	1.000	
(3) Production innovation	3.330	0.695	0.447***	0.412***	1.000

indicated in the direct effects model (β=0.53, t=3.681), it is also significant in the partial mediation model (H2: β=0.44, t=3.286). Together these three points provide compelling evidence that there exists a partial mediating effect of learning process on the relationship between social control and production innovation.

5. Conclusions and Discussions

This study investigated the role of social control in facilitating learning process and production innovation. Based on 96 respondents from organizations in manufacturing, the service industry of international joint ven tures in China, we found that the functions of social control have a positive impact on the organizational learning processes. We also examined the impact of learning process on the adopting of social control to facilitate production innovation. Two interesting results were found. First, learning process is related to the effect of adopting social control to facilitate production innovation. Second, learning process partial mediates the influ-

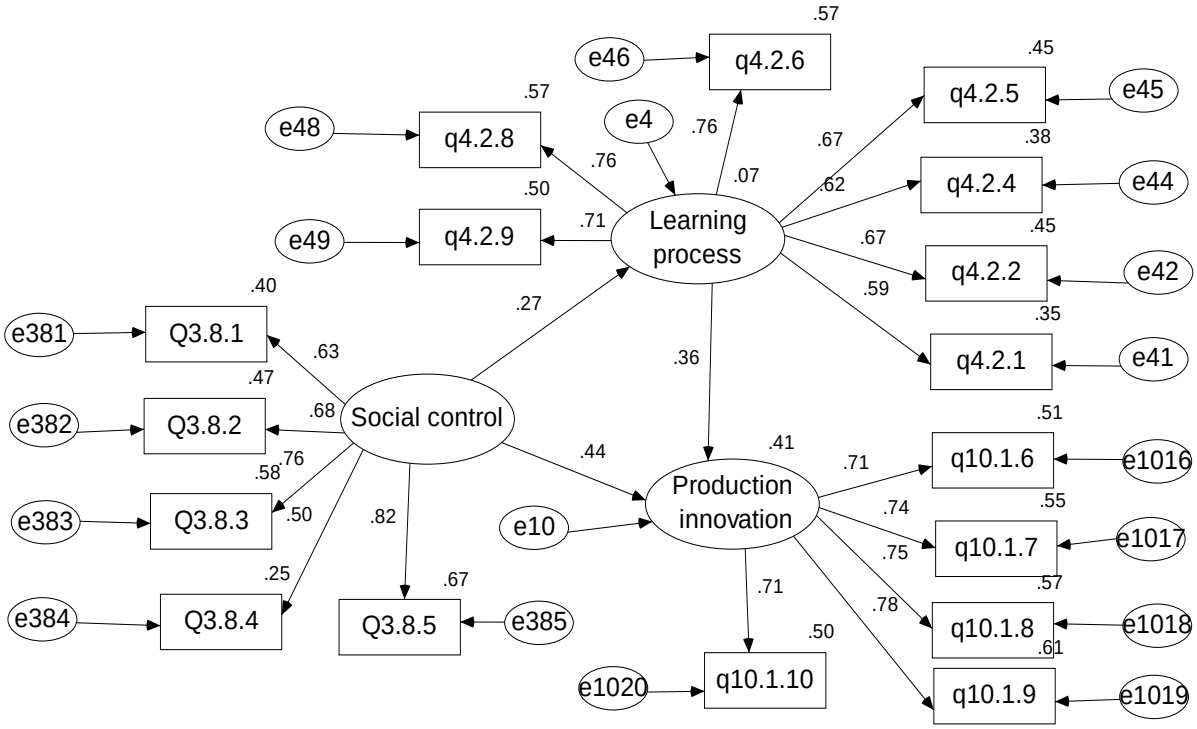

Figure 2. The overall theory model and parameter

Table 2. Test results that learning process partial mediate the influence of social control on production innovation

Parameter	Direct effects model	Partial mediation model
Hypothesized paths		
Social control →Learning process	—	0.27(2.099)
Social control →Product innovation	0.53(3.681)	0.44.(3.286)
Learning process →Product innovation	—	0.36(2.913)
Goodness-of-fit statistics		
χ^2	=53.624	=204.397
d.f.	=34	=116
GFI	=0.903	=0.809
Comparative fit index(CFI)	=0.947	=0.871
Tucker-Lewis index(TLI)	=0.930	=0.849
RMR	=0.044	=0.051
RMSEA	=0.078	=0.090

ence of social control on production innovation. Unlike previous research, this paper examines the impact of social control on production innovation in a more comprehensive way. First, we integrate social control, learning process and production innovation in a theoretical model. Second, we conducted an empirical study to specify the learning process that mediates the influences of the adoption of social control on facilitating production innovation. By emphasizing the features of social control, we may facilitate the effect of learning process. On the other hand, we cannot overlook learning process, since it may increase or decrease the effect of production innovation. The implications of this study are three fold.

First, this research explores the social control features that may facilitate learning process. Some of the features are new in supporting learning process. Understanding the features of learning process that facilitate organizational learning is very helpful for management. Managers should emphasize the learning process.

Second, understanding the characteristics of the organizational learning process in an in international joint ventures environment may help us to design the facilitating mechanisms accordingly. The difference of the organizational learning process between in international joint ventures and regular environment is therefore worthy of future research.

Finally, our empirical study identified social control and learning process synergistic affect production innovation. In order to obtain the most effective way of production innovation, it is crucial that managers develop social control adopting strategy. This study further supports the belief that production innovation should be congruent with the learning process international joint ventures. To promote production innovation, firms should actively manage their learning processes. Whenever a new IJV is set up, its learning potential has to be carefully evaluated. Management should examine wheth- er the IJV provides novel and useful management experience and whether knowledge can be acquired from other partners of the IJV. The evaluation step has to be followed by allocating appropriate resources for the purpose of learning. To exploit the learning potential identified above, clear objectives with respect to learning have to be set for managers who are sent to work in the IJV.

Like all research, this study has left some unanswered questions that suggest avenues for future research. First, our sample of South China may not be representative of China. Thus, a direction for future research would be to test the various hypotheses in this study using a broader sample of China. Further, the sample size, 96 IJV, was relatively small, which might account for some of the marginal effects that we observed; future research should aim to replicate this study with a larger sample. Second, the use of data and analysis do not allow the possibility of bidirectional (feedback) effects to be explored. Future studies can collect longitudinal data to assess such bidirectional (feedback) effects. Longitudinal data may also be sought in future studies in order to explore the causal link between social control, learning process and production innovation. The model proposed and tested in this study is based on the notion that learning process partial mediates the influence of social control on production innovation. However, as a relationship develops, the model may also affect the evolution of a mix of control

mechanisms used to achieve desired production innovation. Hence, an interesting approach for future research would be a longitudinal analysis of how control mechanisms used in production innovation. Third, we discuss only social control; we believe researchers should include control mechanism other than just social control when studying the production innovation of international joint ventures in China.

6. Acknowledgement

This research was supported by National Natural Science Found of China (Key Program) under Grant 70832003, National Natural Science Fund of China (NSFC) under Grant 70572095 and 70972108, Program for New Century Excellent Talents in University under Grant NCET-07-0308.

REFERENCES

[1] K. M. Eisenhardt, "Control: Organizational and economics approaches," Management Science, Vol. 31, No. 2, pp. 134–149, 1985.

[2] W. G. Ouchi, "Markets, bureaucracies, and clans, administrative science quarterly," Vol. 25, pp. 129–141, 1980.

[3] W. G. Ouchi and M. A. Maguire, "Organizational control: Two functions," Administrative Science Quarterly, Vol. 20, pp. 559–569, 1975.

[4] G. Hame, "Competition for competence and interpartner learning within international strategic alliances," Strategic Management Journal, Vol. 12, pp. 83–103, 1991.

[5] B. Kogut, "Joint ventures: Theoretical and empirical perspectives," Strategic Management Journal, Vol. 9, No. 4, pp. 319–322, 1988.

[6] J. F. Hennart, "A transaction costs theory of equity joint ventures: An empirical study of Japanese subsidiaries in the United States," Management Science, Vol. 37, No. 4, pp. 483–497, 1991.

[7] A. C. Inkpen and P. W. Beamish, "Knowledge, bargaining power and international joint venture instability," Academy of Management Review, Vol. 22, No. 1, pp. 177–202, 1997.

[8] M. Kotabe and K. S. Swan, "The role of strategic alliances in high-technology new product development," Strategic Management Journal, Vol. 16, No. 8, pp. 621–636, 1995.

[9] J. Wind and V. Mahajan, "Issues and opportunities in new product development: An introduction to the special issue," Journal of Marketing Research, Vol. 34, pp. 1–12, 1997.

[10] A. M. Jaeger, "The transfer of organizational culture overseas: An approach to control in the multinational corporation," Journal of International Business Studies, Vol. 04, pp. 91–114, 1983.

[11] B. R. Baliga and A. M. Jaeger, "Multinational corporations: Control systems and delegation issues," Journal of International Business Studies, Vol. 04, pp. 25–40, 1984.

[12] R. D. Hisrich and P. Szirmai, "Developing a market-oriented economy: A Hungarian perspective," Entrepreneurship and Regional Development, Vol. 5, pp. 61–71, 1993.

[13] M. A. Lyles and I. S. Baird, "Performance of international joint ventures in two Eastern European countries: The case of Hungary and Poland," Management International Review, Vol. 34, No. 4, pp. 313–330, 1994.

[14] C. Schoonhoven, K. Eisenhardt, and K. Lyman, "Speeding products to market: Waiting time to first product introduction in new firms," Administrative Science Quarterly, Vol. 35, pp. 177–207, 1990.

[15] F. Damanpour, "Organizational innovation: A meta-analysis of effects of determinants and moderators," Academy of Management Journal, Vol. 34, No. 3, pp. 555–590, 1991.

[16] S. Gopalakrishnan and F. Damanpour, "A review of innovation research in economics, sociology and technology management," The International Journal of Management Science, Vol. 25, No. 1, pp. 15–28, 1997.

[17] J. C. Nunnally, "Psychometric theory," New York: McGraw-Hill, 1978.

[18] J. L. Price and C. W. Mueller, "Handbook of organizational measurement," Marshfield, MA: Pitman Publishing Inc. 1986.

[19] J. F. Hair, R. E. Anderson, R. L. Tatham, and B. J. Grablowsky, "Multivariate data analysis," Tulsa, OK: PPC Books, 1979.

[20] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," Psychometrika, Vol. 16, No. 3, pp. 297–334, 1951.

[21] A. M. Jaeger and B. R. Baliga, "Control systems and strategic adaption: Lessons from the Japanese experience," Strategic Management Journal, Vol. 6, pp. 115–134, 1985.

[22] S. W. Chou, "Computer systems to facilitating organizational learning: IT and organizational context," Expert Systems with Applications, Vol. 24, pp. 273–280, 2003.

[23] H. M. Xie, C. Wang, X. C. Li, and Y. Y. Ou, "An empirical research of influencing technical innovation's factor," Studies in science of science, Vol. 25, No. 5, pp. 963–970, 2007.

Appendix

Table 3. Social control (five indicator variables; Cronbach's α = 0.803

Factor	loadings
1. We often have social meetings where our firm managers and foreign agents interact with each other	0.745
2. We frequently send our managers to this foreign country to update our agents about product changes	0.763
3. Our firm managers make frequent trips to this foreign country to meet with our agents	0.810
4. Our foreign agents make frequent trips and visit our business headquarters	0.605
5. We provide training to our foreign agents in our business headquarters	0.834

Table 4. Learning process (five indicator variables; Cronbach's α = 0.852

Factor	loadings
1. Employees are wiling to contribute their expertise	0.668
2. I can contribute my working experience easily	0.744
3. A learning community may eliminate temporal and spatial constraints	0.675
4. My problems will be solved by adopting colleagues' opinions	0.725
5. I can obtain the solution for a specific problem from a learning community	0.796
6. To solve the problems that may have in work practice, I need more innovative ideas	0.783
7. Learning community will facilitate the creation of innovative ideas	0.752

Table 5. Product innovation (five indicator variables; Cronbach's α = 0.854

Factor	loadings
1. This organization's new products acquire many innovation prizes	0.768
2. This organization develops abundance products	0.804
3. This organization has more patent than others	0.805
4. This organization mostly profit result from new products and service	0.831
5. This organization seeks out and acquires new technologies of improving process	0.781

J. Service Science & Management, **2010, 3: 91-97**
doi:10.4236/jssm.2010.31011 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Scientific Research

Padé Approximation Modelling of an Advertising-Sales Relationship*

M. C. Gil-Fariña, C. Gonzalez-Concepcion, C. Pestano-Gabino

Department of Applied Economics, Faculty of Economics and Business Administration, University of La Laguna (ULL), Campus of Guajara, Tenerife, Spain.
Email: {mgil, cogonzal, cpestano}@ull.es

Received December 18th, 2009; revised January 21st, 2010; accepted February 20th, 2010.

ABSTRACT

Forecasting reliable estimates on the future evolution of relevant variables is a main concern if decision makers in a variety of fields are to act with greater assurances. This paper considers a time series modelling method to predict relevant variables taking VARMA and Transfer Function models as its starting point. We make use of the rational Padé-Laurent Approximation, a relevant type of rational approximation in function theory that allows the decision maker to take part in the building of estimates by providing the available information and expectations for the decision variables. This method enhances the study of the dynamic relationship between variables in non-causal terms and allows for an ex ante sensibility analysis, an interesting matter in applied studies. The alternative proposed, however, must adhere to a type of model whose properties are of an asymptotic nature, meaning large chronological data series are required for its efficient application. The method is illustrated through the well-known data series on advertising and sales for the Lydia Pinkham Medicine Company, which has been used by various authors to illustrate their own proposals.

Keywords*: *Time Series Modelling, Expectations, Economics, Numerical Methods, Padé Approximation

1. Introduction

The possibility of forecasting reliable estimates for the future behaviour of relevant variables is main concern to decision making in numerous fields (including business, industry, energy, environmental, government agencies and medical and social network fields). Consequently, from a scientific standpoint, it is necessary to investigate alternative methods that can provide estimates while also introducing them into a technological system that allows the decision maker himself to participate in the composition of said predictions. Many researchers have attempted to satisfy this requirement from different perspectives, such that the prediction problem is always present in any generic data-mining task. And yet, the technique selected for use depends on the availability and type of data in relation to the hypotheses of the methods that sustain the desired technique and which yield different degrees of accuracy, time horizons and different computational and social costs. The use of a relevant technique within the scope of rational approximation, namely the Padé Approximation (PA), has had a gratifyingly enriching and stimulating effect on the study of the dynamic relationship structure between variables, especially within the context of identifying univariate and multivariate ARMA models (see, among others, [1–5]).

In particular, within the scope of rational model formulation in causal terms, this technique acquires a special relevance for characterizing simple models at a computational level with which to specify the deterministic part in certain time series models. At the same time, it provides reliable initial estimates for obtaining a definitive model by using more efficient, iterative methods once the random component has been identified.

The choice of a particular VARMA model, namely the Transfer Function [6] model which constitutes one of the most widely used representations in the input-output context of dynamic stochastic systems through the use of polynomial rational expressions, serves to highlight the influence that the expectations or forecasts for an explanatory variable (input) exerts on the model under consideration.

In this sense, the formulation of the time series analysis instruments that encompass the available sampling

*This research was partially funded by *Plan Nacional* Project no. MTM2008-06671, MTM2006-14961-C05-03 and the *Centro de Investigación Matemática de Canarias* (CIMAC).

information and the establishment of non-causal models into which the forecasts are incorporated, that is, the knowledge that by way of *ex ante* or *ex post* expectations is provided by economic theory or by the empirical evidence, if any, about the model's explanatory variable(s), provide a basic framework for tackling the study of the rational identification of time series models from a more general context.

The method that allows for a sustained study of the time series identification from an evolutionary, but not necessarily causal, standpoint of the variables involved is based on a generalization of the PA concept to the study of formal Laurent series [7]. The use of this approach allows for the study of new dynamic identification procedures by means of a single model that simultaneously approximates both time directions in a formal Laurent series while encompassing the classical case as a specific one when the expectations are not included (see, for example, [8,9]).

The consideration of this broader dynamic framework, however, into which future information on the model's variables is incorporated, favours a continuous feedback process through which the formulated expectations are replaced period by period by updated information as the empirical evidence modifies or confirms the predictions made.

It is worth noting that the models are in no way insensitive to changes in the way the expectations are made; on the contrary, these changes lead to different dynamic behaviours insofar as it is the precise nature of the expectations that determines the explicit dynamics of the forecasted variables.

In this sense, the possibility of offering different models for a single data series by simply modifying the *ex ante* expectations made by economic agents allows us to obtain future knowledge about their influence on the model and therefore to contrast and compare different dynamic specifications so as to yield an optimum model.

In short, the performance of a sensibility analysis will permit for a contrast of the extent to which the forecasts of a variable's future behaviour affect the mode's predictions and, as a consequence, its adequacy to the empirical data.

So as to illustrate the rational Padé Approximation in modelling time series, we develop an application for the study of data from the advertising-sales series involving the activity of the Lydia E. Pinkham Medicine Company for the period 1907–1960 [10]. This series, which has been analyzed by numerous authors to illustrate their proposed methods, presents, as noted in [11,12], various characteristics which make it the ideal example for studying the relationship between both variables. Some of the reasons that justify the prominent role of this series in studies conducted to date are, among others, the importance of advertising as the company's sole marketing instrument, as well as the elevated advertising costs to sales ratio (above 40%) which, in some instances, even exceeded 80%.

This paper is structured into three sections. The first two present certain theoretical foundations for the properties of the PA that allow for the identification of the orders in VARMA models and TF models with expectations. We note the last section, which is devoted to the empirical results of the study on the aforementioned advertising-sales series. We conclude the work with the more relevant conclusions and the main references.

2. VARMA Models

A non-deterministic, k-dimensional centred process can be expressed as a vector autoregressive moving average model (VARMA(*p, q*)) if $A_p(B)X_t = B_q(B)a_t$, where

$$A_p(B) = I - A_1B - ... - A_pB^p$$

and

$$B_q(B) = I - B_1B - ... - B_qB^q$$

are *k*x*k* polynomial matrices in the lag operator *B*, that is, the coefficients A_i and B_j $(i = 1,..., p; j = 1,..., q)$ are *k*x*k* matrices.

The *k* vector a_t is a series of i.i.d. random variables with a zero mean multivariate norm and covariance matrix Σ.

The series X_t is said to be stationary when the zeros of $|A_P(B)|$ are outside the unit circle and invertible when the same can be said for those of $|B_q(B)|$.

This type of model is useful for understanding the dynamic relationships between the components of the series X_t. In this sense, one series can cause another or there may be a feedback relationship or they may be contemporaneously related.

In the case of the Lydia Pinkham advertising-sales series, the joint modelling of these effects by means of the procedures described in [13] allows the type of dynamic relationship existing between both variables to be determined [14].

A consideration of the PA matrix method yields the following theorem, which can be used in the first steps of the VARMA model identification.

Theorem 1 [15].- Let X_t be a second-order stationary *k*-dimensional process. Let $R(h) = Cov(X_t\ X_{t-h})$ be the covariance matrices of the process. Let

$$M1(i, j) = ((R(i - j + 1 + k + h))_{k,h=0}^{j-1}.$$

Then,

$$X_t \sim \text{VARMA}(p, q) \Leftrightarrow \text{rank}\, M1(i, j) =$$

$$\text{rank}\ \ M1(i + 1, j + 1)\ \ \forall i, j / i \geq q, j \geq p$$

3. Transfer Function Models with Expectations

Based on the PA definitions for a formal power series [16] and its extension to the study of formal Laurent series [7], we can provide a characterization for the identification of a TF model with expectations by means of the Toeplitz determinants

$$T_{f,g}(c_i) = det\left[(c_{f+k-j})_{k,j=1}^{g}\right]$$

Given two stationary time series y_t, x_t, let us assume the existence of a unidirectional causal dynamic relationship $x_t \to y_t$ given by the combination of simultaneous and shifted effects of the input variables (including the presence of expected values that may or may not follow the same distribution as the data), and let us consider a generalized TF model of the form:

$$y_t = v(B)x_t + N_t; \quad v(B) = \sum_{i=-\infty}^{\infty} v_i B^i$$

in which the x_{t-i} refers to the present and past of the input series (data) if $i \geq 0$ and to the expected values of the input series (expectations) if $i < 0$, and such that the exogenous variables represented by x_t satisfy a VARMA model.

A finite-order representation for the Impulse Response Function (IRF), namely, $v(B) = \sum_{i=-\infty}^{\infty} v_i B^i$ that simultaneously approximates both directions in time and enables the estimation of a finite number of observations will be characterized by the following result:

Theorem 2 [17].- Given the series $v(B) = \sum_{i=-\infty}^{\infty} v_i B^i$, the following conditions are equivalent:

a) $$v(B) = \frac{\sum_{i=H}^{K} a_i B^i}{(1+\sum_{i=1}^{N} q_{-i} B^{-i})(1+\sum_{i=1}^{U} q_i B^i)}$$

b)

$$T_{H,N}(v_i) \neq 0,\ T_{K,U}(v_i) \neq 0;\ T_{J,M}(v_i) = 0\ \forall J < \mathrm{H} \wedge M > \mathrm{N};$$

$$T_{J,M}(v_i) = 0\ \forall J > K \wedge M > U$$

Due to the properties of the lag operator, the following equivalency holds:

$$y_t = \frac{A_{H,K}(B)}{Q_{-N,U}(B)} x_t + N_t \equiv \frac{A_{I,L}(B)}{Q_M(B)} x_t + N_t$$

where

$$I = H + N,\ L = K + N,\ M = N + U, A_{I,L}(B) \text{ y } Q_M(B)$$

are polynomials of the form:

$$A_{I,L}(B) = \sum_{i=I}^{L} a_i B^i \qquad Q_M(B) = \sum_{i=0}^{M} q_i B^i$$

Assuming the non-existence of common roots and the conditions for ensuring the stability of the model hold, the problem will consist of determining, in keeping with the sample information available and by means of a direct estimation of the IRF, the best orders I, L, M that describe $v(B)$.

With this in mind, the steps to follow in studying the dynamic identification for a generalized formulation of the TF model are:

1) Obtain the estimates $\hat{v}_i$ for the weights v_i of the IRF $v(B)$, approximating $v(B)$ by a finite number of terms, that is, $v(B) \cong \sum_{i=k}^{k'} \hat{v}_i B^i$ with $k, k' \in Z$, normally $k < 0, k' > 0$.

2) Calculate the Toeplitz determinants associated with the series of estimated relative weights $\hat{\eta}_i = \frac{\hat{v}_i}{\max_{k \leq i \leq k} |\hat{v}_i|}$ and arrange them in a tabular form (T-table) as a generalization of the C-Table for the classic case.

3) Estimate the model parameters.

4. Empirical Results

The first model for the Lydia Pinkham advertising-sales for the period 1907–1960 involving the TF models method was proposed by [10], which assumed a unidirectional causal dynamic relationship from advertising to sales. The possible existence of a relationship in the other direction has, however, been mentioned by other authors. Various subsequent papers, including [14,18,19], have illustrated the application of this method using the analyzed series as reference. This assumption of unidirectional causality has been questioned on several occasions, however, as some maintain there is a feedback relationship [11,14,19,20]. That is why, in what follows, we analyze two cases, one involving the identification of a VARMA model, and another in which, assuming the existence of a unidirectional causal dynamic relationship from advertising to sales, we present various TF models and analyze the influence on the sales trends of different behaviour schemes or *ex ante* forecasts for the variable input, which in this case is advertising.

In any case, and given that both series exhibit non-stationarity, we present the models by taking first differences.

4.1 VARMA Models

Once the ranks for $M1(i, j)$, $0 \leq i \leq 5,\ 0 \leq j \leq 5$ are obtained, applying Theorem 1 to the aforementioned data

provides the results shown in Tables 1 and 2, depending on the sample size used in the estimates. This indicates, as per the above theorem, that the differentiated data follow a VARMA (0, 1) model exchangeable with (1, 0), according to Table 1, and a VARMA (0, 2) interchangeable with (2, 0), according to Table 2.

Note how in Table 1, even though square (1, 1) does not have the same value as squares (i, i), i≥2 due to rounding errors in the calculations, using theoretical Padé Matrix Approximation results we know that these values have to coincide.

Once the model orders are calculated, the maximization of the likelihood function allows for a determination of efficient estimators. Using the Time Series Processor (TSP) software package [21], the results for the estimated VARMA (1, 0) and VARMA (2, 0) models yield:

$$X_t - \begin{pmatrix} -0.20 & 0.40 \\ -0.05 & 0.45 \end{pmatrix} X_{t-1} = \varepsilon_t$$

$$\Sigma = \begin{pmatrix} 44132 & \\ 23893 & 46790 \end{pmatrix}$$

and

$$X_t - \begin{pmatrix} -0.25 & 0.55 \\ -0.09 & 0.51 \end{pmatrix} X_{t-1} - \begin{pmatrix} -0.49 & -0.04 \\ -0.32 & 0.08 \end{pmatrix} X_{t-2} = \varepsilon_t$$

$$\Sigma = \begin{pmatrix} 33326 & \\ 18249 & 45672 \end{pmatrix}$$

Table 1. Orders of VARMA models. Alternative 1

	0	**1**	**2**	**3**	**4**	**5**
0	0	0	0	0	0	0
1	2	1	0	0	0	0
2	4	2	2	0	0	0
3	5	4	2	2	0	0
4	6	5	4	2	2	0
5	7	6	5	4	2	2

Table 2. Orders of VARMA models. Alternative 2

	0	**1**	**2**	**3**	**4**	**5**
0	0	0	0	0	0	0
1	2	2	1	0	0	0
2	3	3	3	1	0	0
3	4	3	3	3	1	0
4	5	4	3	3	3	1
5	6	5	4	3	3	3

where $X_t = (X_{1t,} X_{2t})^t$ and X_{1t} and X_{2t} are the first differences of the advertising and sales data, respectively.

As for the relationship between variables, we can conclude that a) the value of element (1, 2) in coefficient A_1 in both models, namely 0.40 and 0.55, indicates the existence of a relationship between sales and future advertising in a period, b) the negligible value of element (2,1) in coefficient A_1, namely -0.05 and -0.09, indicates a weak relationship between advertising and future sales, and c) the estimated correlation between residuals $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ indicates that advertising and sales are contemporaneously related.

Figures 1 and **2** show both models.

4.2 FT Models with Expectations

Taking into account the relationship between sales and future advertising noted in the above VARMA models, we now build transfer function models associated with three specific cases, depending on whether the advertising expectations follow a increasing or decreasing trend or whether they respect the predictions of the ARMA model for the input series (advertising), that is,

$$(1+0.269B^2 - 0.325B^4)\Delta x_t = N_t .$$

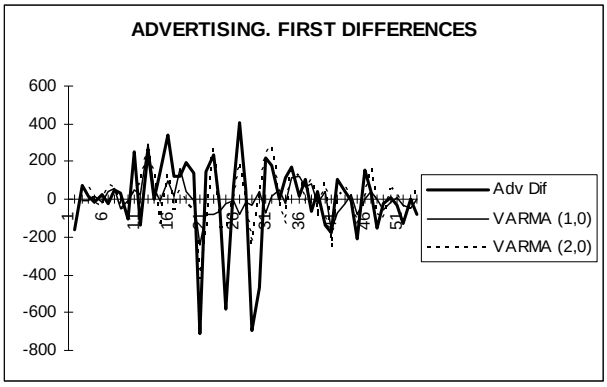

Figure 1. Advertising. First differences

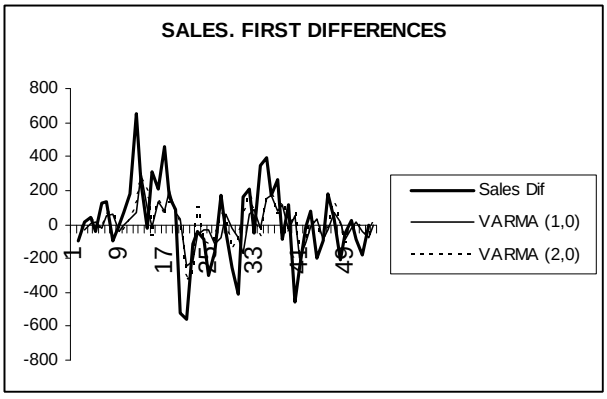

Figure 2. Sales. First differences

Starting from a generalized TF model for the advertising (x_t) and sales (y_t) variables, and keeping in mind that the series are first-order integrable, we formulate the model

$$\Delta y_t = v(B)\Delta x_t + a_t$$

in which $v(B)$ is approximated by a finite number of terms from k to k', that is, $v(B) \cong \sum_{i=k}^{k'} \hat{v}_i B^i$, normally $k < 0, k' > 0$.

Following the steps suggested by the method proposed yields the following results:

Case a: The expectations follow an increasing trend

After estimating the weights $\hat{v}_i$, the resulting T-table for the series of estimated relative weights is as shown in Table 3.

The behaviour of this table suggests a model of the form $\frac{A_{-3,0}(B)}{Q_{-1,1}(B)} \equiv \frac{\hat{A}_{-2,1}(B)}{\hat{Q}_{0,2}(B)}$

Specifically, once the parameters are estimated and the noise process is analyzed using the SCA software [22], the resulting model is:

$$\Delta y_t = \frac{.4334B^{-1} - .3267B}{1 - 1.5010B + .6518B^2}\Delta x_t + a_t$$

Table 3. T-Table. Orders of TF models with expectations in Case a

	1	2	3	4	5	6	7
-12	-.29	.09	-.02	.01	.00	.00	.00
-11	-.11	.21	-.05	.05	-.02	.01	.00
-10	.68	.45	.32	.24	.14	.09	.05
-9	-.11	-.07	.12	.24	.07	.04	.08
-8	.12	.04	.10	.21	-.03	-.05	.09
-7	.23	.14	.01	.19	.16	.14	.06
-6	-.76	.47	-.28	.16	.04	-.18	.25
-5	.46	.05	.16	.20	.20	.17	.45
-4	-.22	-.15	-.06	.06	.15	.34	.27
-3	.43	.23	.08	.06	.01	.44	.56
-2	.21	-.13	.013	.05	-.17	.57	.12
-1	.40	-.05	.31	.44	.51	.78	.82
0	1.00	.93	.87	.75	.51	.35	-.01
1	.16	-.16	.16	.28	-.01	.17	.02
2	.18	.00	.08	.11	-.09	.08	-.02
3	.22	.09	.04	.07	.01	.03	.03
4	-.26	.08	-.06	.04	.02	.00	-.03
5	-.05	-.09	.01	.04	.01	.02	.02

Case b: The expectations follow a decreasing trend

In this case, the method yields the following model:

$$\Delta y_t = \frac{.4331B^{-1} - .3285B}{1 - 1.5016B + .6514B^2}\Delta x_t + a_t$$

as suggested by Table 4.

Note the analogy between the models for cases a) and b).

Case c: The expectations are generated by the ARMA model of the input series

In this case, the T-table obtained by the series of estimated relative weights is shown in Table 5 and suggests a model of the form $\frac{A_{-3,0}(B)}{Q_{-2,1}(B)} \equiv \frac{\hat{A}_{-1,2}(B)}{\hat{Q}_{0,3}(B)}$ which, once estimated, yields

$$\Delta y_t = \frac{.5413B^{-1} - .1775}{1 - 1.3754B + .6612B^2}\Delta x_t + a_t$$

whose parameters differ significantly from those obtained in the first two cases.

The results for the three cases considered are shown in **Figures 3, 4** and **5**.

As we can see, the incorporation in the model of the *ex ante* advertising forecasts as well as the various formative mechanisms that not only include the information contained in the available historical data but also that derived from the company's desires and the strategies and outlooks devised by the economic agents in the decision-making process, facilitate obtaining valid dynamic formulations from a data fitting perspective. In addition,

Table 4. T-Table. Orders of TF models with expectations in Case b

	1	2	3	4	5	6	7
-12	-.21	.04	-.01	.00	.00	.00	.00
-11	-.14	.19	-.06	.04	-.02	.01	.00
-10	.83	.68	.58	.51	.42	.34	.28
-9	-.14	-.13	.20	.47	.10	.01	.25
-8	.18	.07	.18	.39	.01	-.07	.21
-7	.25	.21	.02	.32	.28	.25	.14
-6	-.82	.56	-.38	.25	.04	-.30	.51
-5	.46	.02	.16	.24	.27	.27	.76
-4	-.23	-.13	-.05	.06	.15	.44	.40
-3	.39	.20	.07	.05	-.02	.50	.69
-2	.22	-.11	.10	.06	-.16	.60	.18
-1	.41	-.05	.26	.41	.50	.78	.83
0	1.00	.91	.83	.71	.47	.33	-.02
1	.21	-.13	.16	.26	-.03	.15	.01
2	.18	-.02	.07	.10	-.08	.07	-.01
3	.23	.10	.04	.06	.01	.02	.03
4	-.23	.08	-.06	.03	.02	.00	-.03
5	-.10	-.07	.01	.03	.01	.02	.03

Table 5. T-Table. Orders of TF models with expectations in Case c

	1	2	3	4	5	6	7
-12	-.25	.06	-.02	.00	.00	.00	.00
-11	-.11	.21	-.05	.05	-.02	.01	-.01
-10	.79	.61	.50	.43	.33	.26	.20
-9	-.12	-.12	.19	.40	.12	.03	.22
-8	.17	.06	.17	.32	.01	-.10	.22
-7	.27	.21	.04	.25	.26	.27	.15
-6	-.79	.50	-.29	.15	.12	-.34	.53
-5	.49	.08	.18	.23	.26	.20	.69
-4	-.20	-.16	-.04	.06	.15	.39	.38
-3	.41	.21	.07	.04	-.01	.47	.66
-2	.23	-.12	.11	.04	-.14	.58	.15
-1	.41	-.06	.27	.40	.49	.77	.83
0	1.00	.92	.83	.70	.46	.31	-.04
1	.21	-.14	.17	.27	-.01	.15	.02
2	.18	-.02	.08	.11	-.09	.07	-.02
3	.25	.10	.05	.07	.01	.02	.03
4	-.24	.08	-.06	.04	.02	.00	-.03
5	-.07	-.08	.01	.04	.02	. 02	.02

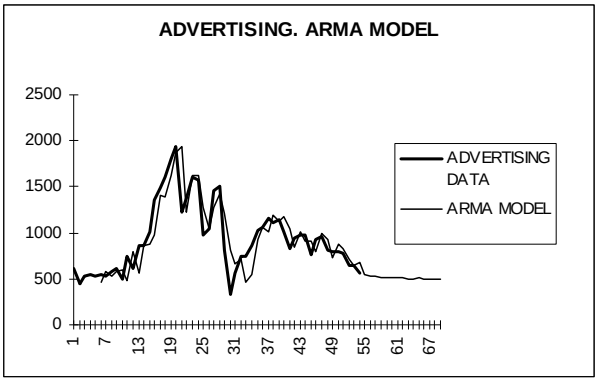

Figure 3. Advertising. ARMA model

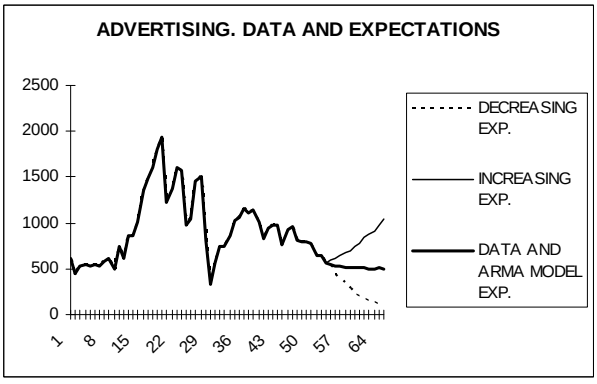

Figure 4. Advertising. Data and expectations

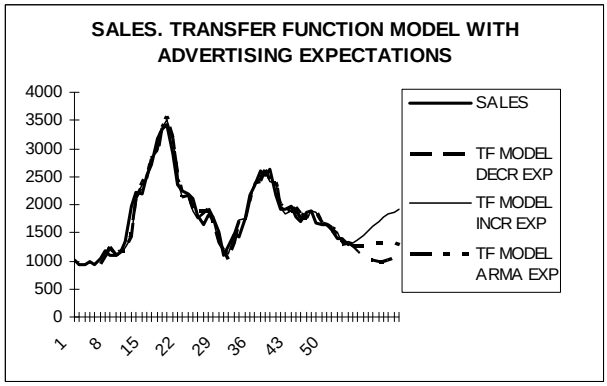

Figure 5. Sales. Transfer function model with advertising expectations

to the extent that a knowledge can be had of the future influence of said forecasts on the relationship under consideration, it is possible to evaluate its sensitivity to alternative dynamic specifications and, as a consequence, determine the optimum model.

5. Conclusions

In this paper we show the usefulness of the Padé-Laurent Approximation to the study of the deterministic part in the dynamic relationship between various variables, since it allows for the introduction of expectations or expected future values for certain variables. We illustrate the technique described by modelling the dynamic relationship between advertising and sales for the available data in the references consulted for the Lydia E. Pinkham Company. Also included is a sensitivity analysis of the estimates as a function of the expected future values or expectations of the decision maker for the advertising variable.

It would be interesting to combine this approach with the use of hybrid models that include a proper combination of linear and/or non-linear models of the type studied in [23].

REFERENCES

[1] J. M. Beguin, C. Gourieroux and A. Monfort, "Identification of a mixed autoregressive-moving average process: The corner method," In Time Series, O. D. Anderson, Ed., Amsterdam, North-Holland, pp. 423–436, 1980.

[2] G. M. Jenkins and A. S. Alavi, "Some aspects of modelling and forecasting multivariate time series," Journal of Time Series Analysis, Vol. 2, pp. 1–47, 1981.

[3] K. Lii, "Transfer function model order and parameter estimation," Journal of Time Series Analysis, Vol. 6, No. 3, pp. 153–169, 1985.

[4] P. Claverie, D. Szpiro and R. Topol, "Identification des modèles à fonction de transfert: La méthode Padé-transformée en z," Annales D'Economie et de Statistique, Vol. 17, pp. 145–161, 1990.

[5] A. Berlinet and C. Francq, "Identification of a univariate ARMA model," Computational Statistics, Vol. 9, pp. 117–133, 1994.

[6] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Time series analysis: Forecasting and control," (3rd ed.), Englewood Cliffs, Prentice-Hall, New Jersey, 1994.

[7] A. Bultheel, "Laurent series and their Padé approximations," Birkhaüser, Basel/Boston, 1987.

[8] C. González-Concepción and M. C. Gil-Fariña, "Padé approximation in economics," Numerical Algorithms, Vol. 33, No. 1–4, pp. 277–292, 2003.

[9] C. González-Concepción, M. C. Gil-Fariña and C. Pestano-Gabino, "Some algorithms to identify rational structures in stochastic processes with expectations," Journal of Mathematics and Statistics, Vol. 3, No. 4, pp. 268–276, 2007.

[10] R. M. Helmer and J. K. Johansson, "An exposition of the Box-Jenkins transfer function analysis with an application to the advertising-sales relationship," Journal of Marketing Research, Vol. 14, pp. 227–239, 1977.

[11] D. M. Hanssens, "Bivariate time-series analysis of the relationship between advertising and sales," Applied Economics, Vol. 12, pp. 329–339, 1980.

[12] A. Aznar and F. J. Trivez, "Métodos de predicción en economía. Análisis de series temporales," Editorial Ariel, Barcelona, Vol. 1–2, 1993.

[13] G. C. Tiao and R. S. Tsay, "Multiple time series modeling and extended sample cross-correlations," Journal of Business and Economics Statistics, Vol. 1, pp. 43–56, 1983.

[14] J. F. Heyse and W. W. Wei, "Modelling the Advertising-sales relationship through use of multiple time series techniques," Journal of Forecasting, Vol. 4, pp. 165–181, 1985.

[15] C. Pestano-Gabino and C. González-Concepción, "Rationality, minimality and uniqueness of representation of matrix formal power series," Journal of Computational and Applied Mathematics, Vol. 94, pp. 23–38, 1998.

[16] G. A. Baker (jr.) and P. Graves-Morris, "Padé approximants. Part I: Basic theory. Part II: Extensions and applications," Encyclopedia of Mathematics and Its Applications, Addison-Wesley, Reading, Vol. 13–14, 1981.

[17] M. C. Gil-Fariña, "La aproximación de Padé en la determinación de modelos racionales para series temporales doblemente infinitas," Estudios de Economía Aplicada, Universidad de Palma de Mallorca, Servicio de Publicaciones, Vol. 2, pp. 61–70, 1994.

[18] W. Vandaele, "Applied time series and Box-Jenkins models," Academic Press, New York, 1983.

[19] W. W. Wei, "Time series analysis: Univariate and multivariate methods," Addison-Wesley, Redwood City, 1990.

[20] M. N. Bhattacharyya, "Lydia Pinkham data remodelled," Journal of Time Series Analysis, Vol. 3, pp. 81–102, 1982.

[21] L.-M. Liu, G. B. Hudak, G. E. P. Box, M. E. Muller and G. C. Tiao, "Forecasting and time series analysis using the SCA statistical system," Scientific Computing Associates Corporation, Vol. 2, 2000.

[22] B. H. Hall and C. Cummins, "Time series processor user's guide," TSP International, Palo Alto, California, USA, 2005.

[23] R. Sallehuddin, S. M. H. Shamsuddin, S. Z. M. Hashim, and A. Abrahamy, "Forecasting time series data using hybrid grey relational artificial neural network and autoregressive integrated moving average model," Neural Network World, Vol. 6, pp. 573–605, 2007.

J. Service Science & Management, **2010, 3: 98-105**
doi:10.4236/jssm.2010.31012 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Analytic Approximations of Projectile Motion with Quadratic Air Resistance

R. D. H. Warburton[1], J. Wang[2], J. Burgdörfer[3]

[1]Department of Administrative Sciences, Metropolitan College, Boston University, Boston, USA; [2]Department of Physics, University of Massachusetts Dartmouth, North Dartmouth, USA; [3]Institute for Theoretical Physics, Vienna University of Technology, Vienna, Austria.
Email: rwarb@bu.edu

Received November 6th, 2009; revised December 11th, 2009; accepted January 18th, 2010.

ABSTRACT

We study projectile motion with air resistance quadratic in speed. We consider three regimes of approximation: low-angle trajectory where the horizontal velocity, u, is assumed to be much larger than the vertical velocity w; high-angle trajectory where $w \gg u$ *; and split-angle trajectory where* $w \sim u$ *. Closed form solutions for the range in the first regime are obtained in terms of the Lambert W function. The approximation is simple and accurate for low angle ballistics problems when compared to measured data. In addition, we find a surprising behavior that the range in this approximation is symmetric about* $\pi/4$ *, although the trajectories are asymmetric. We also give simple and practical formulas for accurate evaluations of the Lambert W function.*

Keywords**:** *Projectile Motion, Air Resistance Quadratic*

1. Introduction

In a previous paper on projectile motion with air resistance linear in speed, we presented closed form solutions for the range in terms of the Lambert *W* function [1]. Amid the growing list of problems that benefited from using the *W* function [2–7], one question naturally arises as to whether a similar approach exists if air resistance is quadratic in speed, the more realistic case in practice.

We have studied this problem and found that solutions exist for low-angle trajectories using the *W* function. In this paper, we report our findings, starting with an overview of the regimes of approximation in Section 2. We focus on the low-angle regime in Section 3 and discuss the dynamics, which leads to a remarkable property that the range is symmetric about $\pi/4$, even in the presence of air resistance. For completeness, high-angle and split-angle regimes are briefly discussed in Section 4 and 5, respectively, followed by a comparison with observed data and discussions in Section 6. To make the *W* function easily accessible, in the Appendix we give simple and practical formulas for the accurate evaluation of this function.

2. Regimes of Approximation

We assume the net force $\vec{F}$, including air resistance and gravity on the projectile of mass, m, to be $\vec{F} = -mbv\vec{v} - mg\vec{j}$. This leads to the equations of motion as

$$\frac{du}{dt} = -bvu, \quad \frac{dw}{dt} = -buw - g, \tag{1}$$

where *b* is the drag coefficient with the dimensions m^{-1}. The components of the velocity are $\vec{v} = u\vec{i} + w\vec{j}$, with $u = dx/dt$ and $w = dy/dt$, where *x* and *y* are the usual horizontal and vertical positions of the projectile. The initial position of the projectile is at the origin. Throughout the paper, we use the following notations for frequently occurring terms:

u_0, w_0 = initial horizontal and vertical velocities;

$$R_0 = \frac{2u_0 w_0}{g} = \text{ range with no air resistance;}$$

$$\beta = b\frac{2u_0 w_0}{g} \equiv bR_0 \tag{2}$$

To our knowledge, closed form solutions to (1) are known only for special initial conditions, not for arbitrary initial conditions [8]. The difficulty is with the coupling of *u* and *w* in the speed $v = \sqrt{u^2 + w^2}$ which makes the problem inseparable. Further approximations are necessary to solve (1) analytically.

We consider three approximations aimed at linearizing the speed according to the relative magnitudes of u and w: low-angle trajectory (LAT) approximation, $u \gg w$; high-angle trajectory (HAT) approximation, $w \gg u$; and split-angle trajectory (SAT) approximation, $u \sim |w|$, as,

$$v = \sqrt{u^2 + w^2} \approx \begin{cases} u & if\ u \gg w, LAT \\ |w| & if\ |w| \gg u, HAT \\ \sqrt{2}u\left(or\ |w|\right) & if\ u \gg |w|, SAT \end{cases} \quad (3)$$

Each of the approximations will be discussed below, with the emphasis on LAT that is important in practice and in ballistics.

3. Low-Angle Trajectory (LAT) Approximation

3.1 The Trajectory

We assume in this case the horizontal velocity is on average much greater than the vertical velocity, $u \gg w$. This will be the case if the firing angle is small, a case also discussed by Parker [9]. Here the speed may be approximated as $v \approx u$ according to (3), so that the equations of motion in the LAT approximation are

$$\frac{du}{dt} = -bu^2, \quad \frac{dw}{dt} = -buw - g. \quad (4)$$

One can solve for u first, after which w, x, and y can be obtained. Hereafter, we will omit non-essential intermediate steps. The solutions may be verified by substitution into the equations of motion. The solutions are

$$u = \frac{u_0}{1+at}, \quad w = \frac{w_0 - gt/2}{1+at} - \frac{1}{2}gt, \quad a \equiv bu_0, \quad (5)$$

$$x = \frac{1}{b}\ln(1+at), \quad y = \left(w_0 + \frac{g}{2a}\right)\frac{1}{a}\ln(1+at) - \frac{1}{4}gt^2 - \frac{gt}{2a}. \quad (6)$$

Equation 6 gives the trajectory in the LAT approximation.

The trajectories computed from (6) are plotted in **Figure 1** for three angles: 20°, 45°, and 70°. The initial speed is 9.8 m/s for all angles. The drag coefficient is $b = 0.1 m^{-1}$.

For comparison, we also show the trajectory from ideal projectile motion with no air resistance and the trajectory from the solutions with the full v^2 air resistance, Equation 1. We will simply refer to the former as the ideal motion, and the latter as the full solution.

The full solution is carried out by numerically integrating the equations of motion with the full v^2 resistance (1), using the Runge-Kutta method. (The two curves labeled HAT and SAT in **Figure 1** are discussed in Sections 4 and 5.)

The agreement between the LAT approximation and the full solution is good at 20° (nearly indistinguishable in **Figure 1(a)**), and it becomes worse for larger angles. This is as expected since the assumption was that LAT is valid only at small angles. The range is much reduced compared to the ideal motion. Air resistance introduces in the trajectories a well-known backward-forward asymmetry. The ascending part of the trajectory is shallower and longer, and the descending part is steeper and shorter.

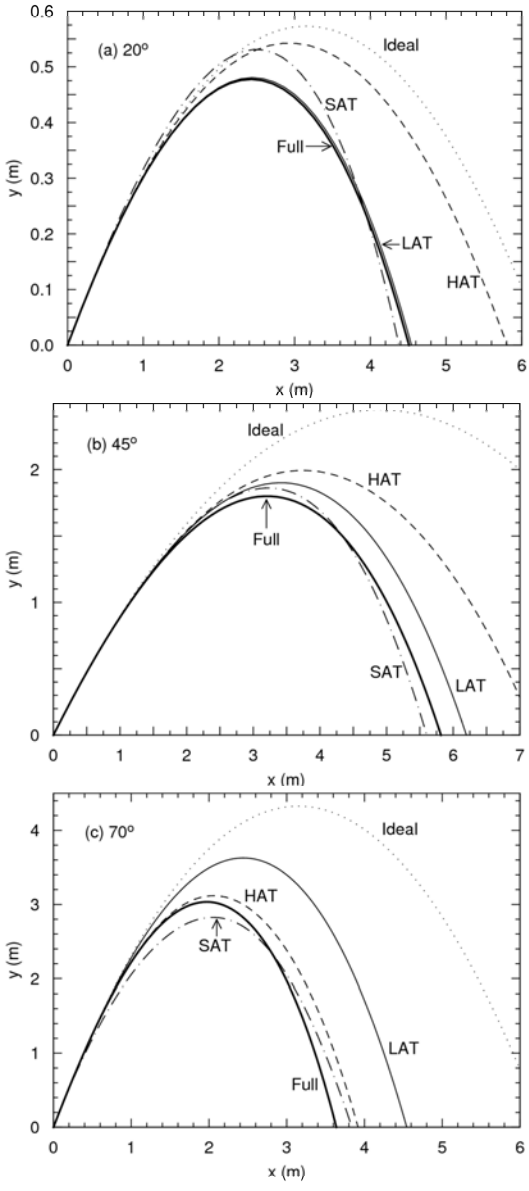

Figure 1. The trajectories for an initial speed of $v_0 = 9.8\text{m/s}$ and drag coefficient $b = 0.1\text{m}^{-1}$ at three firing angles, 20° (a), 45° (b), and 70° (c). The labeled curves are (see text): Full (thick solid line) - the numerical solution with the full v^2 resistance; LAT (solid line); HAT (dashed line, see Section 4); SAT (dash-dotted line, see Section 5), Ideal (dotted line) - motion without air resistance

3.2 The Range and Height in Comparison with Measurements

To find the range of the projectile we eliminate *t* from (6) to obtain

$$y = \left(w_0 + \frac{g}{2a} \right) \frac{x}{u_0} - \frac{g}{4a^2} \left(\exp(2bx) - 1 \right). \tag{7}$$

The range, R, is the value of x when $y = 0$. It satisfies

$$\left(w_0 + \frac{g}{2a} \right) \frac{R}{u_0} - \frac{g}{4a^2} \left(\exp(2bR) - 1 \right) = 0. \tag{8}$$

This is a transcendental equation that until now had been customarily solved numerically or graphically [9]. But as we reported [1], equations of this type can also be solved analytically, with the results written in closed form in terms of the Lambert *W* function.

The Lambert *W* function [10] is defined, for a given value *z*, as the (inverse) function satisfying

$$W(z) \exp\left(W(z) \right) = z. \tag{9}$$

Our discussions below refer closely to the properties of this relatively "new" function. For readers unfamiliar with this function, we give a brief review in the Appendix. There the reader will also find some practical formulas for evaluating *W*.

To solve for the range *R* in terms of *W*, (8) needs to be rearranged in the form of (9) such that the multiplicative prefactor to the exponential, $W(z)$, is the same as the exponent. This can be achieved by following the steps in [1]. The result is

$$\left[-2bR - \frac{1}{1+\beta} \right] \exp\left[-2bR - \frac{1}{1+\beta} \right] = -\frac{1}{1+\beta} \exp\left[-\frac{1}{1+\beta} \right], \tag{10}$$

with defined in (2). We identify from (9) and (10) that

$$W(z) \equiv \left[-2bR - \frac{1}{1+\beta} \right], \text{or} \quad R = -\frac{1}{2b} \left[W(z) + \frac{1}{1+\beta} \right], \tag{11}$$
$$z = -\frac{1}{1+\beta} \exp\left[-\frac{1}{1+\beta} \right]$$

This is the closed form expression for the range, R, in the LAT approximation.

The height, *H*, can be obtained by maximizing *y* in (7), and is given by

$$H = \frac{w_0}{2bu_0} \left[\frac{1+\beta}{\beta} \ln(1+\beta) - 1 \right]. \tag{12}$$

We compare in **Table 1** measured range and height data with calculations from Equations 11 and 12 for firing angles less than one degree. (The details of the calculations are given in the next subsection.)

Table 1 shows that the measured data [11] and the calculations agree well once the firing angle is above about 10 minutes of arc. The discrepancy between measurement and theory is due to the uncertainty in the value *b* (see **Table 1** caption), and not the approximation itself. Because the largest angle is still less than 1, higher order corrections are negligible. The relative error for the height is usually much larger than the relative error for the range. This is probably due to the difficulty in accurately measuring the relatively small height in this case.

The large errors in height and in range at the two smallest angles need not cause concern. It is due to the combination of exceeding difficulty in determining the small angle and the small height at the same time. Note that the diameter of the projectile and the height are of the same order of magnitude here. Overall, this example shows that if a reasonable *b* can be obtained, the LAT approximation should work well for low angle ballistics problems.

3.3 Analytic Properties of the Range

3.3.1 The Symmetry of Range in Firing Angle

The analytic Formula 11 enables us to immediately draw several surprising conclusions on the general properties of the range, *R*. Since *z* is a function of β which depends on the product $u_0 w_0 \propto \sin(2\theta_0)$ (θ_0 is the firing angle), we conclude that i) the range *R* is symmetric about $\pi/4$, that is to say, two firing angles θ_1 and θ_2 will lead to the same range if $\theta_1 + \theta_1 = \pi/2$ (see the LAT curves in

Table 1. Measured and calculated results for a projectile of mass *m* = 9.7 g, diameter *d* = 0.76 cm, and muzzle velocity *823* m/s. The measured data are taken from [11]. Results are calculated from the LAT approximation (11,12). The drag coefficient *b* = 1.05×10^{-3} m^{-1} is determined from $b = C\rho d^2 / m$, where *C* = 0.15 is the recommended value [11], and ρ = 1.2 kg/m^3 is the air density

Firing Angle (min)		2	5	8	12	16	20	26	33	40	49
Range, R (m)	Measured	91	183	274	366	457	549	640	732	823	914
	Calculated	76	177	265	367	456	534	636	738	826	923
	Error, %	16	3.3	3.3	0.27	0.22	2.7	0.63	0.82	0.36	0.98
Height (m)	Measured	0.02	0.1	0.19	0.33	0.61	1.0	1.5	2.3	3.2	4.5
	Calculated	0.011	0.068	0.17	0.36	0.62	0.93	1.5	2.3	3.2	4.5
	Error, %	45	32	11	9.1	1.6	7.0	0.0	9.5	3.2	2.2

Figures 1(a) and **(c)**); and ii) because *W(z)* is a monotonic function, the maximum range occurs at $\pi/4$ [12]. Remarkably, these properties i) and ii) with air resistance in the LAT approximation are exactly the same as for ideal projectile motion *without* air resistance.

Given initial conditions u_0 and w_0, the range, R, can be computed from (11). Since *z* is negative and *W(z)* is multi-valued (see Appendix) for $z<0$, we have a choice of the branches W_0 or W_{-1}. Comparing *z* in (11) and (9) we identify one trivial solution, namely the primary branch,

$$W_0(z) = -\frac{1}{1+\beta} \tag{13}$$

This solution, although mathematically correct, is unphysical because it gives a zero range when substituted into (11). The physical choice must be W_{-1}. We note that for linear resistance [1] the choices were the opposite, where W_0 was the physical solution and W_{-1} the unphysical one.

The range, *R*, calculated from (11) using W_{-1} is shown in **Figure 2** as a function of the firing angle. It clearly demonstrates that *R* is symmetric about, and maximum at, $\pi/4$. Also shown in **Figure 2**, for comparison, are the ranges for ideal projectile motion with no air resistance and with the full v^2 resistance.

Unlike the LAT case, the range of the full solution in **Figure 2** is not symmetric about $\pi/4$. It reaches maximum at an angle below $\pi/4$, a fact that is well known.

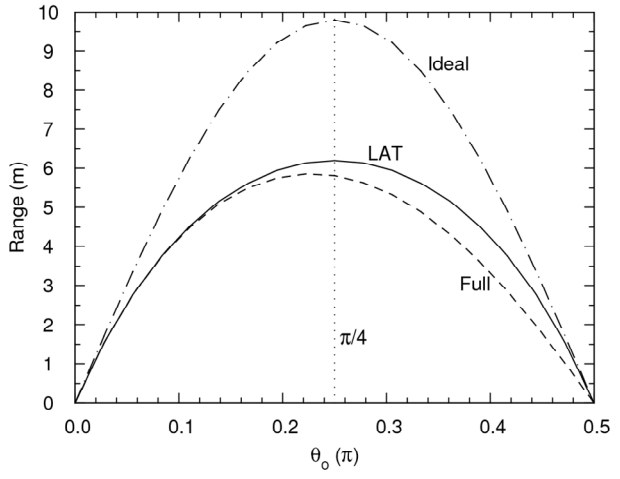

Figure 2. The range of projectile motion as a function of the firing angle θ_0. The initial speed is v^0 = 9.8 m/s and the drag coefficient is *b* = 0.1 m^{-1}. The LAT approximation (solid line) is maximum at and symmetric about $\pi/4$, as is the ideal motion (dash-dotted line). The range with the full v^2 resistance (dashed line) peaks before $\pi/4$, and is asymmetric

The LAT range is in good agreement with the full solution for low firing angles as expected, up to around $\pi/6$. Compared to the ideal case, the maximum ranges in the LAT and the full solutions are substantially reduced, by about 40% for this particular set of parameters.

We note that the LAT approximation produces asymmetric trajectories (**Figure 1**) but symmetric ranges. The physical reason can be traced to two factors influencing the range: the time of flight and the average horizontal velocity, as discussed below.

3.3.2 The Time of Flight and the Average Velocity

The average horizontal velocity, $\langle u \rangle$, and the time of flight, *T*, are related by

$$R = \langle u \rangle T. \tag{14}$$

The time of flight *T* can be obtained from (6) by setting $x = R$, at $t = T$. Together with (11), this gives

$$T = \frac{\exp(bR)-1}{bu_0} = \frac{1}{bu_0}\left\{\exp\left[-\frac{1}{2}\left(W(z)+\frac{1}{1+\beta}\right)\right]-1\right\}. \tag{15}$$

The average horizontal velocity can be expressed from (14) and (15) as

$$\langle u \rangle = \frac{u_0}{2}\left[W(z)+\frac{1}{1+\beta}\right]\left\{1-\exp\left[-\frac{1}{2}\left(W(z)+\frac{1}{1+\beta}\right)\right]-1\right\}^{-1}. \tag{16}$$

Quantitatively, as the firing angle increases, the time of flight *T* increases, but the average horizontal velocity $\langle u \rangle$ decreases. However, before $\pi/4$, the increase in *T* is more than the decrease in $\langle u \rangle$ so that the range as governed by (14), increases. After $\pi/4$, however, the opposite happens so that the range decreases. The symmetric range is a result of the balance between *T* and $\langle u \rangle$.

3.3.3 The Range for Small and Large Air Resistance

In the limit of small air resistance, $b \to 0$, the dimensionless parameter β will be small, $\beta \to 0$. The argument *z* to the *W* in (11) approaches $z \to -1/e$, and $W(z) \to -1$ (see Appendix). Using the properties of *W*(*z*) and after some algebra (we leave the details as an exercise to the interested reader), the first order correction to the range, *R*, in (9) is

$$R = R_0\left(1-\frac{2}{3}\beta\right) = R_0\left(1-\frac{2}{3}bR_0\right), \quad b \to 0, \tag{17}$$

where R_0 is the range of ideal projectile motion, $b = 0$.

For large air resistance, $b \gg 1$ and $\beta \gg 1$. As $\beta \to \infty$, $z \to -1/\beta$ according to (11). Using the as-

ymptotic expressions for $W_{-1}(z \to 0) \approx \ln\left(z / ln|z|\right)$ (see Appendix), we have from (11)

$$R = \frac{R_0}{2}\frac{\ln(\beta ln\beta)}{\beta} \quad \text{with} \quad b, \beta \gg 1. \tag{18}$$

It is interesting to compare the scaling behavior with our earlier study [1] for linear resistance, where we found the range to scale as $1/b$, the inverse of the resistance. For quadratic resistance, (18) indicates $R \propto \ln(b\ln b)/b$. This shows that the logarithm term is characteristic of the quadratic resistance.

4. High-Angle Trajectory (HAT) Approximation

When the firing angle is large (close to $\pi/2$), we expect that, on average, the vertical velocity *w* will be much larger than the horizontal velocity, $|w| \gg u$. The speed is approximated as $v \approx |w|$ from (3). The equations of motion in the HAT approximation are

$$\frac{du}{dt} = -b|w|u, \quad \frac{dw}{dt} = -b|w|w - g. \tag{19}$$

The solutions are broken into two parts because of |w|. In the ascending part of the trajectory, the solutions are

$$u = \frac{u_0 \cos\varphi}{\cos(\varphi - \omega t)}, \quad w = \sqrt{\frac{g}{b}}\tan(\varphi - \omega\tau), \omega = \sqrt{bg}. \tag{20}$$

$$x = \frac{u_0 \cos\varphi}{\omega}\ln\left[\tan\left(\frac{\pi}{4} - \frac{\varphi}{2} + \frac{\omega\tau}{2}\right)\Big/\tan\left(\frac{\pi}{4} - \frac{\varphi}{2}\right)\right]$$

$$y = \frac{1}{b}\ln\left[\cos(\varphi - \omega t)/\cos\varphi\right], \quad \text{with} \quad \varphi = \tan^{-1}\left(\sqrt{b/g}\,w_0\right). \tag{21}$$

The time it takes to reach the top is $t = \varphi/\omega$. With the values of *u, w, x, y* in (20,21) at the top as the initial condition for the descending trajectory, the solutions for descent are

$$u = u0\cos\varphi\frac{\exp(-\omega\tau)}{1+\exp(-2\omega\tau)}, \quad w = -\sqrt{\frac{g}{b}}\frac{1-\exp(-2\omega\tau)}{1+\exp(-2\omega\tau)}, \tag{22}$$

$$x = \frac{u_0\cos\varphi}{\omega}\left[\frac{\pi}{4} - \tan^{-1}\left(e^{-\omega t}\right) - \ln\tan\left(\frac{\pi}{4} - \varphi/2\right)\right]$$

$$y = \frac{1}{b}\left\{\ln\left[\frac{2\sqrt{1+b\omega_0^2/g}}{1+\exp(-2\omega\tau)}\right] - \omega\tau\right\}. \tag{23}$$

The time τ starts from zero (at the top) in (22, 23).

The trajectories in the HAT approximation are shown in **Figure 1** The best agreement with the full solution is seen at the highest angle 70°, consistent with the underlying assumptions. We note that the agreement is considerably worse descending than ascending (**Figure 1**, 70° (c)). The reason is that near the top, $w \sim 0$, and the validity of the HAT approximation breaks down, causing the large discrepancy while falling back down. By contrast, the LAT approximation (**Figure 1**, 20° (a)) is valid globally as long as the firing angle is small, giving a much better agreement on both parts of the trajectory.

5. Split-Angle Trajectory (SAT) Approximation

In Sections 3 and 4 we discussed low and high angle trajectories. To be complete, we consider in this section the split angle $\sim \pi/4$, between the LAT and HAT approximations. We assume $u \sim w$. and take the symmetric approach: setting $v \approx \sqrt{2}u$ in the horizontal direction and $v \approx \sqrt{2}|w|$ in the vertical direction. The equations of motion read

$$\frac{du}{dt} = -\sqrt{2}bu^2, \quad \frac{dw}{dt} = -\sqrt{2}b|w|w - g. \tag{24}$$

Note that upon replacing $\sqrt{2}b \to b$ in (24), du/dt is the same as that in (4) of LAT, and dw/dt is the same as that in (19) of HAT. The solutions for *u, x* will be the same as those in (5,6), and the solutions for *w, y* will be the same as for *w, y* in (21,23), so they will not be repeated here.

Similarly, the trajectories can be computed as before (with *b* replaced by $\sqrt{2}b$, of course). They are also shown in **Figure 1** at the same angles with the same parameters. Here, we see the best agreement with the full solution at 45° as it should. But, unlike the LAT or HAT curves, where after certain point in time (just before reaching the top) the differences keep increasing, the SAT curve crosses the full solution during the course of motion. This is due to the balance of the horizontal and vertical resistance forces.

Because of this balance, the SAT behaviors are interestingly different at low versus high angles. At 20° (**Figure 1(a)**), the SAT curve is "squeezed" horizontally in comparison with the full solution, resulting in a shorter range and a higher height. This is because for low angles where $u > w$, the horizontal resistance is over-estimated in the equations of motion (24). Conversely, at 70° (**Figure 1(c)**), the SAT curve is compressed vertically, causing a lower height but a longer range. The reason is similarly due to the over-estimation of the resistance in the vertical direction. As a result, curve crossing occurs.

6. Conclusions

In summary, we have presented a detailed discussion of projectile motion with quadratic air resistance in three approximations. Our focus was on the low-angle trajec-

tory approximation where we found closed form solutions for the range and the time of flight in terms of the secondary branch of the Lambert *W* function, W_{-1}. The approximation is simple and accurate for low angle ballistics problems.

Various analytic properties were readily analyzed with these solutions. Together with projectile motion with linear air resistance [1,5], the example studied here serves two educational purposes: i) It is possible to introduce the use of special functions in physics at early undergraduate levels in a familiar, more realistic problem; and ii) It represents a good complimentary case where the physical solution required the secondary branch, W_{-1}, rather than the principal branch W_0 as in the case of linear air resistance.

One interesting and closely-related question remains open, i.e., whether both branches of *W* might be required simultaneously in a physical solution, say when the air resistance contains both linear and quadratic terms, $-av-bv^2$, under some forms of approximation, presumably.

REFERENCES

[1] R. D. H. Warburton and J. Wang, "Analysis of asymptotic projectile motion with air resistance using the Lambert *W* function," American Journal of Physics, Vol. 72, pp. 1404–1407, 2004.

[2] S. R. Valluri, D. J. Jeffrey, and R. M. Corless, "Some applications of the Lambert *W* functions to physics," Canadian Journal of Physics, Vol. 78, pp. 823–830, 2000.

[3] S. R. Cranmer, "New views of the solar wind with the Lambert *W* function," American Journal of Physics, Vol. 72, pp. 1397–1403, 2004.

[4] D. Razansky, P. D. Einziger, and D. R. Adam, "Optimal dispersion relations for enhanced electromagnetic power deposition in dissipative slabs," Physical Review Letters, Vol. 93, 083902, 2004.

[5] S. M. Stewart, Letters to the Editor, American Journal of Physics, Vol. 73, pp. 199, 2005. "A little introductory and intermediate physics with the Lambert *W* function," Proceedings of 16th Australian Institute of Physics, Vol. 005, pp. 194–197, 2005.

[6] E. Lutz, "Analytical results for a Fokker-Planck equation in the small noise limit," American Journal of Physics, Vol. 73, pp. 968–972, 2005.

[7] P. Hövel and E. Schöll, "Control of unstable steady states by time-delayed feedback methods," Physical Review E, Vol. 72, 046203, 2005.

[8] There is a solution presented in terms of the slope angle by A. Tan, C. H. Frick, and O. Castillo, "The fly ball trajectory: An older approach revisited," American Journal of Physics, Vol. 55, pp. 37–40, 1987. But the slope angle is unknown except at the top (zero) of the trajectory, and can be found only numerically or graphically. Therefore, the solution is not in closed form.

[9] G. W. Parker, "Projectile motion with air resistance quadratic in the speed," American Journal of Physics, Vol. 45, pp. 606–610, 1977.

[10] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert *W* function," Advances in Computational Mathematics, Vol. 5, pp. 329–359, 1996.

[11] H. R. Kemp, "Trajectories of projectile motion in air for small times of flight," American Journal of Physics, Vol. 55, pp. 1099–1102, 1987.

[12] The angle maximizing the range can be derived directly by $\frac{dR}{d\vartheta_0}=0$, from [1], yielding $\left[W(z)+\frac{1}{1+\beta}\right]\cos(2\theta_0)=0$. This implies either $\left[W(z)+\frac{1}{1+\beta}\right]=0$, the trivial unphysical solution, (13); or $\cos(2\theta_0)=0$, i.e., $\theta_0=\pi/4$.

[13] B. Hayes, "Why W?" American Science, Vol. 93, pp. 104–108, 2005.

[14] J. Wang, To be published.

[15] M. Abramowitz and I. A. Stegun, "Handbook of mathematical functions," Dover, New York, pp. 18, 1970.

Appendix

1. The Lambert *W* function

In this appendix, we briefly summarize some relevant elements of the Lambert *W* function. The reader can find an extensive review in [10]. In lieu of the growing interest in this function in the physics community [2–7], we also present several practical formulas for the evaluation of the *W* function.

The Lambert *W* function as defined by (9) is generally complex-valued. For our purpose we are interested in the real-valued *W(x)*, namely the principal branch W_0, and the secondary branch W_{-1}. The two branches are shown in **Figure 3**.

The behaviors of W_0 and W_{-1} for small and large *x* are

$$W_0(x\to 0)=x-x^2 \quad W_{-1}(x\to 0^-)=\ln\left(\frac{x}{\ln x}\right),\ W_0(x\to\infty)=\ln\left(\frac{x}{\ln x}\right). \tag{A.1}$$

2. Evaluation of *W*

During the course of our study, we were unable to find in the literature simple and practical algebraic expressions for evaluating *W* over the whole range, and for embedding it in our code. We therefore devised several readily usable approximate formulas given here for this purpose. It should give the interested reader a good starting point in using this function.

W is not (yet) available on a calculator like some elementary functions. It is debatable whether it should be elevated to the status of an elementary function. (See [13] for an enchanting account.) However, one can easily implement it on a programmable calculator with the formulas given here. The accuracy is at least 8 digits, in fact it is usually much better.

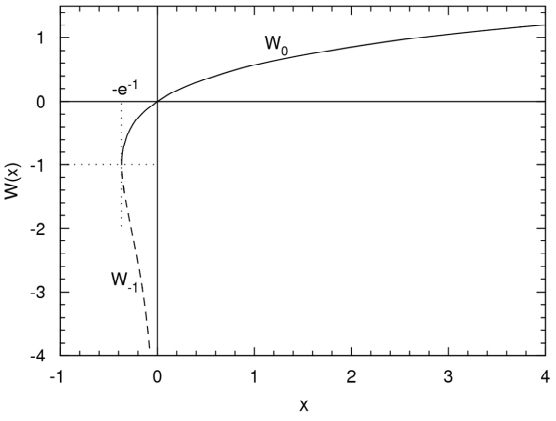

Figure 3. The Lambert *W* function for real argument *x*. The two real branches are the principal branch W_0 (solid line), and the secondary branch W_{-1} (dashed line). The branch point is at $x=-e^{-1}$ where $W=-1$

1) *W* in the Regular Regions

We first give the formulas optimized in the regular regions. Owing to space limitations, we will explain the methods used elsewhere [14]. The general form of the expressions is

$$W_{0,-1}(x)=C+rP(\mathrm{r}) \tag{A.2}$$

where *C* is a constant and $P(\mathrm{r})$ is the Padé approximant defined as

$$P(\mathrm{r})=\frac{a_0+a_1r+a_2r^2+a_3r^3+a_4r^4}{1+b_1r+b_2r^2+b_3r^3+b_4r^4}. \tag{A.3}$$

The expansion variable, *r*, is related to the independent variable, *x*. We give in **Table 2** the constant *C*, the variable *r*, and the coefficients a_i and b_i. To utilize the table, locate which function and region to use, then evaluate (A.2) with the corresponding *C*, *r*, a_i, and b_i.

2) *W* in the Asymptotic Regions

Both W_0 and W_{-1} have the same form in the asymptotic regions

$$\begin{aligned}&W_{0,-1}(x)=q+(q+q^2)r\left(1+a_1r+a_2r^2+a_3r^3+a_4r^4\right),\\&p=\ln\left(\frac{x}{\ln|x|}\right),\quad q=\ln\left(\frac{x}{p}\right),\quad r=\frac{\ln(p/q)}{(1+q)^2}\\&a_1=1/2,\quad a_2=(1-2q)/6,\quad a_3=\left(1-8q+6q^2\right)/24,\\&a_4=\left(1-22q+58q^2-24q^3\right)/120.\end{aligned} \tag{A.4}$$

This expression (A.4) should be used for W_0 in the region $x\in[2.2,\infty)$ and for W_{-1} in $x\in[-0.12,\infty)$. Note that it is the subtle difference in *p*, namely, $\ln(x/\ln x)$ for W_0 and $\ln[x/\ln(-x)]$ for W_{-1} that automatically selects the correct branch.

3) *W* with a Little Programming

If the reader wishes to calculate *W* with arbitrary precision, one can use Newton's rule [15] which, for a given *x*, is the root, *w*, to $f(w)=w\exp(w)-x=0$. The root-finding process is as follows: Starting with an initial guess, w_1, the successive iterations, w_n, approach rapidly to the true value as

$$w_{n+1}=\frac{w_n^2+x\exp(-w_n)}{1+w_n},\quad n=1,2,3... \tag{A.5}$$

It only remains to determine the initial guess, w_1. One way to choose w_1 is

Table 2. The approximate formulas for the evaluation of *W* with (A.2) and (A.3). Spaces are inserted in the coefficients for readability. For optimal precision, all the digits should be used

Function		W_0		W_{-1}
Region	$x \in \left[-e^{-1}, -0.16\right)$	$x \in \left[-0.16, 0.32\right)$	$x \in \left[0.32, 2.2\right)$	$x \in \left[-e^{-1}, -0.12\right)$
Const. *C*	-1	0	0.3906 4638	-1
Variable *r*	$\sqrt{-2\ln(-ex)}$	x	$\ln\left(\sqrt{3}x\right)$	$\sqrt{-2\ln(-ex)}$
a_0	1	1	0.2809 0993	-1
a_1	-0.8040 7820	4.674 4173	0.1116 7016	-0.8178 4020
a_2	0.2802 9706	6.577 4227	0.0 3529 1013	-0.2889 3422
a_3	-0.0 4785 3103	2.730 6731	0.00 5498 1613	-0.0 5003 8980
a_4	0.00 3355 7735	0.1057 7423	0.000 4245 7974	-0.00 3566 1458
b_1	-0.4707 4486	5.674 4173	0.1389 8485	0.4845 0686
b_2	0.0 9560 4321	10.75 1840	0.0 7995 0768	0.0 9965 4140
b_3	-0.00 6612 4586	7.637 5538	0.00 4515 2166	0.00 7066 1014
b_4	0.7961 1402x10^{-5}	1.539 0142	0.000 6368 7954	0.5596 5023x 10^{-5}

$$w_1 \begin{cases} x & if\ -e^{-1} < x \le 1.5 \quad W_0 \\ \ln(x) & if\ 1.5 < x < \infty \quad W_0 \\ \ln(-x) & if\ -e^{-1} < x \le 0 \quad W_{-1} \end{cases} \qquad \text{(A.6)}$$

It can be shown [14] that (A.5) with the seed from (A.6) always converges to the correct branch W_0 or W_{-1}. The convergence is fast, usually to machine accuracy in a few iterations.

Summarizing, accurate values for W_0 and W_{-1} over all regions of x can be found by using **Table 2** plus (A.4) for fixed precision (8 digits or better), or (A.5) and (A.6) for arbitrary precision.

J. Service Science & Management, **2010, 3: 106-109**
doi:10.4236/jssm.2010.31013 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Multivariate Quality Loss Model and its Coefficient Determination

Shuhai Fan, Xia Cao

Department of Industrial Engineering, Nanjing University of Technology, Nanjing, China.
Email: fanshuhai@tsinghua.org.cn

Received August 28th, 2009; revised October 10th, 2009; accepted November 27th, 2009.

ABSTRACT

In the early 1970s*, based on single index deducted from absolute quality deviants, Genichi Taguchi proposed the quality loss module. This module builds the foundation of his three-stage design theory, e.g. system design, parameter design and tolerance design. In actual production process, nevertheless, it is multiple quality indices that influence the total quality. Consequently, the interaction of the quality indices should be imported into the module as a key factor. Accordingly, based on several indices of relative quality deviation, introduce a* 2*-order multivariate quality module at first. Next, extend the module to* 3 *or even higher orders. Then, improve the previous quality module by simplifying the* 2*-order module as a multivariate quality loss module. Finally, bring forward a significant solution to determinate all the coefficients in the multivariate quality loss module and describe its work flow as well.*

Keywords**:** *Quality Loss, Multiple Variables, Model Building*

1. Introduction

In the early 1970s, Dr. Genichi Taguchi, the famous quality management expert of Japan, carried on an innovation research into the theories and methods of quality management [1]. He established the famous "Taguchi Three-stage Design Methods", the system design, the parameter design and the tolerance design. The core of his theory was his quality loss model [2,3]. And the quality loss model was based on the absolute quality deviant, which suited for the single quality index. However, in the actual production process, many partial quality indices often affected the total quality in a coactions mode. After these interactions being considered, a multivariate quality loss mode can be brought forward [4].

To this many partial quality indices, Chan and Ibrahim studied a quality evaluation model using loss function for multiple S-type quality characteristics [5]. To determinate the coefficients of the multivariate quality loss mode, AHP method or tolerance method can be used [6,7].

2. Model Building

The total quality is a synthetic quality, which is often a coaction of some machining qualities and some assembling qualities of every part [1]. And it can be regarded as a function of every partial quality (including the machining qualities and the assembling qualities):

$$Q_{\text{product}} = f(Q_1, Q_2, \ldots Q_h \ldots, Q_H)$$

Assuming that f has continuous partial derivatives till 3-order in a neighborhood D of the origin, $P_0(0,0\ldots0)$.

There are H partial quality indices altogether, q_1, q_2, … q_H. These partial quality indices are all continuous types; their values are very small and stand for relative quality deviants. When the values of q_1, q_2,…q_H equal 0, it stands for no quality deviant (Here we use the relative quality deviants to get dimensionless. Thus a minor error can be avoided in the derivation process of Taguchi's model.) [8]. Then the product quality can be expressed as follows:

$$f(q_1, q_2 \ldots, q_h \ldots, q_H) = f(0,0\ldots,0) + \sum_{h=1}^{H} f'_{Q_h}(0,0\ldots,0) q_h + \frac{1}{2!} \sum_{h_1,h_2=1}^{H} f''_{Q_{h_1}Q_{h_2}}(0,0\ldots,0) q_{h_1} q_{h_2} + R_2$$

while, R_2 is the 2-order remainder term.

$$R_2 = \frac{1}{3!} \sum_{h_1,h_2,h_3=1}^{H} f'''_{Q_{h_1}Q_{h_2}Q_{h_3}}(\theta q_1 \ldots \theta q_h \ldots \theta q_H) q_{h_1} q_{h_2} q_{h_3} \quad (0 < \theta < 1)$$

A brief proof is as follows:
considering a single variant function $Q(t)$

$$Q(t) = f(q_1 t, q_2 t \ldots, q_h t \ldots, q_H t), \qquad (-1 \le t \le 1)$$

$Q(t)$ has continuous partial derivatives till 3-order for $(-1 \le t \le 1)$.

Obeying the 2-order single variant Talylor's formula with Lagrange remainder term, we get:

$$Q(t) = Q(0) + Q'(0)t + \frac{1}{2!}Q''(0)t^2 + \frac{1}{3!}Q'''(\theta t)t^3, (0 < \theta < 1)$$

Let t=1, we get

$$Q(1) = Q(0) + Q'(0) + \frac{1}{2!}Q''(0) + \frac{1}{3!}Q'''(\theta) \quad (1)$$

Follow the definition of $Q(t)$ and the differential method of composite function. We get

$$Q'(t) = \sum_{h=1}^{H} f'_{Q_h}(q_1 t, ...q_h t, ..., q_H t)q_h$$

$$Q''(t) = \sum_{h_1,h_2=1}^{H} f''_{Q_{h_1}Q_{h_2}}(q_1 t, ...q_h t, ..., q_H t)q_{h_1}q_{h_2}$$

$$Q'''(t) = \sum_{h_1,h_2,h_3=1}^{H} f''_{Q_{h_1}Q_{h_2}Q_{h_3}}(q_1 t, ...q_h t, ..., q_H t)q_{h_1}q_{h_2}q_{h_3}$$

Then, when t =0,

$$Q'(0) = \sum_{h=1}^{H} f'_{Q_h}(0, ...0, ..., 0)q_h$$

$$Q''(0) = \sum_{h_1,h_2=1}^{H} f''_{Q_{h_1}Q_{h_2}}(0, ...0, ..., 0)q_{h_1}q_{h2}$$

$$Q'''(0) = \sum_{h_1,h_2,h_3=1}^{H} f''_{Q_{h_1}Q_{h_2}Q_{h_3}}(0, ...0, ..., 0)q_{h_1}q_{h_2}q_{h_3}$$

while

$$Q(1) = f(q_1, q_2 ..., q_h ..., q_H),$$

$$Q(0) = f(0, ..., 0..., 0)$$

Substitute the results into (1)

The expansion equation can get proved.

To estimate the error term, we have:

f has continuous partial derivatives till 3-order; in the neighborhood of (0,…,0,…0); the 3-order partial derivative has a bound M.

denote

$$\rho = \sqrt{q_1^2 + q_2^2 + ...q_H^2} = \sqrt{\sum_{h=1}^{H} q_h^2}, \quad (\rho \to 0)$$

then

$$|R_2| \le \left| \frac{M}{3!} \sum_{h_1,h_2,h_3=1}^{H} q_{h_1}q_{h_2}q_{h_3} \right| = \left| \frac{M}{3!}(q_1 + q_2 + ... + q_H)^3 \right|$$

$$= \frac{M}{3!}\left[\sqrt{(q_1 + q_2 + ... + q_H)^2} \right]^3 = \frac{M}{3!}\left[\sqrt{\sum_{h_1,h_2=1}^{H} q_{h_1}q_{h_2}} \right]^3$$

$$\le \frac{M}{3!}\left[\sqrt{\sum_{h_1,h_2=1}^{H} \frac{1}{2}({q_{h_1}}^2 + {q_{h_2}}^2)} \right]^3$$

$$= \frac{M}{3!}\left[\sqrt{H({q_1}^2 + {q_2}^2 + ... + {q_H}^2)} \right]^3$$

$$= \frac{M}{3!}\left(\sqrt{H}\right)^3 \left[\sqrt{\sum_{h=1}^{H} {q_h}^2} \right]^3 = \frac{M}{3!}\left(\sqrt{H}\right)^3 \rho^3$$

$\therefore R_2 = o(\rho^2), \quad (\rho \to 0)$

It is the general form that is used to describe the total quality index in the above formula. If the quality loss form index is adopted, the model can be much more simplified. Imply that when all the partial quality indices equal 0, the total quality index takes the minimum, 0. It is easy to see that the 0-order and 1-order derivative can be gotten rid of. Then a new product quality loss model can be built as follows:

$$L = Q_{\text{product}} = \sum_{h=1}^{H} w_{h^2} {q_h}^2 + \sum_{h_1=1,h_1<h_2}^{n} w_{h_1h_2} q_{h_1} q_{h_2} \quad (2)$$

Compared it with Taguchi's model, we can see that (2) is actually the extended multivariate form of Taguchi's quality loss model.

Here, Qproduct is a target total quality index and described as the quality loss. All the n partial quality indices influence Qproduct in a coaction mode, and qh is the relative quality deviant. Every partial quality deviant can be gotten by statistic. Here we use the relative partial quality deviant, and it can be expressed by the relative quality value of the deviant from the target value and severalfold tolerances. If the value of quality characteristic lies beyond the tolerance, it should be multiplied by a punishment factor.

where, $\sum_{h=1}^{H} w_{h^2} {q^2}_h$, $\sum_{h_1=1,h_1<h_2}^{n} w_{h_1h_2} q_{h_1} q_{h_2}$ are separately the self-action item and the inter-action item of the partial quality deviants. w_{h^2} is the self-action influence weight of the quality deviant ${q^2}_h$; $w_{h_1h_2}$ is the inter-action weight of deviant $q_{h_1}q_{h_2}$.

Compared with our expansion equation, it can be gotten

$$w_{h^2} = \frac{1}{2!} f''_{Q_h^2}(0, ..., 0, ...0) = \frac{1}{2} f''_{Q_h^2}(0, ..., 0, ...0)$$

$$w_{h_1h_2} = \frac{1}{2!} f''_{Q_{h_1}Q_{h_2}}(0, ..., 0, ...0)$$

$$= \frac{1}{2}\left[f''_{Q_{h_1}Q_{h_2}}(0, ..., 0, ...0) + f''_{Q_{h_2}Q_{h_1}}(0, ..., 0, ...0) \right]$$

$$= f''_{Q_{h_1}Q_{h_2}}(0, ..., 0, ...0) \quad (h_1 < h_2)$$

If necessary, we can also expand the product quality model further more to U-order.

$$Q_{\text{product}} = \sum_{u=2}^{U}\left\{\frac{1}{u!}\sum_{h_1,h_2,\ldots h_u=1}^{H}\left[f^{(u)}{}_{Q_{h_1}Q_{h_2}\ldots Q_{h_u}}(0,0\ldots,0)q_{h_1}q_{h_2}\ldots q_{h_u}\right]\right\}+R_U$$

$$R_U = o(\rho^U)$$

When the quality model is expanded to 2-order, the inter-actions of partial quality indices have been fully considered. In fact, the impactions of items higher than 2-order are very little [10,11]. Thus for the simplification of computation and the actual requirement, further higher orders are never needed.

The above work suits the small-is-better and the nominal-is-best problems. The large-is-better problem can also be reformed into a small-is-better problem.

3. Coefficient Determination

3.1 The Tolerance Limits Method

The quality loss model (2) can also be written as

$$Q_{\text{product}} = \sum_{i=1}^{H} w_{ii}q^2{}_i + \sum_{i=1,i<j}^{n} w_{ij}q_iq_j \qquad (3)$$

the coefficients w_{ij} can be determined using the following method all these w_{ij} ($i\leq j$) can constitute a upper triangular matrix like Figure 1.

1) Determine the elements in Diagonal for an arbitrary i (i=1,2,...,H), when all the q_k=0, ($k\neq i$) Δ_i is the tolerance limit of defective for self quality and A_i is the self quality loss when defective happens.

Then we have

$$Q_{\text{product}} = A_i = \sum_{i=1}^{H} w_{ii}q^2{}_i + \sum_{i=1,i<j}^{n} w_{ij}q_iq_j = w_{ii}(\Delta_i)^2$$

So

$$w_{ii} = A_i/(\Delta_i)^2$$

2) Determine other elements for an arbitrary i,j (i,j=1,2,...,H, $i\leq j$), when all the q_k=0, ($k\neq i$, $k\neq j$), Δ_{ij}^i is the tolerance limit of quality i for coaction quality i and j, Δ_{ij}^j is the tolerance limit of quality j for coaction quality i and j and A_{ij} is the coaction quality loss when coaction defective happens.

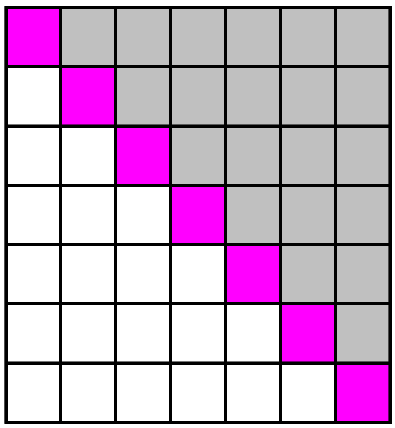

Figure 1. Upper triangular matrix of coefficients then, we can determine these coefficients

Then we have

$$Q_{product} = A_{ij}^0 = \sum_{i=1}^{H} w_{ii}q^2{}_i + \sum_{i=1,i<j}^{n} w_{ij}q_iq_j$$
$$= w_{ii}(\Delta_{ij}^{0,i})^2 + w_{jj}(\Delta_{ij}^{0,j})^2 + w_{ij}\Delta_{ij}^{0,i}\Delta_{ij}^{0,j}$$

Thus we have

$$w_{ij} = \left[A_{ij}^0 - w_{ii}(\Delta_{ij}^{0,i})^2 - w_{jj}(\Delta_{ij}^{0,j})^2\right]/\Delta_{ij}^{0,i}\Delta_{ij}^{0,j}$$

where

$$w_{ii} = A_i^0/(\Delta_i^0)^2,\quad w_j = A_j^0/(\Delta_j^0)^2$$

This coefficient determination process can be described in Figure 2.

Remark:

In some particular circumstances,

$$\Delta_{ij}^{0,i} = \Delta_i^0 \quad , \quad \Delta_{ij}^{0,j} = \Delta_j^0$$

we have a simplified form

$$w_{ij} = \left[A_{ij}^0 - A_i^0 - A_j^0\right]/\Delta_i^0\Delta_j^0$$

3.2 Other Methods

To determinate the weight coefficients, the least square method can also be adopt. However linear neural networks are more suitable [11–14]. The input weights of the trained neural network are just the weight coefficients of the model.

4. Conclusions

Based on the quality loss model of Taguchi, the total quality model which is based on multivariate relative quality deviants was studied. A 2-order product quality model, which was based on several indices of relatively quality deviation, was built. Then, the model was successfully extended to 3 or even higher orders. To deduce

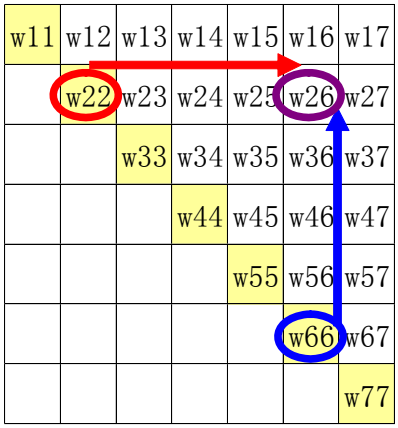

Figure 2. From the diagonal elements to other elements

a simplified formula, a new simplified product quality model (a multivariate quality loss model) was built and the error of it was also analyzed. Finally, a useful method of the coefficient determination was brought forward to determinate all the coefficients in the multivariate quality loss model and the work flow of it was also described.

However, in the complex product, the coupling factor, structure layer and the number of parts may have a further increase, and the quality loss model become more and more complex. Then there are no enough data for the determination of the coefficients and the calculation also become more and more complex. To make full use of historical data and similar data, a quick-response incremental model of multivariate quality loss is needed. So it should be in our further research.

5. Acknowledgment

This project was supported by National Natural Science Foundation of China (70801036), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (07KJB460045).

REFERENCES

[1] G. Taguchi, "Quality engineering [M]," China Translation and Publishing Corporation, December 1985.

[2] H. J. Powell, "Quality tools for high value fabrications [J]," Welding in the World, No. 46 (SPEC), pp. 275–280, July 2002.

[3] R. Rai and V. Allada, "Modular product family design: Agent-based pareto-optimization and quality loss function-based post-optimal analysis [A]," International Journal of Production Research, Vol. 41, No. 17, pp. 4075–4098, November 20, 2003.

[4] S. H. Fan, "Studies on Grey theory based intelligent methods and their applications to MCMQC [D]," Tsinghua University, Beijing, June 2004.

[5] W. M. Chan, R. N. Ibrahim, and P. B. Lochert, "Quality evaluation model using loss function for multiple S-type quality characteristics [J]," International Journal of Advanced Manufacturing Technology, Vol. 26, No. 1–2, pp. 98–101, 2005.

[6] R. Khorramshahgol and G. R. Djavanshir, "The application of analytic hierarchy process to determine proportionality constant of the Taguchi quality loss function [J]," IEEE Transactions on Engineering Management, , Vol. 55, No. 2, pp. 340–348, May 2008.

[7] S. H. Fan, "The coefficient determination of multivariate quality loss model," IEEE EMS'07, 2007.

[8] G. X. Zhang, "Quality management [M]," Higher Education Press, Beijing, 1998.

[9] S. H. Fan, T. Y. Xiao, *et al*., "Building the dynamic network of mass customization [J]," Journal of System Simulation, Vol. 16, No. 7, pp. 1597–1599, 2004.

[10] S. H. Fan, T. Y. Xiao, *et al*., "Model building of the dynamic network in MCMQA [J]," Journal of System Simulation, Vol. 16, No. 12, pp. 2767–2769, 2004.

[11] P. Melin and O. Castillo, "A general method for surface quality control in intelligent manufacturing of materials using a new fuzzy-fractal approach fuzzy [A]," Information Processing Society, NAFIPS, 18th International Conference of the North American, pp. 10–12, June 1999.

[12] M. Q. Zhu, "Artificial neural network approach of quality control for intelligent manufacturing system [J]," Acta Aeronautica et Astronautica Sinica, Vol. 18, No. 5, pp. 603–606, 1997.

[13] S. Fan, "Studies on the application of grey genetic algorithm in dynamic network of MCMQA," Mechanical Science and Technology, Vol. 25, No. 8, pp. 984–988, August 2006.

[14] S. Fan, "E-factory based agile-maintenance simulation-platform," Journal of System Simulation, Vol. 19, No. 5, pp. 1151–1153, May 2007.

J. Service Science & Management, **2010, 3: 110-116**
doi:10.4236/jssm.2010.31014 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Quantitative Assessment of Football Web Sites: An Empirical Study of the Best European Football Club

Francisco Javier Miranda[1], Antonio Chamorro[2], Víctor Valero[2], Jesús Maestre[3]

[1]Professor of Operation Management, University of Extremadura, Badajoz, Spain; [2]Professor of Marketing, University of Extremadura, Badajoz, Spain; [3]Researcher, University of Extremadura, Badajoz, Spain.
Email: fmiranda@unex.es

Received October 1st, 2009; revised November 12th, 2009; accepted December 20th, 2009.

ABSTRACT

Football and the Internet have shown early signs of a prosperous marriage. Web sites with football content are among the most popular, with football followers and web users sharing remarkably similar demographics. Despite these compelling observations, limited empirical research has been undertaken exploring how to maximise the opportunities for competitive advantage that the Internet can provide to football clubs. This research was a response to this void, and was undertaken by manually accessing and evaluating the web sites of European Football Clubs. Quality of web sites was determined using an original Web Assessment Index, which focuses on four categories: *accessibility, speed, navigability and content. A detailed report of the results arising from this investigation is presented and systematically analyzed. These findings will be useful for both researchers and practitioners who seek to understand the issues relevant to football club management.*

Keywords*: *World Wide Web, Football, Content Analysis, Web Design, Internet, Research Paper

1. Introduction

The number of Internet users has grown significantly over the last few years, from virtually nothing to an estimated 1.04 billion users that are 16% of the world's population. This rapid growth in the number of Internet users has promoted a belief in many business circles that the Web represents a huge marketing opportunity.

1) With increasing interest in the Internet's role for business marketing, sport marketers throughout the world and at all levels of the sport industry are rapidly working to incorporate emerging technologies into their marketing strategies. Not surprisingly, the Internet has become a significant marketing tool for many sport organisations, including professional sports teams.

2) The same market forces do not drive European Football Clubs websites as most other sectors. They have a loyal fan base that they can rely upon. They are able to provide a lot of exclusive content that can't be found elsewhere. And they have brand names easily recognisable, in many cases, worldwide. Despite these apparent advantages, each club website should not be seen as operating in a vacuum. Whilst the temptation must be to consider that official club websites operate in an uncompetitive marketplace, there are authoritative and comprehensive websites from newspapers and TV channel and others, including "unofficial" fan websites, that provide alternatives for users to turn to. So, loyalty to one team should not be confused with loyalty to the team's official website.

The results of our survey should be considered within this broader context. Whilst the clubs doubtless enjoy large volumes of traffic, and many successful transactions, there is clear scope for improvement. European Clubs benefit from their high profile brands with regards to the large numbers of incoming links that boost their website's visibility. Clubs will doubtless be able to boast of large numbers of visits originating from search engines. This appears to owe more to their brand name or unique offerings than to any clear evidence of optimisation on their sites. Yet, clubs should not feel that they should rely only on their successful brands. As good as this sector's visibility is, there are still steps that could be taken by the majority of the clubs to improve that. And any improvement in visibility should ultimately result in increased revenue streams.

This research evaluated the current practice of European Football Clubs, assessing the effectiveness of their web sites. For this, we develop a web site assessment

index that can be employed to compare the current usage of the Internet by football clubs. We begin by identifying the main factors considered as determinants of web site quality, as mentioned in previous studies. Secondly we discuss the design of the web assessment index. Then, web sites of most popular European Football Leagues are subsequently analyzed using this index. Finally, the main results of this analysis are discussed and future research directions are outlined.

2. Web Sites Assessment Tool

The quality of a user's visit to a website is not simply determined by the strength of content on that website. The way a website is constructed and maintained can have a positive or negative impact on that experience. Most website visitors have a specific task in mind when they arrive at a website. Placing obstacles in their way from completing those tasks can lead to frustration, and consequently to website abandonment and undermine brand reputation.

Evaluating the performance of web sites has been a constant concern of researchers in different fields. A review of the recent literature on web site assessment reveals some attempts to measure web site quality [1–9].

Some researchers have tried to provide ways of evaluating web sites specifically [10,11]. Most of the previous approaches have focused either on basic management content or a specific set of web site outcomes, using subjective factors, such as easy-access, text clearness, presentation quality, attractiveness of colours, sounds, etc. To minimize this subjectivity site evaluators should be given precise guidelines to rate each factor and a large group of evaluators is needed [4].

Trying to avoid the main weaknesses of previous models [12] developed a new web site assessment index that can be employed to compare the current use of the Internet by different organizations. This model has been previously employed to compare the Internet usage from the 200 largest Spanish companies [13] and more recently applied to Spanish banks [14].

According to Evans and King [4], a web assessment tool must have five main components: categories, factors, weights, ratings and total score. The first step is to choose the categories and factors that are critical to web site effectiveness. Our instrument for evaluating Football Club web sites (Web Assessment Index or WAI) selects four broad categories as the basis for a quality web site: accessibility, speed, navigability and site content (**Figure 1**).

The key factors within each category are chosen based on the literature and the researchers' experience, and must reflect what users are generally considered to be important components and features of web sites.

The first category in the WAI is accessibility. It is clear that the quality of a web site is increased if the site

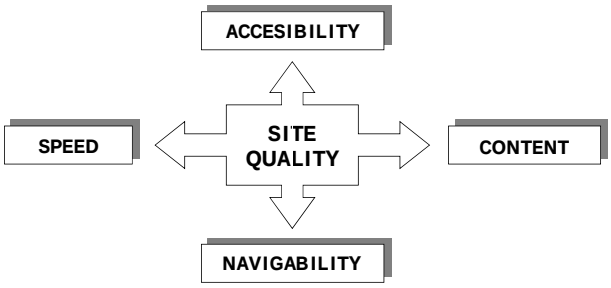

Figure 1. Web site quality assessment model

is easily identifiable and accessible to the users. However, merely counting "hits" on a page is not an accurate measurement of quality or success of a web site [15]. In order to actually evaluate the accessibility of a web site, we need to define objective measurable quantities. Therefore, the authors have employed two indicators to measure this category: search engines presence and link popularity.

1) Higher search engines rankings translate into greater traffic to the site and therefore, increase its degree of accessibility. Search engines are responsible for generating approximately 80% of new visitors to the majority of websites. Appearing in the top results for the most appropriate phrases is key to maximizing the benefit derived from the search engines. In the present work, to evaluate this factor we have chosen Google, because this search engine is the most frequently used by internet users.

2) The second indicator used to measure accessibility is the sitepopularity. The most common measure of web performance is the number of "hits" a site generates. However, there is accuracy problems associated with this type of measurement, because there are some methods to artificially increase the number of hits. Therefore, the total number of hits does not necessarily correspond to the actual number of visits to the site.

Taking this into account, we decided to employ a different kind of measurement: we defined link popularity as the number of external links on the web that point to the web site being analyzed. The advantages of a large number of links to a site are evident: firstly, the more sites that link to you, the more traffic you can expect to receive, and secondly, major search engines will improve your page ranking when you have more links to your web site [13]. For example, Google uses link popularity as the most important factor when ranking sites, so if you want to have a successful web site, you must have high link popularity. In this study we have used the Link Popularity Check, a freeware program that checks the link popularity status of a web site on several search engines and compares it to other web sites on the Internet.

Access speed and response time are obviously very significant, because time is always a critical factor. Some studies have revealed that there is a significant correla-

tion between web site download speed and web user satisfaction [16,17]. The time it takes for a page to download is important for making it easily accessible to every user interested in accessing online services.

The access speed has been measured with a chronometer, but this recording is influenced by a great number of factors such as hardware employed, connection time, web traffic etc. In order to minimize these sources of error, the tests were carried out at the same time with the same computer. Web browsing was undertaken by using the most popular browser, Internet Explorer 6.0. The sites were repeatedly accessed on consecutive days to obtain more representative average speed measurements.

The third category in our index is what we call navigability. Poor web design will result in a loss of potential sales due to users being unable to find what they want, and a loss of potential repeated visits due to an initial negative experience. Given that users should never feel lost, each page should be selfsufficient and provide links to the main contents. In terms of navigation, our analysis focused on the consistency of the navigation style. Navigation bars should be present on every website, in order to allow quick access through the entire site, particularly for those web sites with large amounts of content. So, the hallmark of a good site is that the site index should always be on display, thereby making it very easy for anybody to reach the desired location fast enough. So, following the research of Miranda *et al.* [14], the factors used to assess this category are the following:

1) Permanent site menu allowing a rapid access to the different sections from every page.

2) Web site map, for users to locate available interesting items within the company home page.

3) Keyword search option.

Content is a critical component of any website. No matter how technologically advanced a website's features, if its content is not current or if the information provided is not correct, then it is not fulfilling its purpose. The content quality of the web site will be measured assessing the presence of information relevant to the users. A site must have contents that satisfy users' needs and it should be frequently updated. Football Clubs web sitescan contain various features, including tickets information, history, events, online shop, etc.

Important information should be immediately accessible. If 80% of your users are seeking 20% of your information, then that information should be the most visible and the easiest accessible. Basic contact information of the organization should be on the main menu page and related information should be grouped together rather than scattered in different sections of the site.

The factors selected to quantify content quality were based on site contents identified in previous studies [6,7, 12–14] practitioner journals and the researchers' experience. Our methodology for evaluating football clubs web sites includes such components; however, we have added some additional factors. We have considered three sets of factors to assess the content of a web site using a binary no/yes scale:

1) *Informational factors.*-Providing online information to potential users is the most important part of an effective football club website. So, football club sites are largely informational. We have considered the following informational factors:

- Club History
- Corporate Identity
- Seasonal tickets information
- Shop Information
- Restaurant information
- Events information
- Travel information
- News
- External links
- Corporate Information
- Prices information
- Languages
- Organizational structure

2) *Transactional factors.*-A critical component of football club site is the provision of services online. These options operate under a secure server, which means that information therein is protected once you have logged in.

So, the Football clubs content features that are found in each of the searched web sites can be included in the following categories:

- Tickets online
- Online shop (merchandising)

3) *Communicational factors.*-Given that web sites are often entry points to the football club, visitors typically want to obtain access to information about the football club, contact information for the club should be on the main menu page and therefore easily accessible. More broadened contact information is a benefit for users. In order to evaluate the communicational capacity of each web site we identified whether the site contains the following items:

- Complaints and comments e-mail
- Telephone
- Users feedback tools (e-mail bulletin)
- Reserved area

To ensure reliability, each Football Club Web site was assessed by two evaluators, and in those cases where significant variations (over or under 10%) were shown on the raw score between evaluators, websites were analyzed a third time.

After defining them, all the categories and factors were weighted (based on a total of 100 points). These weightings have been assigned from the analysis of previous studies [4,8,12–14] and have taken into account the views of different Internet expert users. Moreover, to

ensure the reliability of this assignment, 10 web sites users were recruited as evaluators. Each web site user provided the relative importance of the different categories in the instrument. First, users distributed 100 points among the 4 major categories. A Delphi analysis allowed us to establish the final list of items and their weights.

The use of these subjective weightings may be considered as the main limitation of our study; however, some recent studies [14] have employed similar weightings. Mean values of the weights obtained for the different categories and subcategories of Football club sites are shown in **Table 1**.

3. Football Club Web Sites Assessment

We have assessed the 76 best Football Club in the 4 more important leagues in Europe: Premier (England), LFP (Spain), Calcio (Italy) and Portuguese League (Portugal). These leagues are not only ranked within the top five European football leagues from a point of view of performance on the field, but are also widely considered to be the most businesslike football leagues in the world.
This research evaluated the official Web sites of each Football club selected in their native languages in September of 2008. **Table 2** is a list of the Football clubs selected.

According to the WAI, the best web sites correspond to Lazio (Italy) and Chelsea (England), followed by two Spanish clubs Ath. Bilbao and Barcelona. However, the most valuable output from our study is not the ability to identify the best sites, but to see how each Football club site is compared to related sites and to spot ideas and practices that can improve theses sites.

Table 3 shows an outline of the results derived from the study for European Football clubs (according to WAI). The main overall results, grouped by categories (accessibility, speed, navigability and site content), are summarized and discussed in the following sections.

Figure 2 shows the average score by country. England Football club have higher WAI values compared to the rest, although the difference is not highly significative.

Table 1. Web assessment index

CATEGORIES	Weights	CATEGORIES	Weights
ACCESSIBILITY	**15%**	**NAVIGABILITY**	**15%**
Presence in search engines		Site map	
Popularity		Permanent site menu	
SPEED	**15%**	Keyword search option	
Access speed (in seconds)			
CONTENTS QUALITY	**55%**		
Informational content	18,33%	**Transactional content**	18,33%
Club History		Online Shop	
Corporate Identity		Tickets online	
Corporate Information			
Events information			
External links			
Languages			
News		**Communic. content**	18,33%
Organizational structure		Complaints and comments e-mail	
Prices information		Reserved area	
Restaurant information		Telephone	
Seasonal tickets information		Users feedback tools	
Shop Information			
Travel information			
Club History			
Corporate Identity			
Corporate Information			
Events information			
External links			
		TOTAL	**100**

Table 2. List of football club analyzed

Spain	Italy	Portugal	England
Almeria F.C.	Atalanta	Académica Coimbra	Arsenal
Atlethic de Bilbao	Cagliari	Beleneses	Aston Villa
Atlético de Madrid	Catania	Benfica	Black Bum Rover
Barcelona	Empoli	Boa Vista	Bolton Wanderers
Betis	Fiorentina	CD Nacional Madeira	Chelsea
C.D. Osasuna	Genoa	Estrela Amadora	Everton
Depor. de la Coruña	Inter de Milán	Leixoes	Fulham
Espanyol	Juventus	Marítimo	Hull City
Getafe	Lazio	Naval F.C.	Liverpool
Numancia	Livorno	O Porto	Manchester City
Racing de Santander	Milán	Pacos Ferreira	Manchester United
Real Madrid	Napoli	Setúbal	Midlesbrough
Recreat. de Huelva	Palermo	Sporting de Braga	NewCastle United
Sevilla F.C.	Parma	Sporting de Lisboa	Portsmouth
Sporting de Gijón	Reggina	Uniao Lira	Stoke City
U.D. Mallorca	Roma	Victoria Guimares	Sunderland
U. D. Levante	Sampdoria		Tottenham H.
Valencia	Siena		West Bromwich
Valladolid	Torino		West Ham
Villarreal	Udinese		Wigan Athletic

a. Sample of a Table footnote. (*Table footnote*)

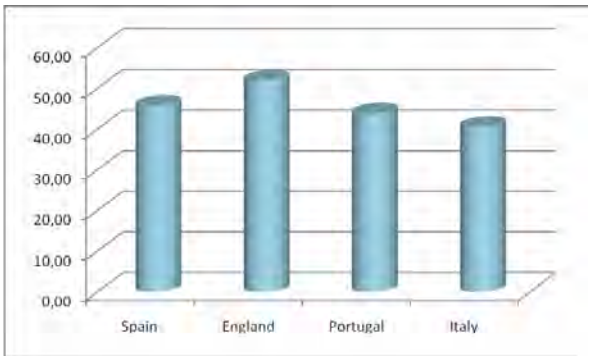

Figure 2. WAI values by country

Table 3. WAI values for European football clubs

Club	WAI	Accesibility	Speed	Navegability	Content
Lazio	73,16	53,21	66,667	66,67	82,14
Chelsea	70,48	63,51	20,000	100,00	78,10
Atlethic de Bilbao	66,87	51,45	70,000	100,00	61,19
Barcelona	65,09	46,08	35,897	33,33	86,90
Sunderland	64,36	28,73	41,176	66,67	79,76
Getafe	64,19	25,91	41,176	66,67	80,24
Bolton Wanderers	64,18	51,47	17,284	66,67	79,76
Atlético de Madrid	63,48	52,25	37,838	100,00	63,57
Liverpool	62,47	49,30	36,842	66,67	71,90
Cagliari	60,79	26,66	50,000	66,67	71,43
Manchester United	58,94	75,55	33,333	66,67	59,29
West Bromwich Albim	57,39	51,27	24,561	66,67	65,48
Hull City	56,99	25,97	34,146	66,67	69,05
Real Madrid	56,21	57,27	43,750	100,00	47,38
Aston Villa	56,00	52,32	8,284	33,33	76,19
Arsenal	55,33	66,39	43,750	66,67	52,38
Inter de Milán	55,11	66,99	45,161	66,67	51,43
Beleneses	53,10	50,53	73,684	33,33	53,57
Sporting de Lisboa	53,04	52,77	73,684	33,33	52,86
Atalanta	52,39	51,37	30,435	66,67	54,76
Sporting de Braga	52,30	50,70	24,561	33,33	65,48
Benfica	52,22	53,39	51,852	33,33	57,14
Victoria Guimares	51,64	0,59	32,558	66,67	66,67
O Porto	51,47	54,09	63,636	33,33	52,38
Wigan Athletic	50,05	26,57	20,588	33,33	69,05
Sevilla F.C.	49,90	54,26	45,161	33,33	54,52
Manchester City	49,59	57,12	63,636	66,67	39,05
Boa Vista	49,02	50,58	20,290	33,33	60,71
Black Burn Rover	47,77	51,80	29,787	66,67	46,43
Valencia	47,51	54,77	63,636	33,33	45,00
Sampdoria	47,22	52,10	28,571	33,33	54,76
Unión Deportiva Levante	45,79	50,25	23,333	66,67	45,00
West Ham	45,05	55,31	21,212	66,67	42,86
Juventus	44,57	57,04	58,333	33,33	40,48
Almeria F.C.	44,53	25,57	73,684	66,67	35,71
Villarreal	44,45	51,68	35,000	33,33	48,10
Stoke City	44,19	13,65	22,222	66,67	52,38
Fulham	44,13	53,67	25,455	66,67	40,48
Portsmouth	44,04	26,95	13,333	33,33	60,00
Milán	43,89	26,56	37,838	66,67	44,05
Midlesbrough	43,83	26,83	19,718	66,67	48,81
Everton	43,27	60,00	17,722	66,67	39,29
Tottenham H.	41,76	30,28	40,000	66,67	38,57
Sporting de Gijón	41,37	50,19	28,000	66,67	35,71
Deportivo de la Coruña	40,90	27,14	53,846	100,00	25,00
NewCastle United	40,84	52,87	18,421	33,33	45,71
Racing de Santander	40,75	25,11	50,000	33,33	44,52
Torino	40,60	52,47	20,588	66,67	35,71
Numancia	40,56	50,55	22,222	66,67	35,71
Fiorentina	39,94	51,18	33,333	33,33	40,48
Valladolid	39,57	25,95	9,655	66,67	44,05
C.D. Osasuna	39,43	51,89	46,667	33,33	35,71
Reggina	39,35	51,42	29,167	33,33	40,48
Siena	38,80	51,14	100,000	33,33	20,24
Uniao Lira	38,52	50,24	46,667	33,33	34,52
Udinese	38,49	51,77	35,000	100,00	19,05
Roma	38,02	100,00	15,385	33,33	28,57
Marítimo	37,48	0,66	29,787	66,67	41,67
Leixoes	37,47	50,20	42,424	0,00	42,86
Betis	36,66	51,95	50,000	33,33	29,76
Empoli	35,71	50,90	14,141	33,33	38,10
Estrela Amadora	34,32	50,13	31,818	33,33	30,95
Palermo	33,33	27,89	51,852	33,33	29,76
Recreativo de Huelva	32,00	50,89	63,636	33,33	17,86
Livorno	31,91	1,79	46,667	33,33	35,71
Académica Coimbra	30,91	0,62	41,176	33,33	35,71
U.D. Mallorca	29,33	51,08	26,415	33,33	23,10
C.D. Nacional Madeira	29,17	0,03	38,889	33,33	33,33
Genoa	27,12	51,16	4,651	33,33	25,00
Espanyol	26,90	51,85	40,000	33,33	14,76
Napoli	25,86	13,71	20,588	33,33	28,57
Parma	24,36	25,78	37,838	33,33	17,86
Catania	21,11	13,16	20,000	33,33	20,24

a. Each factor was measured on a 1-100 scale.

3.1 Accessibility

Overall website visibility for this sector is good. The biggest factor in this success is the number of incoming links to the websites. For the majority of websites these are high and this sector's average number of incoming links exceeds the average across the internet as a whole. These links will help to drive traffic to the website, act as an endorsement and boost brand reputation, and help to improve search engine performance. There are pros and cons to having multiple domain names, as a majority of the clubs does, however one of the drawbacks is that by not having all the incoming links pointing at a single domain name, it dilutes the positive effect that they have on search engine performance.

Given that clubs achieve better profit margins selling their merchandise directly, it is surprising not to see more evidence of optimization despite the relatively strong showing by some of the clubs in the search engines. Roma and Manchester United show the greater values in this category.

3.2 Speed

Typically, those sites whose pages fully loaded quickly were also easily navigable. The term speed is often used synonymously as data rate in networking. Technically speaking, speed refers to the user-perceived performance of the network application.

Siena, Sporting de Lisboa, Almeria and Beleneses were the sites that received the highest possible rating in this area. The range of values measured varied from 4 to 100 seconds.

Football websites are as likely to be browsed at home as at work. This is a difficult issue to resolve, as website visitors doubtless want to see images of the team, players and matches, whilst the clubs want to harness the popularity of their websites to place advertising. A large number of websites employ pop-up windows for advertising. Use of popup windows are generally considered to be poor practice from a usability point of view, as they tend to irritate users reducing the download speed and with pop-up blockers becoming more common, the effectiveness of this approach is also open to question. So, more consideration of connection speed may provide a better user experience for website visitors.

3.3 Navigability

We assessed how easy it was to navigate around the site, to return to the home page or to find relevant information. Links to components within the site should be available from every page and the security must be appropriate for the interactions conducted at the site [4].

Given the size of the websites, and the amount of content available on them, it is disappointing to report a shortage of alternative navigation options. Almost 96% of the sites provided a permanent menu but only 42% of the web websites have an internal search engine. Additionally, only 15% have a site map.

Four Spanish clubs, Atlethic de Bilbao, Deportivo de la Coruña, Real Madrid, Atlético de Madrid, with Udinese (Italy) and Chelsea (England) achieved the higher rates in this category.

3.4 Site Content

Site content has been assessed by considering the relevant information that must appear in a Football club site according to the needs of potential users. Barcelona, Lazio and Getafe achieved the higher rates in this category.

External links, Club History and News are included in almost all web sites analyzed (see **Figure 3**). Only 53,42 % offer seasonal tickets information.

Gaining the ability to communicate with new international audiences in their own language will boost sales and create awareness of a brand, service or product. With every language added to a Web site, there is the potential for an increase of more than 100% in sales.

Much has been made of football clubs attempting to promote themselves worldwide, and with players coming from all over the world to play for their clubs, interest in the European Football Leagues has never been greater. Therefore, it can be seen as a missed opportunity that relatively few websites (30%) currently offer content limited or otherwise in another language.

However Clubs like Barcelona, Milán, Chelsea, Inter, Manchester United, Sevilla F. C., Real Madrid, Valencia and Manchester City deserve a special mention, for providing content in several languages, including Chinese and Japanese.

With respect to the transactional content features that are found in each of the searched web sites, the most interesting area involves using the Internet to facilitate users online tickets. Only 13,7% Football Clubs offer this option. And only 46% include an online shop to buy merchandising articles.

Finally, with respect to communicational factors, most sites provide an e-mail for complaints and comments (75,3%) and a contact telephone number (80,82%), with only about 17,8% of them providing an e-mail bulletin.

One change we have noticed is the increased use of password protected online areas that provide information to register users (47,9%). Making users register might not be a huge issue to those clubs concentrating on their traditional fan base, but for those clubs looking to attract new audiences – for instance from abroad – this approach could also have a negative impact on their brand, affect customer acquisition and ultimately merchandise sales. It is also worth noting that the registration forms for most of these websites are also quite long and detailed.

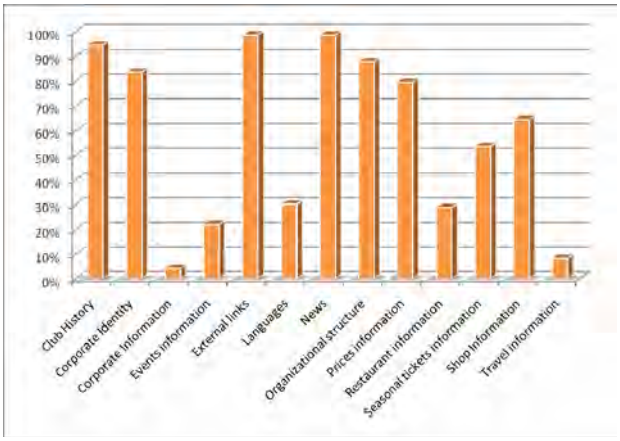

Figure 2. Informational content

4. Discussion and Conclusions

The web environment offers Football Clubs the opportunity to deliver information and services, to enhance communication to and among its stakeholder groups.

This paper proposes and tests a model, the Web Assessment Index (WAI), for evaluating the potential of Football Club web sites, allowing researchers and managers to compare attributes and components of Internet sites, in order to determine the drawbacks and opportunities.

Although the quality of content generally cannot be faulted, the quality of user experience could be improved. Fans having to wait for pages to load, not so much because of pictures of the team, but because of adverts, are entitled to feel frustrated. Website abandonment must still be a factor on these websites, as well as perhaps generating negative feelings towards the brand. Some home pages, for example could take as long as a minute or more to fully load on a standard connection.

For some websites, advertising can take up as much as a third of the screen. And with space for navigation, content sometimes seems squeezed into the middle, almost as an afterthought.

The main challenge in the elaboration of the index was to avoid subjective factors, which have been predominant in previous assessment tools. Our index is based on four broad categories: site content, speed, accessibility and navigability which are quantified in an objective and logical way. The results of the application of this index to the analysis of Football Club web sites have demonstrated the high flexibility of the WAI and have detected the main weaknesses of the web pages assessed.

It is worth noting that there are several limitations for this analysis. Firstly, all the data in the survey was collected from a limited number of visits to each site at a certain time, despite the fact that the web is a highly dynamic and changeable medium. Similar studies at different times are likely to show different results. Such an evaluation over time will also shed some light on whether there is a divergence or convergence of web activities.

In addition, it is prudent to note that differences in club structures, culture, financial circumstances and indeed sporting codes, mean that the results are not necessarily generalisable to every club, or indeed to the wider European Football industry. Nevertheless, every effort has been made to include league-based comparisons in order to identify any differences.

A second concern was the subjective nature of factors weightings, which although are based on the results of previous studies and personal interviews with Internet experts.

Nevertheless, the strength of this study lies in its foundation for future research: Having identified the critical categories and factors in the WAI, the next stage is to test this tool in different contexts and measure the relationship between the index value and organizational success. Further research in this important area is currently under way.

Evaluators can use this instrument easily without specific training or knowledge and, besides, the evaluation time is lower than in other assessment models [11].

As further information on web design and usability methods becomes available, the assessment index presented here can be refined into an empirically validated toolkit for the design of functional corporate sites. The proposed index constitutes a suitable method for evaluating web sites and making a comprehensive analysis of the usage of the new medium.

REFERENCES

[1] D. Selz and P. Schubert, "Web assessment: A model for the evaluation and the assessment of successful electronic commerce applications," Electronic Markets, Vol. 7, No. 3, pp. 46–48, 1997.

[2] C. Liu, K. Arnett, L. Capella, and R. Beatty, "Web sites of the fortune 500 companies: Facing customers through home pages," Information and Management, Vol. 31, pp. 335–345, 1997.

[3] J. Ho, "Evaluating the world wide web: A global study of commercial sites," Journal of Computer Mediated Communication, Vol. 1, No. 3, 1997.

[4] J. R. Evans and V. E. King, "Business to business marketing and the world wide web: Planning, managing and assessing web sites," Industrial Marketing Management, Vol. 28, pp. 343–358, 1999.

[5] R. Simeon, "Evaluating domestic and international website strategies," Internet Research: Electronic Networking Applications and Policy, Vol. 9, No. 4, pp 297–308, 1999.

[6] E. Huizingh, "The content and design of web sites: An empirical study," Information and Management, Vol. 37, pp. 123–134, 2000.

[7] D. Young and J. Benamati, "Difference in public web sites: The current state of large U.S. firms," Journal of Electronic Commerce Research, Vol. 1, No. 3, pp. 94–105, 2000.

[8] C. Bauer and A. Scharl, "Quantitative evaluation of web site content and structure," Internet Research: Electronic Networking Applications and Policy, Vol. 10, No. 1, pp. 31–43, 2000.

[9] J. W. Palmer, "Web site usability, design, and performance metrics," Information Systems Research, Vol. 13, No. 2, pp. 151–167, 2002.

[10] A. Boyd, "The goals, questions, indicators, measures (GQIM) approach to the measurement of customer satisfaction with e-commerce web sites," Aslib Proceed- ings, Vol. 54, No. 3, pp. 177–187, 2002.

[11] R. V. D. Merwe and J. Bekker, "A framework and methodology for evaluating e-commerce web sites," Internet Research: Electronic Networking Applications and Policy, Vol. 13, No. 5, pp. 330–341, 2003.

[12] M. Buenadicha, A. Chamorro, F. J. Miranda, and O. R. González, "A new web assessment index: Spanish universities analysis," Internet Research: Electronic Networking Applications and Policy, Vol. 11, No. 3, pp. 226–234, 2001.

[13] F. J. Miranda and T. M. Bañegil, "Quantitative evaluation of commercial web sites: An empirical study of Spanish firms," International Journal of Information Management, Vol. 24, No. 4, August, pp. 313–318, 2004.

[14] F. J. Miranda, R. Cortés, and C. Barriuso, "Quantitative evaluation of e-banking web sites: An empirical study of spanish banks," The Electronic Journal Information Systems Evaluation, Vol. 9, No. 2, pp. 73–82, 2006.

[15] M. Murray, "Evaluating web impactthe death of the highway metaphor," Direct Marketing, Vol. 59, pp. 36–39, 1997.

[16] S. Muylle, R. Moenaert, and M. Despontin, "Introducing web site user satisfaction: An integration of a qualitative pilot study with related MIS research,'' Working paper, Owen Graduate School of Management, Vanderbilt University, Nashville, 1998.

[17] D. L. Hoffman and T. P. Novak, "Marketing in hypermedia computer mediated environments: Conceptual foundations," Journal of Marketing, Vol. 60, pp. 50–68, July 1996.

J. Service Science & Management, **2010, 3: 117-129**
doi:10.4236/jssm.2010.31015 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

A Review on Deteriorating Inventory Study

Ruxian Li[1], Hongjie Lan[1]*, John R. Mawhinney[2]

[1]School of Economics and Management, Beijing Jiaotong University, Beijing, China; [2]Business School, Duquesne University, Pittsburgh, USA.
Email: *hjlan@bjtu.edu.cn

Received September 2nd, 2009; revised October 21st, 2009; accepted November 29th, 2009.

ABSTRACT

To provide a comprehensive introduction about the deteriorating items inventory management research status, this paper reviews the recent studies in relevant fields. Compared with the extant reviews (*Raafat* 1991; *S.K.Goyal* 2001), *this paper reviews the recent studies from a different perspective. First, this paper proposes some key factors which should be considered in the deteriorating inventory studies; then, from the perspective of study scope, the current literatures are distinguished into two categories: the studies based on an enterprise and those based on supply chain. Literatures in each category are reviewed according to the key factors mentioned above. The literature review framework in this paper provides a clear overview of the deteriorating inventory study field, which can be used as a starting point for further study.*

***Keywords*:** *Deteriorating Items, Inventory Model, Literature Review*

1. Introduction

Deteriorating items are common in our daily life; however, academia has not reached a consensus on the definition of the deteriorating items. According to the study of Wee HM in 1993 [1], deteriorating items refers to the items that become decayed, damaged, evaporative, expired, invalid, devaluation and so on through time. According to the definition, deteriorating items can be classified into two categories. The first category refers to the items that become decayed, damaged, evaporative, or expired through time, like meat, vegetables, fruit, medicine, flowers, film and so on; the other category refers to the items that lose part or total value through time because of new technology or the introduction of alternatives, like computer chips, mobile phones, fashion and seasonal goods, and so on. Both of the two categories have the characteristic of short life cycle. For the first category, the items have a short natural life cycle. After a specific period (such as durability), the natural attributes of the items will change and then lose useable value and economic value; for the second category, the items have a short market life cycle. After a period of popularity in the market, the items lose the original economic value due to the changes in consumer preference, product upgrading and other reasons.

The inventory problem of deteriorating items was first studied by Whitin [2], he studied fashion items deteriorating at the end of the storage period. Then Ghare and Schrader [3]concluded in their study that the consumption of the deteriorating items was closely relative to a negative exponential function of time. They proposed the deteriorating items inventory model as stated below:

$$\frac{dI(t)}{dt} + \theta I(t) = -f(t)$$

In the function, θ stands for the deteriorating rate of the item, $I(t)$ refers to the inventory level at time t and then $f(t)$ is the demand rate at time t. This inventory model laid foundations for the follow-up study. Raafat [4] and Goyal and Giri [5] made comprehensive literature reviews on deteriorating inventory items in 1991 and 2001 respectively.

From a different perspective, this paper reviews the recent trends in deteriorating inventory studies. **Figure 1** presents the factors used to analyze and organize this review. From the perspective of scope, we make a distinction between the studies which focus on the deteriorating items inventory study in a single enterprise from those studies whose focus is on studying the deteriorating items inventory problems across a supply chain. The former is the focus of the early stage in the deteriorating items study and the latter now is attracting more and more attention from the researchers. From the perspective of the factors which should be taken into consideration in deteriorating items inventory study, we involve the important factors such as demand, deteriorating rate and

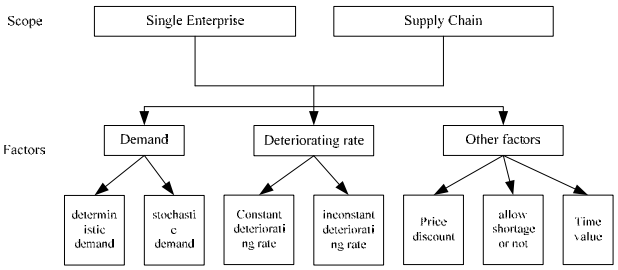

Figure 1. Deteriorating items inventory literature and its relation in the review

other factors such as price discount, allow shortage or not, inflation, time-value of money and so on in our study. By way of integrating different factors in different scope (in a single enterprise or across a supply chain), different models can be established.

This paper is organized as follow. Section two discusses key factors which should be taken into consideration in the deteriorating inventory study. Section three presents an overview on deteriorating items inventory study in a single enterprise. For the purpose of classification, three lines of research can be distinguished: the first and the second lines study the inventory problems in the enterprises which sell deteriorating items. The main difference between the two lines is that the first line involves a warehouse while the second one involves two warehouses. The third line reviews the produce and inventory problems in the deteriorating items manufacturers. The studies on the deteriorating items inventory problems from the perspective of the supply chain are sketched in section four. Then this paper discusses the directions of the future research.

2. Key Factors in the Deteriorating Inventory Study

Factors such as demand, deteriorating rate, and so on should be taken into consideration in the deteriorating inventory study. Among them, demand acts as driving force of the entire inventory system and the deteriorating rate stands for the characteristics of the deteriorating items. Other factors like price discount, allow shortage or not, inflation, and the time-value of money are also important in the study of deteriorating items inventory. By making different combinations of these factors stated above, we can get different inventory models.

2.1 Demand

Acting as the driving force of the whole inventory system, demand is a key factor that should be taken into consideration in an inventory study. There are mainly two categories demands in the present studies, one is deterministic demand and the other is stochastic demand. Constant demand [6,7], time-dependent demand [8,9] inventory level-dependent demand [10–13] and price-dependent demand [14] are all deterministic demand. Among them, ramp type demand is a special type of time-dependent demand. Hill [15] was the first to introduce the ramp type demand to the inventory study. Then Mandal and Pal [16] introduced the ramp type demand to the inventory study of the deteriorating items. After that, many researchers have extensively studied this type of demand [17,18]. Stochastic demand includes two types of demands: the first type characterized by a known demand distribution and on the contrary the second type characterized by arbitrary demand distribution.

2.2 Deteriorating Rate

Deteriorating rate is another key factor in the study of deteriorating items inventory, which describes the deterioration nature of the items. When it comes to the study of deteriorating rate, there are several situations. In the early stage of the study, most of the deteriorating rates in the models are constant, such as Ghare and Schrader [3], Shah and Jaiswal [19], Aggarwal [20], Padmanabhana and Vratb [21], and Bhunia and Maiti [22]. In recent research, more and more studies have begun to consider the relationship between time and deteriorating rate. In this situation there are several scenarios; including deteriorating rate is a linear increasing function of time [23,24], deteriorating rate is two-parameter Weibull distributed [25,26], deteriorating rate is three-parameter Weibull distributed [27], and deteriorating rate is other function of time [28].

2.3 Other Factors

Beside demand and deteriorating rate, other factors like price discount, allow shortage or not, inflation, and time-value of money are also important factors. Price discount is an important strategy which the seller always uses to encourage the buyer to purchase in large quantities; many researchershave taken this factor into consideration in deteriorating items inventory modeling. Allowing shortage or not is another factor which researchers always focus on. Among them, some studies supposed that shortage is not allowed [29], the rest supposed that shortage is allowed and then corresponding inventory strategy can be made according to the two assumptions respectively. In fact, shortages usually happen in our daily life and what's more, in the circumstance of high deteriorating rate, the demand may need to be backlogged to reduce cost due to deterioration, so there are more studies that concentrate on the assumption that shortage is allowed. There are two cases when dealing with the shortage, one case supposes that the shortage items are totally backlogged [30,31] and the other case supposes that the shortage items are partly backlogged, that is to say, the customers are only willing to accept part of the items that are out of stock this period and can only be supplied by

the seller in the next period. There are some different considerations in dealing with the backlogging function. In some of the studies, the backlogging function was assumed to be closely relative to the amount of demand backlogged [32], therefore the more the amount of demand backlogged, the smaller the demand to accept backlogging would be. While in our real life, for deteriorating items with short life time, the waiting time for the next replenishment is the main factor for deciding whether the backlog will be accept or not and when we are willing to accept the backlog, what is the accepting proportion? It is easy for us to know that the willingness of a customer to wait for backlogged items during a shortage period declines with the length of the waiting time. In order to reflect this phenomenon, some studies [33–35] developed inventory models in which the backlogging rate is a function of waiting time.

Trade credit is another factor. Trade credit (permissible delay in payment) is a widely used business strategy. To suppliers, trade credit helps to expand sale but it also adds to the risk of bad debts at the same time. To buyers, trade credit provides a very big advantage, due to the fact that they do not have to pay the seller immediately after receiving the items, but instead can delay their payment until the end of the allowed period. The buyer pays no interest during the agreed time for payment, but it increases the risk of deterioration cost when large amounts are purchased. However, both sellers and buyers need to reach trade-off between the advantages and disadvantages they can get from this inventory policy. At present, there are three kinds of trade credit, in the first case, the length of the credit period is fixed [36–38], in the second case, the length of the credit period is closely relative to the ordering quantity [39–41]. In the third case, the sellers provide trade credit to part of the ordering quantity and the purchase cost of the rest quantity should be paid immediately after receiving the items.

Inflation and time value of money have also attracted attention of researchers. Taking the two factors into consideration is of vital importance. With the integration of the global economy, the economic relationships among countries are closer and the mutual influences are deep. Currency's purchasing power will change from time to time and inflation should not be neglected. Many researchers [14,44–47] have made complementally comprehensive consideration in deteriorating items modeling.

3. Deteriorating Items Inventory Study in a Single Enterprise

Inventory cost is an important part of the enterprise operation cost. For deteriorating items, especially those with high deteriorating rate, deterioration is a key characteristic and its impact on modeling of inventory systems cannot be neglected. So the deterioration rate should be taken into consideration in the development of inventory strategy. For different kinds of enterprises, the emphasis on the deteriorating items inventory study is different. For the seller of deteriorating items, the current studies can be divided into two types; the first type emphasizes the inventory strategies for the retailer of the deteriorating items, the second type focuses on the inventory policy under a two-warehouse system. For the manufactures of deteriorating items, the current emphasis is on developing an optimal production-inventory strategy. So, this paper has divided the present studies on deteriorating items inventory for a single enterprise into three categories as stated above.

3.1 Deteriorating Inventory Strategy

The sellers of deteriorating items such as retailers must frequently assess the replenishment strategy for deteriorating items. Deteriorating items have a shorter life time compare with other items. In order to avoid the loss due to damage or expiration, it is of vital importance to develop a proper inventory strategy. Many researchers have conducted extensive studies on the deteriorating inventory strategy. Compared with normal inventory models, besides demand, deteriorating rate is another key factor that impacts significantly on an inventory management system. So when discussing models below, this paper will mainly concentrate on demand and deteriorating rate, and other factors, such as price discount, inflation, and so on, will also be taken into consideration.

3.1.1 Deterministic Demand

Sarker, Jamal and Wang [48], Chang [49], Chung and Liao [50], Huang and Liao [51]), ZHANG, DAI and HAN [42], ZHANG, DAI, HAN and LI [52] all developed inventory models in which both the demand and deteriorating rate are constant. In their research, the buyer is allowed a delay period to pay for the items purchased. The purpose of their studies is to help the buyers to make economic inventory strategy decisions under the influence of trade credit. Although the constant demand assumption helps to simplify the problem, it is far from the actual situation where demand is always in change. In order to make research more practical, many researchers have studied other forms of demand. Among them, time-dependant demand has attracted considerable attention. WANG Sheng-dong and WANG jun-ping [53] developed a model to determine optimal ordering policy for deteriorating items under inflation, partial backlogging, and time-dependant demand. The effect of the time value of money was also considered in the paper. Panda, *et al.* [54] discussed an inventory model for a seasonal product. In the model, the demand rate is represented by a ramp-type time dependent function and the deteriorating rate is a constant. With the assumption that shortages are not allowed, the paper aimed to develop an optimal replenishment policy for retailer. In addition, Papachristos and Skouri [55], Chu and Chen [56], Khanra and Chaud-

huri [57], Yang [58], Dye *et al.* [59] all conducted research on deteriorating items inventory under the premise that the demand is time-dependent.

Besides time, the inventory level is another factor that has a close relationship with demand. For some item types, the demand rate may fluctuate with the on-hand stock level. It is a common phenomenon that a large amount of goods displayed in the supermarket will lead the customers to buy more, so the demand is closely dependant on the inventory level. Gupta and Vrat [60] first proposed an inventory model in 1986 in which demand rate is stock-dependent. Then Mandal and Phaujdar [61] developed a production-inventory model for deteriorating items with the assumption that demand is a linear function of inventory level. Padmanabhan and Vrat [62] developed an EOQ model for deteriorating items with stock-dependant demand, they proposed three models: completely backlogging, partial backlogging and without backlogging, for the purpose of maximizing profit. In this paper, the backlogging function was assumed to be closely relative to the amount of demand backlogged. As we discussed before, in real life it is more common that the willingness of a customer to wait for a backlog during a shortage period declines with the length of the waiting time, that is to say, the longer the waiting time is, the less likely the customers willing to wait. Then in 2005, Dye and Ouyang [63] extend Padmanabhan's model (Padmanabhan and Vrat, 1995) by proposing a time-proportional backlogging rate. While the opportunity cost due to lost sales was not taken into consideration in Padmanabhan's model, Chung-Yuan Dye took considerable consideration in his improved model. Hou [64] established an inventory model for deteriorating items with stock-dependent demand rate and shortages under inflation and time discounting. There are also some who studied other types of demand, like price-dependent demand [65–67], inflation-dependent demand [68] and so on.

In fact, the real situation is complex and the demand is always affected by several factors such as time, inventory level, price, and so on. Balkhi and Benkherouf [69], Pal *et al* [70], Hsu *et al* [71] combined several of the factors together and considered the impact of the combination on the demand, in this premise the optimal inventory strategy was discussed. WEN *et al* [72] developed an inventory model in which the demand is affected by time and inflation. The aim of the paper is to determine the optimal replenishment strategy including the replenishment time and order quantities which minimize the present value of the total cost.

All the models discussed above have the characteristics of a constant deteriorating rate. While the constant rate simplifies the problems, it cannot reflect the real situation of the deterioration; in fact the deteriorating rate varies with time. Wee and Law [73] proposed a deteriorating inventory model which took into consideration the time-value of money and price-dependent demand. Papachristos and Skouri [74] considered a model where the demand rate is a function of the selling price and the backlogging rate is a time-dependent function. Both of them assumed that the deteriorating rate is two-parameter Weibull distributed. Dye *et al* [75] developed an inventory model for deteriorating items with price-dependent demand. In the model the deterioration rate is a function of time and shortages are allowed. The unsatisfied demand is partially backlogged and the backlogging rate is a negative exponential rate of the waiting time. The aim of the model was to maximize the profit, and according to the model, the inventory strategy, including price, economic ordering quantity and so on, can be determined. Then Dye *et al* [76] took the time-value of money into consideration and extended their earlier research. Chang, *et al* [77] established an EOQ model for deteriorating items. In the model the supplier provided trade credit to the purchaser on the condition that the order quantity was greater than or equal to the predetermined quantity. The paper proposed Taylor's series approximation to obtain the optimal order quantity and replenishment time. Liao [78] then amended the solution for the model proposed by Chun-Tao Chang, etc and developed an alternative approach to determine the optimal ordering policy. HUANG Wei-lai and HUANG song [79]studied the economic order quantity problem with deteriorating items taking time-value of money into consideration. In the problem, the deteriorating rate was two-parameter Weibull distributed and the demand rate is linear function of the inventory level. An optimal order quantity model with deteriorating items based on the minimization of total cost was established, taking account of time value of money.

3.1.2 Stochastic Demand

In fact, with the growing uncertainty in the modern business environment, the assumption of deterministic demand is far from truth, so stochastic demand has attracted more and more attentions.

It is common that when dealing with the stochastic demand, a large number of researchers make the assumption that the demand is Poisson distributed. S. KALPAKAM and S.SHANTHI [80,81] analyzed inventory systems with Poisson demands for deteriorating items. With the assumption of constant demand, in solving the model, Zhu [82] proposed a data processing method in which the demand was accumulated, making the model easily solved by computer and then enhancing the practicality of the model. Li *et al.* [83] introduced the three-parameter weibull function to describe the characteristics of deterioration. In this paper, demand is supposed to be closely relative to deterioration, that is to say, the more items deteriorating, the less the demand will be. Beside this, shortages were allowed and totally backlogged. According to the model, the optimal inventory

policy including the replenishment time, order quantity can be determined.

With the increasing complexity in the business environment, it is of vital importance for enterprises to take factors such as demand, price discounting, inflation, time-value of money, and so on into consideration in inventory decision making. The current studies have given considerable consideration to the factors stated above and so they play an important role in the inventory decision-making.

3.2 Two Warehouse Deteriorating Inventory Study

Two-warehouse inventory problem for deteriorating items is a main category in the current inventory study for deteriorating items. The storage capacity of an enterprise is limited and in some real life situations, when the supplier provides a price discount for bulk purchases, or when the item is a seasonal product, or when the order cost is high, or the demand goes up quickly, or the cost of out-stock is high, the buyer may purchase a large amount of an item [84]. When the amount of the purchased items exceeds the storage capacity of the enterprise's own warehouse (OW), the excess quantities have to be stored in a rented warehouse (RW). This is the real background of the two-warehouse inventory problem. There are some common assumptions in the two-warehouse inventory problem. It is often believed that the storage capability of the OW is limited but the storage capability of the rented warehouse (RW) is unlimited, and that the RW often provides a better preserving facility than the OW. So compared with the OW, the inventory cost in RW is higher but the deterioration rate is lower. According to these assumptions, it is easy for us to understand that in order to reduce the inventory cost; it will be a good choice to consume the items in the RW first and store items in OW before RW (last in first out policy, LIFO). A large number of researchers have studied in this area according to the assumptions stated above, such as [85], yang, S.L. and Zhou, Y.W. [86], Benkherouf [87] and so on.

Sarma [88] first studied the two-warehouse inventory problem for deteriorating items. In his study, the deteriorating item was first stored in the OW and the excess quantity was stored in the RW. Then in the inventory model an infinite replenishment rate was considered and shortages were allowed. Then Pakkala and Achary [89] extended Sarma's study to the situation of an infinite replenishment rate. Yang [90] considered a two-warehouse inventory problem for deteriorating items with constant demand rate and shortages under the circumstances of inflation. Then Yang [91] extended the former study and considered partial backlogging in the two–warehouse inventory model, in which the backlogging rate is closely relative to the waiting time. Then Hui-Ling Yang compared the two two-warehouse inventory models based on the minimum cost approach and the result showed that the model 2 was less expensive than model 1 if partial backlogging and inflation are considered. QIU and LIANG [92]) developed a two-warehouse inventory model on the basis of minimum cost. In the model the factors such as a constant deteriorating rate and demand trade credit were taken into consideration. The numerical results can aid the manager to decide the order quantity, order cycle, and so on.

All the two-warehouse inventory models discussed above have the characteristic of constant demand; in fact many other researchers have studied other types of demand, like time-dependent demand, etc. XU and LI [93] proposed an optimal inventory policy for a two-warehouse inventory model with time-dependent demand and constant deterioration rate. In the model, shortages are not allowed and according to the model the total order quantity and the quantity stored in the rented warehouse in a cycle time can be determined. Dey *et al.* [94] studied the two-warehouse inventory problem from the perspective of retailers. In their study, the storage capability of OW which is always located at a busy market place is limited. So the excess items should be stored in a RW which may be located away from the market place. The paper assumed that the inventory cost in the RW is greater than OW and so the items are first stored in OW and only excess items are stored in RW. This is the same as we discussed before. In order to reduce the inventory cost, the RW was emptied first by transporting the stock from RW to OW in a continuous release pattern. That is to say, the demand of items was met at OW only. Then the paper developed two-warehouse inventory model with the consideration of inflation, time-dependent demand, partially backlogging, and time-value of money. Rong *et al.* [95] studied the similar problem as Dey *et al.* [94] did.

As we stated at the beginning of 3.2, in order to reduce inventory cost, a considerable number of studies on two-warehouse inventory problems for deteriorating items follow the rule of LIFO (last in first out). But Lee [96] held a different opinion in this respect; he believed that in the RW, especially in the public warehouse, the vendor of the warehouse carried a lower operating cost because of well equipment set ups, learning effect of trained workers, and economies of scale. What's more, in the buyer's market more and more warehouses have to offer valued added service with completive lower prices to attract customers. So the assumption that the holding costs are higher in the rented warehouses is not so reasonable. Besides all stated above, the researcher still believed that for deteriorating items, under the rule of LIFO, the cost of item deterioration and related opportunity cost may far exceed the cost saving benefit derived from the rent (on the condition that the holding costs in the rented warehouses are really higher than those in the OW). So it may

be more reasonable to operate under the rule of FIFO. In his study, Chun Chen Lee reconsidered the LIFO model proposed by Pakkala and Achary [97] and then developed a FIFO model. Comparison of the two models showed that the FIFO model was less expensive than the LIFO model, on the condition that the mixed effects of deterioration and holding cost in RW are less than that of the OW. Niu and Xie [98]then amended the model proposed by Chun Chen Lee and so the modified LIFO model always had a lower cost than the FIFO model proposed by Pakkala and Achary.

It is a common phenomenon that seasonal or cyclical items are often purchased in large quantities and so the study of two-warehouse inventory is of important practical significance for these items.

3.3 Production-Inventory Study for Deteriorating Items

For deteriorating items manufacturing, decision makers should take inventory into consideration in production policy making. Proposing an efficient production-inventory decision is conducive to enhancing the competitiveness of enterprises [43]. Many researchers have extensively studied in this area, such as Goyal and Gunasekaran [99], Jiang and Du [100], Chen *et al.* [101], Gong and WANG [102], Maity *et al.* [103] and so on. With the assumption that the demand rate, production rate, and deteriorating rate are all constant, Jui-Jung Liao [104,105] established a production-inventory model for deteriorating items under the condition of the supplier providing the retailer with trade credit. Teng and Chang [106] established an economic production quantity (EPQ) model for deteriorating items when the demand rate was dependant on both the stock level and the selling price per unit, and in which the deteriorating rate was constant. The paper also provided the necessary conditions to decide the optimal solution which maximized profit of the EPQ model. From the model, decision makers could get aid in developing the optimal price and in production run time decisions.

In all the models stated above, shortages are not allowed. In conditions of shortages, each production- inventory cycle in the planning horizon can be described in more detail. Generally speaking, each production- inventory cycle can be divided into four stage, there are production starts——consumption stage; production stops——consumption stage; production stops——shortage stage, and production starts——shortage stage. The order of the four stages depends mainly on the shortage occurring at the beginning or the end of the cycle. Lin Hsinyi *et al.* [107] established a production——inventory model with constant production rate, demand rate and deteriorating rate. Shortages are allowed and occur at the start of each cycle. So the production——inventory cycle can be divided into production stops——shortage stage; production starts——shortage stage; production——consumption stage; production stops——consumption stage. From the model, we can get the number of production-inventory cycles and the starting point of the four stages in every cycle. Incorporating these parameters with the known constant production rate, we can get the optimal production-inventory policy. Zhou *et al.* [108] also considered production-inventory problem in which each cycle of a production-inventory schedule starts with replenishment and ends with a shortage. In this model, time-dependent demand was taken into consideration. In the research of Sana *et al.* [109] and Zhou and Gu [110], shortages are allowed and occur at the end of a cycle. So the cycle consists of four stages as below: production starts——consumption stage; production stops——consumption stage; production stops——shortage stage and production starts——shortage stage. With the consideration of time variable demand and constant deteriorating rate, the optimal production-inventory policy was studied.

All the deteriorating rates in the models stated above are constant and some researchers have studied time-dependent deteriorating rates. With the time-varying demand and deteriorating rate, both Skouri and Papachristos [111] and Chen *et al.* [112] developed a production-inventory model in which the storages are allowed at the beginning of the cycle. In contrast, in Manna and Chaudhuri [114] and Balkhi's research [113], shortages are also allowed but occur at the end of each cycle. Abad [116] studied the pricing and lot-sizing problem for deteriorating items under the conditions of finite production, partial backlogging, and lost sale. Teng *et al* [115]extend P.L. Abad's model by adding the backlogging cost and the cost of lost goodwill. Then, the paper made a comparison between Abad's model [116] and Goyal's model [117] in which shortages occur at the beginning of the cycle. The numerical results show that there is no dominant one between these two models. The paper also provides certain conditions under which one model had more net profit per unit time than the other.

4. Deteriorating Items Inventory in the Supply Chain

With the integration of the global economy, there is more and more diversity and uncertainty in the market. In order to deal with the uncertainties in the market and respond quickly to the diverse and personal demand of customer, enterprises need to cooperate with each other in the form of an integrated supply chain. The traditional inventory theory can not adapt to the current situation any more, the inventory problem should be considered in the supply chain. The deterioration of the deteriorating items makes it is important for the relative firms in the supply chain to make the optimal inventory policy together to minimize the total inventory cost across the supply chain. In the following part of the paper, we will discuss the current studies on the deteriorating items inventory in the supply chain.

4.1 Deterministic Deteriorating Items Inventory in the Supply Chain

Yang and Wee [118,119] have conducted research on the inventory policy for deteriorating item in the supply chain including a single-vendor and multi-buyers. Yang and Wee [118] developed a multi-lot-size production and inventory model for deteriorating items with constant production and demand rates. In this paper, a supply chain including a single-vendor and multi-buyers was established and the supplier was dominant in the supply chain. The supplier delivered the goods in the form of JIT and the aim of the paper was to find the optimal production-inventory policy on the basis of minimizing the inventory cost in the supply chain. According to the numerical results, the optimal policy using the integrated approach can reduce the total cost for the producer and the buyer. However, the producer's cost is higher in the form of supply chain integration than the independent decision without considering the buyer's perspective. So in order to maintain the stability of the supply chain, the buyers have to supply some kinds of incentives to the supplier. Then Yang and Wee [119] established a collaborative inventory system of single vendor and single buyer with the price-dependent demand, shortages are not allowed in the system. The paper developed different inventory models under the cases of vendor and buyer making decisions respectively; vendor and buyer making decisions together but no incentives are incorporated in the decisions; the two parties making decisions together and the vendor providing trade credit to the buyer. The numerical results showed that among the three cases, the profit in the first case was the smallest. The profit in the second case was the same with the one in the third case and the profit in the second and third case was larger than the one in the first case. As is well known to us all, the optimal solution for the whole system is not always beneficial to all the players in the system and in fact in the second case, the buyer's profit was seriously damaged and so the buyer will refuse to cooperate with the supplier. While in the third case, the trade credit was incorporated in the system and it helped to share profit between the two parties. Compared to the first case, both vendor and buyer can get the extra profit and the supply chain was stable in this case.

Compared to the deteriorating items inventory study in a single enterprise, the inventory studies in the supply chain have to emphasize how to maintain the stability of the whole system while achieving the minimum of inventory costs in the supply chain. Therefore, researchers focused on supply chain coordination mechanism in the deteriorating items inventory problem study. LI and HUANG [120] established a three-level supply chain including vendor, manufacture, and retailer in which the manufacture is dominant. The paper developed two different models; in first case all the firms in the supply chain make decisions independently, and in the second the firms make decisions together. In the case of non-integrated decision-making, the manufacturer was the dominant member in the supply chain; the ordering policies of vendor and supplier were subject to the policy made by manufacturer, so the aim of the model is to minimize the costs of the manufacturer. While in the case of integration, the aim of the optimal production and inventory policy making is to minimize the costs of the whole supply chain. According to the numerical results, the costs in the case of integration is lower than that in the case of non-integration, but the optimal solution for the whole system is not beneficial to all the members in the supply chain. In order to realize integration in the supply chain, the paper introduced the cost sharing mechanism in the supply chain to guarantee the implementation of integration and win-win among members. Both An and Luo [121], Qin and Guo [122] studied the effect of quantity discount across the supply chain and their studies showed that quantity discount is a useful coordination method among supply chain members. In Qin and Guo's study [122], quantity discount was used as a coordination mechanism to achieve the stability of supply chain. As the leader of the supply chain, the supplier develops an optimal discount price policy and shares the policy with the customer. Then the customer determines his unit selling price and sales volume. The numerical example showed that when the supplier determined the quantity discount, both the supplier and customer's profit increase.

4.2 Stochastic Deteriorating Items Inventory in the Supply Chain

The studies on stochastic deteriorating items inventory in the supply chain at present are much less than the ones on the deterministic deteriorating items inventory. Du *et al.* [123] studied the deteriorating item stock replenishment and shipment policy for vendor-managed inventory (VMI) system with the assumption that the demand process follows a typical Poisson process. This research is important to practical applications and we will discuss the paper in detail.

In the circumstances of vendor-managed inventory, the vendor acts as a supply chain manager and plans inventory and shipment scheduling from the perspective of the whole supply chain rather than obey orders passively and arrange shipments respectively. The aim of VMI is to develop an inventory and shipment policy in which the vendor can meet the demand of the retailers and determine their own inventory replenishment policy at the same time. The policy can realize the aim of minimizing the total costs in the supply chain. This is the so called integrated shipment which means that gathering the small amount of order quantities from downstream enterprises into a large quantity and then make integrated shipping scheduling. The

integrated shipment policy can reduce relative costs in the inventory and shipment delivery [124].

Du *et al*. [123] made a comprehensive study on the integrated shipment policy under the circumstances of vendor-managed inventory for the deteriorating items. The paper established a supply chain system which includes a manufacturer, a vendor, and many retailers located in different regions and the demand process follows a typical Poisson process. In the mode of VMI, the vendor is responsible for replenishing the retailers' inventory and arranging the corresponding shipment schedules. The purpose of the study is to make a joint decision in those two areas stated above. There are three parameters (Q, q, T) in the integrated shipment policy, among them the Q stands for the order-up to level, q stands for planned shipments and T is the planned delivery period. The policy operates as below: vendor delivery order quantities which accumulate during the predetermined replenish cycle T is shipped to the retailer at the end of the cycle. However, if the accumulated order quantity reaches the amount of q, which is also a predetermined parameter, before the end of the cycle, the vendor will also, delivery items to the retailers. Every time a delivery is planned, the vendors have to decide whether the inventory for each item needs to be replenished, and if it needs replenishment, what is the replenishment amount. It must be guaranteed that the inventory level of the vendor is equal to or more than the order-up to level parameter Q after delivery.

Integrated delivery is an important issue which must be considered when implementing vendor-managed inventory. The studies on the integrated delivery problem for deteriorating items can aid the vendors to make relevant decisions about delivery and replenishment and so can help reduce the costs in the supply chain.

5. Conclusions

The inventory problems for deteriorating items have attracted more and more attention and many researchers have conducted extensive studies in this area. In this paper, from a different perspective, we have tried to make a review on deteriorating inventory literatures after Raafat [4]'s and Goyal's reviews [5]. According to the literatures discussed in this paper, we can draw some useful conclusions. Subsequent subsections present significant findings, the gaps identified in the research and future directions of the research in the relevant area.

5.1 Significant Findings

1) According to the research scope, the current studies can be divided into two categories, the first category studies the deteriorating items inventory problems in a single enterprise and the second category studies the deteriorating items inventory in supply chain. From the perspective of quantity, there are much less studies in the second category compared with those belong to the first category. The study on deteriorating items inventory problems is a new research area compared with the inventory research on common items and so the total research amount is far less than the traditional one.

2) Also, studies in the single enterprise category are less complex than the second category of the integrated supply chain. Two situations are included: first, from the perspective of the number of key factors involved in the model, there are always fewer factors involved in the supply chain deteriorating inventory models. For example, in the first category inventory models, besides demand and deteriorating rate, other factors such as price discount, allow shortage or not, inflation, time-value of money and so on are always combined in the model; then in the second category, in order to simplify the situation, always less factors considered.

3) Then when it comes to a single factor, the models in supply chain circumstances always consider easier situation of the same factor. Take demand for example, Many studies in the second category especially those under the circumstances of multi-echelon supply chain involve constant demand rate in the inventory model. Only few of the studies in the second category take stochastic demand into consideration for it is really hard to deal with in these kinds of models. For the first category, constant demand is the focus of the early stage and now more and more studies begin to research other kinds of demand, such as time-dependent demand, inventory level- dependent demand, price-dependent demand, and even stochastic demand.

4) Deteriorating rate is another key factor in the study of deteriorating items inventory and there are also several kinds of deteriorating rate in the present study, such as constant deteriorating rate, deteriorating rate is a linear increasing function of time, deteriorating rate is two- parameter Weibull distributed, and deteriorating rate is three-parameter Weibull distributed. Among them, constant deteriorating rate is the easiest one and the three-parameter Weibull distribution deteriorating rate is more complex. Some studies which belong to the first category have made extensive study in this factor and some studies even involve the three-parameter Weibull distribution deteriorating rate; the relatively complex one in the model. For the second category, constant deteriorating rate is the most common one in model establishment.

5.2 Future Directions of the Research

1) As we all know, in order to respond quickly to the diversity and uncertainty in the market, all the enterprises should be part of a supply chain and so it is more reasonable to consider inventory problems in a supply chain. The diversity and uncertainty that a supply chain will face are much more complex than a single enterprise so more factors and more complex situation of a single factor should

be taken into consideration of deteriorating inventory problems in a supply chain. Only in this way, can the studies supply practical measurements to managers.

2) In practical situation, the information about inventory is not always precise, most of the time it is vague or imprecise. So it is more reasonable to develop some fuzzy, stochastic and dynamic research methods and this is also the future trend of the deteriorating inventory study.

3) Luckily, many researchers have studied inventory under fuzzy circumstances. Katagiri and Ishii [125] established inventory models for deteriorating items with fuzzy shortage cost and due to the fuzziness of relative parameters; the expected profit function in the study is represented with a fuzzy set. Chen and Ouyang [126] established a fuzzy inventory model for deteriorating items with permissible delay in payment. The study involved the fuzzy carrying cost rate, fuzzy interest paid rate, and fuzzy interest earned rate simultaneously in the model. With the assumption of constant demand and deteriorating rate, Arindam Roy, Samarjit Kar and Manoranjan Maiti [127], Arindam Roy, Manas Kumar Maiti, Samarjit Kar and Manoranjan Maiti [128] established an EOQ model which involved imprecise parameters, such as fuzzy inventory cost, fuzzy storage area, and so on. The purpose of their research was to minimize the total inventory cost.

It is recommended that more and more researchers begin to study deteriorating items inventory problems in supply chain with fuzzy, stochastic and dynamic research methods, only in this way, can the researches be applied to practice. It is hoped that this paper can provide an overview of the deteriorating items inventory in the recent years and so act as a cornerstone for future study in this field.

6. Acknowledgements

This research was supported by The Research on Process Optimization of Cold Chain Logistics Based on Food Safety of Foundation of Beijing Jiaotong University (2007XM018).

REFERENCES

[1] H. M. Wee, “Economic production lot size model for deteriorating items with partial back ordering [J],” Compute~& Industrial Engineering, Vol. 24, No. 3, pp. 449–458, 1993

[2] T. M. Whitin, “Theory of inventory management,” Princeton University Press, Princeton, NJ, pp. 62–72, 1957

[3] P. M. Ghare and G. P. Schrader, “A model for an exponentially decaying inventory [J],” Journal of Industrial Engineering, Vol. 14, No. 5, 1963.

[4] F. Raafat, survey of literature on continuously deteriorating inventory model [J],” Journal of the Operational Research Society, Vol. 42, No. 1, pp. 27–37, 1991.

[5] S. K. Goyal and B. C. Giri, “Recent trends in modeling of deteriorating inventory [J],” European Journal of Operational Research, Vol. 134, No. 1, pp. 1–16, 2001

[6] K.-J. Chung and C.-N. Lin, “Optimal inventory replenishment models for deteriorating items taking account of time discounting [J],” Computers & Operations Research, Vol. 28, No. 1, pp. 67–83, 2001.

[7] L. Benkherouf, A. Boumenir, and L. Aggoun, “A diffusion inventory model for deteriorating items [J],” Applied Mathematics and Computation, Vol. 138, No. 1, pp. 21–39, 2003.

[8] B. C. Giri, T. Chakrabarty, and K. S. Chaudhuri, “A note on a lot sizing heuristic for deteriorating items with time-varying demands and shortages [J],” Computers & Operations Research, Vol. 27, No. 6, pp. 495–505, 2000.

[9] J.-T. Teng, H.-J. Chang, C.-Y. Dye, and C.-H. Hung, “An optimal replenishment policy for deteriorating items with time-varying demand and partial backlogging [J],” Operations Research Letters, Vol. 30, pp. 387–393, 2002.

[10] K.-J. Chung, P. Chu, and S.-P. Lan, “A note on EOQ models for deteriorating items under stock dependent selling rate [J],” European Journal of Operational Research, Vol. 124, pp. 550–559, 2000

[11] B. C. Giri and K. S. Chaudhuri, “Deterministic models of perishable inventory with stock-dependent demand rate and nonlinear holding cost [J],” European Journal of Operational Research, Vol. 105, No. 3, pp. 467–474, 16 March 1998.

[12] D. K. Bhattachaya, “On multi-item inventory [J],” European Journal of Operational Research, Vol. 162, pp. 786–791, 2005.

[13] K.-S. Wu, L.-Y. Ouyang, and C.-T. Yang, “An optimal replenishment policy for non-instantaneous deteriorating items with stock-dependent demand and partial backlogging [J],” International Journal of Production Economics, Vol. 101, pp. 369–384, 2006.

[14] H.-M. Wee and S.-T. Law, “Economic production lot size for deteriorating items taking account of the time-value of money [J],” Computers & Operations Research, Vol. 26, pp. 545–558, 1999.

[15] R. M. Hill, “Inventory model for increasing demand followed by level demand,” Journal of the Operational Research Society, Vol. 46, pp. 1250–1259, 1995.

[16] B. Mandal and A. K. Pal, “Order level inventory system with ramp type demand rate for deteriorating items,” Journal of Interdisciplinary Mathematics, No. 1, pp. 49–66. 1998.

[17] P. S. H. Deng, R. H.-J. Lin, and P. Chu, “A note on the inventory models for deteriorating items with ramp type demand rate [J],” European Journal of Operational Research, Vol. 178, No. 1, pp. 112–120, 2007.

[18] K. Skouri, I. Konstantaras, S. Papachristos, and I. Ganas, “Inventory models with ramp type demand rate, partial backlogging and Weibull deterioration rate [J],” European

Journal of Operational Research, 2007.

[19] Y. K. Shah and M. C. Jaiswal, "An order-level inventory model for a system with constant rate of deterioration," Opsearch, Vol. 14, pp. 174–184, 1977.

[20] S. P. Aggarwal, "A note on an order-level inventory model for a system with constant rate of deterioration," Opsearch, Vol. 15, pp. 184–187, 1978.

[21] G. Padmanabhana and P. Vratb, "EOQ models for perishable items under stock dependent selling rate [J]," European Journal of Operational Research, Vol. 86, No. 2, pp. 281–292, 19 October 1995,

[22] A. K. Bhunia and M. Maiti, "An inventory model of deteriorating items with lot-size dependent replenishment cost and a linear trend in demand [J]," Applied Mathematical Modelling, Vol. 23, pp. 301–308, 1999.

[23] A. K. Bhunia and M. Maiti, "Deterministic inventory model for deteriorating items with finite rate of replenishment dependent on inventory level [J]," Computers & Operations Research, Vol. 25, No. 11, pp. 997–1006, 1 November 1998.

[24] S. Mukhopadhyay, R. N. Mukherjee, and K. S. Chaudhuri, "Joint pricing and ordering policy for a deteriorating inventory [J]," Computers & Industrial Engineering, Vol. 47, pp. 339–349, 2004.

[25] H.-M. Wee, "Deteriorating inventory model with quantity discount, pricing and partial backordering [J]," International Journal of Production Economics, Vol. 59, pp. 511–518, 1999

[26] N. K. Mahapatra, "Decision process for multiobjective, multi-item production-inventory system via interactive fuzzy satisficing technique [J]," Computers and Mathematics with Applications, Vol. 49, pp. 805–821, 2005.

[27] T. Chakrabarty, B. C. Giri, and K. S. Chaudhuri, "An EOQ model for items with weibull distribution deterioration, shortages and trended demand: an extension of philip's model [J]," Computers & Operations Research, Vol. 25, No. 7–8, pp. 649–657, July 1998.

[28] P. L. Abad, "Optimal price and order size for a reseller under partial backordering [J]," Computers & Operations Research, Vol. 28, pp. 53–65, 2001.

[29] F. W. Gu and S. Y. Zhou, "A replenishment policy for deteriorating items which are permissible delay in payment [J]," Logistics Technology, Vol. 26, No. 9, pp. 48–51, 2007.

[30] M. K. Salameh, S. A. Fakhreddine, and N. Noueihed, "Effect of deteriorating items on the instantaneous replenishment model with backlogging [J]," Computers & Industrial Engineering, Vol. 37, pp. 261–264, 1999.

[31] Y.-W. Zhou and H.-S. Lau, "An economic lot-size model for deteriorating items with lot-size dependent replenishment cost and time-varying demand [J]," Applied Mathematical Modelling, Vol. 24, pp. 761–770, 2000.

[32] Y.-W. Zhou, H.-S. Lau, and S.-L. Yang, "A new variable production scheduling strategy for deteriorating items with time-varying demand and partial lost sale [J]," Computers & Operations Research, Vol. 30, pp. 1753–1776, 2003.

[33] K. Skouri and S. Papachristos, "A continuous review inventory model, with deteriorating items, time-varying demand, linear replenishment cost, partially time-varying backlogging [J]," Applied Mathematical Modelling, Vol. 26, pp. 603–617, 2002.

[34] S.-P. Wang, "An inventory replenishment policy for deteriorating items with shortages and partial backlogging [J]," Computers & Operations Research, Vol. 29, pp. 2043–2051, 2002.

[35] C.-Y. Dye, "Joint pricing and ordering policy for a deteriorating inventory with partial backlogging [J]," Omega, Vol. 35, pp. 184–189, 2007.

[36] A. M. M. Jamal, B. R. Sarker, and S. J. Wang, "Optimal payment time for a retailer under permitted delay of payment by the wholesaler [J]," International Journal of Production Economics, Vol. 66, pp. 59–66, 2000.

[37] X. P. Song and X. Q. Cai, "On optimal payment time for retailer under permitted delay of payment by the wholesaler [J]," International Journal of Production Economics, Vol. 103, pp. 246–251, 2006.

[38] L.-Y. Ouyang, K.-S. Wu, and C.-T. Yang, "A study on an inventory model for non-instantaneous deteriorating items with permissible delay in payments [J]," Computers & Industrial Engineering, Vol. 51, pp. 637–651, 2006.

[39] K.-J. Chung and J.-J. Liao, "Lot-sizing decisions under trade credit depending on the ordering quantity [J]," Computers & Operations Research, Vol. 31, pp. 909–928, 2004.

[40] K.-J. Chung and J.-J. Liao, "The optimal ordering policy in a DCF analysis for deteriorating items when trade credit depends on the order quantity [J]," International Journal of Production Economics, Vol. 100, pp. 116–130, 2006.

[41] L.-Y. Ouyang, J.-T. Teng, S. K. Goyal, and C.-T. Yang, "An economic order quantity model for deteriorating items with partially permissible delay in payments linked to order quantity [J]," European Journal of Operational Research, 2008.

[42] C. Zhang, G.-X. DAI, G.-H. Han, and M. Li, "Study on optimal inventory model for deteriorating items based on linear trade credit [J]," Journal of Qingdao University (Natural Science Edition), Vol. 20, No. 3, pp. 70–74, 2007.

[43] X. W. Wen and L.-J. Zuo, "A review of development of deteriorating inventory theory research [J]," Journal of South China Agricultural University: Social Science Edition, Vol. 6, No. 4, pp. 46–52, 2007.

[44] H.-C. Liao, C.-H. Tsai, and C.-T. Su, "An inventory model with deteriorating items under inflation when a delay in payments is permissible [J]," International Journal of Production Economics, Vol. 63, pp. 207–214, 2000.

[45] K.-J. Chung and S.-F. Tsai, "Inventory systems for deteriorating items with shortages and a linear trend in demand-taking account of time value [J]," Computers & Operations Research, Vol. 28, pp. 915–934, 2001.

[46] J.-M. Chen, "An inventory model for deteriorating items with time-proportional demand and shortages under inflation and time discounting [J]," International Journal of

Production Economics, Vol. 55, No. 1, pp. 21–30, 10 June 1998,

[47] I. Moon, B. C. Giri, and B. Ko, "Economic order quantity models for ameliorating/deteriorating items under inflation and time discounting [J]," European Journal of Operational Research, Vol. 162, pp. 773–785, 2005.

[48] B. R. Sarker, A. M. M. Jamal, and S. J. Wang, "Supply chain models for perishable products under inflation and permissible delay in payment [J]," Vol. 27, pp. 59–75, 2000.

[49] C.-T. Chang, "An EOQ model with deteriorating items under inflation when supplier credits linked to order quantity [J]," International Journal of Production Economics, Vol. 88, pp. 307–316, 2004.

[50] K.-J. Chung and J.-J. Liao, "The optimal ordering policy in a DCF analysis for deteriorating items when trade credit depends on the order quantity [J]," International Journal of Production Economics, Vol. 100, pp. 116–130, 2006.

[51] K.-N. Huang and J.-J. Liao, "A simple method to locate the optimal solution for exponentially deteriorating items under trade credit financing [J]," Computers and Mathematics with Applications, Vol. 56, No. 4, pp. 965–977, August 2008.

[52] C. Zhang, G.-X. Dai, G.-H. Han, and M. Li, "Study on inventory model for deteriorating items based on trade credit and cash discount [J]," Operations Research and Management Science, Vol. 16, No. 6, pp. 33–37,41, 2007.

[53] S.-D. Wang and J.-P. Wang, "A multi-stage optimal inventory model for deteriorating items by considering time value and inflation rate [J]," Operations Research and Management Science, Vol. 14, No. 6, pp. 142–148, 2005.

[54] S. Panda, S. Senapati, and M. Basu, "Optimal replenishment policy for perishable seasonal products in a season with ramp-type time dependent demand [J]," Computers & Industrial Engineering, Vol. 54, pp. 301–314, 2008.

[55] S. Papachristos and K. Skouri, "An optimal replenishment policy for deteriorating items with time-varying demand and partial-exponential type-backlogging [J]," Operations Researh Letters, Vol. 27, pp. 175–184, 2000.

[56] P. Chu and P. S. Chen, "A note on inventory replenishment policies for deteriorating items in an exponentially declining market [J]," Computers & Operations Research, Vol. 29, pp. 1827–1842, 2002.

[57] S. Khanra and K. S. Chaudhuri, "A note on an order-level inventory model for a deteriorating item with time-dependent quadratic demand [J]," Computers & Operations Research, Vol. 30, pp. 1901–1916, 2003.

[58] "A comparison among various partial backlogging inventory lot-size models for deteriorating items on the basis of maximum profit [J]," International Journal of Production Economics, Vol. 96, pp. 119–128, 2005.

[59] C.-Y. Dye, H.-J. Chang, and J.-T. Teng, "A deteriorating inventory model with time-varying demand and shortage-dependent partial backlogging [J]," European Journal of Operational Research, Vol. 172, pp. 417–429, 2006.

[60] R. Gupta and P. Vrat, "Inventory model with multi-items under constraint systems for stock dependent consumption rate [J]," Operations Research, Vol. 24, pp. 41–42.

[61] B. N. Mandal and S. Phaujdar, "An inventory model for deteriorating items and stock-dependent consumption rate [J]," Journal of the Operational Research Society, Vol. 40, pp. 483–488, 1989.

[62] G. Padmanabhan and P. Vrat, "EOQ models for perishable items under stock dependent selling rate [J]," European Journal of Operational Research, Vol. 86, No. 2, pp. 281–292, 19 October 1995.

[63] C.-Y. Dye and L.-Y. Ouyang, "An EOQ model for perishable items under stock-dependent selling rate and time-dependent partial backlogging [J]," European Journal of Operational Research, Vol. 163, pp. 776–783, 2005.

[64] K.-L. Hou, "An inventory model for deteriorating items with stock-dependent consumption rate and shortages under inflation and time discounting [J]," European Journal of Operational Research, Vol. 168, pp. 463–474, 2006.

[65] Z.-H. Peng and T. Peng, "Pricing and inventory model based on quantity discounts of deteriorating goods [J]," Journal of University of Shanghai For Science and Technology, Vol. 26, No. 6, pp. 565–568,574, 2004.

[66] J.-T. Teng, C.-T. Chang, and S. K. Goyal, "Optimal pricing and ordering policy under permissible delay in payments [J]," International Journal of Production Economics, Vol. 97, pp. 121–129, 2005.

[67] Y.-Y. Qin, H.-W. Tang, and C.-H. Guo, "On an inventory model for deteriorating items of price-sensitive demand and quantity discounts [J]," Operations Research and Management Science, Vol. 15, No. 4, pp. 22–26, 2006.

[68] C. K. Jaggi, K. K. Aggarwal, and S. K. Goel, "Optimal orser policy for deteriorating items with inflation induced demand [J]," International Journal of Production Economics, Vol. 103, pp. 707–714, 2006.

[69] Z. T. Balkhi and L. Benkherouf, "On an inventory model for deteriorating items with stock dependent and time-varying demand rates [J]," Computers & Operations Research, Vol. 31, pp. 223–240, 2004.

[70] A. K. Pal, A. K. Bhunia, and R. N. Mukherjee, "Optimal lot size model for deteriorating items with demand rate dependent on displayed stock level (DSL) and partial backordering [J]," European Journal of Operational Research, Vol. 175, pp. 977–991, 2006.

[71] P.-H. Hsu, H. M. Wee, and H.-M. Teng, "Optimal ordering decision for deteriorating items with expiration date and uncertain lead time [J]," Computers & Industrial Engineering, Vol. 52, pp. 448–458, 2007.

[72] Z.-L. Wen, Y. Zhu, J. Li, and C.-M. Xu, "An inventory model for deteriorating items with backorder dependent lost rate and time value of money [J]," Journal of Guangxi University of Technology, Vol. 18, No. 3, pp. 28–32, 2007.

[73] H.-M. Wee and S.-T. Law, "Replenishment and pricing policy for deteriorating items taking into account the time-value of money [J]," International Journal of Production Economics, Vol. 71, pp. 213–220, 2001.

[74] S. Papachristos and K. Skouri, "An inventory model with deteriorating items, quantity discount, pricing and time-dependent partial backlogging [J]," International Journal of Production Economics, Vol. 83, pp. 247–256, 2003.

[75] C.-Y. Dye, T.-P. Hsieh, and L.-Y. Ouyang, "Determining optimal selling price and lot size with a varying rate of deterioration and exponential partial backlogging [J]," European Journal of Operational Research, Vol. 181, pp. 668–678, 2007.

[76] C.-Y. Dye, L.-Y. Ouyang, and T.-P. Hsieh, "Inventory and pricing strategies for deteriorating items with shortages: A discounted cash flow approach [J]," Computers & Industrial Engineering, Vol. 52, pp. 29–40, 2007.

[77] C.-T. Chang, L.-Y. Ouyang, and J.-T. Teng, "An EOQ model for deteriorating items under supplier credit linked to ordering quantity [J]," Applied Mathematical Modelling, Vol. 27, pp. 983–996, 2003.

[78] J.-J. Liao, "A note on an EOQ model for deteriorating items under supplier credit linked to ordering quantity [J]," Applied Mathematical Modelling, Vol. 31, pp. 1690–1699, 2007.

[79] W.-L. Huang and S. Huang, "An inventory model for items with weibull distribution deterioration rate under stock-dependent demand [J]," Industrial Engineering and Management, Vol. 2, pp. 72–75,101, 2007.

[80] S. Kalpakam and S. Shanthi, "A perishable system with modified base stock policy and random supply quantity [J]," Vol. 39, pp. 79–89, 2000.

[81] S. Kalpakam and S. Shanthi, "A perishable inventory system with modified(S-1,S)policy and arbitrary processing times [J]," Computers & Operations Research, Vol. 28, pp. 453–471, 2001.

[82] G.-P. Zhu, "Optimal inventory model for single perishable item [J]," Journal of Ningxia University (Natural Science Edition), Vol. 22, No. 1, pp. 15–16, 2001.

[83] L.-F. Li, P.-Q. Huang, and J.-W. Luo, "A study of inventory management for deteriorating items [J]," Systems Engineering, Vol. 22, No. 3, pp. 25–30, 2004.

[84] X.-H. Xu and R.-J. Li, "A two-warehouse inventory model for deteriorating items with time-dependent demand [J]," Logistics Technology, No. 1, pp. 37–40, 2006.

[85] Z. L. Yang, Z. Y. Wu, and S. L. Yang, "An economic model and its optimal solution for deteriorating items with two warehouses and parabollic demand [J]," Systems Engineering Theory Methodology Applications, No. 4, pp. 376–378, 2005.

[86] S. L. Yang and Y. W. Zhou, "Two-warehouse inventory model: Considering time-varying demand and price discounts [J]," Journal of Systems Engineering, Vol. 18, No. 6, pp. 498–505, 2003.

[87] L. Benkherouf, "A deterministic order level inventory model for deteriorating items with two storage facilities [J]," International Journal of Production Economics, Vol. 48, No. 2, pp. 167–175, 24 January 1997.

[88] K. V. S. Sarma, "A deterministic order level inventory model for deteriorating items with two storage facilities [J]," European Journal of Operational Research, Vol. 29, No. 1, pp. 70–73, April 1987.

[89] T. P. M. Pakkala and K. K. Achary, "A deterministic inventory model for deteriorating items with two warehouses and finite replenishment rate [J]," European Journal of Operational Research, Vol. 57, No. 1, pp. 71–76, 25 February 1992.

[90] H.-L. Yang, "Two-warehouse inventory models for deteriorating items with shortages under inflation [J]," European Journal of Operational Research, Vol. 157, pp. 344–356, 2004.

[91] H.-L. Yang, "Two-warehouse partial backlogging inventory models for deteriorating items under inflation [J]," International Journal of Production Economics, Vol. 103, pp. 362–370, 2006.

[92] H. Qiu and L. Liang, "A two-warehouse inventory model with delay in payment [J]," Systems Engineering, Vol. 25, No. 2, pp. 46–50, 2007.

[93] X.-H. Xu and R.-J. Li, "A two-warehouse inventory model for deteriorating items with time-dependent demand [J]," Logistics Technology, No. 1, pp. 37–40, 2006.

[94] J. K. Dey, S. K. Mondal, and M. Maiti, "Two storage inventory problem with dynamic demand and interval valued lead-time over finite time horizon under inflation and time-value of money [J]," European Journal of Operational Research, Vol. 185, pp. 170–194, 2008.

[95] M. Rong, N. K. Mahapatra, and M. Maiti, "A two warehouse inventory model for a deteriorating item with partially/fully backlogged shortage and fuzzy lead time [J]," European Journal of Operational Research, Vol. 189, pp. 59–75, 2008.

[96] C. C. Lee, "Two-warehouse inventory model with deterioration under FIFO dispatching policy [J]," European Journal of Operational Research, Vol. 174, pp. 861–873, 2006.

[97] T. P. M. Pakkala and K. K. Achary, "A deterministic inventory model for deteriorating items with two warehouses and finite replenishment rate [J]," European Journal of Operational Research, Vol. 57, No. 1, pp. 71–76, 25 February 1992.

[98] N. Bo and J. X. Xie, "A note on 'Two-warehouse inventory model with deterioration under FIFO dispatching policy' [J]," European Journal of Operational Research, Vol. 190, pp. 571–577, 2008.

[99] S. K. Goyal and A. Gunasekaran, "An integrated production-inventory-marketing model for deteriorating items [J]," Computers & Industrial Engineering, Vol. 28, No. 4, pp. 755–762, October 1995.

[100] D. L. Jiang and W. Du, "A study on two-stage production systems of perishable goods [J]," Journal of Southwest Jiaotong University, Vol. 33, No. 4, pp. 430–435, 1998.

[101] S.-J. Chen, S.-P. Chen, and X.-B. Liu, "A production-inventory model for deteriorating items [J]," Journal of Zhaoqing University, Vol. 23, No. 2, pp. 66–68, 2002.

[102] Z.-J. Gontg and C.-Q. Wang, "A production-inventory arrangement model for deteriorating items in a linear in-

creasing market [J]," Logistics Technology, No. 10, pp. 6–8, 2005.

[103] A. K. Maity, K. Maity, S. Mondal, and M. Maiti, "A chebyshev approximation for solving the optimal production inventory problem of deteriorating multi-item [J]," Mathematical and Computer Modelling, Vol. 45, pp. 149–161, 2007.

[104] J.-J. Liao, "On an EPQ model for deteriorating items under permissible delay in payments [J]," Applied Mathematical Modelling, Vol. 31, pp. 393–403, 2007.

[105] J.-J. Liao, "An EOQ model with noninstantaneous receipt and exponentially deteriorating items under two-level trade credit [J]," International Journal of Production Economics, Vol. 113, pp. 852–861, 2008.

[106] J.-T. Teng and C.-T. Chang, "Economic production quantity models for deteriorating items with price- and stock-dependent demand [J]," Computers & Operations Research, Vol. 32, pp. 297–308, 2005.

[107] H. Lin, X. W. Wen, and Q. L. Da, "Optimal production policy for deteriorating items with backlogging demand in stochastic production process [J]," Journal of Southeast University (Natural Science Edition), Vol. 37, No. 4, pp. 731–736, 2007.

[108] Y.-W. Zhou, H.-S. Lau, and S.-L. Yang, "A new variable production scheduling strategy for deteriorating items with time-varying demand and partial lost sale [J]," Computers & Operations Research, Vol. 30, pp. 1753–1776, 2003.

[109] S. Sana, S. K. Goyal, and K. S. Chaudhuri, "A production-inventory model for a deteriorating item with trended demand and shortages [J]," European Journal of Operational Research, Vol. 157, pp. 357–371, 2004.

[110] S.-Y. Zhou and F.-W. Gu, "An optimal production-inventory policy for deteriorating items [J],' Logistics Technology, Vol. 26, No. 4, pp. 43–46, 2007.

[111] K. Skouri and S. Papachristos, "Optimal stopping and restarting production times for an EOQ model with deteriorating items and time-dependent partial backlogging [J]," International Journal of Production Economics, Vol. 81–82, pp. 525–531, 2003.

[112] J.-M. Chen andn L.-T. Chen, "Pricing and production lot-size/scheduling with finite capacity for a deteriorating item over a finite horizon [J]," Computers & Operations Research, Vol. 32, pp. 2801–2819, 2005.

[113] Z. T. Balkhi, "On a finite horizon production lot size inventory model for deteriorating items: An optimal solution [J]," European Journal of Operational Research, Vol. 132, pp. 210–223, 2001.

[114] S. K. Manna and K. S. Chaudhuri, "An EOQ model with ramp type demand rate, time dependent deterioration rate, unit production cost and shortages [J]," European Journal of Operational Research, Vol. 171, pp. 557–566, 2006.

[115] J.-T. Teng, L.-Y. Ouyang, and L.-H. Chen, "A comparison between two pricing and lot-sizing models with partial backlogging and deteriorated items [J]," International Journal of Production Economics, Vol. 105, pp. 190–203, 2007.

[116] P. L. Abad, "Optimal pricing and lot-sizing under conditions of perishability, finite production and partial backordering and lost sale [J]," European Journal of Operational Research, Vol. 144, pp. 677–685, 2003.

[117] S. K. Goyal and B. C. Giri, "The production-inventory problem of a product with time varying demand, production and deterioration rates [J]," European Journal of Operational Research, Vol. 147, pp. 549–557, 2003.

[118] P. C. Yang and H. M. Wee, "A single-vendor and multiple-buyers production -inventory policy for a deteriorating item [J]," European Journal of Operational Research, Vol. 143, pp. 570–581, 2002.

[119] P. C. Yang and H. M. Wee, "An integrated multi-lot-size production inventory model for deteriorating item [J]," Computers & Operations Research, Vol. 30, pp. 671–682, 2003.

[120] L.-F. Li and P.-Q. Huang, "Study of production-inventory cooperation for deteriorating items on supply chain [J]," Journal of Shanghai Jiaotong University, Vol. 39, No. 3, pp. 464–467, 2005.

[121] Q. An and J.-W. Luo, "Study on coordination in supply chain inventory of perishable item with price discount [J]," Journal of Industrial Engineering, Vol. 21, No. 4, pp. 80–84, 2007.

[122] Y. Y. Qin and C. H. Guo, "Discount pricing model for easy-deteriorated items under elastic demand [J]," Chinese Journal of Management, Vol. 4, No. 2, pp. 163–168, 2007.

[123] S.-F. Du, L. Liang, J.-J. Zhang, and Z.-G. Lu, "Hybrid replenishment and dispatching policy with deteriorating item for VMI: Analytical model, optimization and simulation [J]," Chinese Journal of Management Science, Vol. 15, No. 2. pp. 64–69, 2007.

[124] L.-W. Liu and J.-R. Yuan, "Inventory and dispatch Models in VMI systems [J]," Chinese Journal of Management Science, Vol. 11, No. 5, pp. 31–36, 2003.

[125] H. Katagiri and H. Ishii, "Fuzzy inventory problems for perishable commodities [J]," European Journal of Operational Research, Vol. 138, PP. 545–553, 2002.

[126] L.-H. Chen and L.-Y. Ouyang, "Fuzzy inventory model for deteriorating items with permissible delay in payment [J]," Applied Mathematics and Computation, Vol. 182, pp. 711–726, 2006.

[127] A. Roy, S. Kar, and M. Maiti, "A deteriorating multi-item inventory model with fuzzy costs and resources based on two different defuzzification techniques [J]," Applied Mathematical Modelling, Vol. 32, pp. 208–223, 2008.

[128] A. Roy, M. K. Maiti, S. Kar, and M. Maiti, "An inventory model for a deteriorating item with displayed stock dependent demand under fuzzy inflation and time discounting over a random planning horizon [J]," Applied Mathematical Modelling, 2008.

J. Service Science & Management, **2010, 3: 130-137**
doi:10.4236/jssm.2010.31016 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Study of the Instant Incentive Mechanism for Zero-Time Enterprise

Yang Hu[1,2], Xinrong Yang[1]

[1]School of Economics and Management, Hunan Institute of Science and Technology, Yueyang, China; [2]School of Economics and Management, Wuhan University, Wuhan, China.
Email: drhuyang@sina.com

Received December 19th, 2009; revised January 12th, 2010; accepted February 11th, 2010.

ABSTRACT

The zero-time enterprise's mission is response to the market and customer needs in "zero time" in order to gain more customers and competitive advantages. To realize zero-time response to market and customers, it should inspire the employees' work enthusiasm and creativity, so that they can do their work conscientiously to ensure quick and smooth going. This paper discussed how to establish an effective instant incentive mechanism for zero-time enterprise, which is a problem of rich practical significance. Firstly, based on the briefly explanation of zero-time enterprise, it analyzed the characteristics and principles of zero-time enterprise; and then elaborated the important roles of instant incentive. Finally, it proposed the basic contents of instant incentive and the "Trine" realization mechanism.

Keywords*: *Zero-Time Enterprise, Instant, Incentive

1. Introduction

Enterprise competitions and changes in business environment prompting the mode of competition to transfer from price-based competition to quality and variety-based competition, and now to the time-based competition (Time-based competition, TBC). TBC doesn't mean to take no account of price, quality and variety, but under the premise of meeting all these factors, it is to immediately respond to customer needs [1]. Meanwhile, due to customer needs have become personalized and diversified. These factors have posed a challenge to the modern enterprises' operations. They have to strive for customers in order to survive and develop in current rapidly changing time characterized as TBC, the instant response to customer needs of personalized and diversification has become an irrefutable fact. According to the actual requirement that instant response to customer needs, the concept of zero-time [2] has been proposed, and the study of zero-time enterprises begins to be concerned by scholars.

Modern enterprise management is attributed to the management of human resources. It hopes that through the effective deployment of personnel to get immediate response to customer needs, so that realizing the strategic development. Human resources are strategic resources of modern enterprises, but also the most critical factor in the enterprise survival and development, and incentive is an important part of human resources. The most fundamental purpose of incentive mechanism proposed by enterprises is to correctly direct the employees' work motivation, enabling them achieve business goals while achieving their own needs, increasing their satisfaction, so that their enthusiasm and creativity could be maintained and carried forward. This is one of the necessary conditions that immediate response to customer needs of individual and various. In fact, nowadays in enterprises, quite a few of staff would have been able to achieve even greater success, but they did not. This is not due to they lack of skills, but a lack of motivation and passion to work, it is the enterprises lack of immediate and effective incentives for them. Therefore, how to establish an effective incentive mechanism is a very important problem for each enterprise facing and urgent needing to solve. This paper is for zero-time enterprise, to explore the important issue how to establish an effective instant incentive mechanism.

2. The Connotation of Zero-Time Enterprise and Characteristics of Incentives

2.1 The Connotation of Zero-Time Enterprise

The concept of zero-time, firstly proposed by Raymond T. Yeh and Keri Pearlson in 1998, in their paper, "Zero time: A Conceptual Architecture for 21st Century Enterprises". In 2000, Raymond T. Yeh, Keri Pearlson and

George Kozmetsky co-authored of "Zero Time: Providing Instant Customer Value—Every time, All the Time!" [2] in detail discussing the zero-time philosophy and 5 rules. In China, Bosheng Zhou and Dongping Fan (2000) [3] gave an explanation of introduction for the zero-time concept. Briefly, zero-time is a philosophy or a concept, refers to the shortest time applied to respond to customer needs. It is similar to zero-defect for quality control and zero-inventory for production management. The proposal of zero-time is an important contribution to the theory of TBC, it extremely promote the concept of enterprise competition, and gives the most brilliant general of the concept.

The 21st century is a rapid changing era characterized as explosive expansion of information, the changes of competitive environment and affecting competition factors have made the birth of zero-time enterprise to be a necessity. So called zero-time enterprise, in short, refers to the enterprise operated in accordance with zero-time concept and rules, "within the organization, every business process based on response to customer needs, should be implemented immediately, if needed with automatic learning. Managers and employees possess the knowledge and capabilities to make police (decisions) and vendors can provide the required parts and services immediately", its essential characteristic is to instantly respond to customer's personalized requirements [4]. The ultimate goal for zero-time enterprise is not only "enable customer to trust, do as product leader, make operation excellent", also far from being market share, but rather to "permanently occupy the market", that is to say, by providing instant personalized products and services in the existing untapped market to occupy a dominant position, and through the constant release of innovative products to gain market dominance, which is performed as spiral process of innovation, new products, new services and market, is similar to the waves in the sea, constantly alternating emerge, and never cease. Traditional enterprise must overcome some deficiencies firstly to become zero-time enterprise. The reference [3] point that there are 5 kinds of gaps in operations between the traditional enterprise and zero-time enterprise, namely: zero value gaps, zero learning gaps, zero management gaps, zero process gaps and zero inclusion gaps. In the course of business operations, if any one of the 5 gaps was delayed, it would impede the enterprise to make quick response to customer needs and changes of business environment.

This paper is based on the special operation and architecture concept for zero-time enterprise, discusses the issue how to effectively motivate internal staff to instant respond to customer needs.

2.2 The Characteristics and Principles of Zero-Time Enterprise Incentive

The zero-time enterprise's mission is response to the market and customer needs in "zero time" in order to gain more customers and competitive advantages. To respond to the market and customers in zero-time, it must ensure the employees with high enthusiasm and creativity, and have a strong sense of responsibility or urgency, so that they can do their work conscientiously to ensure quick and smooth implementation. All behaviors of employees are in pursuit of some kinds of benefits or avoid some disadvantages, thus physically and psychologically inevitably effects in the corresponding preferences and aversion. Incentive is to promote the common tendency of preferences and aversion in order to promote the sustainable development of enterprise. Of course, this tendency is different for different people, but one thing is in common, that is to be instant and appropriate. To this end, it is necessary to give employees the appropriate incentives, but also to motivate them immediately. Inappropriate incentives, overweight or too light incentives can not really realize the purpose; while the incentives are out of time, too early or too late, it would lose the original meaning of incentives. It follows those instant and appropriate incentives critical. Therefore, the zero-time enterprise is characterized as instantaneous. The appropriation is the core principle of enterprise incentives, manifested in details as follows:

1) The appropriate intensity. Incentives are divided to positive and negative one, positive incentives shall be rewarded, and negative incentives shall be punished. Rewards and punishments will directly influence the incentive effects. The overweight reward will make staff generate the feelings of pride and satisfaction, inducing loss of desire to further enhance themselves; the too light reward will not induce the incentive effects, or to make them get the feeling of not being taken seriously. The excessive punishment will make staffs feel injustice, or loss the identity to enterprise, and even induce the emotions of cacanny or damage; too light punishment will make staff underestimate the seriousness of the error, which may also continue to make the same mistake.

2) The appropriate fair. The appropriate fair is one of the most important principles in staff motivation, any unfair treatment will affect their work efficiency and emotion, so that affect the incentive effects. The staff with the same achievements should receive the same level of incentives; similarly, the staff committing the same mistakes should subject to the same level of punishment. If it is failure to do so, the managers prefer not to award or not to punish.

3) The appropriate time. The core of instant incentives is "quick", as the saying in the ancient book of "Si Ma Art of War", which promoted "never miss the right opportunity to reward, and cannot wait to punish until soldier leaves the group." That emphasized only instant incentives could make people be quickly aware of the results of doing good or bad. "Execute an example of rewards and punishments to a hundred", which generates

the shock and sensational effects, in order to achieve the goal that "found credit of rewards, and retain force of punishments".

Enterprise always strives for benefits, maximum benefits is pursuit, while the employees' maximum achievements itself is the basis for enterprise to maximize benefits. Therefore, managers must grasp the instant character of incentives, so that to maximize employees' achievements. Managers are required to be familiar with the specific characteristics of opportunity in order to identify and apply at all times. One character is that opportunity is always hidden, it is impossible for staffs to expose all their desires, but often to be hidden. Through the tortuous courses and complicated mental activities, reflecting in words, behaviors and expressions, and then it is observed by others. Therefore, it is required that managers should learn to carefully weigh up employees' words and watch their expressions, gain an insight into their mentality. Second, opportunity is transient. The desires do not keep unchanging, but always go with people's needs and social values changes. Third, opportunity is changeable, it is similar to the transient, for example, a person takes more emphasis on material motivation at one time, however, when his family's economic condition has been greatly improved, his needs would be more transferred to the spirit (that is, honor awards). According to these three characteristics, managers can more easily and more accurately grasp the principles of appropriate incentives.

4) The appropriate method. There are several kinds of incentives. It is required to apply appropriate incentive method according to the different objects. The objects and method should be matched to get the incentive effects. As the different requirements of different employees, the same incentives would play a variety of incentive effects. Even the same person at different times or environment, would also have different needs. As the incentive is depended on the internal feelings of staff, the incentives must vary in person. During the process of formulation and implementation incentive method, firstly, it required to make a thorough investigation what are their actual needs. Organize, classify, and then determine the appropriate incentives.

5) The appropriate object. This principle mainly suggests that incentives need to seek truth from facts, requiring the subjected object should be the right person or team who has actually made contribution. In other words, reward those who should be subjected to reward, punish those who should be punished.

6) The appropriate place. This principle mainly proposes that incentives should be based on specific conditions. It can not mechanically carry out the positive incentives and negative incentives. For any positive incentives, reward on the right spot can execute an example of rewards to a hundred; for some negative incentives, punishment on the spot can not execute an example of punishments to a hundred, oppositely, it will bruise staffs' self-respect and enthusiasm.

Generally, instantaneous and appropriation are interrelated and complementary. The core of appropriation principle is required that incentives are consistent with merits and demerits. It is undesirable that neither award is greater than or less than achievements, nor punishment is greater than or less than mistakes, but only appropriate instantaneous and instant appropriation can maximize the effects of incentives. There is moderation for everything, inappropriate control may induce going too far or not getting the expected result, both of which should be avoided in management. In short, it should pay attention to the following six points: First, never reward without any achievements, or never punish without any mistakes; second, cannot lightly reward great achievements, or cannot lightly punish big mistakes; third, cannot grandly reward small achievements, or cannot heavily punish little fault; fourth, the number of incentives should not be too much, nor too low; fifth, can not reward mistakes and punish the achievements; sixth, appropriate incentives should be based on specific situations, including different person, issues and conditions. It can not mechanically carry out rewards and punishments.

3. Instant Incentive Mechanisms for Zero-Time Enterprise

3.1 The Important Roles of Instant Incentives

Indubitably, instant incentive plays an extremely important role in promoting the immediate response to the market and customer needs. As the lack of instant incentives will make employees unable to judge whether they work correctly or not, it is difficult for them to develop good habits, eventually leading to lower efficiency and therefore it would be impossible to immediately respond to the market and customers. Instant incentive is not only an incentive approach, but also a business strategy. Thus its roles are significant, which are as follows:

1) Instant incentive is the enterprise's power system. Under the established business strategy, it can be accelerated by instant incentives. Incentive to employees just like accelerating running car on the highway that it is required to step on it in order to make the car able to accelerate to a specific speed. Similarly, employees also need incentives to encourage them work hard, allowing have full of enthusiasm and strong innovation. If you can not give an immediate incentive, employees will gradually become slack, would be difficult to effectively implement business strategies, and thus it's unable to achieve instant response to customer and market. Therefore, instant incentive is enterprise's power system and accelerator as well.

2) Instant incentive is the correction system for business strategy implementation. Under the established busi-

ness strategy, employees' hard working play important role in achieving business's objectives, thus it's highly necessary to ensure them positively and efficiently throw in to gear. Instant incentives indicate that: What is right and what is wrong. Instant incentive is charged with the correction system functions for business strategy implementation, employees will be more concentrated on work when their acts have been affirmed and strengthened, while some of their acts can be immediately rectified if they did not meet the requirements, so as to ensure the smooth implementation of the strategy.

3) Instant incentive is the activator used to make employees change behaviors. Effective instant incentives can form collective memories, prompt them to develop good habits, and improve work efficiency as well. Every one has the instinct of self-motivation, which can fully satisfy his demands, while stimulate the enthusiasm and drive to improve efficiency. Therefore, the instant incentive can be used to activate staff's passion.

In conclusion, instant incentive has an extremely important role in motivation, however, how to fully play the role of incentives, which required establishing an effective implementation mechanism for instant incentive.

3.2 The Basic Contents and Realization Mechanisms of Instant Incentives

The realization mechanisms of instant incentives are determined by their basic contents, only in accordance with the contents can build practical and concrete realization mechanisms. The basic contents include: understanding the actual needs of employees, formulate reasonable incentive systems, grasp the appropriate opportunity to motivate staff, establish incentive model for the knowledge team based on the psychological contract, and form "Trine" implementation mechanism.

1) Understand the actual needs of employees.

The managers need to really take consideration on the employees' position, to think about their work motivation and hardship paid for work; from their points of view to consider their work and should be paid remuneration and reward issues. For the staff with success on a particular position, certainly there would be other motives and needs other than his daily responsibilities set by organization. For example, he may want to make more money, to travel abroad with his girlfriend, as well as to repay his family; or if he wants to promote himself and to be a manager in order to win the admiration and so on. All these are his motives and needs outside of the daily responsibilities; managers must dig out these motives, so as to truly understand the employees in order to establish instant and effective incentive schemes.

Most of time, people have been accustomed to using tower type "Need-hierarchy theory", which was proposed by Abraham · H · Maslow in the 50 years of 20th century [5], to explain the employee needs and guide the formulation of incentives. In fact, if we change the perspective of the employee needs, perhaps it may be more realistic and more practical. We believe that whether employees can keep mind on work and even work hard, depending on the ability of the business or how to meet employees in the following four aspects: material benefits, career ideals, sense of belonging, and marginal pursuit.

Material benefits: To determine a person whether is willing to work in the enterprise, an important variable is the level of material benefits provided by the enterprise. Obviously, the employees not only want to have accumulation under the premise of feeding themselves and families, but also need access to a cultural-sociological identity, as in a particular social environment, the salary is an important aspect reflecting social status. In addition, quite high proportions of senior human talents tend to judge how much respect and the value of their cognition they would get in the enterprises by the level of material benefits provided by their business. However, material benefits it is not the one and only factor for all the staffs to choose business and to fulfill their responsibilities. Obviously, if the employee is highly devoted to business, but he was not placed in appropriate positions closed to his profession, even if be paid much better material benefits, he could also hardly concentrate on work.

Career ideals: it is the person's wish that he would like to make achievements in his favorite area, being enviable and satisfactory or proud of himself, even if this wish may be just a staged achievement. The person's career ideals are commensurate with his education, working and living environment, personal hobbies, personality and personal pursuit, and related to the particular social values as well. For the person with strong career ideals, the enterprise may not be attracted to him by the high wages, but merely provide him with training opportunities for learning and displaying his talent. For such employee, it's just required to provide him with a platform to achieve personal ambition, he might be satisfied, and then throw himself into work.

Sense of belonging: the sense of belonging is an integrated sense for the employees' own businesses, such as whether they feel security in enterprise, whether it is worthwhile to throw himself into work, with or without a sense of honor, whether be respected, whether there are harmonious interpersonal relations, whether the enterprises treat employees friendly, with or without long-term development and so on. In essence, every one will hope to work in the enterprise with a sense of belonging. Once the staff feels that he is working for the enterprise without any sense of belonging, he would not keep his mind on the work, and also can not effectively complete his own duties.

Marginal pursuit: Someone work in enterprise, simply because of taking the enterprise as springboard, that is to accumulate capitals/create conditions/pass time for his

further new job opportunities, we call this pursuit as "marginal pursuit". In fact, not only the person, who takes enterprise as springboard, is possessed of marginal pursuit, but also all of whom would have the possibility of the similar pursuit. For example, someone may prefer to living in a certain city, so as to choose a company in this city; someone may not change another company just because his current located one is closed to his family; someone may not be attracted by another company's higher wage just because he could access to higher social assessment now. The concept of marginal pursuit is proposed for the meaning that, if companies are good at digging and encourage their staffs to obtain the marginal pursuit, it will be conductive to improving the relationship between staffs and companies, that is to increase the attraction of companies to staffs, and staffs' loyalty to company. For example, if a senior worker cannot keep his mind on work due to the distance from his family, the enterprise could just try to solve the decantation of his wife, which comes up to satisfy his marginal pursuit, so that he could throw himself into work afterwards.

In practice, when company managers try to understand the actual needs of employees, they need to considerate from the vertical and horizontal two angles. From the vertical view, the employees with different levels of knowledge and salaries would also possess different demands. As for the employees with lower salary, it will focus on meeting their material benefits and sense of belonging, that is, to improve their living conditions. As for the employees with higher salary, it should pay more attention to meet their career ideals and marginal pursuit. From the horizontal view, for the employees in same levels, the focuses of their demands are also different because of the differences of their personality and living environments. Someone may highly focus on material benefits due to personality characteristics; someone may much like to be praised, and enjoy the spiritual sense of achievement; and also someone may be good at studying techniques, be possessed of strong learning desires, while not deeply need money. Therefore, it's required to "tailor-made" list of requirements for different employees.

In conclusion, as the staff needs are complex and diversified, it's really not easy to truly understand everyone's needs. The managers need to make more effort, from the vertical and horizontal angles of view, in order to summarize the real demands of each staff, then to find the two or three dominant demands and give targeted incentives for them.

2) Reconstruct the balanced model for organizational structure under zero-time competition.

Management contingency principle indicated that various kinds of management methods are effective, but the most suitable one is the best one. To accord with the fundamental principles of contingency management, it's required to appropriately combine the organization strategic goals, organizational structure and employee conditions, etc., to carry through self-summarization, self-creation, so as to develop a suitable organizational assessment model. TBC environment has broken up the original balance of the organizational structure, which urgently needs to be restructured, through changing control method and control structures. In a new balance, all levels of staffs should be clear about themselves roles and responsibilities. It is encouraged to achieve "free hand" management and bottom-up participation, and it will be applied to a new performance management system, so as to enhance the flexibility of performance management. However, it's useless to just orally appeal "free hand" management and bottom-up participation, but also needs some quick and effective methods. Such as the power down, so that making decisions by the most qualified person, and entrusting the people in different levels with different freedom of making discretions (As for the general small proposals can be assessed by the staff themselves to take action, while the more important proposals are assessed and determined by sector managers, and the major ones are handled by the leaders in charge of the whole department, and so on). Meanwhile, the proposal can also step across the traditional bottom-up approving process, establishing the proposals management department, which makes the acceptance, evaluation, submission and the numbers of proposals adopted by company be included to the assessment of proponents.

3) Construct the integrated evaluation model for performance results and behaviors under zero-time competition.

Usually, the performance evaluation, which is focusing on results, is results-oriented, focusing on work's final performance, and its evaluation contents are mainly concentrated on the actual outcomes; while the performance evaluation, which is process-oriented, is concentrated on employees' behaviors, effort intensity and work attitudes in process of their working. As the working process in Zero-Time Enterprise is much more complicated, the employees' performances have significantly compounded characteristics, it should synthetically evaluate their work process and behavioral outcomes. The performance evaluation for non-procedural worker, it should be result-oriented as far as possible; while for procedural knowledge worker, it can use behavior-oriented approach. If it is uncertain to find the employees' performance results, the outcome itself is not easy to be quantified, the value reflecting is lagging, and it is of strong professionalism and innovation, then its performance evaluation needs to combine the behavior-oriented approach and result-oriented one.

In recent years, objective management, key performance indicators, balanced scorecard and other performance assessment methods become very popular in many

companies, but they often ignore the behavioral anchors titration, behavioral observation scales such as the use of technology for performance evaluation. In fact, from the view of development of performance management in western enterprises, objective management, key performance indicators, balanced scorecard, etc., are the tools for strategic performance assessment, which would be able to connect employee performances to the strategy of the whole organization. However, the tools for strategic performance assessment are inseparable from the performance evaluation methods and technology support. Many performance evaluation techniques such as behavioral anchors titration and behavioral observation scales can directly provide specific behavioral degrees and evaluation scales, for the evaluator to establish a uniform evaluation standard, will not only help managers to make objective evaluation, but also help to guide staffs' performance behaviors.

4) Construct double evaluation model of individual-team under zero-time competition.

The staffs in Zero-Time Enterprise in many cases are in the form of teamwork, the character of this work determines that the performance evaluation can not only be aimed at individual work, but should also evaluate the working conditions of the entire team. If only consider the individual performance indicators, it would induce employees to focus only on the quality of their own work and performance, so as to lack of team spirit, be poor in collaboration, this will reduce the efficiency of the entire team, and even lead to team conflicts and confusion. Therefore, in the design of evaluation indicators, on the one hand that it should consider the individual behavior and work results; on the other hand, it has to consider the work progress, work quality and team cohesion and other factors, to pay attention to the integration and consistency of individual and team indicators. Including the team indicators into the performance evaluation system of knowledge-based staffs, will help to improve the sprint of teamwork, encourage the cooperation among the staffs, so that the team can efficiently complete the intended objectives.

5) Grasp the appropriate incentives opportunity.

Incentives need to follow the principle of appropriate opportunity. Generally, it's required to be accorded with the strongest stage of the staffs' desires of winning award, so as to get the best incentives efficiency. As to the rewards aspect, Marshall who is the famous general in United States, believes that the troop of outstanding performance in the battle should be praised promptly, he said that awards can run immediately, be announced to the press; while clerical work can be followed up, its requirement of filling in a variety of forms would cause time delay, which would reduce the incentives value to the lowest perception. Any delaying ideas and behaviors always lost the incentive role of reward, resulting in empty result of reward.

Of course, the instant incentive is not simply to be done quickly. It mainly reflects the style of speedy and strict enforcement, but not mechanically implement, its premise is based on the right clear and accurate incentives. If the incentive objective is error, and its character is not exact, it would be not only meaningless, but also cause harmful effects. Therefore, the principle of instant incentives, is not for the sake of speed, but must be strictly controlled.

6) Establish incentive model for knowledge-based team based on psychological contract.

Zero-Time Enterprise should be knowledge-based enterprise, the composition of its team members should be knowledge-based workers. Therefore, incentives for knowledge-based team should focus on the achievement and growth of knowledge-based workers. Only if fully encourage their enthusiasm, it would drive them all together to work hard, so as to make team play greater synergistic effect and to create better performances; as to incentive method, it emphasizes the combination of individual incentives, team incentives and organizational incentives; as to incentive time effect, it should combine their short-term incentives and long-term incentives, which stresses the long-term positive effect of incentives for human talent; as to the design for incentive compensation mechanism, it would have to establish incentive model for knowledge-based team based on psychological contract [6].

The incentive model for knowledge-based team based on psychological contract is shown in **Figure 1** (dashed lines represent the feedback lines, while solid lines represent the causal relations).

The incentive model for knowledge-based team based on psychological contract, transfers from the single economic contract-based model to the model combined with psychological contract and economic contract, with particular emphasis on the important role of psychological contract in the incentives.

Firstly, psychological contract for external compensation and benefits is the basis of incentives, which directly influent the effectiveness of material motivation; second, psychological contract for internal reward values is the core of incentives, which can enhance mutual sense of responsibility, sense of duty, self-discipline and self-control, so as to lead values of both sides to convergence, and form intrinsic motivation; third, the psychological contract for work atmosphere is the guarantee for incentives, which can provide more relaxed atmosphere for organization, form a democratic management style with sufficient authority to built interpersonal relationships full of trust, compassion and respect, so as to construct an internal contract system with the guarantee of mutual support; finally, the key to incentives are the recognition

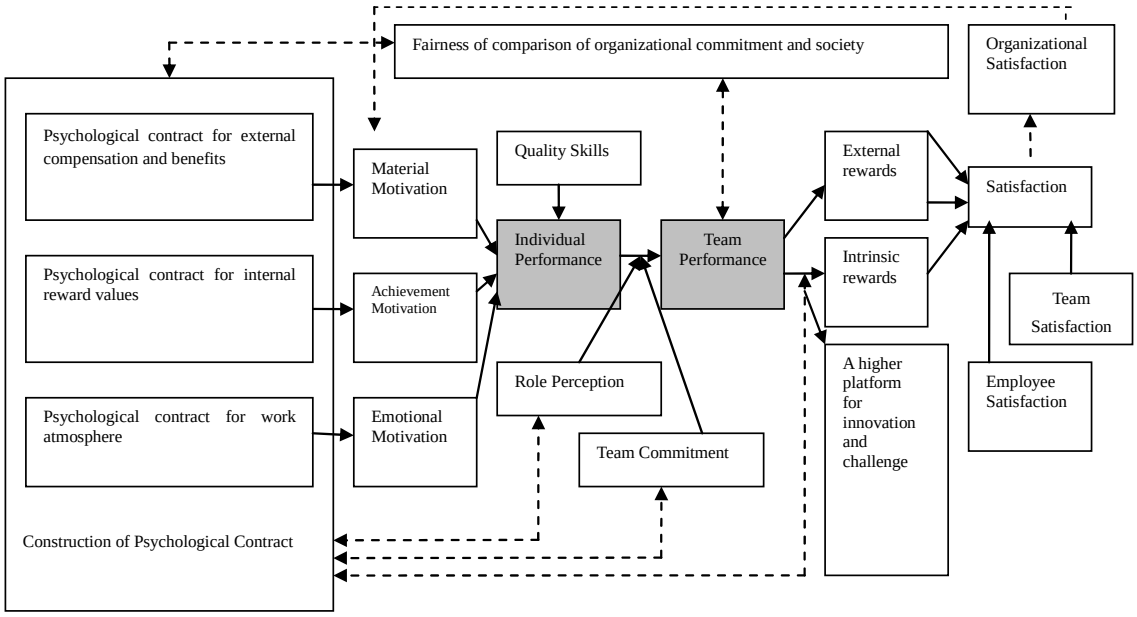

Figure 1. The incentive models for knowledge-based team based on psychological contract

for the employees' value and equitable reward. Knowledge-based worker will pay great attention to the correlation between the values of his own knowledge and skills and his rewards, only with full respect for the value choice of knowledge-based worker and active fulfilling of promise showing a fair, these employees would just take the most advantage of their own talent.

7) Form "Trine" realization mechanism.

Instant incentive mechanism is established based on the immediate characteristics and the appropriate principles. To successfully implement this mechanism, we have to propose the "Trine" realization mechanism, as shown in **Figure 2**. The so-called "triad" represents transposition, orientation and position. First of all, through considering in exchanging positions, managers can understand the actual needs of employees; then, determine the reasonable performance management system according to the actual needs, that is orientation; finally, grasp at the right time to provide instant incentives, that is just in position. As the dynamic demands of employees, it's required to plan the three as a whole. This is the origin of "Trine" realization mechanism. If take instant incentives as a business strategy, the "Trine" realization mechanism is the right process and methods of the strategy implementation.

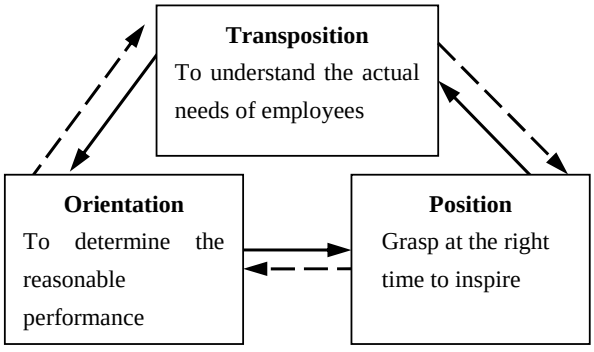

Figure 2. "Trine" realization mechanism

In short, personnel is a very complicated element in enterprise, if want to make them work actively and dutifully, and let the incentives to be the "welcome rain", it's required to spare no efforts to think how to effectively motivate employees. For this reason, instant incentives have already produced good effects in some companies. For example, Cisco attaches great importance to the retention of good staff by material rewards, one of which is "Instant Award". If the staff has exceeded the usual in sales performance, then his master can momentarily reward him within the range of $50 to $2000 USD. In IBM, there is a "Golden Banana Award". One day, a young man walked into IBM founder Tom · Watson's office and told him that he had made a remarkable achievement. Hearing this, Watson was very happy and would like to reward this young man. However, he sought through his desk drawers only to find a banana; therewith he rewarded the boy just with this golden yellow banana. Since then, the "Golden Banana" has become a symbol of the achievements for IBM employees.

4. Conclusions

In modern enterprise management, employee incentive issues have more and more attracted the managers' attentions, and the "incentive" is also one of the key charac-

ters for modern human resources management being different from the traditional one. There is not only regularity, but also strong skill or artistry in incentive management. Managers must be good at deep analyzing and studying the characteristics of enterprise management and the workforce based on specific situations, take the effective and targeted measures, so that employees' internal energy are fully inspired in order to provide powerful driving force for business development. Although the instant incentives were not newly proposed, it is necessary to actively study for effective good use of instant incentive mechanism, and different business practices would be different.

5. Acknowledgements

This paper is supported by China Postdoctoral Science Foundation (supporting ID: 20090460995).

REFERENCES

[1] R. Q. Chen, "Research of new technologies and new methods in operation management time-based competition," (An application report of National Natural Science Foundation), September 2003.

[2] R. T. Yeh, K. Pearlson, and G. Kozmetsky, (translated by D. Q. Tang, W. H. Tang, and H. P. Shao), " Zero time: Providing instant customer value—every time, all the time!" Electronic Industry Press, 6th Edition 2002.

[3] B. S. Zhou and D. P. Fan, "Zero time: The concept of enterprise in 21st century," China Mechanical Engineering, No. 2. 11(1–2): pp. 89–92,113, Ⅻ, 2000.

[4] Y. Hu and G. H. Zhang, "The analysis on performance management modes for zero-time Enterprise," China Human Resources Development, Vol. 12, pp. 26–30, 2009.

[5] H. Mintzberg and Y. Q. Chen, "Strategic skill", IT managers in the world, No. 4, 2004.

[6] N. J. Allen and J. P. Meyer, "The measure meat and antecedents of affective, continuance and normative commitment to the organization," Journal of Occupational Psychology, Vol. 63, pp. 1–18, 1990.

J. Service Science & Management, **2010, 3: 138-142**
doi:10.4236/jssm.2010.31017 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Scientific Research

Service Provision in Knowledge-Based Industries – A Global Study on the Tooling Industry

Günther Schuh, Wolfgang Boos, Fabian Gaus

Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Aachen, Germany.
Email: {g.schuh, w.boos, f.gaus}@wzl.rwth-aachen.de

Received October 2nd, 2009; revised November 15th, 2009; accepted December 24th, 2009.

ABSTRACT

In today's times, more than ever cost competition and high demands of globalized value-adding chains put strong pressure on small- and medium-sized toolmakers. As an exclusive differentiation in price is not an option, new means for achieving sustainable competitive positions have to be found. A promising approach for differentiation is to enhance the existing range of products by offering customer-specific services within so-called industrial product-service-systems. However, the lack of local presence inhibits the toolmakers' abilities to deliver these services to their global customers. To address these challenges the European R&D-project TIPSS has the objective to develop suitable methods, techniques and technologies, for toolmakers to improve their local and global performance thus enabling them to offer industrial product-service-systems on a global scale. The basis of the project is a large scale toolmaker and parts producer survey focusing on service provision. The survey identifies the global footprint of toolmakers regarding the provision of tool-related services. Key findings of the survey with respect to designing industrial product-service-systems are presented in this paper.

Keywords*: *Production Engineering, Product-Service-Systems, Tool & Die Making

1. Introduction

Today's turbulent economic environment confronts the global tooling industry with serious challenges. Cost competition and the high demands of globalized value chains put pressure on small- and medium-sized toolmakers. As a sole focus on price does not lead to a sustainable competitive advantage, toolmakers in high wage economies have to find new means for differentiating themselves from their competitors or else become easily replaceable for their customers [1–3]. One approach for differentiation is the forward integration into the customer's value chain [4,5]. By offering enhanced product related services via so-called industrial product-service-systems (IPS^2) [6], certain aspects of the customer's value-adding chain can be provided by the toolmaker. IPS^2 are hybrid products consisting of the product itself combined with a comprehensive set of product-related services [7–9]. While generating additional cash-flows along the product's entire life-cycle, IPS^2 increase the dependence of the customer towards his toolmaker strengthening the competitive position of the latter.

In order to develop IPS^2 that truly deliver added value, the toolmaker is challenged to acquire a profound knowledge of his customer's needs. Based on this knowledge he can define a portfolio of relevant services and develop the appropriate technological solutions. In the following the concept of IPS^2 is elaborated and, based on an extensive survey, an overview of the capacities and competences of today's toolmakers pertaining to the offering of IPS^2 on a global level is given.

2. Industrial Product-Service-Systems

Linking innovative services to the core product and thereby creating industrial product-service-systems is maybe one of the most promising ways to create more customer value and to gain a true competitive advantage [10–12]. Especially the tooling industry has not yet benefited from this "new manufacturing" that enables producers to profit from higher service profit margins, growth opportunities in mature markets and longer lasting customer relationships.

In the TIPSS project, the concept of industrial prod-

uct-service-systems encompasses the integration of the toolmaker into the customer's value chain. The toolmaker thus becomes an integral part of the customer's production process, increasing the dependence of the customer towards his toolmaker. **Figure 1** illustrates the increasing integration of the toolmaker into the customer's processes as the portfolio of offered services expands. Starting at the core product, the tool, each layer adds another service.

While moving outward in the diagram the degree of connection to the core product decreases, meaning that vertical integration into the customer's processes increases.

The challenge for achieving the optimal added value for both sides is to adequately configure the industrial product-service-system with respect to the service portfolio as well as the technology to enable the service provision. Furthermore offering industrial product-service-systems and making money with them is an issue, which cannot just be carried out on the level of operations. As the success of a company is founded in its business model [13–15], the latter needs to be re-designed to align strategic and operational objectives. Thus offering industrial product-service-systems requires a new business model which addresses the customer's needs by adding value to his processes generating sources of income along the tool's entire life-cycle.

3. Survey Findings

As a base for designing industrial product-service-systems respectively a business model encompassing them, a large scale survey addressing both toolmakers and their customers was conducted. The survey was carried out in two parts, starting with the customers and ending with the toolmakers themselves. In total 278 companies in relevant economies all over the world participated in the survey. The outcome of the survey was on the one hand an evaluation of tool-related services by both toolmakers and parts producers as well as a global footprint of tool makers regarding the provision of these services. The findings referring to the latter are illustrated in the following. The global footprint describes the current structure of the tooling industry regarding four aspects of services provision on a global level:

1) Global customers – Are toolmakers taking advantage of globalized markets?

2) Local on-site presence – How do toolmakers provide services on a global level?

3) Global partnerships – Do toolmakers cooperate to tackle challenges in teams?

4) Global sourcing – Are toolmakers taking advantage of globalized supply markets?

In the following the results of the survey reflecting each of the four dimensions are elaborated.

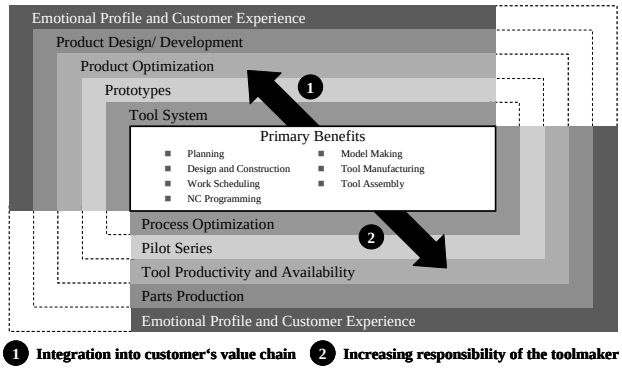

Figure 1. Extended product-services within the TIPSS Business Model

3.1 Global Customers

The first part of the global footprint investigates to which degree toolmakers currently take advantage of globalization by entering new markets to sell their tools. **Figure 2** depicts the structure of the toolmakers' markets in each region.

It can be seen that North American and Western European toolmakers still have a very strong focus on their own region. The regions China/South East Asia, and to a

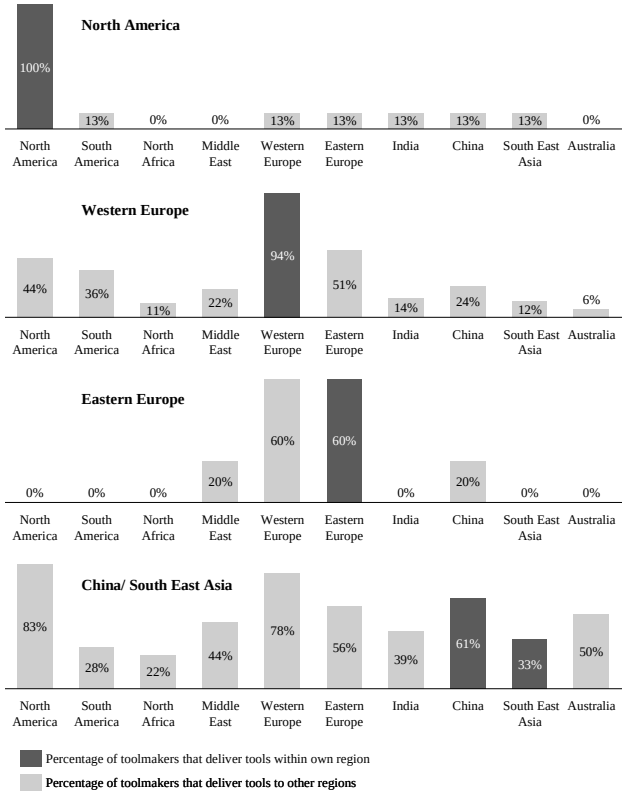

Figure 2. Regions to which toolmakers deliver tools

lesser extent Eastern Europe, have a higher percentage of export to other regions. This shows on the one side, that North America and Western Europe are currently still target markets when it comes to parts production. On the other side it also shows that toolmakers in China/South East Asia and Eastern Europe are using the globalization of markets more consequently for selling their tools.

The customer's point of view is given in **Figure 3**. The illustration shows the ratio of global vs. local purchasing of companies within the different regions. Parts producing companies were asked to rank the top four countries from which they purchase tools. It can be seen that companies in Western Europe and China/South East Asia frequently purchase tools from toolmakers within their own region.

Figure 4 illustrates reasons why parts producing companies are hesitant to buy tools outside their own region:

Customers were asked to rank the importance of several disadvantages they perceive when buying tools from toolmakers in other regions of the world on a one-to-seven scale, seven being the most severe disadvantage.

The evaluation shows, that a lack of services was ranked just after quality issues and delivery time, clearly implicating the importance for a toolmaker to be able to deliver services on a global level.

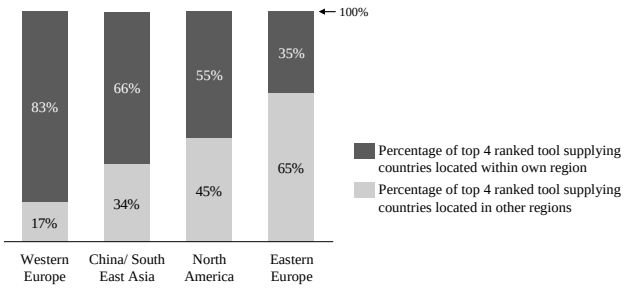

Figure 3. Local vs. global purchasing of tools

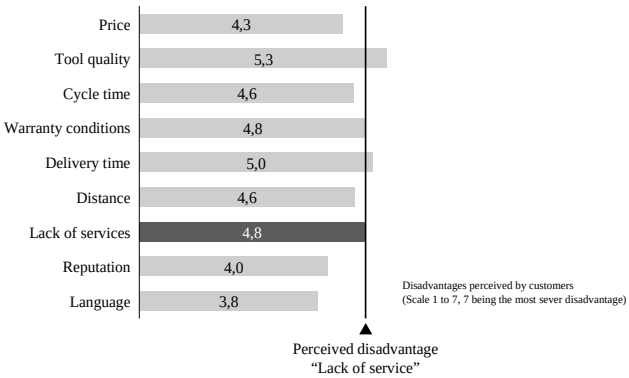

Figure 4. Disadvantages perceived by customers when purchasing tools on a global level

3.2 Local on-Site Presence

To draw a picture of toolmakers' on-site presence, parts producer were asked to estimate the percentage of maintenance work that was conducted by their toolmaker at their production site. **Figure 5** shows that a customer's toolmaker only performs tool related services locally at the customer's production site in rare cases.

This either means that tools are shipped back to the toolmaker for maintenance or that the customer performs maintenance with his on capacities. **Figure 6** proves that indeed a large part of the required maintenance work is performed by the customers themselves. Three quarters of the customers state that they conduct more than 50 percent of the maintenance with their own capacities.

3.3 Global Partnerships

The third perspective of the global footprint investigates partnerships and other relationships between toolmakers in terms of vertical or horizontal integration into each other's processes. Toolmakers where asked whether or not they do have any partnerships regarding the provision of services.

Figure 7 shows that a large part of the toolmakers, especially in China/South East Asia, is currently not cooperating with other toolmakers for improving service provision (left). The right side of the figure shows the reasons why toolmakers choose not to cooperate. It sticks out that while North American and Western European toolmakers are worried about the protection of their know-how, Chinese/South East Asian toolmakers simply

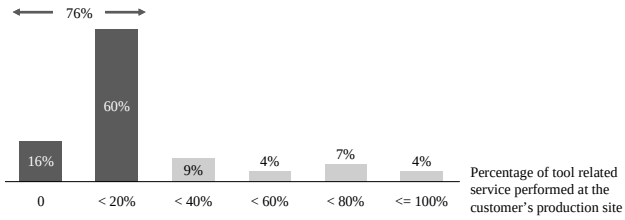

Figure 5. On-site service provision (customer response)

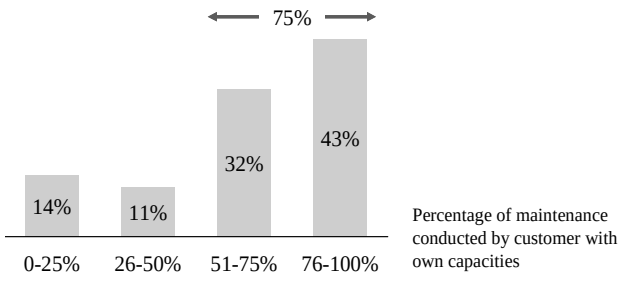

Figure 6. Maintenance service performed by customer

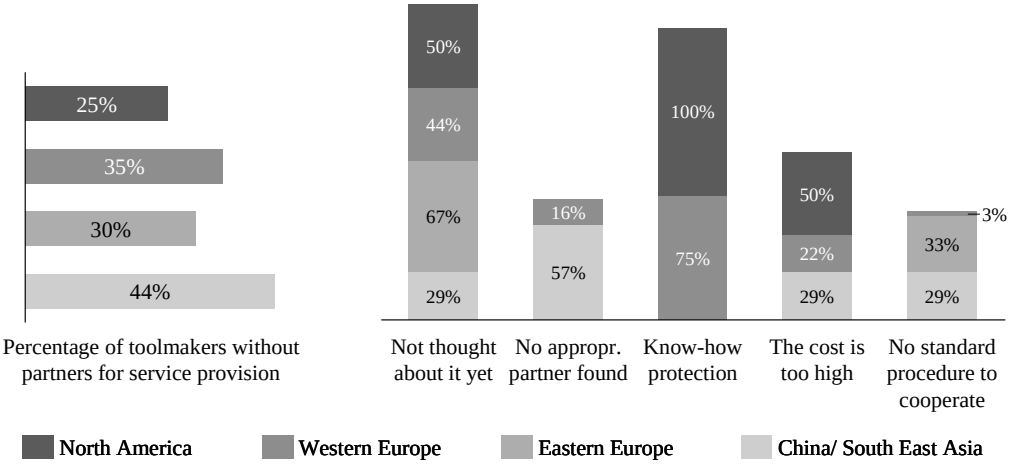

Figure 7. Percentage of toolmakers without service partners (left), reasons for not having service partners (right)

have trouble finding appropriate partners.

3.4 Global Sourcing

The last perspective of the global footprint focuses on the ability of the toolmaker to use globalized markets for sourcing. **Figure 8** list for each region, which part of the value chain is externalized to the global markets. It can be seen that standard tool parts are the largest part of global sourcing. It seems that while Eastern European and Chinese/South East Asian Toolmakers have a stronger focus on selling their tools in other regions than their Northern American and Western European counterparts, they also go further when it comes to opening up their value chain to suppliers from other regions. Furthermore the graphic shows that while standard tool parts are commonly sourced globally, service are still mostly procured locally.

Toolmakers were asked to list advantages and disadvantages they perceive when sourcing globally (**Figure 9**). While North American, Western and Eastern European toolmakers mostly see an advantage in price, Chinese/ South East Asian toolmakers source globally because of quality, reliability, reputation, warranty conditions and, surprisingly, delivery time.

4. Conclusions and Outlooks

The analysis of the results showed that only certain aspects of the challenges regarding global service provision are being adequately addressed by toolmakers in general. When distinguishing between the different regions of the world, it becomes clear that especially emerging markets like China/South East Asia and Eastern Europe are open to take advantage of opportunities like global purchasing and exporting of tools. However, local on-site presence as well as partnerships between toolmakers for supplying

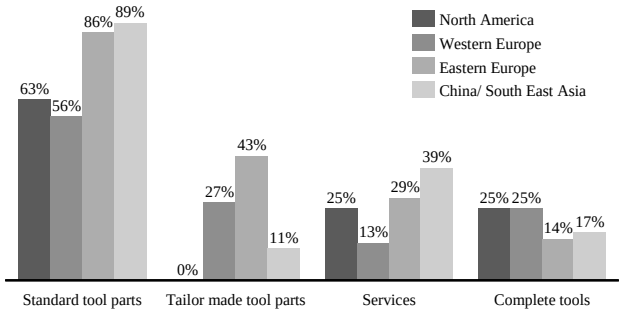

Figure 8. Parts purchased by toolmakers in foreign countries by region

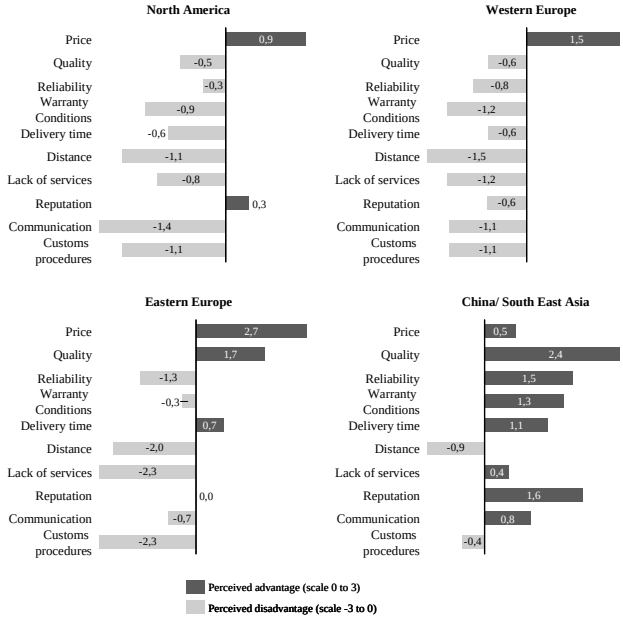

Figure 9. Advantages and disadvantages of global sourcing perceived by toolmakers

services are currently still on a very low level in all regions of the world. In general the results show, that cooperation among toolmakers is still not common and their ability to deliver tool related services on a global level is not sufficient for the current demand for such services.

The results show that in order to be able to offer IPS^2, toolmakers have to adapt new business models, which focus on four major topics:

1) Service provision
2) Cooperation with partners
3) Customer integration
4) Strategies for identifying relevant customer needs

Currently each of these four topics is being addressed insufficiently by toolmakers on the whole. Especially smaller toolmakers that do not have the capacities to offer adequate services on a global level will have to adjust their business models accordingly. Only through close cooperation with other toolmakers as well as their customers will they be able to strengthen their position in the vast competition of the globalized markets.

The development of an appropriate business model for toolmakers is currently being addressed within the FP7 project TIPSS. Further information on the project as well as the complete evaluation of the TIPSS toolmaking survey can be found on the project's website (www.tipss-fp7.eu).

5. Acknowledgements

The research and development project "TIPSS" (Tools for Innovative Product-Product-Service Systems in Global Tool and Die Networks) is co-funded by the European Commission within the Seventh Framework Programme (FP7).

REFERENCES

[1] T. Ittner and J. Wüllenweber, "Tough times for toolmakers," The McKinsey Quarterly, Vol. 2, pp. 14–16, 2004.

[2] G. Schuh, C. Klotzbach, and F. Gaus, "Business models for technology-supported, production-related services of the tool and die industry," In S. Takata, and Y. Umeda, (eds.), Advances in Life Cycle Engineering for Sustainable Manufacturing Businesses, Springer, London, pp. 207–211, 2007.

[3] G. Schuh, C. Klotzbach, and F. Gaus, "Service provision as a sub-model of modern business models for tool and die companies," Proceedings of the 15th CIRP Conference on Life Cycle Engineering, Vol. 2, No. 1, pp. 570–574, 2008.

[4] S. L. Vargo, "Customer integration and value creation," Journal of Service Research, Vol. 11, No. 2, pp. 211–215, 2008.

[5] C. W. Lamb, J. F. Hair, and C. McDaniel, "Marketing, 10th edition," Cengage Learning, Mason, 2008.

[6] M. Lindahl and G. Ölundh, "The meaning of functional sales," Proceedings of the 8th International Seminar on Life Cycle Engineering, pp. 211–220, 2001.

[7] J. F. Castellano, K. Rosenzweig, and H. A. Roehm, "How corporate culture impacts unethical distortion of financial numbers," Management Accounting Quarterly, Vol. 5, No. 4, pp. 37–41, 2004.

[8] R. Lifset, "Moving from product to services," Journal of Industrial Ecology, Vol. 4, No. 1, pp. 1–2, 2000.

[9] V. Salminen and P. Kalliokoski, "Challenges of industrial service business development," In B. Hefley, and W. Murphy, (eds.), Service Science, Management and Engineering, Springer, New York, pp. 41–48, 2007.

[10] C. van Halen, C. Vezzoli, and R. Wimmer, "Methodology for product service system innovation," Assen van Gorcum, 2005.

[11] O. Mont, "Clarifying the concept of product-service system," Journal of Cleaner Production, Vol. 10, No. 3, pp. 237–245, 2002.

[12] A. Tan, T. C. McAloone, and C. Gall, "Product/service-system development: An explorative case study in a manufacturing company," Proceedings of the 16th International Conference on Engineering Design, 2007.

[13] H. I. Ansoff and P. A. Sullivan, "Optimizing profitability in turbulent environments: A formula for strategic success," Long Range Planning, Vol. 26, No. 5, pp. 11–24, 1993.

[14] H. Mintzberg and J. Lampel, "Reflecting on the strategy process," Sloan Management Review, Vol. 40, No. 3, pp. 21–30, 1999.

[15] G. Schuh, C. Klotzbach, and F. Gaus, "Service provision as a sub-model of modern business models," Production Engineering Research and Development, Vol. 1, No. 2, pp. 79–84, 2008.

J. Service Science & Management, **2010, 3: 143-149**
doi:10.4236/jssm.2010.31018 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Allocating Collaborative Profit in Less-than-Truckload Carrier Alliance

Peng Liu[1], Yaohua Wu[1]*, Na Xu[2]

[1]School of Control Science and Engineering, Shandong University, Jinan, China; [2]School of Business, Shandong Jian-zhu University, Jinan, China; *Corresponding Author.
Email: ken0211@gmail.com, mike.wu@263.net, xuna1011@hotmail.com

Received September 21st, 2009; revised November 5th, 2009; accepted December 20th, 2009.

ABSTRACT

International Financial Crisis has made the less-than-truckload (LTL) industry face with severe challenges of survival and development. More and more small and medium-sized LTL carriers choose to collaborate as the potential savings are large, often in the range 5–15%. A key question is how to distribute profits/savings among the participants. Since every LTL carriers are guided by their own self-interests and their contributions to the collaboration are quite different, the proposed allocation method should be a collectively and individually desirable solution. In this paper, we firstly analyze the profit opportunities from collaboration and present mechanisms to realize these benefits by two illustrative examples. Based on the cooperative game theory, we formulate the LTL collaboration game and discuss the well-known profit allocation concepts including Proportional Allocation, Shapley value and Nucleolus. We then propose a new al-location method named Weighted Relative Savings Model (WRSM) which is in the core and minimizes the maximum difference between weighted relative savings among the participants. Simulation result for real-life instances shows the effectiveness of WRSM.

Keywords*: *Cooperative Game, Profit Allocation, Collaborative Transportation

1. Introduction

International Financial Crisis causes a huge decrease in transportation requests and has made less-than-truckload (LTL) segment of the trucking industry face with severe challenges of survival and development. Under this circumstance, horizontal collaboration becomes a good choice for small and medium-sized LTL carriers. In the collaborative alliance, a number of complementary transportation resources from the participants could be integrated and thus more profits could be gained for every participant compared with their stand-alone operation. The potential cost savings from collaboration are often range from 5% to 15%.

Although the benefits from collaboration are appealing, the key question is how to distribute the collaborative profits among every participant to ensure the establishment and sustainability of the alliance and realize the potential of collaboration. Since every participant is guided by their own self-interests and their contributions to the collaboration are quite different, the proposed allocation method should be a collectively and individually desirable solution [1]. The challenge is to design mechanisms that are fair, reasonable and easy-to-implement. We will show that the proposed Weighted Relative Savings Model (WRSM) satisfies all these requirements.

The remainder of the paper is organized as follows. In Section 2, we analyze the opportunity to increase every LTL carrier's profit through collaboration and present two illustrative examples to demonstrate the mechanisms to realize these benefits. In Section 3, based on the cooperative game theory, we formulate the LTL collaborative game and discuss the well-known profit allocation concepts. We then propose a new solution method called Weighted Relative Savings Model (WRSM) which is in the core and minimizes the maximum difference between the weighted relative savings among the participants. Simulation result for real-life instances is presented and analyzed in Section 4 to show the effectiveness of WRSM.

2. Profit Opportunities from Collaboration

The construction of LTL carriers' alliance will enable the formulation of collaborative transportation system. In this section, we will analyze the profit opportunities of this system.

2.1 Collaborative Transportation System

As it is shown in Figure 1, the collaborative transportation system is a kind of system in which all participants share the network and transportation resources.

E denotes Terminal Point (TP) which is the boundary point of the carrier's business coverage. N denotes Switch Point (SP), through which the cargos transport to the carrier's adjacent business point. W denotes Exchange Point (EP) where two or more collaborative carriers in the alliance exchange their cargos and transport the exchanged cargos to their own business point. S de-notes Shared Point (SDP) which is shared by two or more collaborative carriers in the alliance. From the system–wide point of view, transportation network and re-sources are shared among all the LTL carriers in the alliance through EP and SDP which expand the business scope of every participant.

Resource sharing will help to build more reasonable transportation plans to better utilize vehicles, reduce travel time, unloaded distance and lower the total transportation cost effectively.

2.2 Benefits of Collaboration

Cruijssen and Salomon [2] analyze the effect of collaboration for an entire coalition and show, using a case study that cost savings may range from 5 to 15% and can be even higher. Ergun *et al.* [3] note that shippers can reduce their "hidden costs" by cooperating, partly due to higher utilization of their less-than-truckload loading and asset repositioning capabilities. In the time-constrained lane covering problem, the savings range is from about 5.5 percent to a little over 13 percent, where the savings tend to be larger when the size of the instance is larger. [4] Krajewska and Kopfer [5] show that, using a case study, cooperation between the two carriers yields a 10% reduction in the number of vehicles and a 12.46% reduction in routing cost. In practice, after forming collaborative partnerships with others in the Nistevo Network, Georgia-Pacific's percentage of empty movements decreased from 18% to 3%, which corresponds to $11,250,000 savings yearly [6].

We demonstrate the potential benefits of LTL carrier collaboration with the following two examples.

1) *Backhauling*

Consider a network with three cities and two carriers A and B. We assume that the cost of traveling between two cities is the same for both carriers and, for simplicity, that there is no difference in cost between traveling loaded or empty. We further assume that carrier A has a contract in place to serve lane (2, 1), (1, 3) and that carrier B has a contract in place to serve lane (3, 2). The cost C and freight F of each lane in the network and other relevant information are given in Figure 2, where a dashed line represents repositioning (or empty travel).

Without collaboration, carrier A and B operate individually and the corresponding profit of them are

Profit A = F21 + F13 – C21 – C13 – C32 = 1300

Profit B = F32 – C32 – C23 = 400

As it is shown in Figure 3, if carrier A and carrier B collaborate and carrier A serve lane (3, 2) instead of carrier B, they significantly increase their total profit to 2100 by reducing two empty trips. Assume that the profit allocation rate is 0.75, then the new profit become 1575 for carrier A and 525 for carrier B. Carrier A and B increase their profits by 21% and 31% respectively.

Through collaboration, carrier A reduces its empty trip and fully utilizes the truck while carrier B does not need to transport the cargos. But they both gain more benefits since the total repositioning cost is much lower.

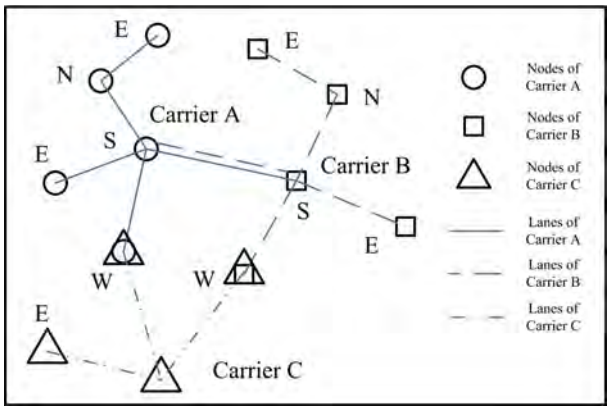

Figure 1. Collaborative transportation system

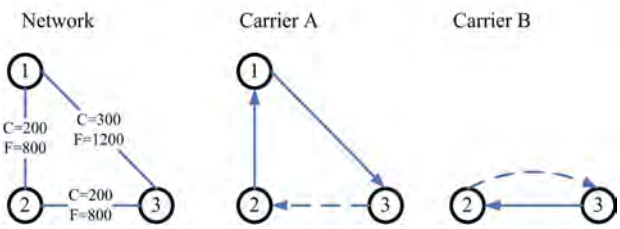

Figure 2. Network information and transportation requests

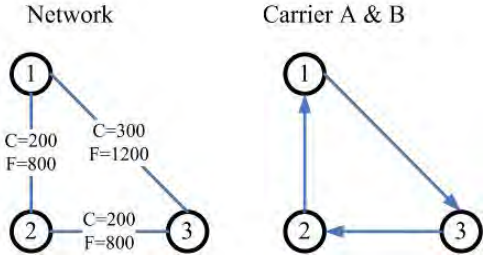

Figure 3. Collaboration between carrier A and B

2) *Lane/Request Exchanging*

Consider a network with four cities and two carriers A and B. We assume that the cost of traveling between two cities is the same for both carriers and, for simplicity, that there is no difference in cost between traveling loaded or empty. We further assume that carrier A has a contract in place to serve lane (2, 1), (1, 3), (3, 4) and that carrier B has a contract in place to serve lane (4, 3), (3, 2). The cost C and freight F of each lane in the net-work and other relevant information are given in Figure 4, where a dashed line represents repositioning (or empty travel).

Without collaboration, carrier A and B operate individually and the corresponding profits are 1900 for carrier A and 1000 for carrier B.

We assume that the existing contracts are not long term contractual agreements so can potentially be exchanged between the carriers [1]. As it is shown in Figure 5, if carrier A and carrier B collaborate and exchange lanes (3, 4) and (3, 2), the corresponding profits are 2100 for carrier A and 1200 for carrier B. Carrier A and B increase their profits by 10.5% and 20% respectively.

Through collaboration, the optimal set of cycles covering the contract lanes are assigned to each carrier. Empty travels are greatly reduced and total profits are redistributed between carrier A and carrier B.

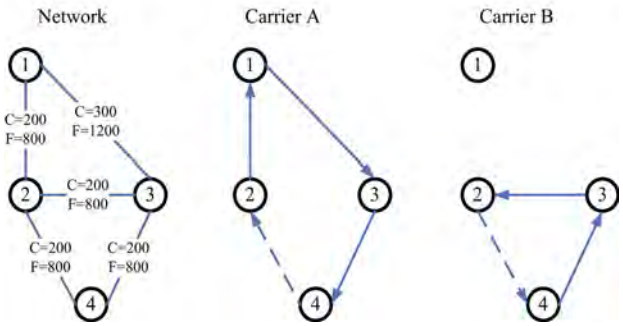

Figure 4. Network information and transportation requests

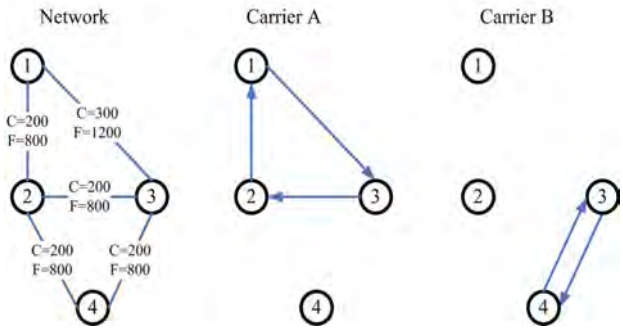

Figure 5. Collaboration between carrier A and B

3. Profit Allocation Problem

Cooperative game theory provides a natural framework for the profit allocation. There are a set of papers that join the transportation related profit or cost allocation problems and cooperative game theory.

Sakawa *et al.* [7] discuss the production and transportation profit and cost allocation based on nucleolus in the fuzzy environment and shows, using actual data, the usefulness of fuzzy programming and the effectiveness nucleolus allocation. Sanchez-Soriano *et al.* [8] study the core of the transportation games, prove the nonemptiness of the core for these games and provide some results about the relationship between the core and the dual optimal solutions of the underlying transportation problem. Sanchez-Soriano *et al.* [9] study the cost allocation of the integrated transportation services provided by Alacant University for students, formulate the problem as tree buses game, propose the aggregated egalitarian solution concept and show it is the core of the game. Engevall *et al.* [10] formulate the traveling salesman game and vehicle routing game, discuss nucleolus, TSP nucleolus, TSP demand weighted nucleolus, Shapley value and τ value respectively. Matsubayashi *et al.* [11] study a cost allocation problem arising from hub-spoke network systems and show that, if the demand across the system has a block structure and the fixed cost is high, allocating the cost proportional to the flow that an agent generates belongs to the core. Ozener [1] study the cost allocation in the collaborative transportation procurement network and discuss the truckload carrier's collaboration. Krajewska *et al.* [5] study the profit sharing problem among carriers in the horizontal collaboration, discuss the possibilities of sharing these profit margins fairly among the partners, apply the Shapley value to determine a fair allocation of the problem and present numerical results for real-life and artificial instances.

These papers in general study the existing profit or cost allocation methods with well-studied properties from cooperative game theory and present the computational results for such allocations. However, to the best of our knowledge, there is no literature on profit allocation for LTL collaborative transportation problem that considers both the relative cost savings and contribution differences, which are very important in the contractual agreement negotiation of the collaboration. In this section, we will search for a new profit allocation method that satisfies these requirements based on the well-known solution concepts from cooperative game theory.

3.1 Problem Definitions and Assumptions

We formulate the profit allocation for LTL collaborative transportation problem as a co-operative game (N, v).

$N=\{1,2,...,n\}(n\geq 2)$ is called the grand coalition which denotes all collaborative carriers. $v(S)$ is the characteristic function which assigns to each possible coalition of carriers $S(S\subseteq N)$ a numerical value to be interpreted as the cost savings realized by the carriers in coalition S. $y_i(i\in N)$ is the profit/cost savings allocated to carrier i. $Y=(y_1,y_2,...,y_n)$ is the profit allocation.

It is assumed that all carriers have the opportunity to form and cooperate in coalition. When coalition S cooperates, the total cost $c(S)$ is generated and we have

$$v(S)=\sum_{i\in S}c(\{i\})-c(S),\ \forall S\subseteq N \tag{1}$$

Below we discuss some of the most commonly used profit allocation properties from cooperative game theory.

A profit allocation method that splits the total profit $v(N)$ among the carriers $i\in N$ is said to be *efficient* or *budget balanced*, that is $\sum_{i\in N}y_i=v(N)$.

A profit allocation is said to be *individual rational* if no carrier gains less than its "stand alone profit/cost saving", which equals to zero. Mathematically, this property is expressed as $y_i\geq v(\{i\}),\forall i\in N$.

The core of the game is defined as those profit allocations $Y=(y_1,y_2,...,y_n)$ that satisfy the conditions

$$\sum_{i\in S}y_i\geq v(S),\forall S\subset N \tag{2}$$

$$\sum_{i\in N}y_i=v(N) \tag{3}$$

That is, no single carrier or coalition of carriers would be better off if they decide to opt out and collaborate only among themselves. A profit allocation in the core is said to be *stable*.

For each coalition S and a given profit allocation $Y=(y_1,y_2,...,y_n)$, we can compute the *excess*

$$e(Y,S)=\sum_{i\in S}y_i-v(S),\ \forall S\subseteq N \tag{4}$$

which expresses the difference between the sum of the profits allocated to its members and the total profit of a coalition. For a given profit allocation, the vector of all excesses can be thought of as a measure of how far the profit allocation is from the core. If a profit allocation is not in the core, at least one excess is negative.

3.2 Well-known Profit Allocation Concepts

3.2.1 Proportional Allocation

In practice, the most commonly used solution is to distribute the collaborative profit/cost savings of the grand alliance $v(N)$ among the carriers equally, weighted with each carrier's stand alone cost. This is expressed as

$$y_i=r_i\times v(N),\ \forall i\in N \tag{5}$$

where r_i is equal to $c(\{i\})\Big/\sum_{i\in N}c(\{i\})$.

Although this method is easy to understand, easy to show and easy to compute, it is not stable from a cooperative game theoretic point of view since a participant will pay, possibly, more than when operating alone [1].

3.2.2 Shapley Value

A well-known cost allocation method is the Shapley Value, which is defined for each player as the weighted average of the player's marginal contribution to each subset of the collaboration [12]. Shapley Value can be interpreted as the average marginal contribution each member would make to the grand coalition if it were to form one member at a time [13]. Mathematically, Shapley Value is expressed as

$$y_i=\sum_{i\in S}\frac{(n-s)!(s-1)!}{n!}\left(v(S)-v(S\setminus\{i\})\right),\forall i\in N$$

where s denotes the number of carriers in coalition S.

Shapley Value is the unique allocation method to satisfy three axioms: dummy, additivity and equal treatment of equals. Although Shapley Value may return cost allocations in the core for some instances, there are many instances where allocations based on Shapley Value are not stable [1].

3.2.3 Nucleolus

Nucleolus, introduced by Schmeidler [14], is the cost allocation that lexicographically minimizes the maximal excess, the difference between the total allocated profit to a subset and the stand alone cost of that subset, over all the subsets of the collaboration. Mathematically, it is expressed as

$$\begin{aligned}&\textit{Minimize}\ \ \varepsilon\\ &\textit{s.t.}\quad \sum_{i\in S}y_i-v(S)\leq\varepsilon \qquad \forall S(S\subset N,S\neq\phi)\\ &\qquad \sum_{i\in N}y_i=v(N)\\ &\qquad y_i\geq v(\{i\}) \qquad\qquad \forall i\in N\end{aligned}$$

The nucleolus exists and is unique. However it does not take into account each carrier's contributions to the coalition and the relative cost savings.

3.3 Weighted Relative Savings Model

As discussed above, the existing solutions are not always stable, which keeps the sustainability of the LTL collaboration, and different to show that some participants can gain more if they contribute more and all participants have a similar relative profit or cost savings. In a nego-

tiation situation it would be beneficial to have an initial allocation where the relative savings are as similar as possible for all participants.

We therefore propose the Weighted Relative Savings Model (WRSM) which is completely new and motivated by finding a stable allocation that minimizes the maximum difference between relative savings among the participants and also reflects the contribution difference.

The relative savings of carrier i is expressed as $y_i / c(\{i\})$. Thus, the difference in relative savings between two participants i and j is equal to

$$\frac{y_i}{c(\{i\})} - \frac{y_j}{c(\{j\})} \tag{6}$$

The contribution to the collaboration depends on the distribution of power among freight carriers, on their level of interdependency and willingness to make compromises, and on the market within which the freight carriers operate [5]. Following the ideas of the Shapley Value, we define the contribution of carrier i to the grand coalition as

$$\sum_{i \in S} \left(v(S) - v(S \setminus \{i\}) \right), \forall i \in N \tag{7}$$

In order to reflect the contribution difference, we modify the relative savings by adding the contribution ratio weight ω_i which is expressed as

$$\omega_i = 1 - \frac{\sum_{i \in S} \left(v(S) - v(S \setminus \{i\}) \right)}{\sum_{i \in N} \sum_{i \in S} \left(v(S) - v(S \setminus \{i\}) \right)} \tag{8}$$

The weighted relative savings of carrier i is then equal to $\omega_i \cdot y_i / c(\{i\})$ and the difference in relative savings between two participants i and j is equal to

$$\frac{\omega_i y_i}{c(\{i\})} - \frac{\omega_j y_j}{c(\{j\})} \tag{9}$$

The Weighted Relative Savings Model (WRSM) is the following LP problem which we need to solve to find the allocation.

$$\begin{aligned} &\textit{Minimize} \quad f \\ &\textit{s.t.} \quad \frac{\omega_i y_i}{c(\{i\})} - \frac{\omega_j y_j}{c(\{j\})} \le f \qquad \forall i, j \in N \\ &\sum_{i \in S} y_i \ge v(S) \qquad \forall S \subset N \\ &\sum_{i \in N} y_i = v(N) \end{aligned}$$

The first constraint set is to measure the difference between all participants' weighted relative savings. The variable f is used in the objective function to minimize the maximum difference. The other two constraint sets ensure that the allocation is in the core and thus stable.

We add a minimum penalized slack in the constraints defining the core. In the case the core is empty we propose to use the epsilon-core or alternatively seek the maximal number of players present in a game for which the core exists. However, how this subgroup of players should be selected remains to be studied in future research.

Compared with the Proportional Allocation and the Shapley Value, this allocation is stable. Since the objective is a combination between participants and considers the relative savings and the contribution difference, this model is not a weighted nucleolus. In the literature of this field, we have not been able to find an allocation method with similar objective. Therefore, to the best of our knowledge, this allocation concept is new.

4. Simulation Result and Analysis

In order to show the effectiveness of the method we propose, we compare the Weighted Relative Savings Model (WRSM) with Proportional Allocation, Shapley Value and Nucleolus based on the existing test instances in [5].

Table 1 presents the instances used in our test and related calculations. There are three carriers in the grand coalition and the optimal number of vehicles and cost of each subset of the grand coalition is calculated according to the transportation requests in the sub-coalition [5]. Cost Savings of Coalition is calculated using (1). Contribution to the Grand Coalition is calculated using (7) and Contribution Ratio Weight is calculated using (8) respectively.

Table 2 shows the results for test instances and the comparison among Proportional Allocation, Shapley Value, Nucleolus and WRSM. For each allocation concept, Cost Savings allocated to carrier is calculated according to the related definitions and algorithms discussed above. Net Cost equals to Stand-alone Cost minus Cost Savings.

These results show clearly that it is indeed worth pooling the LTL carriers' transportation resources through collaboration to serve customer requests. The cost savings is range from 7.3% to 18.7%.

Although the Proportional Allocation and Shapley Value is stable using our test instances, carrier C will not agree with those allocation methods since he contributes more to the grand coalition but gains the same relative savings as carrier A and B in Proportional Allocation and the smallest savings in Shapley Value allocation. The Nucleolus, which divides the cost savings equally among three carriers, does not take into account the contribution difference among the three and may be rejected by any of them. WRSM which is in the core and considers both relative savings and contribution difference makes the

Table 1. Test instances and related calculations

Carriers in Coalition	# Requests	# Vehicles	Cost	Cost Savings of Coalition	Contribution to the Grand Coalition	Contribution Ratio Weight
A	61	13	16512.6	0.0	13216.5	0.64
B	96	11	17876.0	0.0	8463.7	0.77
C	100	28	38585.4	0.0	14575.7	0.60
A B	157	24	31961.6	2427.0		
A C	161	36	49615.0	5483.0		
B C	196	32	53354.8	3106.6		
A B C	257	38	64560.9	8413.1		

Table 2. Results for test instances

Carrier	Stand-alone Cost	Proportional Allocation			Shapley Value			Nucleolus			WRSM		
		Cost Savings	Net Cost	Savings Ratio	Cost Savings	Net Cost	Savings Ratio	Cost Savings	Net Cost	Savings Ratio	Cost Savings	Net Cost	Savings Ratio
A	16512.6	1903.7	14608.9	11.5%	3087.2	13425.4	18.7%	2804.4	13708.2	17.0%	1920.5	14592.1	11.6%
B	17876.0	2060.9	15815.1	11.5%	1899.0	15977.0	10.6%	2804.4	15071.6	15.7%	1723.5	16152.5	9.6%
C	38585.4	4448.5	34136.9	11.5%	3427.0	35158.4	8.9%	2804.4	35781.0	7.3%	4769.1	33816.3	12.4%
SUM	**72974.0**	**8413.1**	**64560.9**		**8413.1**	**64560.9**		**8413.1**	**64560.9**		**8413.1**	**64560.9**	

weighted relative savings as similar as possible among different participants. It can be accepted by all the carriers and makes the collaboration sustainable.

5. Conclusions

Collaboration is a good choice for small and medium-sized LTL carriers under the background of the international financial crisis. Potential cost savings of the collaborative alliance is large and every participant can gain more profits comparing with stand-alone operation. In order to realize the benefits, collaborative profit allocation mechanism must be able to construct the alliance and make it sustainable.

The underlying profit allocation problem is discussed in this paper. We have demonstrated that collaboration can yield a considerable cost decrease and proposed a new profit allocation method named Weighted Relative Savings Model (WRSM) based on the cooperative game theory. Simulation result for real-life instances shows the effectiveness of the proposed model.

The truck transportation industry has not yet adopted horizontal cooperation on a large scale [5]. So the key challenge in terms of future developments is to adapt the proposed method for practical use so that not all possible coalitions need to be analyzed.

6. Acknowledgements

This research is supported by National Natural Science Foundation of China (No. 50175064). The authors are also grateful to anonymous referees for their helpful comments and insights.

REFERENCES

[1] O. O. Ozener, "Collaboration in transportation," PhD thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2008.

[2] F. Cruijssen and M. Salomon, "Empirical study: Order sharing between transportation companies may result in cost reductions between 5 to 15 percent," CentER Discussion Paper, 2004, http://ideas.repec.org/p/dgr/kubcen/200480.html.

[3] O. Ergun, G. Kuyzu, and M. Savelsbergh, "Shipper Collaboration," Computers & Operations Research, No. 34, pp. 1551–1560, 2007.

[4] O. Ergun, G. Kuyzu, and M. Savelsbergh, "Reducing truckload transportation costs through collaboration," Transportation Science, No. 41, pp. 206–221, 2007.

[5] M. Krajewska and H. Kopfer, "Horizontal cooperation among freight carriers: Request allocation and profit Sharing," Journal of the Operational Research Society, No. 59, pp. 1483–1491, 2008.

[6] P. Strozniak, "Collaborative logistics: Overcoming its challenges can lower transportation and inventory costs and reduce stockouts," Frontline Solutions, 2003, http://www.frontlinetoday.com.

[7] M. Sakawa, I. Nishizaki, and Y. Uemura, "Fuzzy programming and profit and cost allocation for a production and transportation problem," European Journal of Operational Research, No. 131, pp. 1–15, 2001.

[8] J. Sanchez-Soriano, M. A. Lopez, and I. Garcia-Jurado, "On the core of transportation games," Mathematical Social Sciences, No. 41, pp. 215–225, 2001.

[9] J. Sanchez-Soriano, N. Llorce, A. Meca, and E. Molina, "An integrated transport system for Alacant's students," Annals of Operations Research, No. 109, pp. 41–60, 2002.

[10] S. Engevall, M. Gothe-Lundgren, and P. Varbrand, "The heterogeneous vehicle-routing game," Transportation Science, No. 38, pp. 71–85, 2004.

[11] N. Matsubayashi, M. Umezawa, Y. Masuda, and H. Nishino, "A cost allocation problem arising in hub–spoke network systems," European Journal of Operational Research, No. 160, pp. 821–838, 2005.

[12] L. S. Shapley, "A value for n-person games," Annals of Mathematical Studies, No. 28, pp. 307–317, 1953.

[13] H. P. Young, "Cost allocation: Methods, principles, applications," European Journal of Operational Research, No. 27, pp. 254–255, 1986.

[14] D. Schmeidler, "Nucleolus of a characteristic function game," Siam Journal on Applied Mathematics, No. 17, pp. 1163–1170, 1969.

J. Service Science & Management, **2010, 3: 150-158**
doi:10.4236/jssm.2010.31019 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Transformational Leadership of Afghans and Americans: A Study of Culture, Age and Gender

Belal A. Kaifi[1] , Bahaudin G. Mujtaba[2]

[1]Franklin University, Columbus, USA; [2]Nova Southeastern University, Fort Lauderdale, USA.
Email: belalkaifi@yahoo.com, mujtaba@nova.edu

Received October 22nd, 2009; revised November 29th; accepted January 3rd, 2010.

ABSTRACT

Afghans and Americans have been exposed to different leadership styles which may have influenced their orientation on leading others. Age, gender, and/or culture are possible factors for such differences in leadership orientations. This research surveyed the responses of 502 *Americans and* 300 *Afghans to better understand their orientation toward transformational leadership. The Afghan respondents had significantly higher scores for transformational leadership orientation. Female American respondents had a significantly higher transformational leadership orientation than their Afghan female colleagues. Younger Afghans had a significantly higher tendency toward a transformational leadership orientation than their American counterparts. Suggestions for managers and implications for future research are presented.*

Keywords*: *Afghans, Americans, Afghanistan, Gender, Age, Culture, Leadership, Ethics

1. Introduction

Culture, age group, and gender may influence a person's leadership style and overall change management approach. Leaders must believe in change in order for there to be a better future for a society or organization. "In highly competitive, rapidly changing environments, caring and appreciative leaders are the ones to bet on for long-term success" [1]. The 21st century leader must be equipped with the right tools to be effective, empathic, and efficient in all aspects of the workplace. Today's competitive workplace requires more transformational leaders as they tend to influence workers more positively. Rhodes, Walsh, and Lok state, "While leaders initiate and drive organizational change, they manage the change only with the help of other change agents. These change agents operate with different change skills and competencies depending on particular requirements and circumstances" [2]. Pounder explains that the effect of transformational leadership on subordinates centers on three leadership outcomes: a) the ability of the leader to generate extra effort on the part of those being led, b) subordinates' perception of leader effectiveness, and c) their satisfaction with the leader [3]. Furthermore, "Bass and his associates' views on morality relative to transformational and transactional leadership do suggest that transactional leaders would be expected to engage in unethical practices more so than transformational leaders," and further state that "Judgments of a leader's ethical posture may play a particularly strong role in influencing follower satisfaction with the leader" [4]. Franken, Edwards, and Lambert explain how "Business leaders are under constant pressure to comply with their demands while maintaining the organization's competitiveness in increasingly complex markets" [5]. Transformational leaders are expected to not only to control, lead, plan, and organize, but also motivate, empower, and build healthy relationships with their peers throughout an organization.

A study by Mujtaba and Kaifi illuminated how Afghan leaders have higher scores on the relationship orientation which relates to better job performance [6]. The skills of building and maintaining a healthy relationship is an important element of not just performance but also transformational leadership, especially when one has to create a win-win relationship between employees, vendors, and other stakeholders in the community or local governments. "Over the last decade, considerable research effort has been invested into understanding the processes through which transformational leadership positively relates to follower attitudes, behavior, and performance" [7]. When exploring the conditions under which transformational leadership weaves its effects on performance, research results show that transformational leadership relates to follower identification with work unit and self-efficacy, to predict individual performance, thus representing a moderated mediation effect [7]. Another

study demonstrated "that instructors displaying transformational leadership qualities in the classroom had a positive and significant influence on student perception of classroom dynamics measured in terms of the three leadership outcomes: extra effort, effectiveness, and satisfaction" [3]. The results appear to be consistent with other findings on transformational leadership which suggest that employees view this style "positively in terms of effectiveness, satisfaction, and motivation to expend effort" [3]. Other researchers argue that "the proposed association of transformational and transactional leadership has been one of augmentation. The augmentation hypothesis argues that transformational leadership will significantly predict leadership criteria after controlling for transactional leadership" [4]. Of course, effective leaders use a variety of methods to influence their followers and some researchers claim that "employees with higher levels of power distance orientation are less likely to be influenced by transformational leadership behaviors alone and may instead need to be led via different or additional leadership styles" [8]. Kirkman *et al*., further state that "individual-level cultural value orientations, and particularly power distance orientation, should not be ignored in studies of the impact of transformational leadership on followers across cultures" [8]. This study focuses on the leadership style of people in Afghanistan and the United States because an individual's culture might very well be a factor in their influencing style. In both cultures, followers respect managers who are perceived to be just, fair, and ethical. Leaders who do the right thing are able to enhance the morale of their employees because all individuals want to be associated with good causes and institutions that take society's needs into consideration. Transformational leaders must have an ethical intention in order to be followed and respected.

Ricoeur states, "Let us define 'ethical intention' as aiming at the 'good life' with and for others, in just institutions" [9]. The ethical intention is what transformational leaders strive for in organizations of the 21st century because in order for an organization to be successful, there must be unity and collectivity along with just practices. When aiming at the good life, leaders must consider not only themselves but how each person can make a difference by being more understanding, ethical, and concerned for the welfare of all people within an organization. Ricoeur explains, "the good is not contained in any particular thing. The good is rather that which is lacking in all things. This ethics in its entirety presupposes this nonsaturable use of the word predicate good" [9]. Ricoeur recognizes the importance of being ethical and the positive implications it has on daily actions of an organization. In an organization, every conversation can be meaningful, and may bring one that much closer to a good life by learning about the other. Ricoeur states, "To narrate and to follow a story is already to 'reflect upon' events with the aim of encompassing them in successive totalities" [10]. Being able to listen, understand, and communicate to a colleague is important for a transformational leader and will benefit an organization by building trust, friendships, awareness, and respect. As Ricoeur explains, "the good life is for each of us, the nebulus of ideals and dreams of achievements with regard to which a life is held to be more or less fulfilled or unfilled" [9]. In organizations, transformational leaders need to understand others by focusing on solicitude, building friendships, and emphasizing equality because Ricoeur explains, "equality, however it is modulated, is to life in institutions what solicitude is to interpersonal relations" [9]. By having an ethical intention, a transformational leader is able to lead by example. While the controversy regarding whether leaders are products of nature or nurture continue to live on, one element that is clear is how all people are influenced by their specific culture, age group, and gender.

2. Leadership and Culture

Culture, or the way people behave in an organization can influence how a leader evaluates performance. Rhodes, Walsh, and Lok state that the "development of a high-performance culture requires the inspiration, drive, and commitment of the leader, leadership is critical to the change process underpinning the Balanced Scorecard" [2]. Furthermore, according to these authors, cultural values tend to have important implications on management practices, because "there is a positive relationship between a high performance culture and the adoption of best practices" [2]. For example, the United States' culture is considered to be high performance and, if organizations are to remain competitive, they must adopt best practices from around the globe. Some of these best practices may include the training and development of effective ethical leaders through succession planning and development because effective leaders require vision, knowledge, and execution skills.

Leaders come in many shapes, forms, and styles which is why transformational leadership has been studied over the years. For example, Kim claims that "there may be at least two different leadership styles in groupthink: the lack of impartial leadership and laissez-faire" [11]. Others state that "In combination with the provision of autonomy, the team leader functions as a supportive member" [12]. Kaifi explains how a transformation leader should consider "the primary tasks of the organization by figuring out the most efficient, effective, and productive way to complete all tasks [13]. Daniel and Davis emphasize that "Managers who lead high-performance teams in highly competitive industries must

balance complex interpersonal relationships with corporate deadlines" [14]. Teamwork is critical for success in the modern workplace because fast and effective solutions to conundrums require the creativity and talents of all stakeholders. According to Pryor *et al*., "It is imperative that managers, team leaders, and team members understand the elements required for team success and the potential for team failure" [15]. Kearney proposes that transformational leadership in the modern workplace "will engender positive effects on team performance only to the extent that the team members regard it as legitimate and appropriate that one person among them, the leader, occupies a privileged position" [16]. As such, this study proposes the following hypothesis:

Hypothesis 1 – Afghans and Americans will have similar transformational leadership scores.

3. Leadership and Gender

Males and females may lead from different perspectives and these diverse leadership perspectives can lead to different approaches to problem-solving. "Current psychological research on leadership and team interaction suggests that men and women exhibit different leadership styles and interpersonal communication styles in a variety of small-group situations from student problem-solving situations to industry and community situations" [17]. Lantz writes that "women executives are much more likely than males to be a department head or to fill some other staff position, whereas men are much more likely to be chief executive officer (CEO), chief operating officer (COO), president, or vice president" [18]. Many believe that the glass ceiling phenomenon continues to play a role in today's workforce. For example, Babcock explains that "Rather than intentional acts of bias, second-generation gender biases reflect the continuing dominance of traditionally masculine values in the workplace" [19]. Women are perceived to be too emotional and less competitive than men. For example, some authors have written that "Women are significantly more risk averse, tend to be less overconfident and behave less competitively oriented" [20]. Overall, there is a disparity between the equal promotion of men and women to higher positions: "Despite high-profile success stories of female CEOs such as Meg Whitman of eBay, only a handful of Fortune 500 firms in 2008 have a woman in the top spot. Consequently, concern remains about the progress women are making" [21]. Some of the biases are very subtle and not necessarily as blatant as they used to be, which may be a result of human resource professionals doing a more effective and meticulous job of making their managers aware of the existing equal opportunity employment laws in the United States. However, "Workplace inflexibilities and lack of organizational support are driving women away, creating a leaky pipeline of female talent" [22]. If there are biases in the promotion of female candidates, there will be fewer role models for others to follow who are interested in following in their footsteps. This phenomenon is likely to be just as true in Afghanistan as it is in the United States. "In gender studies of public administration, there is significant evidence that women have less organizational power than men, measured in lower pay, fewer career opportunities, and underrepresentation at the highest leadership levels" [23]. As such, this study proposes the following hypotheses:

Hypothesis 2 – Afghan and American males will have similar transformational leadership scores.

Hypothesis 3 – Afghan and American females will have similar transformational leadership scores.

4. Leadership and Perceptions of Age

Leaders and styles of leadership may vary based upon age groups. Some authors have written that "With an older leader, the team may be more open to a leader's transformational behaviours, because the team members may be more accepting of the leader's special status" [16]. Kearney further comments that: Results indicate that it makes a difference whether transformational leadership is provided by a leader who is older than or about the same age as the followers. Only in the former case was there a positive relationship between transformational leadership and team performance. Moreover, the data yielded no evidence for a curvilinear effect. [16].

According to Van Vugt, "Age relates to leadership in a complicated way, according to the psychological literature. Some research finds a positive correlation between age and leadership, whereas others find a zero or negative correlation" [24]. Age makes a difference if the leader is respected for his or her knowledge, expertise, overall wisdom, and communication or listening skills. For example, "Experience in Britain suggests that there is a growing confidence that an effective coach working with senior people can achieve positive results, provided he or she has the ability to listen" [25]. Leaders must demonstrate that they are intelligent and capable of making decisions: The evolutionary game model also suggests that people who are quicker to recognize a situation as a coordination problem, that requires leadership, emerge as leaders more often. Furthermore, they must convince people that following is the best option. From this, I expect that leadership correlates with intelligence, because it helps in identifying coordination problems as well as in coordinating actions of multiple actors [24].

Besides making intelligent personal and professional decisions, it is also important for younger and older leaders to be seen as generous and fair as "leadership is correlated with traits and behaviors that signal generosity and fairness" [24]. The theory of moral development states that as individuals' age, they are likely to gain more knowledge about right and wrong and fairness. So

age has positively contributed to wisdom and knowledge of human beings since individuals are able to learn from personal experiences and the observation of others, and ultimately make positive changes to their habits and behaviors. Van Vugt explains that: In ancestral environments, some situations required the possession of unique and specialized knowledge, for example, where to find a waterhole that has not yet dried up. Knowledge about where to go would have been more likely to be held by older and experienced individuals, and, thus leading is expected to correlate positively with age in this domain. In the present time, evidence for this link between age and leadership can still be found in professions that require a considerable amount of specialized knowledge and experience, such as in science, politics, and arts [24].

Of course, "workforce aging is not a new phenomenon, its confluence with other factors such as early retirement programmes and decreased supplies of skilled craft workers appears to be giving the problem an unprecedented significance" [26]. As such, this study proposes the following hypotheses:

Hypothesis 4 – Afghan and American respondents who are 26 years of age and older will have similar transformational leadership scores.

Hypothesis 5 – Afghan and American respondents who are 25 years of age and younger will have similar transformational leadership scores.

5. Research Methodology

Afghans and Americans who participated in this study completed a modified MLQ 5X-Short (Multifactor Leadership Questionnaire) survey that was originally developed by Bass and Avolio for leadership studies. The survey instrument used for this study had eleven short questions (see Appendix A) designed for the population. Many transformational leadership researchers (Pounder, 2008; Kearney, 2008; Ling *et al*., 2008; and Jansen *et al.*, 2008) have used similar instruments to study the leadership styles of various participants. For example, Pounder used a modified version of the MLQ Form 5X-Short which "involved a sample of instructors and undergraduate students in a Hong Kong university business school [3]. Pounder used a version of the Multifactor Leadership Questionnaire that was modified for a classroom situation to better understand the styles of prospective leaders.

The survey questions were set up in a Likert scale format where a response of 1 means "Never" and a response of 5 means the element is "Always" a characteristic of the responder. One of the questions in the survey asked the respondent the following: "I express with a few simple words what we could and should do", and a different question asked, "I provide others with new ways of looking at puzzling things." The respondent would rate him or herself from a scale of 1 to 5 in regard to how many words he or she might use to actually express his or her views. The higher the overall mean scores, the more chances that he or she is likely to have a stronger orientation toward a transformational leadership style. One's range for being a transformational leader can be expressed with a score of "Very low" to a score of "Very high" as presented in **Table 1**.

The survey instrument was distributed to 2,000 Afghans and Americans by using Facebook as a social-networking instrument to get good participation. A total of 300 surveys were completed successfully by Afghans who live throughout the United States and abroad. Furthermore, a total of 502 surveys were completed successfully by American respondents from the United States. So a total of 802 responses, which represents a 45% response rate, were used for analysis.

The research question for this study is: *Do Afghans and Americans have similar or different transformational leadership scores*? For this survey, the higher the overall sum of the scores, the more likely that the participant is strongly oriented toward a transformational leadership style.

6. Results and Analysis

The responses of 802 Afghans and Americans demonstrate that there is a statistically significant difference in their mean transformational leadership scores ($t = 7.83$; $p < 0.001$), as presented in **Table 2**. Afghan respondents have a significantly higher score on transformational leadership, which means that the first hypothesis, "*Afghans and Americans will have similar transformational leadership scores*," cannot be supported. While the mean score of Afghan respondents fall in the high range, the American respondents mean score falls in the moderately high range for transformational leadership orientation.

Table 1. Transformational leadership orientation range

45–50	Very high range
40–44	High range
35–39	Moderately high range
30–34	Moderately low range
25–29	Low range
10–24	Very low range

Table 2. Transformational leadership score by culture descriptive statistics and T-test of two means

Culture	Mean	Standard Deviation	Sample Size
Afghan	41.266*	4.26	300
American	39.6096*	1.612	502

** $t = 7.83$; $p < 0.001$*

The second hypothesis predicted that "*Afghan and American males will have similar transformational leadership scores*" and, as presented in **Table 3**, this study could not support this supposition because male Afghans have a significantly higher score than their American counterparts.

The third hypothesis predicted that "*Afghan and American females will have similar transformational leadership scores*" and, as presented in **Table 4**, this study could not support this supposition because American females' mean score is significantly higher than their female Afghan counterparts.

The fourth hypothesis predicted that "*Afghan and American respondents who are 26 years of age and older will have similar transformational leadership scores*" and, as presented in **Table 5**, this study could not support this supposition because older Americans have a significantly higher score than their Afghan colleagues.

The last hypothesis predicted that "*Afghan and American respondents who are* 25 *years of age and younger will have similar transformational leadership scores*" and, as presented in **Table 6**, this study could not support this supposition because the scores are significantly different. Younger Afghans have a significantly higher transformational leadership score than their American colleagues.

Table 3. Transformational leadership score of males descriptive statistics and T-test of two means

Males	Mean	Standard Deviation	Sample Size
Afghan Males	44.7933	2.18689	150
American Males	39.54	1.75	280

$t = 27.2$; $p < 0.001$

Table 4. Transformational leadership score of females descriptive statistics and T-test of two means

Females	Mean	Standard Deviation	Sample Size
Afghan Females	37.74	2.56049	150
American Females	39.7	1.41	222

$t = –9.48$; $p < 0.001$

Table 5. Transformational leadership score by age descriptive Statistics and T-test of two means

Age	Mean	Standard Deviation	Sample Size
Older Afghans	37.36	2.097	50
Older Americans	39.59	1.64	302

$t = –8.53$; $p < 0.001$

Table 6. Transformational leadership score by age descriptive statistics and T-test of two means

Age	Mean	Standard Deviation	Sample Size
Younger Afghans	42.048	4.15	250
Younger Americans	39.645	1.57	200

$t = 7.99$; $p < 0.001$

This study has demonstrated that Afghans and Americans have significantly different transformational leadership scores. The Afghan respondents scored in the higher range for having a transformational leadership orientation. This finding corroborates Kaifi's study which explains how most Afghans are natural transformational leaders partly because they have been influenced by leaders who have fought for reform, modernization, and a prosperous vision for Afghanistan [27]. Furthermore, there was a statistically significant difference between the scores of Afghans and Americans based on gender and age. Female American respondents as well as older American respondents had a significantly higher transformational leadership orientation than their Afghan colleagues. Younger Afghan respondents as well as male Afghans had a significantly higher tendency toward a transformational leadership orientation than their American counterparts.

7. Implications of the Study

This study has demonstrated that Afghans seem to be naturally oriented toward transformational leadership tendencies. From an ideological perspective, Afghans are inclined to be transformational leaders because the Afghan culture places a considerable emphasis on respecting elders because of their knowledge, wisdom, and experience, which explains why older transformational leaders are usually more successful in influencing the Afghan population. This research has shown that young Afghan respondents are even more inclined to be transformational leaders than their older counterparts. Global business and human resource management practices can be challenging when an organization is working with people of a different culture [28]. Organizations that are interested in conducting business in Afghanistan or around the globe should feel confident in hiring an Afghan to lead the organization, which contradicts the current frame of thought.

Due to societal conditioning and cross-cultural tendencies of human nature, some managers assume that employees from high-context cultures such as Afghanistan are likely to be more relationship-oriented which may distract employees from not completing their tasks in a timely manner. For example, they may not be assertive enough to pressure their peers toward being more productive when there is a backlog or even to ask for

help when necessary because they do not want to appear "pushy" or "rude." Of course, such assumptions are often wrong as the Afghan respondents in this survey have higher transformational leadership scores when compared to their American counterparts.

These results are important elements for multinational managers, administrators of USAID (the United States of America's International Development) agency, NGOs (non-governmental agencies), and other contractors recruiting professionals for jobs and assignments in and around Afghanistan. As a matter of fact, "Afghan leaders from all professions have started emerging. There are currently many Afghan-American medical doctors, engineers, attorneys, professors, police officers, and many who work in either the private or public sectors of the workforce" [29]. The modern workplace for a multinational firm can be very diverse as the workforce today is filled with people from many different leadership styles. Kaifi explains how using multiple frames to evaluate an organization will help a transformational leader understand complex issues within an organization and will result in continuous improvements [13]. As such, "administrators must be managers of diverse interests" and "this necessity grows out of the relativity of values and the pluralization of society" [30]. Managers should provide a better understanding of their rules and policies along with diversity education for all their employees [31]. It is for certain that Afghans and Afghan-Americans bring diverse views and perspectives with an orientation toward transformational leadership into the workplace which can help make them ethical and respected leaders within their teams, departments, organizations, and communities.

In all organizations, there is a high-demand for transformational leaders because of organizational changes due to technology, globalization, and competition. For example, radiology departments throughout America have implemented digital technology (digital images) that has replaced the analog technology (x-rays). During this organizational change, transformational leaders were tasked with successfully implementing this new technology which is why "organizations are quick to look for leaders who are great communicators, visionary thinkers, and who can also get things done and follow through" [32]. It is important for a transformational leader to first understand the organizational culture and gain the trust of subordinates in order to effectively reframe, implement new strategies, and transform an organization to be able to compete in today's global economy.

A research study by Morhart, Herzog and Tomczak suggests "that managers should make a paradigm shift from a TRL [transactional leadership] to a TFL [transformational leadership] philosophy" and go on to say, "At first glance, specifying behavioral codices and scripts for employees dealing with customers and then monitoring and rewarding appropriate demeanor might seem to be an easy solution for obtaining adequate performance from employees representing the corporate brand" [33]. They further mention that:

However, we found that a highly transactional style was counterproductive in terms of followers' motivational condition. Managers would do much better by opening their minds to a TFL approach, which would entail behaviors such as articulating a unifying brand vision, acting as an appropriate role model by living the brand values, giving followers freedom to individually interpret their roles as brand representatives, and providing individualized support by acting as a coach and mentor. This would allow followers to experience the feelings of relatedness, autonomy, and competence in their roles as brand representatives, which would ultimately spill over into the commitment, authenticity, and proactivity that characterize a real brand champion [33].

While transactional leadership styles might be appropriate for specific organizations and situations such as in hospitals experiencing technological changes, this style is usually effective only in the short-term and may not serve a developmental objective. As such, transformational leadership is usually the best alternative for long-term success in employee development situations and when one is trying to inspire and motivate. Research shows that "although employees were more likely to have higher levels of power distance orientation in the PRC [People's Republic of China] than in the U.S., individuals in *both* countries reacted differently to transformational leadership on the basis of their individual power distance orientation" and furthermore, "irrespective of country-level cultural variation, transformational leadership is especially important for managers whose employees have a low power distance orientation" [8]. A different study explains that, "to be effective, transformational leaders must develop high quality leader-follower relationships, both LMX [Leader-Member Exchange] and interactional justice, with followers. In this sense, leaders must treat followers with kindness and fairness, develop meaningful social exchange relationships (e.g., trust, professional respect) with them, and maintain equitable exchanges with them" [34]. Justice, fairness and good ethical values are important for all leaders, regardless of their cultural orientation. In a highly just and ethical environment, employees will be more committed and motivated to do what is right for the company and their customers. Fisher explains that "Transformational leadership goes beyond the idea that workers are motivated by rewards and punishments by considering other motivators for effective performance" [35]. Transformational leaders must also create an organizational culture where people feel free to think, discuss and express their new ideas. Furthermore, in such an environment:

Managers need to be mindful that selecting employees on the basis of their learning orientation alone will not guarantee creativity. It is building the creative self-efficacy of their employees that will provide the facilitating conditions for the learning orientation to take hold and bring forth creativity. Managers can be instrumental here in terms of providing an environment that stimulates and nourishes creative self-efficacy, through, for example, applying transformational leadership principles [36].

The 21st century requires transformational leaders who are able to help organizations thrive, continuously improve, and accomplish its goals.

8. Limitations of the Study

There are some limitations to this study and one is the modified MLQ 5X-Short (Multifactor Leadership Questionnaire) survey used for this research. This short survey can be combined with other more comprehensive instruments to enhance and confirm the results. Future studies can duplicate the research with a greater number of Afghan participants that are compared to other ethnicities. The fact that this study was conducted with a convenient sample population, living in urban areas and included expatriate Afghans living outside of Afghanistan, was a further limitation. This point is particularly important for the Afghan respondents, who consequently, may have become more "westernized" as a result of living and working in foreign locations. Future studies might control this variable by expanding the research population to include respondents from within Afghanistan. Finally, future researchers should consider translating the survey instrument into Persian and Pashto languages so as to facilitate the test subjects' preferred and dominant reading skills.

9. Conclusions

This study focused on the Afghan and American population to compare and better understand their transforma tional leadership tendencies. The premise of the study was to determine if Afghans and Americans have similar transformational leadership scores and this study has confirmed that their scores are significantly different in regard to culture, age, and gender. From a cultural perspective, the results demonstrated that the Afghan respondents scored in the high range and the Americans scored in moderately high range for transformational leadership orientation. From gender and age perspectives, female and older American respondents had a significantly higher transformational leadership orientation than their Afghan colleagues. Younger, as well as male, Afghans had a significantly higher tendency toward a transformational leadership orientation than their American counterparts.

It is important for transformational leaders to understand the organizational culture, provide fair promotional opportunities for both males and females, remain ethical and expect high moral standards, respect people of all ages, and gain the trust of subordinates in order to successfully compete in today's global economy.

REFERENCES

[1] J. M. Kouzes and B. Z. Posner, "Encouraging the heart," Jossey-Bass, San Francisco, 2003.

[2] J. Rhodes, P. Walsh, and P. Lok, "Convergence and divergence issues in strategic management – Indonesia's experience with the balanced scorecard in HR managementm," The International Journal of Human resources Management, Vol. 19, No. 6, pp. 1170–1185, 2008.

[3] J. Pounder, "Transformational leadership: Practicing what we teach in the Management classroom," Journal of Education for Business, Vol. 84, No. 1, pp. 2–6, 2008.

[4] R. Vecchio, J. Justin, and C. Pearce, "The utility of transactional and transformational leadership for predicting performance and satisfaction within a path-goal theory framework," Journal of Occupational and Organizational Psychology, Vol. 81, No. 1, pp. 71–82, 2008.

[5] A. Franken, C. Edwards, and R. Lambert, "Executing strategic change: Understanding the critical management Elements that lead to success," California Management Review, Vol. 51, No. 3, pp. 49–73, 2009.

[6] B. G. Mujtaba and B. A. Kaifi, "Afghan and American professionals' leadership orientation toward tasks and relationships: Are there tendencies toward convergence or divergence?" Fortune Journal of International Management, Vol. 5, No. 1, pp. 107–125, 2008.

[7] F. Walumbwa, B. Avolio, and W. Zhu, "How transformational leadership weaves its influence on individual job performance: The role of identification and efficacy beliefs," Personnel Psychology, Vol. 61, No. 4, pp. 793–825, 2008.

[8] B. Kirkman, G. Chen, J. Farh, Z. Chen, and K. Lowe, "Individual power distance orientation and follower reactions to transformational leaders: A cross level, cross-cultural examination," Academy of Management Journal, Vol. 52, No. 4, pp. 744–764, 2009.

[9] P. Ricoeur, "Oneself as another," University of Chicago Press, Chicago, 1992.

[10] P. Ricoeur, "The narrative function," John B. Thompson, Ed., Cambridge University Press, Cambridge, pp. 274–296, 1982

[11] Y. Kim, "A comparative study of the 'Abilene Paradox' and 'Groupthink'," Public Administration Quarterly, Vol. 25, No. 2, pp. 168–189, 2001.

[12] A. Laszlo, K. Laszlo, and C. Johnsen, "From high-performance teams to evolutionary learning communities: New pathways in organizational development," Journal of Organizational Transformation and Social Change, Vol. 6, No. 1, pp. 29–48, 2009.

[13] B. A. Kaifi, "21st century leadership in healthcare," Chapter Twelve in the Pharmaceutical Technician Laboratory Manual by Sandeep Bansal, Jones and Bartlett Publications, Boston, Massachusetts, pp. 90–100, 2009

[14] L. Daniel and C. Davis, "What makes high-performance teams excel?" Research Technology Management, Vol. 52, No. 4, pp. 40–45, 2009.

[15] M. Pryor, L. Singleton, S. Taneja, and L. Toombs, "Teaming as a strategic and tactical tool: An analysis with recommendations," International Journal of Management, Vol. 26, No. 2, pp. 320–333, 2009.

[16] E. Kearney, "Age differences between leader and followers as a moderator of the relationship between transformational leadership and team performance," Journal of Occupational & Organizational Psychology, Vol. 81, No. 4, pp. 803–811, 2008.

[17] J. Winter, J. Neal, and K. Waner, "How male, female, and mixed-gender groups regard interaction and leadership differences in the business communication course," Business Communication Quarterly, Vol. 64, No. 3, pp. 43–58. 2001.

[18] P. Lantz, "Gender and leadership in healthcare administration: 21st century progress and challenges," Journal of Healthcare Management, Vol. 53, No. 5, pp. 291–301. 2008.

[19] L. Babcock, "What happens when women don't ask?" Negotiation, Vol. 11, No. 96, pp. 1–4, 2008.

[20] D. Beckman and L. Menkhoff, "Will women be women? Analyzing the gender difference among financial experts," Kyklos, Vol. 61, No. 3, pp. 364–384, 2008.

[21] D. Wyld, "How do women fare when the promotion rules change?" Academy of Management Perspectives, Vol. 22, No. 4, pp. 83–85, 2008.

[22] E. Cabrera, "Fixing the leaky pipeline: Five ways to retain female talent," People and Strategy, Vol. 32, No. 1, pp. 40–45, 2009.

[23] S. Portillo and L. Dehart-Davis, "Gender and organizational rule abidance," Public Administration Review, Vol. 69, No. 2, pp. 339–347, 2009.

[24] M. Van Vugt, "Evolutionary origins of leadership and followership," Personality& Social Psychology Review, Vol. 10, No. 4, pp. 354–371, 2006.

[25] W. Altman, "Executive coaching comes of age," Engineering Management, Vol. 17, No. 5, pp. 26–29, 2007.

[26] M. Ashworth, "Preserving knowledge legacies: workforce aging, turnover and human resource issues in the US electric power industry," International Journal of Human Resource Management, Vol. 17, No. 9, pp. 1659–1688, 2006.

[27] B. A. Kaifi, "A transformational leader," Sitara Magazine, Vol. 1, No. 4, pp. 48–50, 2008.

[28] G. Karadjova-Stoev and B. G. Mujtaba, "Strategic human resource management and global expansion lessons from the Euro disney challenges in France," International Business and Economics Research Journal, Vol. 8, No. 1, pp. 69–78, 2009.

[29] B. A. Kaifi, "The impact of 9/11 on Afghan-American leaders," Xlibris, Bloomington, 2009.

[30] T. Cooper, "The responsible administrator (4th ed.)," Jossey-Bass, San Francisco, CA, 1998.

[31] B. A. Kaifi and B. G. Mujtaba, "Workforce discrimination: An inquiry on the perspectives of Afghan-American professionals," Journal of Business Studies Quarterly, Vol. 1, No. 1, pp. 1–15, 2009.

[32] T. Rath and B. Conchie, "Strengths based leadership," Gallup Press, NY, 2009.

[33] F. Morhart, W. Herzog, and T. Tomczak, "Brand-specific leadership: Turning employees into brand champions," Journal of Marketing, Vol. 73, No. 5, pp. 122–142, 2009.

[34] M. Carter, A. Jones-Farmer, A. Armenakis, H. Field, and D. Svyantek, "Transformational leadership and followers' performance: Joint mediating effects of leader-member exchange and interactional justice," Academy of Management Conference Proceedings, Chicago, United States, 2009.

[35] E. Fisher, "Motivation and leadership in social work management: A review of theories and related studies," Administration in Social Work, Vol. 33, No. 4, pp. 347–367, 2009.

[36] Y. Gong, J. Haung, and J. Farh, "Employee learning orientation, transformational leadership, and employee creativity: The mediating role of employee creative self-efficacy," Academy of Management Journal, Vol. 52, No. 4, pp. 7765–7778. 2009

Appendix A

Transformational Leadership Survey

The modified MLQ (Multifactor Leadership Questionnaire) Form 5X-Short measures your leadership capabilities based on different factors related to Transformational leadership. To determine whether you are a transformational leader, circle one of the following options that best describe how you see yourself (or the person that is being evaluated) regarding each statement. For each statement, you can indicate the degree to which you (or the person being evaluated) engage (s) in the stated behavior. A rating of 1 means Never and a rating of 5 means Always with the person demonstrating the specific behavior.

	Questions	Never…….Always				
1.	I express with a few simple words what we could and should do.	1	2	3	4	5
2.	I provide appealing images about what we can do.	1	2	3	4	5
3.	I help others find meaning in their work.	1	2	3	4	5
4.	I enable others to think about old problems in new ways.	1	2	3	4	5
5.	I provide others with new ways of looking at puzzling things.	1	2	3	4	5
6.	I get others to rethink ideas that they had never questioned before.	1	2	3	4	5
7.	I make others feel good to be around me.	1	2	3	4	5
8.	I tell others the standards they have to know to carry out their work.	1	2	3	4	5
9.	I call attention to what others can get for what they accomplish.	1	2	3	4	5
10.	Others have complete faith in me.	1	2	3	4	5
	Total Score:					

J. Service Science & Management, **2010, 3: 159-164**
doi:10.4236/jssm.2010.31020 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Culture and Knowledge Transfe: Theoretical Considerations

Jing Liu

School of Economics and Management, Northwest University of Politics and Law, Xi'an, China.
Email: ljta.0417@yahoo.com.cn

Received October 30st, 2009; revised December 19th, 2009; accepted January 27th, 2010.

ABSTRACT

Knowledge and culture are indissolubly linked together in organizations. Considerable evidence supports the importance of culture in the success or failure of knowledge management. Then, the effectiveness of knowledge transfer needs both cultural understanding and new considerations in the knowledge transfer of intercultural organizations. This paper identifies the influence of different dimensions of culture on knowledge transfer in different types of knowledge. Based on a topology that classified national culture into four dimensions provided by [12,13]*, power distance and individualism/collectivism are chosen as the representatives of national culture in this work and discuss the effect of national culture on knowledge transfer. The aim of this paper is to propose a theoretical framework for knowledge transfer processes based on differences in national culture for future research.*

Keywords*: *Knowledge, Knowledge Types, Knowledge Transfer, Culture, Cultural Dimensions

1. Introduction

There are diverging opinions the question of whether culture might influence knowledge management. Reference [1] believes that there is a convergence of approaches with regards to knowledge management and that effective knowledge management may evolve to become a universal concept. Several researchers have found no evidence that differences in national culture have an affect on knowledge management practices [2–4]. But there are considerable evidence supports the importance of culture in the success or failure of knowledge management within organizations. Reference [5] is very critical of the viewpoints ignoring culture influence on knowledge management. He point out that these viewpoints gives the impression that knowledge management operates in a kind of unitary vacuum in which diversity in terms of language, cultural and ethnic background are compressed into one giant independent variable which is in any case pushed to the side. Reference [6] agree and state that knowledge management models that exclude the influence of national and regional culture seriously undercut their potential effectiveness particularly in global applications. They suggest that "cultural bias exists in data bases and in all business and innovation" and that "western analytical assumptions about knowledge and information management, dominates both information and knowledge management research and development". Recently, a few researchers have found empirical evidence that differences in national culture do affect knowledge sharing [7,8].

Knowledge and culture are indissolubly linked together in organizations. Recent technological revolution, accompanied by rapid globalization [9], has led to increased cultural heterogeneity within organizations. As the world becomes more and more globalized, western organizations now have access to a pool of job candidates from increasingly diverse cultural backgrounds [10]. National borders no longer preclude individuals of different cultures from working in international organizations. Consequently, organizations today exhibit more cultural diversity among their employees. Simultaneously, advancing globalization is forcing organizations to engage in alliances and networks with partners with widely diverse national or ethnic cultural backgrounds. The differences in the cultural orientation of the collaborating organizations increase the risk of misunderstandings and conflicts, and often lead to failure, but if managed in a balanced manner may also improve performance [11]. To overcome these barriers to success, we need both cultural understanding and new considerations in the knowledge transfer of intercultural organizations.

The effectiveness of knowledge transfer is directly related to the type of knowledge involved in the transfer

process. In addition, the transfer of knowledge is moderated by 1) the nature of transacting cultural patterns and 2) the cognitive styles of the individuals.

The paper comprises five sections. In the next section, the author will introduce a conceptual framework for different types of knowledge and discuss the effect they impose for knowledge transfer. The third section first discusses the characteristics of culture and then proposes a classification scheme based on a topology that classified national culture into four dimensions provided by [12,13]. In the fourth section, power distance and individualism/collectivism are chosen as the representatives of national culture in this work and discuss the effect of national culture on knowledge transfer. A discussion on the theoretical and managerial implications concludes the study.

2. Theoretical Considerations: Knowledge and Knowledge Types

In order to articulate knowledge transfer we need a basic conceptualization of the concept of knowledge. The academic question of how knowledge should best be defined is a subject of a lively epistemological debate. The complex nature of knowledge has been discussed extensively in information technology (IT), strategic management, organizational theory and knowledge management literature. Reviewing crucial literature, principally there are two approaches to defining knowledge. One uses the concept of a value chain or hierarchical structure among data, information, and knowledge. The other focuses on the analysis of the process of knowing.

The most common way to describe knowledge is to distinguish it from data and information [14,15]. Reference [16] suggests that knowledge is authenticated information and information is interpreted data. Reference [17] regards data as carrier of information and knowledge, information as relating to descriptive and historical fact, and knowledge as new or modified insight or predictive understanding. Reference [18] defines data as observation or facts, with information as data in a meaningful context and knowledge as meaningfully organized accumulation of information. Reference [19] regards knowledge as a production that is made from raw material - information. Reference [20] argues that data can be classified as raw numbers, images, words, and sounds derived from observation or measurement. Information represents data arranged in a meaningful pattern. Unlike information, knowledge is about beliefs, commitment, perspectives, intention and action. The common factor of those definitions is that knowledge is located at the top of a hierarchical structure.

Another thought defines knowledge as a process related to application [18–22]. Reference [23] identifies both justified belief and commitment anchored to the overall epistemological structure of the holder as key ingredients of knowledge. Reference [15] further adds to this definition of Nonaka and Takeuchi that to know is to be able to take part in the process that makes the knowledge meaningful. Reference [24] concludes that knowledge is a high-value form of information that is ready to be applied to decisions and actions.

One impact of these definitional differences occurs when discussing knowledge transfer. The differences in viewpoints on knowledge suggest different implications for knowledge transfer. It is common to consider knowledge as arranged in a knowledge hierarchy, where data is transformed into information, and information is transposed into knowledge.

A further key question of knowledge transfer research concerns the relationship and interaction among different types of knowledge. Reference [25] note that there are at least three distinct types of knowledge: human knowledge, social knowledge, and structured knowledge. Human knowledge constitutes what individuals know or know how to do, is manifested in important skills, and usually comprises both explicit and tacit knowledge. It could be conceptual or abstract in orientation. Social knowledge exists in relationships among individuals or within groups. Social or collective knowledge is largely tacit, composed of cultural norms that exist as a result of working together, and its salience is reflected in our ability to collaborate and develop transactional relationships. Structured knowledge is embedded in organizational systems, processes, rules, and routines. This kind of knowledge is explicit and rule based and can exist independently of the knowers [26].

These three types of knowledge work in concert with terms of the three dimensions of knowledge, proposed by [27]: simple versus complex, explicit versus tacit, and independent versus systemic. The first dimension- simplicity versus complexity-is relevant in cross-border knowledge transactions. Complex knowledge evokes more causal uncertainties and conveys such types of knowledge required amount of factual information. Simple knowledge can be captured with little information and is, therefore, relatively easy to transfer. The explicit versus tacit dimension concerns how well articulated or implicit the knowledge is. The transfer of tacit knowledge requires richer context and richer media, because tacit knowledge requires more than just codification. Explicit knowledge, however, can be codified and is transferred with relative ease. The third dimension of knowledge deals with the independent versus systemic character of knowledge-that is, the extent to which the knowledge is embedded in the organizational context. Knowledge that is independent can be described by itself, whereas knowledge that is systemic must be described in relation to a body of knowledge existing in the transferring organization.

Using these dimensions, human knowledge can be conceptualized as either simple or complex, as tacit or explicit (or both), and, generally, as more independent or systemic. Social knowledge can be either simple or complex and is largely tacit and systemic in character. Structured knowledge is either simple or complex, is usually more explicit than tacit, and is largely systemic in character. "Sticky" knowledge [28], which is more complex, tacit, and systemic, is more difficult to transfer, regardless of cultural differences. Some combinations of human, social, and structured knowledge can take on the character of sticky knowledge and become even more difficult to transfer, regardless of the cultural differences involved between the transacting organizations.

Reference [27] argue that the position of knowledge along each of the three dimensions affects the amount of information required to describe it and the amount of effort needed to transfer it. Therefore, it is more difficult to transfer and to absorb if the type of knowledge (human, social, or structured) being transferred is tacit, complex, and systemic.

All of these criteria of effective knowledge transfer are affected when knowledge transfers involve transacting organizations that are located in dissimilar cultural contexts. Cross-border transfer of organizational knowledge is most effective when the type of knowledge (i.e., human, social, or structured) being transferred is simple, explicit, and independent and when such transfers involve similar cultural contexts. In contrast, transfer is least effective when the type of knowledge being transferred is complex, tacit, and systemic and involves dissimilar cultural contexts.

3. Culture and Cultural Dimensions

The type of knowledge being transferred is the most important antecedent of effectiveness. However, it should be noted that there are strong interactions between cultural patterns and cognitive styles. In addition, some cultural contexts might foster some cognitive styles that are uniquely responsible for the evolution and practice of certain types of organizational knowledge, compared to other cultural contexts, which might emphasize different styles.

Culture is "a pattern of basic assumptions -- invented, discovered, or developed by a given group as it learns to cope with its problems of external adaptation and internal integration - that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think and feel in relation to those problems" [29]. Researchers suggest different manifestation of culture. Reference [25] notes that values, norms, and practices are reflections of culture, while [30] categorizes culture into values and practices subsuming symbols, heroes, and rituals. No matter how researchers define culture, however, there is a common view that culture has at least two layers: the inner layer and the core. The core of culture is value, which is described as a fairly stable emotional tendency to respond consistently to some specific object, situation, person or category of people [31]. It's an invisible, unconscious, and embedded basic feeling that is manifested in the outer visible layer of culture, such as attitudes and practices, and in alternatives of behaviors [25–32]. The key role of culture in organizations is creating a consensually validated system of beliefs and values which influences organizational behavior [33].

Culture can be applied to different dimensions, such as nations, organizations, religious groups, and so on. National culture (external culture) and organizational culture (internal culture) are widely accepted as important cultural dimensions for organizations. National (external) culture is national, regional, composed of values, common perceptions, similar views of reality, while organizational (internal) culture is emerging from group mechanics, relevant in understanding the sub-populations who make up the firm [33]. These two dimensions have been regard as a dominant influence on organizational behaviors. National culture is believed to play significant roles in determining the efficacy of knowledge transfer within the same organization that cross different national borders and cultures [34]. Reference [12,13] provided a topology that classified national culture into four dimensions: Power Distance, Individualism/Collectivism, Uncertainty Avoidance, and Masculinity/Femininity. This topology is being adopted for the current study because it provides the most rich and well articulated conceptualization of culture available.

Power distance Power distance can be conceptualized as the degree of separation between individuals at adjacent levels of rank. Individuals who score highly on power distance place a high value on societal hierarchy, while individuals who score low value societal hierarchy less [13–35]. Norms and customs in high power distance cultures include centralized decision making at the top, showing a great deal of respect for individuals with higher rank [36], and a tendency to form bureaucratic organizations [13].

Individualism/Collectivism Several scholars such as [37] regard the individualism-collectivism dimension of cultural variation as the major distinguishing characteristic in the way that different societies analyze social behavior and process information. Reference [38,39] has defined individualism as a cultural pattern consisting of loosely linked individuals who view themselves as independent with their own preferences, needs, rights, and contracts, whereas collectivism refers to a cultural pattern that consists of closely linked individuals who see themselves as belonging to one or more collectives (e.g., family, organizations) and who are motivated by the norms, duties, and obligations thus imposed. People are

inclined to give priority to the goals of these collectives over their own personal goals. Reference [37] argues that the collectivism-individualism dimension strongly influences what kind of information people prefer and are more prepared to process.

Uncertainty Avoidance Uncertainty avoidance can be conceptualized as the propensity of individuals to avoid actions where the outcome is unclear. Customs in cultures with high uncertainty avoidance include dichotomization (conceptualizing people and situations as either good or bad), modularation and compartmentalization of tasks, in an attempt to simplify them [13–40].

Masculinity/Femininity The concept of masculinity is associated with the competitiveness of individuals. Masculine individuals value ambition and the acquisition of wealth, while feminine individuals value nurturing and quality of life. Masculine individuals typically believe that failure is catastrophic, while feminine individuals see failure as common and find it easier to move on.

4. National Culture and Knowledge Transfer

National culture is a crucial factor in knowledge transfer. Cultures shape the value of both managers and employees. Cultural differences evoke subtle yet powerfully different managerial behaviors and leadership styles [41]. Such behaviors and leadership styles provide the organizational context within which employees transfer their knowledge to one another.

As mentioned in above, national culture can be classified into four dimensions, which are Power distance, Individualism/Collectivism, Uncertainty avoidance, Masulinity/Femininity. These dimensions determine assumptions and behaviors of managers and employees in the process of knowledge transfer. According to several scholars [13–42], power distance and individualism/collectivism are the primary distinctions between North America (Canada and US) and China (Hong Kong and Mainland China). Reference [13] indicates that, the power distance scores of people in Hong Kong and China are much higher than that of people in North America, while the individualism scores of people in Hong Kong and China are much lower than that of people in North America. Thus, power distance and individualism/collectivism are chosen as the representatives of national culture in this work and discuss in detail in the following.

Power Distance Power distance deals with leaders' decision power. It affects both the way in which people organize themselves and the way in which they write about organizing [12]. Power Distance is the degree to which people accept and expect unequal authority. Individuals who score high on power distance believe that supervisors should maintain decision making authority, receive credit for success, and that supervisors deserve respect and admiration from subordinates. Conversely, individuals who score low on power distance believe that the supervisor and the subordinate are colleagues, working toward the same goal, and are similar in terms of respectability. The superior position will improve decision power in high-power-distance culture while hard-working, good work and experiences are ways to increase decision power in low-power-distance culture [12–44].

Based on the above analysis, people from different Power Distance societies will act differently toward authority: the larger the power distance is, the more people would accept unequal authority. Hence, managers with different cultural backgrounds might play different roles in the process of inter-organizational knowledge transfer.

Individualism/Collectivism In individualistic societies, members have less respect and loyalty to the group they belong to than members of collectivist societies have. They prefer to stand on their own feet, favor independent work, emphasize competition and achieving specific statuses, and have a calculated involvement in group affairs. In contrast, members of collectivist societies respect and remain loyal to their group and emphasize cooperation and group work. They prefer low internal competition, relationships, harmony, order and discipline [42] and favor cooperation and teamwork [12–45].

Cultures shape the norms that define the context for social interaction [25]. Individualism and collectivism strongly influence ways of thinking. Specifically, they influence how members of a culture process, interpret, and make use of a body of information and knowledge [37]. Collectivists maintain respect, harmony, and loyalty to the groups they belong to and support order, discipline and centralized authority vested at the top. They are more likely to obey managers' orders and go along with their managers' wills. Furthermore, employees in collectivist societies rely on their supervisors while those in individualistic societies prefer to get help from their peers [46]. In addition, workers in individualist societies envision knowledge creation as an intervention of individual effort while workers in collectivist societies think of the integration and modification of existing knowledge as a group effort [46]. Thus, managers in collectivist societies will more thoroughly create the right context for knowledge sharing among different groups and better harmonize differences among the involved groups.

5. Theoretical Conclusions, Limitations, and Future Research

Knowledge and culture are indissolubly linked together in organizations. Considerable evidence supports the importance of culture in the success or failure of knowledge management.

Starting from the basic concept of the culture and knowledge and basing on the type of knowledge and the dimension of culture both influence knowledge-sharing, power distance and individualism/collectivism are cho-

sen as the representatives of national culture in this work and discuss the effect of national culture on knowledge transfer.

This paper proposes two abstract conclusions: 1) people from different Power Distance societies will act differently toward authority: the larger the power distance is, the more people would accept unequal authority. Hence, managers with different cultural backgrounds might play different roles in the process of inter-organizational knowledge transfer. 2) Individualism and collectivism strongly influence ways of thinking. Collectivists maintain respect, harmony, and loyalty to the groups they belong to and support order, discipline and centralized authority vested at the top. Managers in collectivist societies will more thoroughly create the right context for knowledge sharing among different groups and better harmonize differences among the involved groups.

The research is limited by its scope because it focuses on the transfer processes. Further research might examine the culture factor influencing knowledge management in other knowledge management processes. There is also an unclear detailed relationship between knowledge management performance and knowledge management decision based on different culture dimensions. Further research might examine the relationship between organization performance and knowledge management decision based on different culture dimensions, as well as empirical research on the cultural conditions that lead to appropriate and inappropriate adaptation.

REFERENCES

[1] H. Takeuchi, "Towards a universal management of the concept of knowledge, managing industrial knowledge: Creation, transfer and utilization," in I. Nonaka and D. J. Teece, Eds. London: Sage, 2001.

[2] B. L. Simonin, "Ambiguity and the process of knowledge transfer in strategic alliances," Strategic Management Journal, Vol. 20, No. 7, pp. 595–623, July 1999.

[3] A. K. Gupta and V. Govindarajan, "Knowledge flows within multinational corporations," Strategic Management Journal, Vol. 21, No. 4, pp. 473–496, April 2000.

[4] R. Jensen and G. Szulanski, "Stickiness and the adaptation of organizational practices in cross-border knowledge transfers," Journal of International Business Studies, Vol. 35, No. 6, pp. 508–523, November 2004.

[5] N. Holden, "Knowledge management: Raising the specter of the cross-cultural dimension," Knowledge and Process Management, Vol. 8, No. 3, pp. 155–163, July/September 2001.

[6] D. J. Pauleen and P. Murphy, "In praise of cultural bias," MIT Sloan Management Review, Vol. 46, No. 2, pp. 21–22, Winter 2005.

[7] S. C. Voel and C. Han, "Managing knowledge sharing in China: The case of Siemens ShareNet," Journal of Knowledge Management, Vol. 9, No. 3, pp. 51–63, 2005.

[8] N. Finestone and R. Snyman, "Corporate south africa: Making multicultural knowledge sharing work," Journal of Knowledge Management, Vol. 9, No. 3, pp. 128–141, 2005.

[9] C. Grimm and K. Smith, "Strategy as action: 1ndustry rivalry and coordination," Cincinnati, OH: South-Western College Publishing, 1997.

[10] T. L. Friedman, "The world is flat: A brief history of the twenty-first century," New York, NY: Farrar, Straus and Giroux, 2005.

[11] J. Li, L. Karakowsky and K. Lam, "East meets east and east meets west: The case of Sino-Japanese and Sino-West joint ventures in China," Journal of Management Studies, Vol. 39, No. 6, pp. 841–863, September 2002.

[12] Hofdtese, "Management scientists are human," Management Science, Vol. 40, No. 1, pp. 4–13, January, 1994.

[13] G. Hofstede, "Culture's Consequences: International Differences in Work-related Values (2nd Edtion)," Thousand Oaks, CA: Sage, 2001.

[14] I. Nonaka, "A dynamic theory of organizational knowledge creation," Organization Science, Vol. 5, No. 1, pp. 14–37, February 1994.

[15] J. C. Spender, "Making knowledge as the basis of a dynamic theory of the firm," Strategic Management Journal, Special Issues: Knowledge and the Firm, Vol. 17, pp. 45–62, Winter 1996.

[16] D. Vance, "Information knowledge and wisdom: The epistemic hierarchy and computer-based information systems," in Proceedings of the Third America's Conference on Information Systems, Indianapolis, August 1997.

[17] N. Kock and R. McQueen, "Knowledge and information communication in organizations: An analysis of core, support and improvement process," Knowledge and Process Management, Vol. 5, No. 1, pp. 29–40, March 1998.

[18] M. Zack, "Managing codified knowledge," Sloan Management Review, Vol. 40, No. 4, pp. 45–58, Summer 1999.

[19] F. I. Dretske, "Knowledge and the flow of information," Stanford, CA: CLSI Publications, 1999.

[20] I. Nonaka and V. Peltokorpi, "Objectivity and subjectivity in knowledge management-a review of 20 top articles," Knowledge and Process Management, Vol. 13, No. 2, pp. 73–82, 2006.

[21] R. McDermot, "Why information technology inspired but cannot deliver knowledge management," California Management Review, Vol. 41, No.4, pp. 103–117, Summer 1999.

[22] C. Frappaolo and S. Capshaw, "Knowledge management software: Capturing the essence of know-how and innovation," Information Management Journal, Vol. 33 No. 3, pp. 44–48, July 1999.

[23] I. Nonaka and H. Takeuchi, "The knowledge creating company," New York, NY: Oxford University Press,

1995.

[24] T. H. Davenport, D. W. Long, and M. C. Beers, "Successful knowledge management projects," Sloan Management Review, Vol. 39, No. 2, pp. 43–57, Winter 1998.

[25] D. W. De Long and L. Fahey, "Diagnosing cultural barriers to knowledge management," Academy of Management Executive, Vol. 14, No. 4, pp. 113–128, November 2000.

[26] R. Glazer, "Measuring the knower: Towards a theory of knowledge equity," California Management Review, Spring, Vol. 40, No. 3, pp. 175–194, 1998.

[27] R. Garud and P. R. Nayyar, "Transformative capacity: Continual structuring by intertemporal technology transfer," Strategic Management Journal, Vol. 15, No. 5, pp. 365–385, June 1994.

[28] G. Szulanski, "Exploring internal stickiness: Impediments to the transfer of best practice within the firm," Strategic Management Journal, Special Issue: Knowledge and the Firm, pp. 27–44, Vol. 17, Winter 1996.

[29] E. H. Schein, "Organizational culture and leadership," San Francisco, CA: Jossy-Bass Inc, 1985.

[30] G. Hofstede, B. Neuijen, D. D. Ohayv, and G. Sanders, "Measuring organizational cultures: A qualitative and quantitative study across twenty cases," Administrative Science Quarterly, Vol. 35, No. 2, pp. 286–316, June 1990.

[31] G. Johns, "Organizational behavior: Understanding and managing life at work," New York, NY: HarperCollins College Publishers, 1996.

[32] G. Hofstede, "Attitudes, values and organizational culture: Disentangling the concepts," Organization Studies, Vol. 19, No. 3, pp. 477–492, 1998.

[33] P. X. Meschi and A. Roger, "Cultural context and social effectiveness in international joint ventures," Management International Review, Vol. 34, No. 3, pp. 197–216, 1994.

[34] B. L. Kedia and R. S. Bhagat, "Cultural constraints on transfer of technology across nations: Implications for research in international and comparative management," Academy of Management Review, Vol. 13, No. 4, pp. 559–571, October 1988.

[35] R. T. Watson, T. H. Ho and K. S. Raman, "Culture: A fourth dimension of group support systems," Communications of the ACM, Vol. 37, No. 10, pp. 44–55, October 1994.

[36] M. Srite and E. Karahanna, "The role of espoused national culture values in technology acceptance," MIS Quarterly, Vol. 30, No. 3, pp. 679–704, September 2006.

[37] R. S. Bhagat, B. L. Kedia, P. D. Harveston, and H. C. Triandis, "Cultural variations in the cross-border transfer of organizational knowledge: An integrative framework," Academy of Management Review, Vol. 27, No. 2, pp. 204–221, April 2002.

[38] H. C. Triandis, "Individualism and collectivism," Boulder, CO: Westview, 1995.

[39] C. Triandis, "Vertical and horizontal individualism and collectivism: Theory and research implications for international comparative management," in Advances in International Comparative Management, J. L. Cheng and R. B. Peterson, Eds., Stanford, CT and London: JAI Press Inc, Vol. 12, pp. 7–35, 1998.

[40] E. Hall, "Beyond Culture (2"[d] Edtion)," New York, NY: Anchor Books/Random House, 1989.

[41] J. Michael, "A conceptual framework for aligning managerial behaviors with cultural work values," International Journal of Commerce & Management, Vol. 7, No. 3/4, pp. 81–102, 1997.

[42] I. Vertinsky, D. K. Tse, D. A. Wehrung, and K. Lee, "Organizational design and management norms: A comparatives study of managers' perceptions in the People's Republic of China, Hong Kong, and Canada," Journal of Management, Vol. 16, No. 4, pp. 853–867, December1990.

[43] C. W. Chow, T. M. Lindquist, and A. Wu, "National culture and the implementation of high-stretch performance standards: An exploratory study," Behavioral Research in Accounting, Vol. 13, pp. 85–109, 2001.

[44] M. C. Marchese, "Matching management practices to national culture in India, Mexico, Poland, and the U.S," The Academy of Management Executive, Vol. 15, No. 2, pp. 130–132, May 2001.

[45] T. H. Cox, S. A. Lobel, and P. L. McLeod, "Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task," Academy of Management Journal, Vol. 34, No. 4, pp. 827–847, December 1991.

[46] Y. Yoo and B. Torrey, "National culture and knowledge management in a global learning organization: The strategic management of intellectual capital and organizational knowledge," New York, NY: Oxford University Press, Inc., 2002.

J. Service Science & Management, **2010, 3: 165-172**
doi:10.4236/jssm.2010.31021 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Study on Coalmine Safety Behavior Control System Model and its Manager Behavior Choice

Qingren Cao[1], Xiaobing Yang[2]

[1]School of Management, China University of Mining and Technology, Xuzhou, China; [2]School of Art, Xuzhou Institute of Technology, Xuzhou, China.
Email: caoqr-cumt@163.com, xbyang_711@yahoo.com.cn

Received May 17th, 2009; revised December 20th, 2009; accepted January 15th, 2010.

ABSTRACT

To control coal miner's unsafe behavior, miner and manager behaviors were analyzed, and a behavior control system model was constructed. The necessity conditions of optimal management behavior of coalmine manager in both complete information and incomplete information condition were studied. The result established that, only if the estimated value of unsafe behavior by manager was the same as the estimated value of unsafe behavior by miner, namely both miner's behavior control and manager own behavior control were optimal, the optimal behavior control of coalmine manager were achieved. Because some factors influenced manager's behavior unavoidably, it is very difficult for coalmine manager to achieve his optimal behavior choice. It is the underlying reason for many managerial defects, even errors inherent in the coalmine safety management.

***Keywords*:** *Coalmine Manager, Behavior Control System, Optimization Model, Unsafe Behavior, Manager Behavior Choice*

1. Introduction

Recently, major casualty accidents took place repeatedly in China has been paid attention all around the world. By investigating these accidents carefully, it was found that many managers supervised coal miner only by their experiences, and there were few scientific theories supporting their management behaviors. These behaviors are no doubt harmful to the coalmine safety management, but this is very common in China, so it is inevitable there are considerable managerial defects and errors existed in coalmine safety management. It was for this underlying reason that has caused casualty accidents took place repeatedly in China.

As we all know, the characteristic of coal production determines there are full of potentially dangerous in coal production process. In order to prevent possible accidents or reduce their harmful degree to an acceptable level, it is needed to remove the factors that might conduce to serious injury or loss of life by implementing various safety management measures, such as managing the essential factors that influence miner to do unsafe behavior. However, as one composition of coalmine safety management, behavior management is more important to physical management, because miner's behavior is a more active ingredient in the coal production process, and the miner's unsafe behavior is the primary and direct factors leading to coalmine accidents. In addition, there are always behavior mistakes underlying unsafe physical status. So the unsafe behavior management is always a more important method to prevent accidents in coalmine safety management [1].

In the coal production process, miner's behavior and manager's behavior are very different on the view of management. Although miner's unsafe behavior is the direct contributing factors to accidents, the miner's behavior is a management objective of the manager's behavior. So studying manager's behavior and its characteristics is even more important.

For a coalmine manager, there are a lot of available management measures, but different management measures have different effects. If a manager could not properly make use of his limited time, energy and other available resources, his management measures are difficult to achieve some certain or better results [2]. So the coalmine manager will be facing a theoretic problem how to make the right choice of different management measures scientifically.

In fact, many scholars have studied this kind of behavior choice problem for ages. Mintzberg may be the pioneer to study this kind of question. He divided the manager behavior into several work roles, and studied the relationship of time spent on these different roles and the work results [3].

Allan had surveyed 1476 New York City government managers for their more important jobs and time spent on them. He discovered that there are 57 core jobs needed to be accomplished in different positions or different departments. Along with the position moving up, the duty becomes extended accordingly, so the middle managers and the senior managers have 80 and 87 core jobs to be done respectively [4]. Yuk also studied this issue by similar method, but he divided the manager work jobs into 22 items [5].

Similar to these scholars, Schaechtel disclosed that safety management is only one job for a manager in many organizations, so how to force a manager pay more attention to the safety management work all the time is very important [6]. He deemed that it needed to construct a safety management system framework to measure or evaluate manager's safety performance. This framework should meet the requirement that the safety standard is delivered to every employee distinctly, the duty to implement standard is understood and accept, there are documents to record the implemented standard, and there are effective internal control mechanisms. He also brought forward a method to construct this kind of safety management system.

But Jens deemed that a risk management system consists of different members, such as lawmakers, supervisors, planners, and manipulators [7]. In the past, the system problems were studied in different areas respectively. To construct a control system model needed to generalize or sum up all these systems or their specific dangers. In effect, to construct a risk management system model, we must regard the risk management as a control problem and study them crossways. Only in that way can we consider all the relevant problems.

All these studies show that manager has many jobs to do, and he needed to make choice rightly because different behavior choice leads to different management effects. Unlike these studies, a behavior control system model was constructed by summarizing the influence factors of coal miner and manager from the point of behavior control mechanism in this paper [8]. How to make the optimal control behavior choice for coalmine manager was analyzed theoretically, the conditions of making the optimal control behavior choice were induced, and the characteristics of optimal control behavior were summed up. This is the theoretical basis for empirical research of coalmine manager's control behavior.

2. Coal Miner Behavior System

For a long time, many scholars have studied the factors that influence people to make unsafe behavior choice and behavior system. These studies help us to understand all kinds of human behaviors, and provide us with precondition of engaging in the safety management study nowadays.

Specific to the miner behavior, Jason has studied many variables to influence safety management measures put into practice, such as the miner's sensation to effectiveness of behavior safety training, the miner's trust to the ability of manager, the manager's responsibility to performance evaluation, the miner's relevant education experience, the manager's work time in the organization, etc. [9].

After detailing investigation, Michael discovered that several factors influenced the employee's sensation, such as the manager's duty [10]. By making use of these factors carefully, the manager could reduce the accidents to the lowest level. In the meantime, Geller studied the sensation, perception, pressure, suffering and complexity of human being, analyzed the behavior safety and its training on individual, and put forward four behavior intervention avenues in accordance with three kinds of behavior manner [11].

All these studies could help us to understand the miner behavior very well. But after all, the coal production task is fulfilled by the miner's actual production behavior. During his work, the miner's behavior is influenced or stimulated by many extrinsic factors. These factors may make the miner's mental state transferring continually by his perception. Furthermore they may impact the estimated value of behavior factors, which decide he is to adopt safe or unsafe behavior.

There are a lot of extrinsic factors influencing the coal miner behavior. These factors may be summed up five aspects, i.e. work circumstance, tasks and behavior criterion, incentive, rewards and punishment behavior, education and training, and group and organization behavior [12–14]. They also can be regarded as the input variables of coal miner behavior system. The first aspect indicates work objective and work circumstance of coal miner, but manager can change it by work design. The latter four aspects indicate the control measures available to the manager.

Coal miner behavior may be described as behavioral state and state transformation. In the coal production process, many factors stimulated miner behavior continuously, inspired different behavior motivation, and made the choice of different behaviors. It can be regarded as transformation from one state to another, meanwhile completing the coalmine production task.

The inner factors are fundamental to coal miner behavior. They may be summed up five aspects, i.e. miner's mental state, knowledge state, physical state, organizational loyalty and work effort. All of these factors together may be described as miner's safety consciousness and work capability that can be used as the miner's state variables [15]. The output of miner behavior system is the safe or unsafe behavior according to their estimated values.

Suppose x_{1t}, x_{2t} respectively denotes miner's safety consciousness and work capability in time t (t=0,1,2…T), and u_{1t}, u_{2t}, u_{3t}, u_{4t}, u_{5t} respectively denotes task and behavior criterion provided by manager, manager's rewards and punishment behavior, education and training giving by manager, group or organization behavior shaped by manager, and working circumstance designed by manager. Then the miner's behavior state can be described as

$$\boldsymbol{x}(t+1)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \quad (1)$$

There, $\boldsymbol{x}(t)=(x_{1t}, x_{2t})^{\mathrm{T}}$, $\boldsymbol{u}(t)=(u_{1t}, u_{2t}, u_{3t}, u_{4t}, u_{5t})^{\mathrm{T}}$, $\boldsymbol{f}$ is a vector function.

The input variables and state variables of miner behavior system could influence the estimated value of unsafe behavior in the production process. Firstly, miner's safety consciousness could influence his selective perception, including selective perception of influencing factors of unsafe behavior. Work capability also influences his choice of unsafe behavior. Secondly, some factors in work circumstance could influence the estimated value of unsafe behavior, because miner's selective perception of work circumstance factors may strengthen or weaken his judgment of unsafe behavior cost, and the manager's rewards and punishment behavior could influence the miner's value judgment of unsafe behavior. Thirdly, task and behavior criterion, rewards and punishment behavior, education and training, and group or organization behaviors could influence the miner's value judgment of unsafe behavior by his behavior state variables.

Take the miner's estimated value of unsafe behavior as the function of its influencing factors in time t (t=0, 1,2…T), thus

$$w(t)=g(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \quad (2)$$

The above (1) and (2) make for the description of coal miner behavior system.

3. Coalmine Manager Behavior System

Coalmine manager behavior is different from coal miner behavior, because coalmine manager accomplishes his task by influencing or controlling miner behavior using management knowledge and skills. In the management process, manager needs to keep access to a variety of information repeatedly, such as the work circumstance, miner behavior. Based on this information, the manager makes his own behavior choice to control the miner behavior. Manager behavior choice is also influenced by the incentive contract and the coalmine production task.

In order to maintain consistent with the extrinsic influencing factors of coalmine manager, the input of manager behavior system may be summarized as the following aspects, i.e. manager's cognition of work circumstance, manager's cognition of miner unsafe behavior, coalmine production task of manager, and manager's incentive contract [16]. The former two aspects are related to the coal production process, so they can be regarded as the information that the manager needed to complete his task of safety management, and belong to the controllable factors. The latter two aspects are not related to the coal production process, and are not related to completing the task of safety management, belonging to the uncontrollable factors.

The inner factors are fundamental to the coalmine manager. They may be summed up five aspects, i.e. manager's mental state, knowledge state, physical state, organizational loyalty and work effort. All these factors integrated may be described as manager's work initiative and work capability that can be used as the state variables of coalmine manager.

In order to maintain consistent with miner's unsafe behavior control in coal production process, the output of manager behavior system should be constituted by safety management measures that may be carried out by the manager, including the behavior choice of safety management and the work effort in that choice. The manager behavior consists of designing and providing production task and behavior criterion, carrying up rewards and punishment, bringing into effect of education and training, and creating group atmosphere and organizational climate.

Suppose y_{1t}, y_{2t} respectively denotes the manager's work initiative and work capability in time t (t=0, 1,2…T), and v_{1t}, v_{2t}, v_{3t}, v_{4t} respectively denotes manager's production task, manager's incentive contract, manager's cognition of work circumstance, manager's cognition of miner unsafe behavior. Then the manager's behavior state may be described as

$$\boldsymbol{y}(t+1)=\boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t) \quad (3)$$

There, $\boldsymbol{y}(t)=(y_{1t},y_{2t})^{\mathrm{T}}$, $\boldsymbol{v}(t)=(v_{1t},v_{2t},v_{3t},v_{4t})^{\mathrm{T}}$, $\boldsymbol{h}$ is a vector function.

The output of manager behavior system is mainly the safety management measures available to the manager, including designing and providing production task and behavior criterion, carrying up rewards and punishment, bringing into effect of education and training, and creating group atmosphere and organizational climate. Consistent with the miner behavior system, they are denoted as u_{1t}, u_{2t}, u_{3t}, u_{4t} respectively. They are all extrinsic influencing factors for the miner and are influenced by the manager's behavior state variable and input variable. Thus

$$\boldsymbol{u'}(t)=\boldsymbol{h'}(\boldsymbol{y}(t),\boldsymbol{v}(t),t) \quad (4)$$

here, $\boldsymbol{u'}(t)=(u_{1t},u_{2t},u_{3t},u_{4t},)^{\mathrm{T}}$, $\boldsymbol{h'}$ is a vector function.

The above (3), and (4) make for the description of coalmine manager behavior system.

4. Safety Behavior Control System Model

In fact, miner behavior system interacts with manager behavior system in the coal production process. The

miner fulfills the coal production task assigned by the manager, in the same time making choice of safe or unsafe behavior based on his perception and judgment. The manager implements his management actions based on the work circumstance, work objective, and miner behavior, influencing or changing the miner's work circumstance and unsafe behavior.

The safety management measures available to the manager consist of designing and providing production task and behavior criterion, carrying up rewards and punishment, bringing into effect of education and training, and creating group atmosphere and organizational climate. They are all extrinsic factors that may influence the miner's behavior. Thus, directed by the coal production task, under the action of safety management measures available to the manager, miner behavior system and manager behavior system make up of an integrating safety behavior control system. The system model is described as in Figure 1.

There are two inputs in the coalmine safety behavior control system. One comes from the manager, including manager's production task, manager's incentive contract, manager's cognition of work circumstance, and manager's cognition of miner unsafe behavior. Although manager's cognition of miner unsafe behavior is an input of manager behavior system, it is not an input of the coalmine safety behavior control system. Another comes from the miner, which include the stimulation of work circumstance. Others are not the input of the coalmine safety behavior control system. The state variables of coalmine safety behavior control system consisted of both state variables of miner behavior system and manager behavior system. The output variable is safety or unsafe behavior of miner. The model implies that manager behavior control variables are the input coupling variables of miner behavior system, and the output variables of miner behavior system are input coupling variables of manager behavior system.

The output of safety behavior control system model is

$$w(t) = g''(\boldsymbol{x}(t),\boldsymbol{y}(t),u_{5t},v_{1t},v_{2t},v_{3t},t) \tag{5}$$

So the safety behavior control system model can be expressed as

$$\min\sum_{t=0}^{T-1} w(t) = \sum_{t=0}^{T-1} g''(\boldsymbol{x}(t),\boldsymbol{y}(t),u_{5t},v_{1t},v_{2t},v_{3t},t) \tag{6}$$

$$\text{s.t} \quad \boldsymbol{x}(t+1) = \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \quad x(0) \text{ is the known initial value}$$
$$\boldsymbol{y}(t+1) = \boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t) \quad y(0) \text{ is the known initial value}$$

As to the safety behavior control system, the manager's coal production task, manager's incentive contract, manager's cognition of work circumstance, and the work circumstance factors stimulating to miner, are all the extrinsic input variables that can influence the miner behavior. The manager's behavior control variables are the intrinsic variables that can influence the miner behavior. Therefore, although manager's behavior control variables are the important factors influencing miner behavior, the extrinsic variables play a more important role to miner behavior in the safety behavior control system. In order to better control the unsafe behavior of miners, government should strengthen the management of the extrinsic factors such as the coal production task, work circumstance, and manager's incentive contract.

5. The Manager Optimal Control Behavior in Complete Information Condition

The estimated value of unsafe behavior is the foundation basis for miner to make the choice of safe or unsafe behavior. In complete information condition, coal manager is aware of this estimated value of unsafe behavior. So the manager chooses his behavior according to this estimated value, not being influenced by his state factors. The manager's mission mainly is to depress miner's estimated value of unsafe behavior and encourage miner to adopt safe behavior using various management measures. This manager's mission can be decomposed into two aspects. On one aspect, coal manager needs to encourage miner to adopt safe behavior; on the other aspect, he also needs to cut down the total number of miner's unsafe behavior.

Because the miner makes the choice of safe or unsafe behavior based on the estimated value of unsafe behavior $w(t)$, the above management task can be described approximately by depressing the cumulated miner's estimated value of unsafe behavior in a period. Thus, coal mine safety problems faced by managers is to seek the optimal management behavior $\boldsymbol{u}(t)$ (t=0,1,2,…,T) to minimize the cumulated miner's estimated value of unsafe behavior in a period. That is

$$\min\sum_{t=0}^{T-1} w(t) = \sum_{t=0}^{T-1} g(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \tag{7}$$

s. t $\boldsymbol{x}(t+1)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t)$ $x(0)$ is the known initial value. The Hamilton's function of this system is

$$\boldsymbol{H}(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\lambda}(t+1),t) = g(\boldsymbol{x}(t),\boldsymbol{u}(t),t) + \boldsymbol{\lambda}(t+1)\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \tag{8}$$

The optimal necessity condition of this behavior control system is that taking various management measures to drive the Hamilton's function to its optimal value. When there is no constraint of management measures, it is

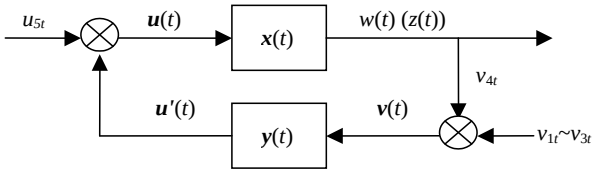

Figure 1. Coalmine safety behavior control system model

$$\begin{cases} \dfrac{\partial g}{\partial \boldsymbol{u}} + \boldsymbol{\lambda}(t+1)\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} = 0 & t = 0,1,2,\ldots,T\text{-}1 \\ \boldsymbol{\lambda}(t) = \dfrac{\partial g}{\partial \boldsymbol{x}} + \boldsymbol{\lambda}(t+1)\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} & t = 0,1,2,\ldots,T\text{-}1 \\ \boldsymbol{x}(t+1) = \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) & \boldsymbol{x}(0)\text{is the known initial value} \end{cases} \tag{9}$$

This demonstrates that, with no resource constraint, manager's optimal management behavior consists of making full use of various management measures to let the marginal utility of various management measures on Hamilton's function equal to zero.

6. The Manager Optimal Control Behavior in Incomplete Information Condition

Different management measures have different effectiveness on the perception of miner's estimated value of unsafe behavior. It is foundational for manager to make his behavior choice on behalf of the coalmine. In complete information condition, the manager is easy to make his behavior choice as he aware of this estimated valuation. But in incomplete information condition, this estimated value is unknown to him, so the manager has a great space of decision-making.

Here, the management measures available to manager still consist of four factors, i.e. designing and providing production task and behavior criterion, carrying up rewards and punishment, bringing into effect of education and training, and creating group atmosphere and organizational climate. Suppose the miner's estimated value of unsafe behavior is $w(t)$, and the manager makes his behavior choice based on the estimated effects of various management measures. Thus, the manager's behavior variables will influence his management behavior, and these effects will eventually embody in miner's estimated value of unsafe behavior [17].

The following three kinds of behavior have an important influence on the behavior choice of coalmine manager:

1) Indifference: There are a lot of works to do for a manager, and controlling miners' unsafe behavior is only one of these works. It is very natural for a manager neglecting to control miners' unsafe behavior when he needs to deal with other seemingly more important works. This phenomenon is the manager indifference to control miner's unsafe behavior. The knowledge level and psychology state variables of a manager could influence this kind of choice behavior.

2) Incorrectness: Although manager thinks much of controlling miner's unsafe behavior, he maybe chooses the incorrect management measures. If so, this will inevitably unable to implement the management measures should be implemented. The knowledge level, psychology variables, and skill of manager influence this kind of behavior.

3) Passiveness: Manager maybe not works hard. This phenomenon is the manager irresponsible for his duties. For example, manager would be work inactively when his individual goals conflict with the coalmine objectives. The organizational loyalty and the expected reward from manager influence this kind of behavior.

On making choice of management measures, manager needs to estimate their different effects to miner's estimated value of unsafe behavior. This estimation is influenced by his inner factors, i.e. manager's mental state, knowledge state, physical state, organizational loyalty and work effort. All the influencing factors integrated may be described as manager's work initiative and work capability [18]. Let $w'(t)$ denote miner's estimated value of unsafe behavior by manager, then it has

$$w'(t)=g'(\boldsymbol{y}(t),\boldsymbol{v}(t),\boldsymbol{x}(t),\boldsymbol{u}(t),t) \tag{10}$$

Similarly, the management task can be described as depressing the cumulated miner's estimated value of unsafe behavior estimated by the manager in a period. Thus, for coalmine manager, the task is to seek the optimal management behavior $\boldsymbol{u}(t)$ (t=0,1,2,…,T) to minimize the cumulated miner's estimated value of unsafe behavior estimated by the manager in a period. So it has

$$\min \sum_{t=0}^{T-1} w'(t) = \sum_{t=0}^{T-1} g'(\boldsymbol{y}(t),\boldsymbol{v}(t),\boldsymbol{x}(t),\boldsymbol{u}(t),t) \tag{11}$$

s.t $\boldsymbol{y}(t+1)=\boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t)$ $\boldsymbol{y}(0)$ is the known initial value.
$\boldsymbol{x}(t+1)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t)$ $\boldsymbol{x}(0)$ is the known initial value.

The optimal necessity condition of this behavior control system can be obtained similarly. Its Hamilton's function is

$$\begin{aligned} &\boldsymbol{H}(\boldsymbol{y}(t),\boldsymbol{v}(t),\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\lambda}_1(t+1),\boldsymbol{\lambda}_2(t+1),t) \\ &=g'(\boldsymbol{y}(t),\boldsymbol{v}(t),\boldsymbol{x}(t),\boldsymbol{u}(t),t)+\boldsymbol{\lambda}_1(t+1)\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \\ &\quad +\boldsymbol{\lambda}_2(t+1)\boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t) \end{aligned} \tag{12}$$

With no resource constraint, the optimal necessity condition of the behavior control system is

$$\begin{cases} \dfrac{\partial g'}{\partial \boldsymbol{u}} + \boldsymbol{\lambda}_1(t+1)\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} = 0 & t = 0,1,2,\ldots,T\text{-}1 \\ \boldsymbol{\lambda}_1(t) = \dfrac{\partial g'}{\partial \boldsymbol{x}} + \boldsymbol{\lambda}_1(t+1)\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} & t = 0,1,2,\ldots,T\text{-}1 \\ \boldsymbol{x}(t+1) = \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) & \boldsymbol{x}(0)\text{ is the known initial value} \\ \dfrac{\partial g'}{\partial \boldsymbol{v}} + \boldsymbol{\lambda}_2(t+1)\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{v}} = 0 & t = 0,1,2,\ldots,T\text{-}1 \\ \boldsymbol{\lambda}_2(t) = \dfrac{\partial g'}{\partial \boldsymbol{y}} + \boldsymbol{\lambda}_2(t+1)\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{y}} & t = 0,1,2,\ldots,T\text{-}1 \\ \boldsymbol{y}(t+1) = \boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t) & \boldsymbol{y}(0)\text{ is the known initial value} \end{cases} \tag{13}$$

This optimal necessity condition maybe considered consists of two parts. The former three equations determine the optimal necessity condition of managing miner's behavior; the management objective is to minimize the cumulated miner's estimated value of unsafe behavior estimated by the manager in a period. The latter three equations determine the optimal necessity condition of managing manager own behavior; the management objective also is to minimize the cumulated miner's estimated value of unsafe behavior estimated by the manager in a period. The optimal necessity condition of the whole behavior control system can be achieved only when these two optimal conditions are met.

In order to facilitate a comparative analysis, the miner's estimated value of unsafe behavior by manager may be described as the form

$$\boldsymbol{w}'(t)=g'(\boldsymbol{y}(t),\boldsymbol{v}(t),g(\boldsymbol{x}(t),\boldsymbol{u}(t),t),t) \tag{14}$$

Then the optimal management measures of manager is decided by the optimal necessity condition of this behavior control system model

$$\min \sum_{t=0}^{T-1} w'(t) = \sum_{t=0}^{T-1} g'\left(\boldsymbol{y}(t),\boldsymbol{v}(t),g\left(\boldsymbol{x}(t),\boldsymbol{u}(t),t\right),t\right) \tag{15}$$

s.t $\boldsymbol{y}(t+1)=\boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t)$ $\boldsymbol{y}(0)$ is the known initial value.

$\boldsymbol{x}(t+1)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t)$ $\boldsymbol{x}(0)$ is the known initial value.

Its Hamilton's function is

$$\begin{aligned}&\boldsymbol{H}(\boldsymbol{y}(t),\boldsymbol{v}(t),\boldsymbol{x}(t),\boldsymbol{u}(t),\lambda'_1(t+1),\lambda'_2(t+1),t)\\ &=g'(\boldsymbol{y}(t),\boldsymbol{v}(t),g(\boldsymbol{x}(t),\boldsymbol{u}(t),t),t)\\ &\quad+\lambda'_1(t+1)\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t)\\ &\quad+\lambda'_2(t+1)\boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t)\end{aligned} \tag{16}$$

The optimal necessity condition of the behavior control system is

$$\begin{cases}\dfrac{\partial g'}{\partial g}\dfrac{\partial g}{\partial \boldsymbol{u}}+\boldsymbol{\lambda}'_1(t+1)\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}=0 & t=0,1,2,\ldots,T\text{-}1\\ \boldsymbol{\lambda}'_1(t)=\dfrac{\partial g'}{\partial g}\dfrac{\partial g}{\partial \boldsymbol{x}}+\boldsymbol{\lambda}'_1(t+1)\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} & t=0,1,2,\ldots,T\text{-}1\\ \boldsymbol{x}(t+1)=\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) & \boldsymbol{x}(0)\text{ is the known initial value}\\ \dfrac{\partial g'}{\partial \boldsymbol{v}}+\boldsymbol{\lambda}'_2(t+1)\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{v}}=0 & t=0,1,2,\ldots,T\text{-}1\\ \boldsymbol{\lambda}'_2(t)=\dfrac{\partial g'}{\partial \boldsymbol{y}}+\boldsymbol{\lambda}'_2(t+1)\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{y}} & t=0,1,2,\ldots,T\text{-}1\\ \boldsymbol{y}(t+1)=\boldsymbol{h}(\boldsymbol{y}(t),\boldsymbol{v}(t),t) & \boldsymbol{y}(0)\text{ is the known initial value}\end{cases} \tag{17}$$

Here $\dfrac{\partial g'}{\partial g}$ describers the relationship between miner's estimated value of unsafe behavior by the manager and by the miner. It gives us the information of what is the effectiveness of various management measures on miner's estimated value of unsafe behavior. The former three equations determine the optimal necessity condition of managing miner's behavior, and the latter three equations determine the optimal necessity condition of managing manager own behavior. The optimal necessity condition of the whole behavior control system is achieved only when these two optimal conditions are met in the same time.

With regard to the effectiveness of various management measures on miner's estimated value of unsafe behavior, there is a theorem.

Theorem: If $\dfrac{\partial g'}{\partial g}=1$, the optimal management behavior in safety behavior control system of (15) is the same as that of (9).

The proof of this theorem is omitted here.

The theorem implies that, when manager's behavior factors influence his understanding of miner's estimated value of unsafe behavior, his choice of management behavior definitely is not the optimal one. Because it is difficult to take the optimal management behavior for the manager, it is inevitable there are managerial defects even errors in the coalmine safety management.

7. Discussion and Conclusions

We propose a model of safety behavior control system to discuss some issues of coalmine manager behavior, and analyze the conditions of optimal management behavior of coalmine manager both in complete information and incomplete Information condition. The following conclusions could be arrived.

1) The management measures of coalmine manager are the important factors influencing miner behavior. However, manager's production tasks, manager's incentive contract, manager's cognition of work circumstance take more important roles on the point of whole safety behavior control system. In order to better control the unsafe behavior of miners, government should strengthen the management of the extrinsic factors such as the coal production task, work circumstance, and manager's incentive contract.

2) The manager's inner factors do not influence his management behavior in complete information condition, but they influence it in incomplete information condition. The necessity condition of optimal manager behavior choice consists of two blocks, one block determines the optimal necessity condition of managing miner's behavior, and the other block determines the optimal necessity condition of managing manager own behavior. The optimal necessity condition of the whole behavior control system is achieved only when these two blocks' optimal

conditions are met in the same time.

3) Only when the miner's estimated value of unsafe behavior by manager is same as the miner's estimated value by himself, the manager's optimal behavior choice is possible. Because the manager's behavior factors influence his understanding of miner's estimated value of unsafe behavior unavoidably, it is difficult for the manager to obtain the optimal management behavior in the real world. It could well explain why it is inevitable there are managerial defects even errors in the coalmine safety management.

As discussed above, in coalmine safety behavior control system, manager is the subject of implementation of management measures, miner is the object of implementation of management measures, both of them have the ability to learn and adapt, and would make choice by themselves judgments.

In the coal production, the purpose of manager and miner is to achieve their respective behavior value. For a miner, this kind of behavior value is achieved by his specific production activities, so what determines the miner's behavior is his perceived value of various production activities. But for a manager, this kind of behavior value is achieved by his management activities, so what determines the manager's behavior is his perceived value of various management activities.

In order to make coal production safe enough, the manager first needs to determine the safety level of coal production based on the coal production task and the actual production condition. Then, according to this safety level, he needs to make the choice of various available management measures. In this way, what determines the manager's behavior choice is the effectiveness of different management measures to the miner's unsafe behavior.

Furthermore, the manager's behavior choice is based on his understanding of miner's estimated value of unsafe behavior, and this cannot but influenced by his own behavior variables, especial influenced by his own specific perception and judgments in the work. So in order to make a proper behavior choice, manager should try to avoid influencing by own state variables. Therefore, he needs to strengthen self-behavior management.

How to strengthen self-behavior management for a manager, and how to inspirit the manager to strengthen self-behavior management, are the problems needed to be studied carefully in the near future.

8. Acknowledgement

This research is supported by the National Natural Science Foundation of China under Grant 70871113 and the Humanities and Social Science Research Foundation of Ministry of Education of the People's Republic of China (Project ID: 07JA630025).

REFERENCES

[1] Z. Y. Lin and L. C. Xu, "Analysis of the human error and its reasons in coal mine accidents," Chinese Journal of Ergonomics, Vol. 2, No. 2, pp. 17–20, 1996.

[2] Q. R. Cao and X. F. Song, "Problems being faced with the unsafe behavior study and its countermeasure," China Coal, Vol. 32, No. 11, pp. 62–63, 2006.

[3] H. Mintzberg, "The nature of managerial work," Harper & Row, New York, 1973.

[4] P. Allan, "Managers at work: A large-scale study of the managerial job in the New York city government," Academy of Management Journal, Vol. 24, No. 3, pp. 613–619, 1981.

[5] G. Yukl, S. Wall, and R. Lepsinger, "Preliminary report on validations of the managerial practices survey," In Measures of leadership, K. E. Clark and M. B. Clark, Eds., West Orange, NJ, 1989.

[6] D. Schaechtel, "How to build a safety management system?" Professional Safety, Vol. 42, No. 8, pp. 22–24, 1997.

[7] R. Jens, "Risk management in a dynamic society: A modeling problem," Safety Science, Vol. 27, No. 2–3, pp. 183–213, 1997

[8] S. R. Sun and L. P. Qi, "Developments of mathematical models of behavior control mechanism and behavior management system," Journal of University of Shanghai for Science and Technology, Vol. 26, No. 1, pp. 47–52, 2004.

[9] P. Jason, D. Pasquale, and E. S. Geller, "Critical success factors for behavior-based safety: A study of twenty industry-wide applications," Journal of Safety Research, Vol. 30, No. 4, pp. 237–249, 1999.

[10] O. Michael, "The relationship between employees' perceptions of safety and organizational culture," Journal of Safety Research, Vol. 33, No. 1, pp. 231–243, 2002.

[11] E. S. Geller, "Working safe: How to help people actively care for health and safety," Second Edition, Lewis Publishers, Inc., 2001.

[12] M. Neal, A. Griffin, and P. M. Hart, "The impact of organizational climate on safety climate and individual behavior," Safety Science, Vol. 34, pp. 99–109, January 2000.

[13] M. Jane, "Investigating factors that influence individual safety behavior at work," Journal of Safety Research, Vol. 35, pp. 275–285, January 2004.

[14] Q. R. Cao, "Analysis of the influencing factors of coal miner unsafe behavior," Mining Safety & Environmental Protection, Vol. 33, No. 6, pp. 80–82, 2006.

[15] Q. R. Cao, "A Study on the unsafe behavior in the coal mining and its control system," Ph. D. dissertation, Department of Science and Technology, China University of Mining and Technology, Xuzhou, 2005.

[16] Q. R. Cao and X. F. Song, "Study on the unsafe behavior of coal miner and its management avenue," Economic Management, Vol. 15, pp. 62–65, July 2006.

[17] M. Grabowski and H. R. Karlene, "Human and organizational error in large scale systems," IEEE Transaction on System, Man, and Cybernetics-part A: System and Human, Vol. 26, No. 1, pp. 2–16, 1996.

[18] Q. R. Cao, "Study on the cognitive diversities between managers and workers in controlling unsafe behavior," China Safety Science Journal, Vol. 17, No. 1, pp. 22–28, 2007.

J. Service Science & Management, **2010, 3: 173-179**
doi:10.4236/jssm.2010.31022 Published Online March 2010 (http://www.SciRP.org/journal/jssm)

Construction of National High-Quality Courses for LIS in China

Changping Hu, Xueqin Zhao

School of Information Management, Wuhan University, Wuhan, China.
Email: zxqwhu@yahoo.com.cn

Received November 11th, 2009; revised December 26th, 2009; accepted January 28th, 2010.

ABSTRACT

The paper first introduces the background of Martin Trow's theory of how to develop mass higher education and then brings up the necessity of Chinese course teaching reform. It introduces the current situation of the construction of the National High-quality Courses (*NHCs*) *in China. And taking the practice of Wuhan University's library and information science* (*LIS*) *course teaching as an example, the paper also discusses the implementation of this project, focusing on two aspects*: *a new system of LIS course teaching and network platform construction. The paper concludes by recommending that, for successful LIS NHCs, it is necessary to transform our educational ideology, strengthen our innovation, and set up a more beneficial teaching environment for innovation cultivation.*

Keywords*: *Mass Higher Education, Course Teaching, National High-Quality Courses, Creative Capability

1. Introduction

Based on his research on the development of higher education in America, the well-known American educational sociologist Martin Trow argued that: higher education systems that enrolled up to 15 percent of the college-age population were best described as elite systems; systems that enrolled between 15 and 40 percent of the age group were mass systems; and those that enrolled more than 40 percent were universal systems [1]. This theory has drawn widespread attention all over the world, and many countries have attempted accelerating the development of higher education and stepping into the mass phase as soon as possible.

Alongside educational change run concern with and the study of innovation. Rogers (1995) has noted that "no other field of behavior science research represents more effort by more scholars in more disciplines in more nations" [2]. In an article discussing the work of UNESCO in supporting the development of library schools, Kereszteti (1982) was especially concerned with the prevalence of Western influence on the LIS curriculum [3]. An article by Ballard (1980), a professor at North Carolina Central University School of Library Science, felt that a core curriculum for library science should be developed and adopted by developing countries which would include courses that respond to the particular conditions of these low-income countries and not be carbon copies of curricula of developed countries [4].

By 2007, when the gross enrollment ratio reached 23%, Chinese higher education stepped into mass education phase [5]. As a country whose development was relatively late and was impelled by external factors, China is exploring its own way of developing higher education. Thus we don't necessarily have to judge the massification course of Chinese higher education strictly according to the traditional criterion (15%). Although many aspects of Chinese higher education (e.g. gross enrollment ratio, educational modal diversity, admission and management, curriculum organization, etc.) have shown some characteristics of massification, its development follows a radical path and lacks sufficient resource preparation, which makes course construction, can not meet the mass higher education standard badly. In recent years, the Ministry of Education has successively constructed the NHCs from 2003 to 2010 in the whole country to implement the higher education quality standard.

2. Review of NHCs Construction in China

The development of National High-quality Courses Construction is using following formats: unprompted school construction, autonomous regions and municipalities recommendation, the Ministry of Education evaluation, honorary title conferring, and finally construction fees assistance. Since China's reform and open, especially in recent years, state and municipal

education bureaus and the colleges are increasing funds for the investments of High-quality Courses Construction. Since 2003, the Ministry of Education is actualizing "Higher education teaching quality and teaching reform project" and improving the performance of these several aspects: teaching contents reform and modernization, management system of high-quality course, and courses system reorganization. These successions of measures have enhanced the teaching quality and research level and created new opportunities of educating high-ability students who have innovation spirit and practice ability [6].

By the end of December 2009, among 6,469 courses at 959 colleges and universities (distributed over 31 provinces, autonomous regions and municipalities directly under the State Council) which are required to undergo evaluation. Of those, 2212 were certified as NHCs at 819 colleges and universities. With the 4.43-billion-Yuan investment (2.21 billionfrom colleges and universities, 1.11 billion from governments at various levels and 1.11 billion from colleges and universities levels), these qualified NHCs should consolidate their achievements after evaluation, continue strengthening their construction, deepen the course teaching reform and further enhance their teaching level and working capability (**Table 1**).

NHCs is not only a significant criterion to evaluate the teaching level but also an important aspect reflecting the overall strength of colleges and universities. Therefore, Wuhan University has paid high attention to and conducted unified planning for the construction and realization of its NHCs. In addition to the construction of basic course system already accomplished over the years, Wuhan University is determined to quicken its NHCs construction and make some of them meet the requirements of the provincial standard, with some qualified under national standards. A batch of college-level demonstration courses are to be established and an course teaching demonstration system at the national, provincial and college levels is to be progressively formed. To accomplish these assignments, we should modernize our ideas, deepen our understanding, and improve the course teaching contents and methods, and found new systems. Furthermore, we should pay attention to staff building and establish an effective management and operating system.

Table 1. NHCs from 2003–2009 in China [7–13]

Year	Courses applied	Courses gained
2003	467	127
2004	720	249
2005	940	258
2006	980	358
2007	1102	411
2008	1172	400
2009	1195	409

3. Constructions of NHCs for LIS

Along with the rapid development of digitalization and network technology, information resources are being digitalized continuously and transmitted to all corners of the world through the Internet. Consequently a globalized and highly effective information transmission system is taking shape and changing the pattern which people follow in producing, transmitting and applying information. All these have brought unprecedented challenges as well as developing opportunities to LIS education. As LIS is a multi-disciplinary intersection, its professional knowledge and skills are widely recommended in this digitalized network era and the innovational reform of LIS education has become a significant subject energetically probed all over the world. Like any other disciplines of higher education, the essential aim of LIS education is to cultivate high-quality synthetic talents. Therefore the whole of LIS teaching should be brought into the scope of course teaching, level upon level and step by step, so that the teaching service system can be reformed in the end.

3.1 Methods to Found NHCs for LIS in Wuhan University

The School of Information Management (SIM) of Wuhan University is an institution for information management education and research which has the longest history and the grandest scale in China. It grew out of the college of library science (established by the American scholar Mary E. Wood in 1920) of Boone University in Wuchang and then was developed as Boone Library School in 1929. In 1953, Boone Library School was merged into Wuhan University and renamed as the department of library science (for college students). The department for undergraduates was then established in 1956. After being approved by the Ministry of Education, the School of Library and Information Science came into being in 1984, and then it was renamed as SIM in 2001 [14].

The school have library science, information management and information systems (whose precursor was the scientific and technological information specialty which had not been founded before 1978 in China), archivistics science (professors were recruited in 1934 and in 1940, when the archivistics specialty was founded, but it ceased in 1947 and rebuilt in 1984) and electronic commerce (founded in 2001).

Form 2003–2009, there are 21 NHCS in LIS subject which belong to 16 high colleges and universities (see **Table 2**). Wuhan University has 6 courses that received the honor, which is the first place in Chinese Colleagues of LIS subject.

Table 2. NHCs of LIS subject (2003–2009)

Subject	NHCs	Colleges
LIS	Introduction to Bibliography	Wuhan University
	Foundations of Information Management	Wuhan University
	Electronic File Management	Wuhan University
	Information Service and User Study	Wuhan University
	Information Metrology	Wuhan University
	Information Resource Construction	Wuhan University
	Introduction to Library Science	Peking University
	Introduction to Archival Science	Renmin University of China
	Information Management Science	Hefei University of Technology
	Computer Network	National university of Defense Technology
	Database system and Application	Jiangxi University of Finance and Economics
	Management Information System	Xiamen University of Technology
	Information Systems Analysis and Design	Liaoning University
	Introduction to Decision Support System	Donghua University
	Management Information Systems Analysis and Design	Huazhong University of Science and Technology
	Documentations	Heilongjiang University
	Information Retrieval and Utilization	Shandong University of Technology
	Information Resource Sharing	Sun Yat-sen University
	Science of Archives Management	Guangxi University for Nationalities
	Archival Science	Shanghai University
	Document Retrieval	Central South University

3.2 Course Teaching Reform Based on Core Course Integration

The core courses of Library science, Archives science and Information management science will be integrated with the development trend of the subject integration. In our school, we reform in 5 NHCs which named as Introduction to Bibliography, Foundations of Information Management, Electronic File Management, Information Service and User Study, Information Resource Construction and Information Metrology in order to realize the core course reform of LIS. In the open environment, we try to construct the new teaching system of core course that can adapt to the new requirements. The teaching reform has laid equal stress on both classical and modern, essential and frontier theories. It has been updated closely by following the discipline's development to create a bigger space and to get a raise in synthesis, design, and research so that course teaching, scientific research, projects and social application practice can be closely integrated, furthering reciprocity and establishing a skillful, personalized, diversified and modernized course teaching system.

The course reform should also fully be brought into playing the supporting role of the two national key disciplines of library science and information science as well as the following organizations: the research center of information resources; the cultivation center for senior publication talents of the Press and Publication Administration unit; the senior research center of intellectual property rights; the research center for science evaluation in China; the Chinese e-commerce research and development center; the Institute of Library and Information Science; the Institute of Publication, the Institute of Digital Libraries; the Institute of Study on Complete Collection in Four Treasuries; and the Institute of Electronic Document and Government Information. Scientific research methods and achievements and scientific thinking patterns should be organically integrated into experimental teaching, and parts of the social science,

natural science, base and horizontal projects (between 2005 and 2009) undertaken by tutors have been divided into smaller research subjects (see **Table 3**) for students to choose freely. From 2005 to 2009, teachers of SIM have altogether implemented 120 social science, natural science, base and horizontal projects, with the project funds of 12,600,000 Yuan. The laboratory is open to all the students and accepts some of them to participate in the research, which offers beneficial support for them to independently accomplish their spare-time scientific research. The close integration of class teaching, scientific research and experimental teaching has greatly promoted the experimental teaching reform and the implementation of innovative education and improved the quality of composite talent cultivation. Meanwhile, the teachers have undertaken the teaching reform projects by Hubei Province and Wuhan University, which promote the development of teaching reform for LIS (see **Table 4**).

Table 3. Projects and subprojects undertaken by tutors (2005–2009)

Projects	Classification	Tutors	Sub-projects	Students
Research of Basic Experimental Teaching on LIS	WuHan University teaching reform projects	Xiangxing Shen	Research of Experimental Teaching Systems on LIS	Shuguang Han
			Research on Experimental Teaching Web-Platform of LIS	Ji Liu
Construction of Information Service Systems and Security Systems for a Country based on Innovation	The Key National Funds of Social Sciences	Changping Hu	Research of Tourism Information Services based on Mobile Internet	Han Zhang
			Optimization for IT Project Management	Yan Tang
Research of Knowledge Information Service Systems for a Country based on Innovation	The Key project for Philesophy and Social Science of Ministry of Education	Changping Hu	Research of Value Chain and Payoff Model in M-Commerce	Rong Sun
			Research of Current Tourism Networking Situation and its Prospect in Our Country	Jingyun Huang
Research of User-oriented Information Integration and Service	The National Funds of Natural Sciences	Changping Hu	Research of Regional Tourism Information Resources Development based on Web2.0	Lei Wang
			Marketing Program Design based on Mobile Platform	Lin Chen
Research of Digital Information Resource Plan, Management and Usage	The Key project for Philosophy and Social Science of Ministry of Education	Feicheng Ma	Information Economic Analyse on Pricing Strategy in E-commerce	Jingzhi Zhao
Promotive Government Strategies for Public Access to Information Resource	The National Funds of Natural Sciences	Xincai Wang	Applied Research on E-government Construction combined with WEB2.0	Qiang Zhang
Empirical Research of Libraries' Approaches to Copyright	The National Funds of Social Sciences	Chuanfu Chen	Research on Barrier of the Users for Library Accession	Woruo Zhao
			Research of Development Model for Community Libraries in a Harmonious Society	Ying Yu
Research of National Strategies for Chinese Culture Digital Resources Preservation	The National Funds of Natural Sciences	Jiazhen Liu	Study on the Publicity and Secrecy for Archives Information	Kui Ye
Research of Information Resource Development and its Technology Implements in E-government	（The Key project of Humanities-Social Sciences of Chief Research Center Ministry of Education ）	Xincai Wang	Evaluation and Analysis for Chinese Government Websites Services	Yuyu Wang
Research of Information Resources Configuation Theory and Model	（The Key project of Humanities-Social Sciences of Chief Research Center Ministry of Education ）	Xianjin Cha	Study on Information Organization for Network Products Based on Ontology	Xuan Yang

Table 4. Teaching reform projects undertaken by tutors (2005–2009)

Projects	Classification	Tutors	Fund	Time
Construction and Realization of opening teaching system for Information Management discipline	Wuhan University teaching reform projects	Changping Hu	10,000	2007–2009
Research of Whole Process Teaching Reform for Information Management discipline	Projects for teaching reform of Hubei Province	Changping Hu	30,000	2008–2010
Construction and Research of Basic Experimental Teaching for LIS	Projects for teaching reform of Hubei Province	Xiangxing Shen	30,000	2005–2007
Development and Research for web class of document information service	Projects for teaching reform of Hubei Province	Lin Yuan	30,000	2002–2005
Research of Structure Adjustment for Specialty Archival Science	Wuhan University teaching reform projects	Jiazhen Liu	10,000	2002–2006
Research and Design of multimedia courseware for Library Science basic class	Wuhan University teaching reform projects	Chuanfu Chen	10,000	2004–2008
Research and Practice of innovation for teaching management mode in school of information management	Wuhan University teaching reform projects	Xincai Wang	10,000	2006–2007
Employability Training and Class System Construction for students in information management	Projects for teaching reform of Hubei Province	Yikun Xia	30,000	2006–2007
Construction and Development for Archival Science based on first key discipline	Projects for teaching reform of Hubei Province	Hai Yan	30,000	2007–2008

Such a system has step-by-step guided students to form scientific thinking patterns and understand the essential contents of information obtainment, organization, transmission and application. It also has enlightened students' innovation awareness, discovered students' potential for self-innovation, and stimulated students' interest in participating in class teaching and scientific research. As a result, students positively carry out spare-time scientific research and science and technology practica during their summer vacation, which has greatly expanded the space of class teaching. Between 2005 and 2009, students published nearly 150 academic papers during their schooling.

3.3 Construction of Information Web Site for NHCs

The School has established an LIS gigabyte network center to connect all the teaching and scientific research organizations and teaching laboratories together. A wireless network covers the whole LIS building. In addition to the teaching management, the courses website constructions mainly apply to assist with class teaching. Students can accomplish selecting courses online and consulting abundant online teaching information, such as syllabi, courseware and course scores through the website. It can facilitate students' preview and review and can also supply a communication platform between teachers and students (see **Table 5**).

The use of web information technology can support the teaching of "equal opportunities for interaction", so that more democratic classroom, learners can participate in the teaching content to the formulation and evaluation. Began in March 2007, "Information Services and User Study "course conducted a curriculum network to support teaching, research and practice, has achieved an full-assisted online teaching, students and teachers to achieve the open web-based interaction. Course website collection class teaching, extra-courser research and social practice in one of the open and interactive, emphasizing teacher-student interaction, students and student exchanges between the units of interaction and practice, the main body of the students as a teaching and attach importance to the participation of students in teaching effect. Such as class discussion site set up the module, students who log on the forum will issue a timely reflection, teachers can answer questions to achieve the interaction with the students [15].

4. Conclusions

Founding the NHCs for LIS is a systematic project rather than merely a kind of conditional construction. Firstly,

Table 5. Web site resources of LIS NHCs

Columns	Introductions to each column
Teacher	A brief introduction to the course teachers.
Brief introduction	A brief introduction to the course Credit.
Teaching syllabi	Syllabi for the courses.
Textbook	Textbooks for the courses.
Synchronous teaching	Power Point of all the courses teachers.
Research Project	Projects that undertaken by teachers.
Academic Result	Teachers teaching papers published statistical tables.
Students' papers	Papers that students published after the class.
Practices out of class	A brief introduction to the practices out of class. Relative practice results and excellent. Practice reports are also available
Teaching news	News for course teaching.
Examinations and tests	General methods of tests for tests courses.
Teaching video	Teaching video for the course.
Teaching references	List bibliographies of all the excellence courses and links of some relative websites for learning. Provide commonly-used software downloads for students.
Teaching blog	Provide students a platform to discuss together and thus promote their enthusiasm for study.

we have the obligation of transforming our modes of thinking, further deepen course teaching reform, and promote the transition of education towards the quality-oriented type so as to create a more beneficial course teaching environment for cultivating innovation ability. Second, besides the hardware construction, we should pay more attention to the software to strive to construct the course teaching center as a high-level course teaching base with real and demonstrative functions, cultivate innovative cross-century talents and make a contribution to the improvement of LIS higher education quality.

The American education system lays more emphasis on freedom, dialectical thinking, and student-centered pedagogy, while China's education system lays more stress on well-knit, highly disciplined, and top-down pedagogy. Since these two kinds of systems can't completely be mixed together, what we could do is to profoundly understand the American education system and enhance our course teaching reform to improve LIS higher education quality and thus cultivate more creative talents for China.

5. Acknowledgment

The paper is the research result of the projects for teaching reform of Hubei Province "Research of Whole Process Teaching Reform for Information Management discipline" (No. 2008004).

REFERENCES

[1] T. Martin, "The expansion and transformation of higher education," International Review of Education, Vol. 1, pp. 61–63, 1972.

[2] E. D. Rogers, "Diffusion of Innovations, 4th ed.," The Free Press, New York, 1995.

[3] M. Keresztesi, "UNESCO's work in the field of library education and training: An overview and assessment," International Information and Library Review, Vol. 14, pp. 349–361, 1982.

[4] R. M. Ballard, "Library education and library problems in developing nations," International Information and Library Review, Vol. 12, pp. 65–70, 1980.

[5] The Ministry of Education of PRC, "The 2007 education development statistical bulletin," Accessed May 3, 2009, Available at: http://www.moe.gov.cn/edoas/website18/54/info1209972965475254.htm.

[6] The Ministry of Education of PRC, "The notice for the higher education teaching quality and teaching reform project of national excellence courses," Accessed May 7, 2009, Available at http://www.moe.edu.cn/edoas/website18/level3.jsp?tablename=1193&infoid=13676.

[7] The Ministry of Education of PRC, "The notice for the national excellence course in 2003," Accessed October 8, 2009, Available at: http:// www.jpkcnet.com/new/zhengce/Announces_detail.asp?Announces_ID=33.

[8] The Ministry of Education of PRC, "The notice for the national excellence courses in 2004," Accessed October 9, 2009, Available at: http://www.jpkcnet.com/new/zhengce/Announces_detail.asp?Announces_ID=43.

[9] The Ministry of Education of PRC, "The notice for the national excellence courses in 2005," Accessed October 11, 2009, Available at: http://www.jpkcnet.com/new/zhengce/Announces_detail.asp?Announces_ID=63,

[10] The Ministry of Education of PRC, "The notice for the national excellence courses in 2006," Accessed October 12, 2009, Available at: http://www.jpkcnet.com/new/zhengce/Announces_detail.asp?Announces_ID=84.

[11] The Ministry of Education of PRC, "The notice for the national excellence courses in 2007," Accessed October 13, 2009, nounces_detail.asp?Announces_ID=115.

[12] The Ministry of Education of PRC, "The notice for the national excellence courses in 2008," Accessed October 14, 2009, Available at: http://www.jpkcnet.com/new/zhengce/Announces_detail.asp?Announces_ID=142.

[13] The Ministry of Education of PRC, "The notice for the national excellence courses in 2009," Accessed October 17, 2009, http://www.jpkcnet.com/new/zhengce/Announces_detail.asp?Announces_ID=158.

[14] Z. Zheng and X. X. Shen "Construction and discussion of LIS experimental teaching center in China," Journal of Education for Library and Information Science, Vol. 4, pp. 321–326, 2007.

[15] C. P. Hu and X. Q. Zhao, "Construction of web-based

learning platform for information management subject—case study on 'information service & user' course website," Proceedings – International Conference on Computer Science and Software Engineering, CSSE'08, Vol. 5, pp. 760–763, 2008.

Call for Papers

Journal of Service Science & Management (JSSM)

ISSN 1940-9893 (Print), 1940-9907 (Online)
www.scirp.org/journal/jssm/

JSSM is an international multidisciplinary journal with the emphasis laid on the service innovation in the global economy and entrepreneurship, the latest management technologies. It also explores the contributions of knowledge discovery and information management research and applications. The goal of this journal is to keep a record of the state-of-the-art research and promote the fast moving service science and management technologies.

Editor-in-Chief

Prof. Samuel Mendlinger **Boston University, USA**

Editorial Board (According to Alphabet)

Subject Coverage

All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed in hard copy. The journal publishes the highest quality, original papers included but not limited to the fields:

- **Service Science**
- **Business Intelligence**
- **Operational Research**
- **Computational Economics**
- **Financial Engineering**
- **Decision Support System**
- **Business Process Re-engineering**
- **Data Mining and Knowledge Discovery**
- **Innovation and Entrepreneurship**
- **Risk Management**
- **Quality Management**
- **Project Management**
- **Supply Chain Management**
- **Software Engineering Management**
- **Environment and Energy Management**
- **Knowledge Management and Semantic Web**
- **Information System Management**
- **Customer Capital Management**
- **Human Resources Management**

We are also interested in:
Short reports—Discussion corner of the journal:
2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data.
Case studies—Rather than present a classical paper.
Book reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Contact Us:

E-Mail: jssm@scirp.org

J. Service Science & Management, **2010, 3**
Published Online March 2010 in SciRes (www.scirp.org/journal/jssm)