JOURNAL OF BUSINESS AND BEHAVIORAL SCIENCES

Volume 28, Number 2 **ISSN 1946-8113** **Fall 2016**

A REFEREED PUBLICATION OF THE AMERICAN SOCIETY OF BUSINESS AND BEHAVIORAL SCIENCES

JOURNAL OF BUSINESS AND BEHAVIORAL SCIENCES
P.O. Box 502147, San Diego, CA 92150-2147: Tel 909-648-2120
Email: mondal@asbbs.org http://www.asbbs.org
ISSN 1946-8113

The Journal of Business and Behavioral Sciences is a publication of the American Society of Business and Behavioral Sciences (ASBBS). Papers published in the Journal went through a blind-refereed review process prior to acceptance for publication. The editors wish to thank anonymous referees for their contributions.

The national annual meeting of ASBBS is held in Las Vegas in February/March of each year and the international meeting is held in June of each year. Visit www.asbbs.org for information regarding ASBBS.

JOURNAL OF BUSINESS AND BEHAVIORAL SCIENCES

ISSN 1946-8113
Volume 28, Number 2 Fall 2016

TABLE OF CONTENTS

Journal of Business and Behavioral Sciences
Vol 28, No 2; Fall 2016

A CONCEPTUAL FRAMEWORK FOR THE DYNAMICS OF CALL SERVICE CENTER OFFSHORING

Ryoji Ito
Niigata University
Kenneth C. Gehrt
San José State University

ABSTRACT: This paper develops a simple but comprehensive conceptual framework to explain the variation in offshoring outcomes. The focus is on the offshoring of call service centers (CSC). Specifically, based on the prototype theory, service-dominant logic, and the problem-solving perspective, this paper aims to consider why some firms can evaluate offshoring outcomes in early stages but others have a difficult time evaluating outcomes even much later. After developing conceptual framework, the paper suggests several future directions of CSC offshoring research including the considerations of the impact of cultural factors, brand, and the need for empirical investigation.

Keywords: Offshoring, Call service center, Prototype theory, Service-dominant logic, Problem-solving perspective

INTRODUCTION

Increasingly, firms have decided to implement call service center (CSC) offshoring. Studies suggest that the primary purposes of CSC offshoring are cost reduction, service improvement, access to skilled people, and establishment of a foothold in a foreign market (Lewin and Peeters, 2006; Manning, Massini, and Lewin, 2008). This has led to consideration of offshoring as a strategic decision to improve service level or reduction of costs associated with sales, marketing, and customer care activities (Robinson, Kalakota, and Sharma, 2005; Sharma, Iyer, and Raajpoot, 2009; Thelen and Shapiro, 2012; Roza, Van den Bosch, and Volberda, 2011). There are variations in the manner in which companies benefit from CSC offshoring. Those who have benefited are likely to maintain their current offshoring and to perhaps launch new offshore endeavors. Those who have not benefited may try to change their offshoring provider, offshore in another location where cultural or geographic distance is closer to their home country (known as "nearshoring"), or withdraw to their home country (known as "inshoring" or "backshoring") (Benito, Dovgan, Petersen, and Welch, 2013; Kinkel, 2014; Tate, 2014). Some firms can evaluate offshoring outcomes in early stages but others have a difficult time evaluating outcomes even much later.

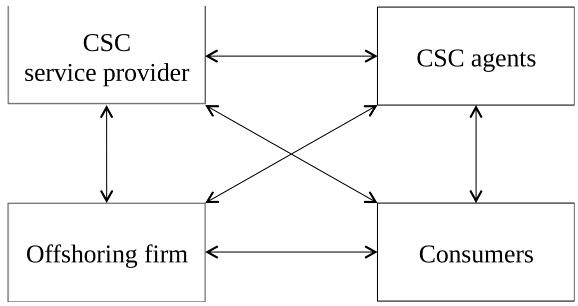
This paper develops a simple but comprehensive conceptual framework to explain the variation in outcomes. Prior studies have identified three factors that affect the success or failure of offshoring (Ito and Gehrt, 2014; Lampel and Bhalla, 2011; Roza, Van den Bosch, and Volberda, 2011; Schmeisser, 2013). These include offshore processes (i.e., back office operation, front office operation, more advanced work, etc.), offshore providers (i.e., subsidiary, joint venture, outsourcing, etc.), and offshore locations. While most prior studies have developed frameworks to understand performance implication of offshoring as a static strategy, there is little research on the dynamic process of identifying the optimal offshoring strategy (e.g., Benito, Dovgan, Petersen, and Welch, 2013; Ito and Gehrt, 2014; Jensen and Pedersen, 2011; Murray, Kotabe, and Westjohn, 2009). This paper develops a conceptual framework to capture the dynamic aspect of offshoring strategy, especially in the context of CSC offshoring.

Youngdahl and Ramaswamy (2008) categorized the services that are the target of offshore outsourcing into two dimensions; the degree of knowledge embedding (solution/transaction) and degree of customer contact (back office/front office). "Knowledge embedding" refers to the degree of routinization in service delivery system. "Customer contact" refers to the degree of presence of the customer in a service delivery system. This is not the physical presence of the customer inside the service delivery system, but instead signifies customer involvement via information communication technology (ICT). The focus of this paper is on front office services. In services of this type, customers directly interact with offshore agents.

CSC offshoring involves four types of stakeholders that are related to each other (see Figure 1) including the offshoring firm, CSC service provider, offshore CSC agents, and consumers (e.g., Jeong, Bekmamedova, and Kurnia, 2012). The offshoring firm is the organization that provides products or services to consumers. This offshoring firm searches for a CSC service provider that can provide CSC services to customers on its behalf. The CSC service provider is an organization that is located in the offshore location and provides the CSC services. The CSC service provider has internal employees who are CSC agents. The agents are the actual providers of the CSC services and directly interact with the offshoring firm's consumers. Finally, consumers are the customers of a given offshoring firm. Consumers generally interact with agents via the telephone. Agents work to resolve the consumers' issues, respond to inquiries from consumers, and encourage consumers to purchase products and services.

Consumers contacting the agents do not know ahead of time the agents with whom they will interact and are sometimes unexpectedly drawn into offshoring (Roggeveen, Bharadwaj, and Hoyer, 2007). Consumers must, therefore, engage in service interactions with foreign agents whether they would like to or not. Consumers only realize that they are being made to participate in offshoring when they receive some kind of clue. Meanwhile, there is a trend of offshoring firms attempting to conceal the fact that they are offshoring their CSC (Honeycutt, Magnini, and Thelen, 2012; Thelen and Shapiro, 2012; Thelen, Yoo, and Magnini, 2011).

Figure 1. Stakeholders of CSC offshoring



Prior research on CSC offshoring has been conducted individually on the arrows shown in Figure 1. However, because the actions of the four stakeholders in CSC offshoring are interrelated, it is difficult to obtain sufficient understanding by focusing on just one relationship. This paper suggests a framework for comprehensively examining these relationships. The foundation for this is the service dominant logic (S-D logic) concept (e.g., Lusch and Vargo, 2006; Vargo and Lusch, 2004, 2008). The S-D logic perspective views the relationship between "provider" and "customer" as a relationship of equals. Provider and customer are viewed as co-creators of value that is created through the exchange of service between the two. S-D logic defines "(a) service" as an exchange that applies the provider's knowledge to the customer and the customer's knowledge to the provider (Lusch and Vargo, 2006; Vargo and Lusch, 2008; Vargo, Maglio, and Akaka, 2008). Co-created value in S-D logic depends on the viability of the beneficiary service system. The service system is a useful abstraction for understanding value and value co-creation (e.g., Maglio and Spohrer, 2008). Value is created collaboratively in interactive configurations of mutual exchange. Research on service science calls these value-creation configurations service systems. Vargo, Maglio, and Akaka (2008) argued that service science is the study of service systems and of the co-creation of value within complex constellations of integrated resources. A viable service system adapts to resolving misalignment that may arise between the customer and provider. S-D logic suggests the important role of learning and subsequent realignment within service systems, leading to value creation (Lusch, Vargo, and Tanniru, 2010).

This paper considers two types of "customer-provider" service system relationships (see Figure 2). In the first relationship the "provider" is the CSC service provider in the offshore location and the "customer" is the offshoring firm that relocates its CSC process to the offshore location. In the second relationship the "providers" are the CSC agents in offshore location and the "customers" are the consumers that are often located in country of the offshoring firm. These two types of relationships are mutually intertwined because CSC service providers and offshoring firms train offshore CSC agents collaboratively and the agents, in turn, affect consumers' attitudes toward offshoring CSC.

According to prior marketing studies, the critical relationship is the interface between offshore CSC agents and consumers (Honeycutt, Magnini, and Thelen, 2012). This often involves communication inefficiencies between consumers and offshore CSC agents caused by issues, such as offshore CSC agents not picking up the call quickly, agents putting calls on hold, consumers' inability to understand calls, increased call time, problems with offshore agents' accents, agents' reliance on scripted answers, and agents' lack of expertise for solving complex problems (e.g., Honeycutt, Magnini, and Thelen, 2012; Sharma, 2012; Sharma, Mathur, and Dhawan, 2009). In this paper, we consider these issues not only from B2C but also B2B perspective. To understand the dynamic nature of CSC offshoring, the foci of this paper are the relationships between the CSC service provider and offshoring firm (left half side of figure 1) and the relationship between offshore CSC agents and consumers (right half side of figure 1). Thus, this paper extends S-D logic theory in the context of interactions between consumer and non-consumer units of analysis. From a managerial perspective, the paper is designed to enhance our understanding of how to increase the chances of success in CSC offshoring endeavors.

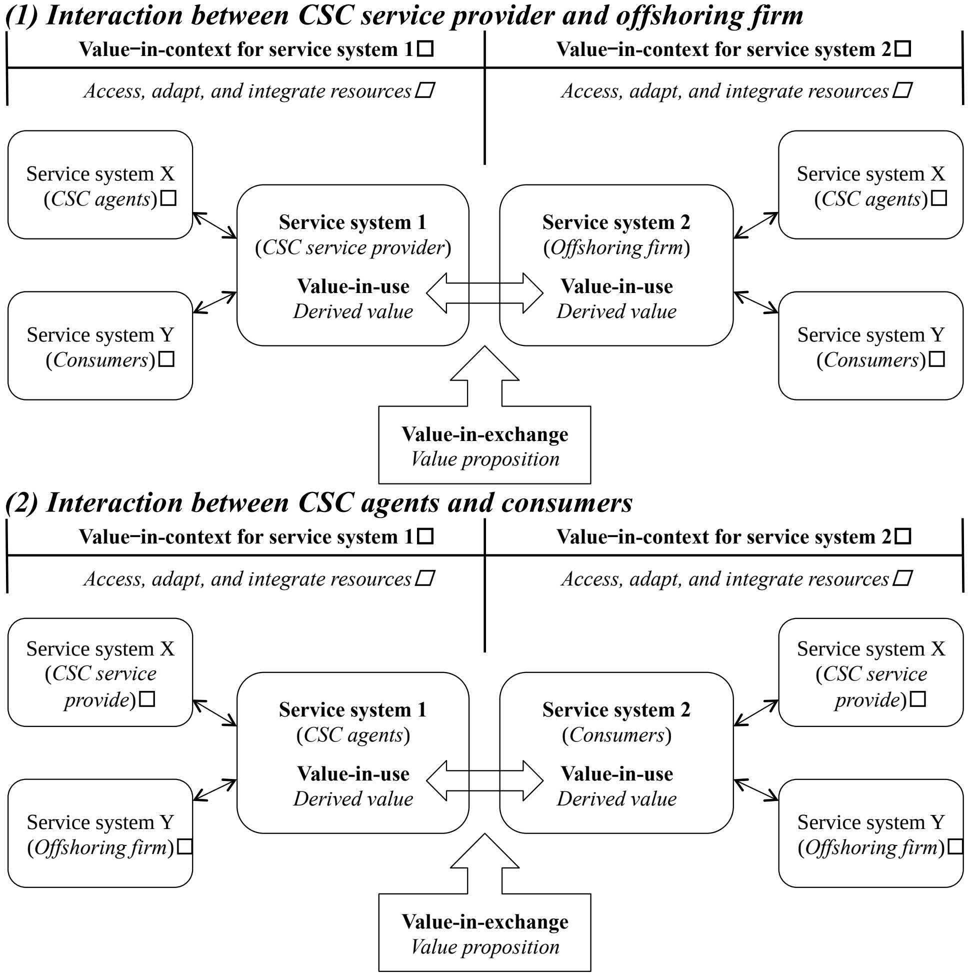
The paper examines three points on an offshoring continuum. Non-offshoring consists of onshore staffing of an onshore service system, offshoring consists of offshore staffing of an offshore service system, and mixed offshoring which is co-created through the interaction between the onshore and offshore service system's agents.

The paper is organized as follows. Section 2 reviews the CSC offshoring literature as well as prototype theory which is used to help understand offshoring relationships. Section 3 discusses S-D logic and the problem-solving perspective (PSP) and outlines a process by which to calculate the value of co-creation within mixed offshoring. Section 4 proposes explanations for variations in viability of mixed offshoring and Section 5 discusses conclusions and limitations of the paper as well as avenues for future research.

OFFSHORING LITERATURE REVIEW

Research on CSC offshoring: Offshoring-related literature has often mentioned monetary cost cutting as a benefit (Ceci and Prencipe, 2013; Roza, Van den Bosch, and Volberda, 2011). Increasingly, the literature has examined invisible costs that sometimes accompany CSC offshoring such as weakened corporate image, reduced customer satisfaction, reduced brand loyalty, and reduced service

Figure 2. Two aspects of the interaction among service systems in CSC offshoring



Source: Adapted with modification, from Vargo, Maglio, and Akaka (2008).

standards (Bharadwaj, 2010; Stringfellow, Teagarden, and Nie, 2008). Corporations must evaluate the degree to which the conservation of short-term, monetary costs are offset by invisible costs that reduce top-line earnings (Bharadwaj, 2010).

According to prior research on service marketing, the components that comprise service quality are tangibles (physical facilities, equipment, and communication materials used when providing services to consumers), reliability (the ability to repeatedly provide the promised service in an effective manner), responsiveness (the desire to assist customers and provision of prompt service), assurance (employees' knowledge, politeness, and ability to inspire trust and confidence), and empathy (individual consideration given to customers) (Parasuraman, Zeithaml, and Berry, 1988; Whitaker, Krishnan, and Fornell, 2008). There are very few cues by which service encounters in CSC offshoring can be judged. For face-to-face service interactions, a consumer can evaluate the service he or she has received based on a broad range of cues (for instance, physical environment and employee attributes and behavior). Furthermore, in CSC offshoring, various types of distance (for example, cultural, linguistic, and geographical) are added between the service agent and consumer (Ceci and Prencipe, 2013; Hopkins, Hopkins, and Hoffman, 2005; Whitaker, Krishnan, and Fornell, 2008). As a result, factors such as the service agent's voice (for example, the consumer judges the sex of the service agent by voice), accent, and customer orientation (to what degree the customer's service-related expectations are fulfilled) serve as some of the only significant cues for the consumer in offshore CSC service encounters (Walsh, Gouthier, Gremler, and Brach, 2012).

Prior surveys and research on CSC offshoring suggest that consumers tend to judge CSC offshoring service agents by comparing them to onshore CSC agents (Bharadwaj and Roggeveen, 2008; Jeong, Bekmamedova, and Kurnia, 2012; Sharma, 2012; Sharma, Mathur, and Dhawan, 2009; Thelen, Magnini, and Thelen, 2009; Walsh, Gouthier, Gremler, and Brach, 2012). In a famous example of CSC offshoring, Dell Inc. relocated its help desk from India back to North America from 2006 to 2007 (Robinson, Kalakota, and Sharma, 2005; Sharma, Mathur, and Dhawan, 2009). This was because consumers felt that the Indian agents overly depended on scripted answers and did not possess the specialized knowledge required to answer complex computer-related questions, compared to North American CSC agents (Sharma, Mathur, and Dhawan, 2009). Prior research also found that consumers found it difficult to engage in conversations with offshore CSC agents who had very different accents than consumers and could not be understood. Other studies (Sharma, 2012; Sharma, Mathur, and Dhawan, 2009) demonstrate that consumers who are highly ethnocentric have elevated negative perceptions of offshored service. Thelen, Magnini, and Thelen (2009) concur with the preceding and go on to assert that with higher dissimilarity between cultures comes more negative feelings about service encounters with the culture. Conversely, homophily theory (Bharadwaj and Roggeveen, 2008) suggests that, when consumers and CSC agents are located in the same country there is an increased degree of homogeneity within the dyad and consumers feel greater satisfaction with the CSC agents' communication skills and problem-solving abilities.

CSC offshoring and prototype theory: In CSC offshoring, the consumer is being made to unexpectedly encounter offshore agents when posing an inquiry. Furthermore, offshoring firms work to conceal the fact that CSC is being offshored (Honeycutt, Magnini, and Thelen, 2012; Thelen and Shapiro, 2012; Thelen, Yoo, and Magnini, 2011). Consumers sometimes realize they are participating in CSC offshoring and sometimes do not. This section will consider how consumers form their realizations.

It is difficult to learn about consumer perceptions of a corporation's CSC agents by investigating individual CSC agents due to individual interactions between consumers and agents and different communication content. Protoype theory (Whittlesea, 1997) may be useful in understanding consumer perceptions. Prototype theory suggests that individuals construct cognitive frameworks representing prototypes via individual experiences (Rosch, 1978; Whittlesea, 1997). The constructed framework is used by the individual to classify and categorize newly obtained information. Prototypes are representative examples for perceiving the most important meanings and characteristics of the experience and are made up of multiple attributes (Baron and Ensley, 2006). Examples of attributes for CSC agents could include "speaks without an accent", "responds promptly to customers", "understands the content of questions posed by consumers", "has specialized knowledge", and "responds flexibly to customers without relying on script answers". Prototype theory supposes that individuals compare new experiences to existing prototypes. If there is a large degree of agreement, the experience is perceived as being positioned inside the prototype (Baron and Ensley, 2006). If there are too many differences, the experience is not perceived as being positioned inside the prototype. When consumers interact with a corporation that offshores its CSC, they classify whether the agents are onshore CSC agents by judging to what degree the CSC agents possess the attributes that comprise the unique qualities of onshore CSC agents.

In research on prototype categorization in the cognitive science field, such as Lakoff (1987), it is thought that people possess category concepts categorized by centrality. Categories are depicted as a series of concentric rings and members who are more distant from the central prototype are less significant (Kogut, MacDuffie, and Ragin, 2004). In prior research on prototypes, this is depicted using a fuzzy set (Kamp and Partee, 1995; Kogut, MacDuffie, and Ragin, 2004). A fuzzy set is a polythetic category that classifies to what degree a matter is affiliated with a prototype (degree of membership) for each attribute (Kogut, MacDuffie, and Ragin, 2004). A value 1 might be typically assigned if a matter is affiliated with a specific group and the value 0 if it is not. The value 0 means the matter does not have membership in the specific group, and 1 means that it has full membership in a specific group. When utilizing a fuzzy set, the matter is expressed in fuzzy algebra and it is possible to use various values from 0 to 1. Thus, it is possible to consider the degree to which the matter is affiliated with the specific group. Degree of membership refers to what degree the fuzzy set components belong to the applicable group (Ragin, 2009). Three qualitative breakpoints are essential for evaluating this (Ragin, 2009). These are full belonging (degree of membership is 1), full not belonging (degree of membership is 0), and branch point (degree of membership is 0.5). A fuzzy set attempts to determine qualitative circumstances and also evaluates the various degrees of membership between full belonging and full not belonging (Smithson and Verkuilen, 2006). The branch point expresses maximum ambiguity, which is judged as a specific matter both belonging and not belonging to a specific group

(Ragin, 2009).

According to prototype research in the field of psychology, prototypes are not stable things but change as a person gains more experience (Knowlton, 1997). There are three noteworthy types of changes: clarity, richness of content, and degree of focus on key attributes of the content domain (Baron and Ensley, 2006). Regarding shifts in clarity, prior research indicates that prototypes become more clearly defined as experience increases (Baron and Ensley, 2006). Specifically, the group of fundamental attributes that comprise the prototype possessed by individuals comes into conformance. For example, some of the attributes of good CSC units that have emerged include speaks fluently, responds promptly and flexibly, understands consumer inquiries fully and correctly, and possesses specialized knowledge. Regarding richness of content, the number of attributes included in a prototype increases as experience increases (Baron and Ensley, 2006). Regarding focus, as consumers gain more experience related to CSC agents, the prototype of CSC agents they possess may come to be focused on the attributes that are most strongly related to the performance of the CSC unit rather than on trivial attributes that are not central to performance.

The discussion up until this point can be summarized as follows. In mixed offshoring, in many cases consumers form a prototype of CSC agents based on onshore CSC agents. Over time, this attribute-based prototype has likely changed to some degree in terms of clarity, richness of content, and degree of focus (Baron and Ensley, 2006). Based on this refined prototype, the consumer has experiences with new CSC agents and then judges whether or not he or she is interacting with an offshore CSC agent.

DEVELOPING A CONCEPTUAL FRAMEWORK

Service-dominant logic: Through the S-D logic lens, the relationship between provider and customer is seen as the relationship of equals in a process of value co-creation (e.g., Lusch and Vargo, 2006; Vargo and Lusch, 2004; Vargo and Lusch, 2008). Value is created through the exchange of a service or knowledge between provider and customer (Vargo and Lusch, 2004; Vargo, Maglio, and Akaka, 2008). A viable co-creation system adapts to resolving dispute that may arise for the customer and/or provider (Maglio, Vargo, Caswell, and Spohrer, 2009). S-D logic suggests the important role of learning and improvement within service systems for effective dispute resolution, leading to value creation (Maglio, Vargo, Caswell, and Spohrer, 2009).

This paper treats offshoring as an attempt to reproduce for offshore as well as onshore CSC agents, an agent image that is close to the onshore agent prototype. These image reproductions are implemented via mutual efforts by the offshoring firm and CSC service provider (left half side of Figure 1). The degree to which the efforts are successful are reflected by how consumers evaluate the interaction (right half side of Figure 1).

A body of research has examined the relationship between the type of offshored process and the corresponding role expected of personnel (e.g., Tambe and Hitt, 2012). A typology of processes for offshorability can be determined by the difficulty of achieving offshorability. In the case of processes that involve a high level of explicit knowledge, learning will be relatively easy since the process can be captured in repeatable routines that can be codified. When it comes to processes that involve a high level of implicit knowledge, learning will be more difficult (e.g., Ceci and Prencipe, 2013; Kalaignanam and Varadarajan, 2012; Kumar, Van Fenema, and Von Glinow, 2009; Le Bon and Huhges, 2009; Liu, Felis, and Scholnick, 2011; Stringfellow, Teagarden, and Nie, 2008; Youngdahl and Ramaswamy, 2008; Wuyts, Rindfleischb, and Citrinc, 2015). The exchange of knowledge between offshoring companies and their offshore agents that is necessary, a co-creation exchange, should lead to a situation in which offshore agents perform at a standard that is more or less equivalent to home country personnel (e.g., Bharadwaj, 2010). As mentioned earlier, there are three types of service/knowledge exchanges with CSC offshoring. There are exchanges 1) between the firm that is offshoring and offshore service provider, 2) between the offshore service provider and the offshore agents that they hire, and 3) between offshore agents and consumers from the home country of the firm that is now offshoring their CSC.

Value co-creation model in mixed offshoring:

Problem-solving perspective in mixed offshoring: The PSP perspective helps to explain knowledge development efforts of a firm. The perspective has been used to understand innovative activities within the firm (e.g., Heiman, Nickerson, and Zenger, 2009; Hsieh, Nickerson, and Zenger, 2007; Macher, 2006). It is also an effective lens through which to examine value-creation (Nickerson, Silverman, and Zenger, 2007). The starting point of PSP is that a firm has an objective that can be achieved by improving its current knowledge state (Leiblein and Macher, 2009; Nickerson, Silverman, and Zenger, 2007; Nickerson and Zenger, 2004). However, the knowledge that the firm seeks may not be readily available. Developing and procuring the necessary knowledge depends on the nature of the search effort (Simon, 1962). PSP analyzes the fit between the identified problem and the knowledge used to solve the problem, subject to constraints such as time and money (Levinthal and Warglien, 1999; Nickerson, Silverman, and Zenger, 2007). In the case of CSC offshoring, the starting point of PSP is that the offshoring firm aspires to manage CSC processes more effectively or efficiently by moving the processes offshore without their offshore CSC agents departing from the CSC prototype. The home country, offshoring firm, and the out-of-country offshore service provider they have chosen will accomplish this via a co-creational, knowledge exchange process (Tate and Ellram, 2012).

PSP applies the NK model, a form of Boolean network, to solution finding efforts in mixed offshoring. The basic NK model allows us to explore ways to improve our current state of knowledge (Levinthal, 1997; Levinthal and Warglien, 1999). In general, N parameter indicates the number of elements in a system and K indicates the degree of connectivity among elements. Applied to CSC offshoring, N is the number of pertinent prototype attributes for onshore

CSC agents and K is the degree of interaction among those attributes. The values of N and K jointly form an attribute landscape. The peaks in the landscape signal important attribute strings. By utilizing this attribute landscape, the adequacy of the mixed offshoring, co-creational service exchange can be compared to the attribute string of the mixed offshoring prototype in an effort to improve the offshoring endeavor.

Knowledge exchange within mixed offshoring behind the establishment of CSC agents' prototype: Now let $\{ATT_{ON\ 1}, ATT_{ON\ 2}, \ldots, ATT_{ON\ m}\}$ and $\{ATT_{OFF\ 1}, ATT_{OFF\ 2}, \ldots, ATT_{OFF\ n}\}$ denote the prototype attributes of the onshore CSC agent image and the attributes of the offshore CSC agent, respectively. These will be referred to as attribute strings. As described previously, consumers evaluate the offshore CSC agents they encounter with the onshore CSC prototype as a reference point.

Generally, therefore, future offshored CSC processes are performed by $\{AG_{ON\ 1}, AG_{ON\ 2}, \ldots, AG_{ON\ i}\}$. $AG_{ON\ i}$ denotes the i position of the onshore CSC agent. Agents perform the CSC process, each with unique knowledge that is collectively essential to performing the process successfully. In the paper, we describe this as $Config\{AG_{ON\ 1\ knowledgeset}, AG_{ON\ 2\ knowledgeset}, \ldots, AG_{ON\ i\ knowledgeset}\}$. This expresses the formation of knowledge possessed by onshore CSC agents. Once a firm commences offshoring, part of $\{AG_{ON\ 1}, A_{ON\ 2}, \ldots, A_{ON\ i}\}$ is replaced with $\{AG_{OFF\ 1}, AG_{OFF\ 2}, \ldots, AG_{OFF\ j}\}$. They have knowledge $Config\{AG_{OFF\ 1\ knowledgeset}, AG_{OFF\ 2\ knowledgeset}, \ldots, AG_{OFF\ j\ knowledgeset}\}$ which differs from that of offshore agents. It should be noted that the number of replaced onshore agents and the number of newly introduced offshore agents are not always equal (Tambe and Hitt, 2012). So, let e denote the number of replaced onshore agents. Although the value of e may be $e < i$, $e = i$, or $e > i$, what is more commonly the case is $e \cong i$. Establishment of valuable mixed offshoring is, therefore, pursued by the following configuration.

$$Config\left\{\begin{array}{c} Config\left(AG_{ON\ 1\ knowledgeset}, AG_{ON\ 2\ knowledgeset}, \ldots, AG_{ON\ i-(e\cong j)\ knowledgeset}\right), \\ Config\left(AG_{OFF\ 1\ knowledgeset}, AG_{OFF\ 2\ knowledgeset}, \ldots, AG_{OFF\ j\ knowledgeset}\right) \end{array}\right\} \quad \text{.......... (1)}$$

The configuration consists partly of common knowledge. Common knowledge is the knowledge that the communicating parties share in common (Cramton, 2001; Srikanth and Puranam, 2011). Frequent interaction among agents within an offshoring firm would help onshore agents to understand the extent of the common knowledge and unique knowledge to some extent (e.g., Ansari, Sidhu, Volberda, and Oshri, 2011; MacDuffie, 2007; Sarker and Sahay, 2004). It is inevitable, however, that they will not be aware of all of what is common and unique knowledge (MacDuffie, 2007). The same will be true for the offshore agents.

Some prior research has investigated the types of offshore provider organizations. These include multinational or local company subsidiaries, multinational or local outsourcing service providers, multinational or local specialists, and multinational or local startups (e.g., Bharadwaj and Roggeveen, 2008; Contractor, Kumar, Kundu, and Pedersen, 2010; Kedia and Mukherjee, 2009; Mudambi and Venzin, 2010). They all have some experience in performing processes that share some similarity with offshored processes. This means offshore agents may partially establish common knowledge regarding offshored CSC processes. Therefore, configuration (1) can be redescribed as the following.

$$Config\left\{\begin{matrix} Config\begin{pmatrix} AG_{ON\ commonknowledge}, AG_{ON\ 1\ uniqueknowledgeset}, AG_{ON\ 2\ uniqueknowledgeset}, \\ \dots, AG_{ON\ i-(e\cong j)\ uniqueknowledgeset} \end{pmatrix}, \\ Config\begin{pmatrix} AG_{OFF\ commonknowledge}, AG_{OFF\ 1\ uniqueknowledgeset}, AG_{OFF\ 2\ uniqueknowledgeset}, \\ \dots, AG_{OFF\ j\ uniqueknowledgeset} \end{pmatrix} \end{matrix}\right\}$$

.... (2)

Prior studies on the intersection of offshoring and distributed work have suggested that in non-offshoring, offshoring, and mixed offshoring, personnel are apt to focus on common knowledge rather than unique knowledge in discussing a problem (e.g., Cramton, 2001; MacDuffie, 2007; Stasser and Titus, 1985; Stringfellow, Teagarden, and Nie, 2008). This paper, therefore, assumes that the efforts to establish mixed offshoring will be pursued in terms of common knowledge (Ceci and Prencipe, 2013; Sidhu and Volberda, 2011; Srikanth and Puranam, 2011). So we rewrite configuration (2) to reduce redundancy.

$$Config\{Config(AG_{ON\ commonknowledge}), Config(AG_{OFF\ commonknowledge})\}$$

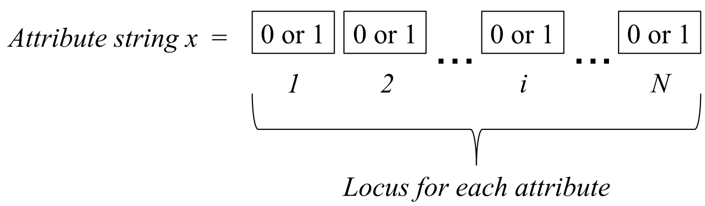
.... (3)

Several studies have viewed the coordination between onshore and offshore personnel as developing common knowledge between the two (Ceci and Prencipe, 2013; Cramton, 2001; Jeong, Bekmamedova, and Kurnia, 2012; Sidhu and Volberda, 2011; Srikanth and Puranam, 2011). This dynamic is compatible with the knowledge exchange, co-creation of value conceptualization of S-D logic. Common knowledge is required to coordinate actions among agents (Aggarwal, Siggelkow, and Singh, 2011; Cramton, 2001; Srikanth and Puranam, 2011). By relying on common knowledge, the onshore agents' expectation for the offshore agents' action converges with actual offshore agents' action. This, may lead to a situation where CSC agents diverge from the prototype of CSC agent. Thus, for offshoring firms, the acceptable minimum level of CSC providers' performance is the same level of performance achieved by onshore agents only (e.g., Bharadwaj, 2010). Knowledge exchange must lead to a level of performance that is equal or greater than the level achieved solely by onshore personnel. Applying PSP to establishment of mixed offshoring, two key efforts are required. First, the offshoring firm and service provider need to evaluate and compare their attribute strings. Second, they should also evaluate their ability to increase value by comparing their attribute strings to that of the prototype attribute string. Otherwise, sub optimal mixed offshoring will result (Rivkin and Siggelkow, 2002, 2003).

NK modeling: As suggested in configuration (3), offshoring firm and service provider need to continue to refresh/develop their common knowledge. This process proceeds as they continue their attempts to identify and co-create the prototypical attribute string/landscape. This can be applied to the NK model which examines the interaction among factors within complex system (Ganco and Hoetker, 2009; Kauffman, 1993; Levinthal and Warglien, 1999; Nickerson and Zenger, 2004). In general, N is the number of factors within a system and the possible states of a factor is described as integer A. K is an indicator of the connectivity or the degree of interaction among factors in a system.

For this study, N is the number of attributes required to establish a new (prototype in mixed offshoring's) attribute string. Let us assume there are two pertinent attributes $A = 2$. In a string, which consists of N loci, if the attribute chosen for locus i takes the value of 1, the indication is that the attribute belongs to the prototype. If the attribute takes the value of 0, the attribute does not belong to the prototype (see Figure 3). Therefore, the number of possible states of attribute string is 2^N.

Figure 3. Applying NK model in PSP to mixed offshoring



- 0 means that this attribute of the agent over the phone doesn't correspond to the attribute of onshore agent.
- 1 means that this attribute of the agent over the phone correspond to the attribute of onshore agent.

Each set of N loci results in attribute string, x, whose performance is modeled as a randomized real number, between 0 and 1. A value close to 1 indicates an attribute string that leads to relatively good performance and a value close to 0 indicates an attribute string that leads to relatively poor performance. In principle, these attribute string values can then be plotted as heights on an attribute landscape where the peaks represent high performance and the valleys represent low performance (e.g., Levinthal and Warglien, 1999; McCarthy, 2004; Nickerson and Zenger, 2004).

NK model assumes that the attribute chosen for locus i makes contribution, $P_i(x)$, to overall performance of the attribute configuration (e.g., Kauffman, 1993; Solow, Vairaktarakis, Piderit, and Ming-chi, 2002). $P_i(x)$ is determined by integer parameter, K. K $(0 \leq K \leq N-1)$ is the degree of interaction among attributes, or the number of other attributes that affect the performance contribution of the attribute chosen for locus i. $K = 0$ indicates that the contribution of each attribute depends on itself. $K = N-1$ indicates that the contribution of each attribute depends on all the rest of the $N-1$

attributes.

The performance of the attribute string x is the average of performance contributions of each attribute, $P_i(x)$. Thus, $P(x) = \sum_{i=1}^{N} P_i(x)/N$. To create valuable mixed offshoring, onshore and offshore agents jointly attempt to optimize the contribution of each attribute to the prototype so that the resulting attribute string has the best performance.

DISCUSSION FOR VARIATIONS IN VIABILITY OF OFFSHORING

As suggested in the previous section, some offshored CSC processes are easy to translate into explicit knowledge and some are not. In the former case, onshore and offshore personnel perform their activities by utilizing rules based on explicit knowledge. In the latter case, things are more difficult. Some sort of prediction procedure to survey the attribute landscape may be necessary (Hsieh, Nickerson, and Zenger, 2007; Nickerson and Zenger, 2004) to successfully replicate the prototypical CSC attribute string.

PSP identifies two distinct searching approaches (Leiblein and Macher, 2009; Nickerson, Silverman, and Zenger, 2007; Nickerson and Zenger, 2004). One is the directional approach by which offshore and onshore personnel use trial and error to evaluate alternative attribute strings. This approach works well when the CSC exchange process can be deconstructed into clearly defined, mutually exclusive components (e.g., Hsieh, Nickerson, and Zenger, 2007; Leiblein and Macher, 2009; Nickerson and Zenger, 2004). The other approach is the heuristic approach and works best when components of the CSC exchange process are interactive and overlapping. The knowledge exchange process in this case tends to be based on a fully integrated overview of the process rather than its component parts (e.g., Hsieh, Nickerson, and Zenger, 2007; Nickerson and Zenger, 2004). With this method, it is useful to be knowledgable of a wide range of fully integrated CSC offshoring models to ensure that an optimal starting point/model is chosen.

Most successful co-creation in the context of mixed offshoring is likely to combine directional search with heuristic search (e.g., Leiblein and Macher, 2009; Nickerson, Silverman, and Zenger, 2007; Nickerson and Zenger, 2004). This might possibly involve a process by which, first, the CSC attribute string is altered by jumping from one fully constituted, comprehensive attribute string configuration to another (heuristic approach). Then, observing the resulting change, individual components that can be isolated may be tweaked to improve the CSC attribute string/landscape (directional approach) (e.g., Leiblein and Macher, 2009; Levinthal, 1997; Nickerson and Zenger, 2004).

Previous studies have indicated that ongoing communication of knowledge can facilitate the continuous change for a CSC system (e.g., Ceci and Prencipe, 2013; Jensen, Larsen, and Pedersen, 2012; Larsen, 2016; Schmeisser, 2013; Srikanth and Puranam, 2011; Sharma, Iyer, and Raajpoot, 2009; Vivek, Richey, and Dalela, 2009). Ongoing communication updates common knowledge about heuristic and directional changes and their impact on CSC offshoring performance. There are two interrelated hurdles, stemming from physical distance between onshore and offshore personnel, that impede the establishment of new common knowledge through ongoing communication (e.g., Ceci and Prencipe, 2013; Ghemawat, 2001; Hinds, Liu, and Lyon, 2011; Larsen, Manning, Pedersen, 2013; MacDuffie, 2007; Sartor and Beamish, 2014). First, physical distance leads to low interaction between offshore and onshore personnel who consequently tend to develop different thought worlds in onshore and offshore. Thought world refers to the community of people who have shared understanding about a certain domain of activity (Dougherty, 1992; Homburg and Jensen, 2007). Specifically, physical distance leads to different cognitive frames that arise naturally in functionally differentiated settings such as onshore and offshore service systems. Second, physical distance can also result in weak shared identity (Hanes, 2013; MacDuffie, 2007; Sidhu and Volberda, 2011). Shared identity is defined as the extent to which the group members are attracted to the core idea of the group (Hanes, 2013; MacDuffie, 2007; Sidhu and Volberda, 2011). In other words, it is the extent to which the members feel identification with the group to which they belong. In order to establish shared identity, individual personnel must not only be familiar with the surrounding context but also the surrounding contexts for other personnel (Cramton, 2001; Hanes, 2013; MacDuffie, 2007; Sarker and Sahay, 2004; Sidhu and Volberda, 2011). Due to these two interrelated hurdles, offshore service system and onshore service system tend to experience some degree of isolation and identifiable routes to CSC service improvement become isolated and “Galapagosized”.

CONCLUSION AND LIMITATIONS

This paper uses service dominant logic (S-D Logic) as a platform for understanding the importance of knowledge exchange as a critical ingredient in achieving efficiency and effectiveness in offshore CSC service systems at the B2C as well as the B2C interfaces. The paper also uses the problem solving perspective (PSP) as a framework for converting less than satisfactory CSC service systems into systems that approach the ideal, prototype system.

Although the perspective of these considerations is broad, there are still several facets that could not be fully covered in this paper. For instance, even after intensive training, it is difficult to change the cultural characteristics of the offshore CSC agents. It is also difficult to change consumers’ negative perceptions of other countries. These problems impact the quality of communication between the consumer and offshore agent. Thus, special consideration should be given in how the pervasive impact of cultural factors (Boussebaa, Sinha, and Gabriel, 2014; Hahn and Bunyaratavej, 2010; Metters, 2008; Peeters, Dehon, and Garcia-Prieto, 2015; Raassens, Wuyts, and Geyskens, 2014; Stringfellow, Teagarden, and Nie, 2008; Youngdahl, Ramaswamy, and Dash, 2010).

The performance of offshored CSC can also be affected by the brand name of the CSC offshoring company. Roggeveen, Bharadwaj, and Hoyer (2007) showed that if the offshoring firm is well known, the offshored CSC's location does not impact the consumer's expectations of the voice-to-voice service encounter; however, if the offshoring firm is not well known, the consumer tends to feel dissatisfied with services provided by a CSC if the he or she thinks the CSC is not in the US or a similar country. Thus, the impact of brand equity on offshore CSC performance should undergo further examination (Bharadwaj, 2010; Chen, 2009; Sharma, 2012; Sharma, Mathur, and Dhawan, 2009). This conceptual paper should also be extended into an empirical investigation to determine whether the concepts covered are pertinent and to examine the extent to which the process that was outlined is capable of guiding the improvement of offshore CSC systems.

An empirical test should also consider various performance outcome dependent variables (Thelen, Yoo, and Magnini, 2011). These could include criteria such perceived quality of service (e.g., Sharma, Mathur, and Dhawan, 2009), perceived value of service (e.g., Sharma, 2012; Sharma, Mathur, and Dhawan, 2009), customer satisfaction (e.g., Whitaker, Krishnan, and Fornell, 2008), desire to make repeat purchases (e.g., Sharma, 2012; Sharma, Mathur, and Dhawan, 2009), customer dissatisfaction (e.g., Honeycutt, Magnini, and Thelen, 2012), boycotts by customers, (e.g., Thelen and Shapiro, 2012) and generation of positive/negative world of mouth (e.g., Walsh, Gouthier, Gremler, and Brach, 2012), to name a few.

REFERENCES

Aggarwal, V., Siggelkow, N., and Singh, H. (2011). Governing collaborative activity: Interdependence and the impact of coordination and exploration. Strategic Management Journal, 32(7), 705-730.

Ansari, S., Sidhu, J.S., Volberda, H., and Oshri, I. (2011). Managing globally disaggregated teams. In Farok J. Contractor, Vikas Kumar, Sumit K. Kundu, and Torben Pedersen (Eds.), Global outsourcing and offshoring: An integrated approach to theory and corporate strategy. (pp.297-323). Cambridge University Press, Cambridge.

Baron, R., and Ensley, M. (2006). Opportunity recognition as the detection of meaningful patterns: Evidence from comparisons of novice and experienced entrepreneurs. Management Science, 52(9), 1331-1344.

Benito, G., Dovgan, O., Petersen, B., and Welch, L. (2013). Offshore outsourcing: A dynamic, operation mode perspective. Industrial Marketing Management, 42(2), 211-222.

Bharadwaj, N. (2010). Offshoring and Marketing. In J. Sheth, and N. Malhotra (Eds.), Wiley International Encyclopedia of Marketing (Part 6). (pp.173-174). Wiley-Blackwell.

Bharadwaj, N., and Roggeveen, A. (2008). The impact of offshored and outsourced call service centers on customer appraisals. Marketing Letters, 19(1), 13-23.

Boussebaa, M., Sinha, S., and Gabriel, Y. (2014). Englishization in offshore call centers: A postcolonial perspective. Journal of International Business Studies, 45,1152-1169.

Ceci F., and Prencipe A. (2013). Does distance hinder coordination? Identifying and bridging boundaries of offshored work. Journal of International Management, 19(4), 324-332.

Chen, S. (2009). A transaction cost rationale for private branding and its implications for the choice of domestic vs offshore outsourcing. Journal of International Business Studies, 40, 156-175.

Contractor, F., Kumar, V., Kundu, S., and Pedersen, T. (2010). Reconceptualizing the firm in a world of outsourcing and offshoring: The organizational and geographical relocation of high-value company functions. Journal of Management Studies, 47(8), 1417-1433.

Cramton C. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. Organization Science, 12(3), 346-371.

Dougherty, D. (1992). Interpretive barriers to successful product innovation in large firms. Organization Science, 13, 77-92.

Ganco, M., and Hoetker, G. (2009). NK modeling methodology in the strategy literature: Bounded search on a rugged landscape. In D.J. Ketchen, and D. Bergh (Eds.), Research methodology in strategy and management (Vol.5). (pp.237-268). Elsevier, Oxford.

Ghemawat, P. (2001). Distance still matters: The hard reality of global expansion. Harvard Business Review, 79(8), 137-147.

Hahn, E., and Bunyaratavej, K. (2010). Services cultural alignment in offshoring: The impact of cultural dimensions on offshoring location choices. Journal of Operations Management, 28, 186-193.

Hanes, D. (2013). Geographically distributed audit work: Theoretical considerations and future directions. Journal of Accounting Literature, 32(1), 1-29.

Hatonen, J. (2009). Making the location choice: A case approach to the development of a theory of offshore outsourcing and internationalization. Journal of International Management, 15, 61-76.

Heiman, B., Nickerson, J., and Zanger, T. (2009). Governing knowledge creation: A problem-finding and problem-solving perspective. In Nicolai J. Foss, and Snejina Michailova (Eds.), Knowledge governance: Process and perspectives. (pp.25-46). Oxford University Press, Oxford.

Hinds, P., Liu, L., and Lyon, J. (2011). Putting the global in global work: An intercultural lens on the practice of cross-national collaboration. The Academy of Management Annals, 5(1), 135-188.

Homburg, C., and Jensen, O. (2007). The thought worlds of marketing and sales: Which differences make a difference? Journal of Marketing, 71(3), 124-142.

Honeycutt, E., Magnini, V., and Thelen, S. (2012). Solutions for customer complaints about offshoring and outsourcing services. Business Horizons, 55, 33-42.

Hopkins, S., Hopkins, W., and Hoffman, K. (2005). Domestic inter-cultural service encounters: An integrated model. Managing Service Quality, 15, 329-43.

Hsieh, C., Nickerson, J., and Zenger, T. (2007). Opportunity discovery, problem solving and a theory of the entrepreneurial firm. Journal of Management Studies, 44(7), 1255-1277.

Ito, R., and Gehrt, K. (2014). Identification of optimal offshoring strategies for small and medium Japanese consumer service companies. Journal of Business and Behavioral Sciences, 26(1), 75-87.

Jensen, P., Larsen, M., and Pedersen, T. (2013). The organizational design of offshoring: Taking stock and moving forward. Journal of International Management, 19, 315-323.

Jensen, P., and Pedersen, T. (2012). Offshoring and international competitiveness: Antecedents of offshoring advanced tasks. Journal of the Academy of Marketing Science, 40(2): 313-328.

Jeong, J., Bekmamedova, N., and Kurnia, S. (2012). Exploring the current status of call centre offshoring research: A research agenda and future directions. 23rd Australasian Conference on Information Systems. Retrieved 25 January, 2016, from https://dro.deakin.edu.au/eserv/DU:30049083/jeong-exploringthecurrent-2012.pdf.

Kauffman, S. (1993). The origins of order: Self organization and selection in evolution. Oxford University Press, New York.

Kalaignanam, K., and Varadarajan, R. (2012). Offshore outsourcing of customer relationship management: conceptual model and propositions. Journal of the Academy of Marketing Science, 40, 347-363.

Kamp, H., and Partee, B. (1995). Prototype theory and compositionality. Cognition, 57, 129-191.

Kedia, B., and Mukherjee, D. (2009). Understanding offshoring: A research framework based on disintegration, location and externalization advantages. Journal of World Business, 44(3), 250-261.

Kinkel, S. (2014). Future and impact of backshoring: Some conclusions from 15 years of research on German practices. Journal of Purchasing & Supply Management, 20(1), 63-65.

Knowlton, B. (1997). Declarative and nondeclarative knowledge: Insights from cognitive neuroscience. In K. Lamberts, and D. Shanks (eds.), Knowledge, concepts, and categories (pp.215-246). MIT Press, Cambridge, MA.

Kogut, B., MacDuffie, J., and Ragin, C. (2004). Prototypes and strategy: assigning causal credit using fuzzy sets. European Management Review, 1, 114-131.

Kumar, K., van Fenema, P., and von Glinow, M. (2009). Offshoring and the global distribution of work: Implications for task interdependence theory and practice. Journal of International Business Studies, 40, 642-667.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Lampel, J., and Bhalla, A. (2011). Living with offshoring: The impact of offshoring on the evolution of organizational configurations. Journal of World Business, 46, 346-358.

Larsen, M. (2016). Failing to estimate the costs of offshoring: A study on process performance. International Business Review, 25(1), 307-318.

Larsen, M., Manning, S., and Pedersen, T. (2013). Uncovering the hidden costs of offshoring: The interplay of complexity, organizational design, and experience. Strategic Management Journal, 34, 533-552.

Le Bon, J., and Hughes, D. (2009). The dilemma of outsourced customer service and care: Research propositions from a transaction cost perspective. Industrial Marketing Management, 38, 404-410.

Leiblein, M., and Macher, J. (2009). The problem solving perspective: A strategic approach to understanding environment and organization. In Jackson A. Nickerson, and Brian S. Silverman (Eds.), Economic institutions of strategy (Advances in Strategic Management, Volume 26). (pp. 97-120). Emerald Group Publishing.

Levinthal, D. (1997). Adaptation on rugged landscapes. Management Science, 43(7), 934-950.

Levinthal, D., and Warglien, M. (1999). Landscape design: Designing for local action in complex worlds. Organization Science, 10(3), 342-357.

Lewin, A., and Peeters, C. (2006). Offshoring work: Business hype or the onset of fundamental transformation? Long Range Planning, 39: 221-239.

Liu, R., Feils, D., and Scholnick, B. (2011). Why are different services outsourced to different countries? Journal of International Business Studies, 42, 558-571.

Lusch, R., Vargo, S., and Tanniru, M. (2010). Service, value networks and learning. Journal of the Academy of Marketing Science, 38(1), 15-31.

Lusch, R., and Vargo, S. (Eds.). (2006). The Service-dominant logic of marketing: Dialog, debate, and directions. M.E. Sharpe: New York.

MacDuffie, J. (2007). HRM and distributed work (Chapter 12). The Academy of Management Annals, 1(1), 549-615.

Macher, J. (2006). Technological development and the boundaries of the firm: A knowledge based examination in semiconductor manufacturing. Management Science, 52(6), 826-843.

Maglio, P., Vargo, S., Caswell, N., and Spohrer, J. (2009). The service system is the basic abstraction of service science. Information Systems E-Business Management, 7, 395-406.

Manning, S., Massini, S., and Lewin, A. (2008). A dynamic perspective on next-generation offshoring: The global sourcing of science and engineering talent. Academy of Management Perspectives, 22(3): 35-54.

McCarthy, I. (2004). Manufacturing strategy: Understanding the fitness landscape. International Journal of Operations and Production Management, 24(2), 124-150.

Metters, R. (2008). A case study of national culture and offshoring services. International Journal of Operations & Production Management, 28(8), 727-747.

Mudambi, R., and Venzin, M. (2010). The strategic nexus of offshoring and outsourcing decisions. Journal of Management Studies, 47(8), 1510-1533.

Murray, J., Kotabe, M., and Westjohn, S. (2009). Global sourcing strategy and performance of knowledge-intensive business services: A two stage strategic fit model. Journal of International Marketing, 17(4), 90-105.

Nickerson, J., Silverman, B., and Zenger, T. (2007). The ‘problem’ of creating and capturing value. Strategic Organization, 5(3), 211-225.

Nickerson, J., and Zenger, T. (2004). A knowledge-based theory of the firm: The problem-solving perspective. Organization Science, 15(6), 617-632.

Parasuraman, A., Zeithaml, V., and Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. Journal of Retailing, 64(1), 12-40.

Peeters, C., Dehon, C., and Garcia-Prieto, P. (2015). The attention stimulus of cultural differences in global services sourcing. Journal of International Business Studies, 46, 241-251.

Raassens, N., Wuyts, S., and Geyskens, I. (2014). The performance implications of outsourcing customer support to service providers in emerging versus established economies. International Journal of Research in Marketing, 31, 280-292.

Ragin, C. (2009), Qualitative comparative analysis using fuzzy sets (fsQCA). In B. Rihoux, and C. Ragin (Eds.), Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques. (pp.87-121). Sage Publications.

Rivkin, J., and Siggelkow, N. (2002). Organizational sticking points on NK landscapes. Complexity, 7(5), 31-43.

Rivkin, J., and Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. Management Science, 49(3), 290-311.

Robinson, M., Kalakota, R., and Sharma, S. (2005). Global outsourcing: Executing an onshore, nearshore or offshore strategy. Mivar Press.

Roggeveen, A., Bharadwaj, N., and Hoyer, W. (2007). How call center location impacts expectations of service from reputable versus lesser known firms. Journal of Retailing, 83(4), 403-410.

Rosch, E. (1978). Principles of categorization In E. Rosch and B.B. Lloyd (eds.), Cognition and categorization (pp.27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Roza, M., Van den Bosch, F., and Volberda, H. (2011). Offshoring strategy: Motives, functions, locations, and governance modes of small, medium-sized and large firms. International Business Review, 20, 314-323.

Sarker, S., and Sahay, S. (2004). Implications of space and time for distributed Work: An interpretive study of US-Norwegian systems development teams, European Journal of Information Systems, 13, 3-20.

Sartor, M., and Beamish, P. (2014). Offshoring innovation to emerging markets: Organizational control and informal institutional distance. Journal of International Business Studies, 45, 1072-1095.

Schmeisser, B. (2013). A systematic review of literature on offshoring of value chain activities. Journal of International Management, 19(4), 390-406

Sharma, A., Iyer, G., Raajpoot, M. (2009). A framework for offshoring marketing processes in business-to-business marketing relationships. Industrial Marketing Management, 38, 419-425.

Sharma, P. (2012). Offshore outsourcing of customer services: Boon or bane? Journal of Services Marketing, 26(5), 352-364.

Sharma, P., Mathur, R., and Dhawan, A. (2009). Exploring customer reactions to offshore call centers: Toward a comprehensive conceptual framework. Journal of Services Marketing, 23, 289-300.

Sidhu, J., and Volberda, H. (2011). Coordination of globally distributed teams: A co-evolution perspective on offshoring. International Business Review, 20 (3), 278-290.

Simon, H. (1962). The architecture of complexity. Proceedings of the American Philosophical Society, 106, 467-482.

Smithson, M., and Verkuilen, J. (2006). Fuzzy set theory: Application in the social sciences. Sage Publications.

Solow, D., Vairaktarakis, G., Piderit, S., and Ming-chi, T. (2002). Managerial insights into the effects of interactions on replacing members of a team. Management Science, 48(8), 1060-1073.

Srikanth, K., and Puranam, P. (2011). Integrating distributed work: Comparing task design, communication, and tacit coordination mechanisms. Strategic Management Journal, 32(8), 849-875.

Stasser, G., and Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. Journal of Personality and Social Psychology, 48, 1467-1478.

Stringfellow, A., Teagarden, M., and Nie, W. (2008). Invisible costs in offshoring services work. Journal of Operations Management, 26(2), 164-179.

Tambe, P., and Hitt, L. (2012). Now IT's personal: Offshoring and the shifting skill composition of the US information technology workforce. Management Science, 58(4), 678-697.

Tate, W. (2014). Offshoring and reshoring: U.S. insights and research challenge. Journal of Purchasing & Supply Management, 20(1), 66-68.

Tate, W., and Ellram, L. (2012). Service supply management structure in offshore outsourcing. Journal of Supply Chain Management, 48(4), 8-29.

Thelen, S., Magnini, V., and Thelen, T. (2009). Identifying elements of offshore service ethnocentrism. Retrieved 25 January, 2016, from http://www.hofstra.edu/pdf/BIZ_mlc_workingpaper6.pdf.

Thelen, S., and Shapiro, T. (2012). Predicting negative consumer reactions to services offshoring. Journal of Services Marketing, 26(3), 181-193.

Thelen, S., Yoo, B., and Magnini, V. (2011). An examination of consumer sentiment toward offshored services. Journal of the Academy of Marketing Science, 39, 270-289.

Vargo, S., Maglio, P., and Akaka, M. (2008). On value and value co-creation: A service systems and service logic perspective. European Management

Journal, 26(3), 145-152.
Vargo, S., and Lusch, R. (2004). Evolving to a new dominant logic for marketing. Journal of Marketing, 68(1), 1-17.
Vargo, S., and Lusch, R. (2008). Service-dominant logic: Continuing the evolution. Journal of the Academy of Marketing Science, 36(1), 1-10.
Vivek, S., Richey, R., and Dalela, V. (2009). A longitudinal examination of partnership governance in offshoring: A moving target. Journal of World Business, 44, 16-30.
Walsh, G., Gouthier, M., Gremler, D., and Brach, S. (2012), What the eye does not see, the mind cannot reject: Can call center location explain differences in customer evaluations? International Business Review, 21(5), 957-967.
Whitaker, J., Krishnan, M., and Fornell, C. (2008). Does offshoring impact customer satisfaction? AMCIS 2006 Proceedings, Paper 394.
Whittlesea, B. (1997). The representation of general and particular knowledge. In K. Lamberts, D. Shanks (eds.), Knowledge, concepts, and categories (pp.335-370). MIT Press, Cambridge, MA.
Wuyts, S., Rindfleischb, A., and Citrinc, A. (2015). Outsourcing customer support: The role of provider customer focus. Journal of Operations Management, 35(5), 40-55.
Youngdahl, W., and Ramaswamy, K. (2008). Offshoring knowledge and service work: A conceptual model and research agenda. Journal of Operations Management, 26(2), 212-221.
Youngdahl, W., Ramaswamy, K., and Dash, K. (2010). Service offshoring: The evolution of offshore operations. International Journal of Operations & Production Management, 30(8), 798-820.

Journal of Business and Behavioral Sciences
Vol 28, No. 2; Fall 2017

EXPLORING OPERATIONAL RESILIENCE IN THE CONTEXT OF MILITARY AVIATION: FINDING THE RIGHT MODE AT THE RIGHT TIME

Timothy J. Pettit
National University
Natalie C. Simpson
University at Buffalo (SUNY)
Philip G. Hancock
Glenochil Associates LLC
Hayden Clark
Trevor Haydel
Jordan Pierce
United States Air Force Academy

ABSTRACT: Operational resilience is the ability of an organization to avoid disruption, and if unsuccessful in avoidance, to recover as quickly as possible. We begin with an a priori assumption that operational resilience is dependent in part on an organization's ability to change its mode of operation in response to disruptive conditions, and that the antecedents of this ability will be most apparent in organizations that routinely change mode out of necessity. The United States Air Force is one such example of a routinely multi-modal organization, thriving on rapid problem solving and effectiveness in wartime, while striving for cost savings and efficiency in home station operations and peace-time. This project focuses on uncovering those policies and procedures associated with quick and effective transition between operational modes, by examining and contrasting the practices of two Air Force aircraft systems. This exploratory work is the initial phase of a project dedicated to both developing new methods to improve operational resilience and identifying those generalizable principles and practices that may enhance resilience in all types of operations and organizations. Initial results will be presented as several themes of what enhances resilience and what limits resilience with respect to effective transition between modes. Further investigation and analysis are forthcoming.

Key Words: Resilience, Operations management, Responsiveness, Flexibility, Operating modes

INTRODUCTION

'Constant change is the new normal' is a theme of the modern business world. Organizations are evaluating recent natural disasters, world health threats, global financial crises, cyber attacks and terrorism threats to learn from the past and seek robustness, agility and resiliency despite decades of cost reductions and streamlining under lean philosophies (Fiksel et al., 2015). But it seems that understanding current environmental conditions and vulnerabilities still lags in many firms (Pettit et al., 2013). Even then, in a dynamic world simply knowing your current environment must be enhanced with sensing it for changes and repeatedly adapting operations to match, to avoid or at least mitigate supply chain disruptions (Ambulkar, Blackhurst and Grawe, 2015). This study leverages current theory on resilience attributes to evaluate how organizations can quickly and effectively change their operational modes to more accurately match their changing environment.

LITERATURE REVIEW

Operational resilience in the field of business is a new and evolving concept. Initial definitions were created from resilience applications in engineering as, "the tendency of a material to return to its original shape after the removal of a stress that has produced elastic strain" (Merriam-Webster, 2007), from ecology as, "the ability for an ecosystem to rebound from a disturbance while maintaining diversity, integrity, and ecological processes" (Folke et al., 2004) and from psychology as, "an individual's stability or quick recovery (or even growth) under significant adverse conditions'' (Leipold & Greve, 2009). These underlying objectives have been translated to the enterprise of business as, "the ability to survive, adapt and grow in the face of turbulent change" (Fiksel, 2006). Recent works have explored the characteristics that develop resilience, focusing on the capabilities that enhance resilience and the vulnerabilities that make an enterprise susceptible to disruptions (Christopher and Peck, 2004; Sheffi, 2005; Pettit, Fiksel and Croxton 2010; Blackhurst et al., 2011; Brandon- Jones et al., 2014; Habermann, Blackhurst, and Metcalf, 2015 among others). Much of this work has focused on the application of resilience toward preventing, or at least mitigating, operational disruptions caused by natural disasters, supplier problems, market change or terrorist attacks. Hence, these concerns primarily arise in the domain of emergency and disaster response. However, there is often a poor understanding of disaster conditions in this earlier body of work, and many use ill-fitting assumptions from outside disaster domain (Green and Kolesar, 2004; Altay and Green, 2006; Boin et al., 2010; Charles et al., 2010; Pettit and Beresford 2009; Pedraza Martinez et al., 2011; Holguin-Veras, et al., 2014). Just as the field of resilience is evolving, operational management during emergencies or disasters is perceived as new, young, and/or underdeveloped (Boin, 2009; Kovacs and Spens, 2009; Boin et al., 2010; Janssen et al., 2010; Tatham and Pettit, 2010). Despite advances, operational resilience in the area of emergency and disaster response still lacks the theoretical understanding needed to integrate case study findings (Altay and Green, 2006; Boin et al., 2010). Compounding this problem, emergency and disaster response management also lacks performance metrics to properly evaluate success or failure (Thomas and Kopczak, 2005; Altay and Green, 2006; Beamon and Balcik, 2008; Boin et al., 2010; Janssen et al., 2010). With those limited metrics, existing emergency and disaster-related research heavily favors mathematical programming (Altay and Green, 2006; Simpson and Hancock, 2009; Galindo and Batta, 2013), which generally render the system under management as a given set of processes that interact according to a single mode of operation.

MODELING RESILIENCE AS MULTIPLE MODES

Moving beyond resilience as a single operational mode, there is a lack of understanding of how organizations can adjust their operations between modes or states, or even know when to adjust, let alone prepare for quick and effective operational adjustments. The foundation of our current understanding is largely built upon the iconic IPO model (Input-Process-Output), featured heavily in the teaching of operations management (Fowler, 1999). The IPO model is grounded in control theory, suggesting that any operation can be framed as a set of inputs passing through a transformation stage to produce outputs, all governed by some controlling element (Astrom and Murray, 2012).

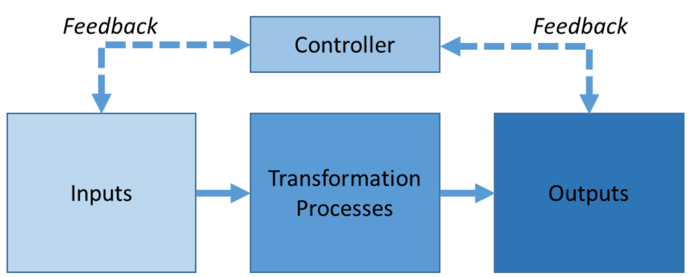


Figure 1: The Input-Process-Output Model

The assumption here that the controlling element is sensing any relevant change or threat to the system and adjusting its regulation accordingly. Beyond adjusting levels of input and output, however, the IPO model itself does not illuminate the controller's options in guiding the operation, particularly toward appropriate changes to the *modality* of the transformation processes. Choices of modality, or the essence of how something is done, have long been outlined in another highly cited conceptual framework appearing in Figure 2, the 'product-process matrix' of Hayes and Wheelwright (1979). Modern presentation often appends project management to the lower right-hand end of the original model's efficient diagonal, but even this revision does not distinguish an emergency response from routine projects such as building construction or graphic design. This is a significant limitation of this framework's inability to draw strategic distinctions between any activity in that category (Oltra, et al., 2005). The service-oriented taxonomy of Schmenner (1986) (Figure 3) clarifies emergency response and acts of resilience as 'professional service', grouping this domain with the activities of accountants, lawyers, consultants and architects.

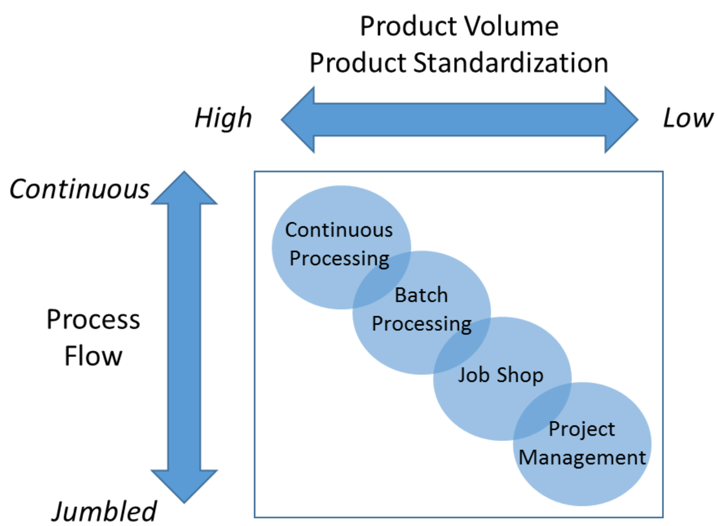


Figure 2: The Product-Process Matrix (Hayes and Wheelwright, 1979)

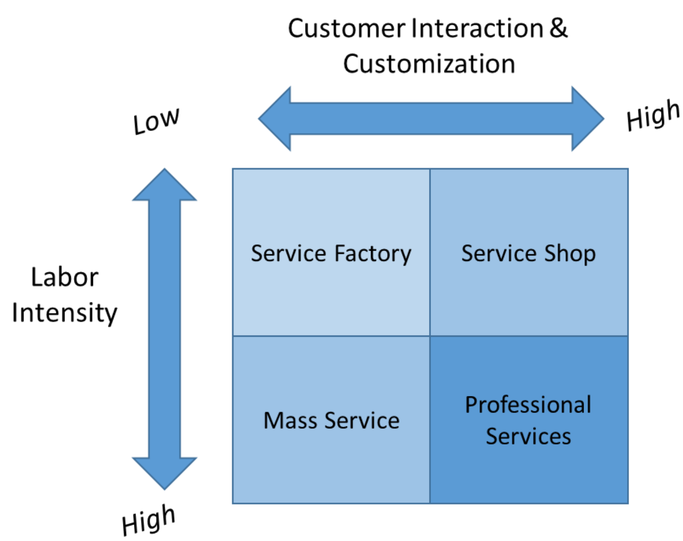


Figure 3: Service Process Matrix (Schmenner, 1986)

Figures 1 through 3 define the essence of operations management, or OM. Although they provide little insight into operational resilience, we nonetheless hypothesize that building a better model of resilience begins with these constructs, as the domain the OM itself has a compelling association with turbulent conditions. For example, the influential work of Weick and Sutcliffe (2001) cite 'attention to operations' as one of five practices central to successful crisis management. Meredith (2001) observes that the popularity of OM moves counter to economic stability and prosperity, in that OM captures the most attention during turbulent periods and languishes otherwise.

Consistent with the IPO illustration, we assume that any operation has a controller, defined as the agent, group, algorithm or mechanism assigned the role of keeping the operation's activities aligned with its objectives through time. However, operational resilience cannot be fully related to the foundational frameworks of OM until the issue of that entity's *degree of control* is recognized. This aggregate measure is the hypothesized result of the interaction of multiple factors specific to the operation, some of which are illustrated in Figure 4. As the divisions within this illustration suggest, these factors can be grouped into three sets that reflect the original IPO model – input and output (external) and process (internal). First, to involve another organization in an operation reduces the degree of control enjoyed by its resident controller (Hayes, 2002). Thus, the left-hand side of Figure 4 suggests a major driver of input-side control is the number of external contributors to the operation, embedded in a set of moderating factors. This powerful dynamic has been the subject of increasing attention in OM literature, as researchers call for better understanding of supply chain partnering and for reframing procurement as a strategic issue (Gunasekaran and Ngai, 2012). Pursuing such goals bolsters control even as the growing complexity of global supply chain networks works against it. Moderating factors such as incentive schemes and greater investment in communication and coordination are characterized broadly as `liaison devices' supporting the control mechanism of mutual adjustment (Mitzenberg, 1980). Other factors like environmental interference are readily recognized as vulnerabilities of the operation, such as supply chain turbulence from natural causes or the threat of intentional interference (Fiksel et al., 2015).

Second, a major factor in degree of control on the output-side is the degree of customization of the deliverable. Product standardization, its inverse, has long been recognized in organization science as a powerful coordination mechanism, one which boosts internal control through predictability of activity and further enhances control by enabling the standardization of skills and work processes (Mitzenberg, 1980). The complexity of customization is then compounded by the number of clients the operation is attempting to serve, and may be further complicated by time pressure and/or adverse environmental conditions for delivery. Finally, the visual complexity at the center of Figure 4 is created by factors associated with transformation, often influencing both input-side and output-side control simultaneously. As an example, shortening production lead times increases control, a known benefit from lean operation, but the introduction of uncertainty into those lead times detracts from this benefit. That uncertainty may be the cascading result of increasing environmental interference felt through an increase in the permeability of the working environment, such as an operation that is forced to move to a temporary location. The relative labor intensity of an operation is assumed to be inversely related to control because people cannot be controlled directly, unlike machinery.

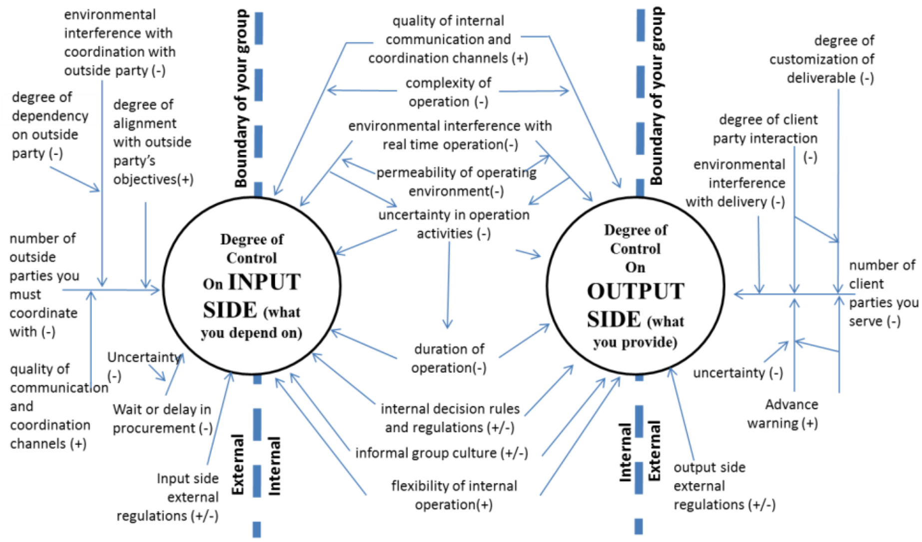


Figure 4: Potential Factors Influencing Control over an Operation

Contrasting the degree of control of the input- versus output-side of a given operation produces an alternate OM taxonomy, appearing in Figure 5. We present the OM domain as a four-box model because this is the minimum level of complexity necessary to distinguish emergency response from commercial

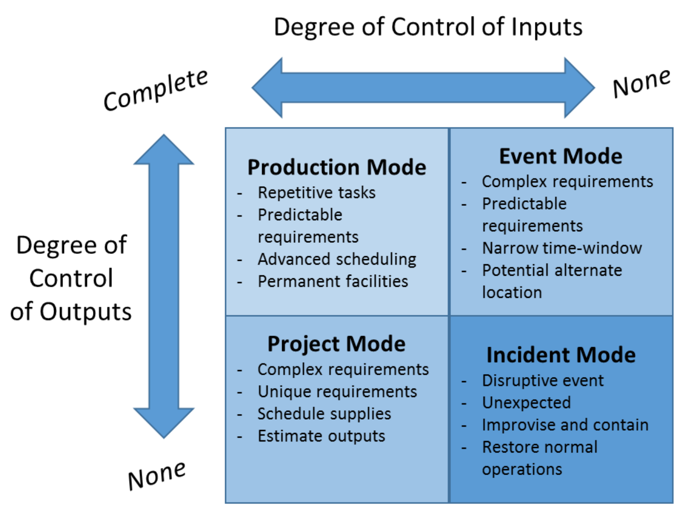


Figure 5: Modes of Operation (Simpson and Hancock, 2012)

businesses, one of the original goals of its development. Again, despite the long history of managing operational modes, there remains a lack of recognition in the business world that an operation is not necessarily dedicated to a single mode, and thus there is a limited understanding of when to alter mode and how best to move between modes quickly. The goal of quickly and efficiently transitioning between modes – at the right time and place – would create what is defined here as operational resilience.

METHODOLOGY

In this pilot study, we explore the conceptualization of resilience as the ability and necessity of transition between known modalities through the study of two distinct component organizations within the Air Force, known as Commands. First, Air Force Special Operations Command represents the potentially more resilient operation of flying and sustaining the new CV-22 Osprey, an advanced and complex hybridization of a fixed-wing aircraft and a helicopter. Second, Air Combat Command provides contrast with the operation of its aging bomber, the B-1B Lancer. Although operational policies for both organizations are mandated through Air Force Instructions, each Command has evolved specific procedures to best suit their mission requirements. Each of the operations described above are also referred to as ‘weapons systems’ in Air Force vernacular, a term referring to the totality of the aircraft itself, the supply chains operating to maintain its readiness, the personnel required to maintain and fly the aircraft, and the pipeline of training required to keep the system operating.

Differing elements in the weapons systems’ vulnerabilities, or potential for disruption, include:

- Age of aircraft design and technology (Advantage: CV-22 Osprey. The CV-22 is the newer design first deployed in 2008, versus B-1B initial deployment in 1985; however, initial designs were from 1970s.)
- Age of aircraft fleet (Advantage: CV-22 Osprey. The average age of the CV-22 fleet is 5 years, while the B-1B fleet averages 27 years.)
- Complexity of aircraft design (Advantage: B-1B Lancer. Even with the B-1B’s offensive and defensive systems and a swing-wing design, the highly complex CV-22 Osprey is capable of both horizontal and vertical flight, with all the associated vibration dynamics and stresses.)
- Environmental exposure (Advantage: B-1B Lancer. The B-1B operates exclusively from bases with improved runways, whereas the CV-22 may operate independent of landing facilities.) (U.S. Air Force, 2015)

To gather data on operational resilience, small group interviews were conducted on-site with key functional experts drawn from various areas of the weapon system – pilots, maintenance crews and logisticians – to gather data on their operational procedures as related to this model of operational resilience. To facilitate this conversation, each group was asked to study the graphic in Figure 5, and confer on which modes describe their operation best at various times. Given that more than one mode was cited during discussion, participants were asked to identify which mode was their 'normal' or 'default' operation and what distinguished transition from one mode to another. Keeping with the systems perspective, participants were then asked to think of their operations in terms of its inputs and outputs, the foundational framework of control theory. At this point, each member of the group was provided a paper copy of Figure 4, a graphic of common influences over the degree of control inherent in the governance of their inputs and outputs. Participants were first invited to cross-out or revise any portion of the diagram that was not applicable or varied from the perspective of their operation. Next, participants were asked to circle any element they felt was particularly important to effectively completing their mission. In all, 6 group interviews were conducted with a total of 28 individuals participating in this initial study. Responses were transcribed and analyzed using qualitative coding for pattern matching (Yin, 2003).

ANALYSIS

To determine the important components of operational resilience, two methods were used to gain a broad understanding of the topics related to resilience: word clouds and pattern matching. Word clouds were created using the text from the CV-22 and B-1 focus groups. In particular, we used word clouds to quickly identify the most common words as a proxy for the most important topics in each group.

The second method was pattern matching, the process grouping common themes and then assigning a relevance of a given input or output based on the responses from each subject (see Figure 4), sorted by weapon system. Each factor was marked as a major factor, relevant, irrelevant, or unknown. This data allowed the extraction of the most critical factors as related to operational resilience through a comparison of the two weapons systems. Although not yet generalizable nor statistically verified, the combined responses follow several patterns.

From the initial discussion of the four modes of Figure 5 (Production, Project, Event and Incident), the data supports the following:

- At least two modes of operation were identified as applicable to participants' weapons system.
- Home-station activities and home-station training were strongly associated with 'normal' or 'default' operation, in turn associated with 'Production Mode.'
- Home-station exercises, deployed humanitarian missions and deployed combat operations were identified with 'Project Management', 'Event Management' and 'Incident Management,' respectively.

Within the visual complexity of Figure 4, the element, 'number of outside parties you must deal with' and its negative influence on the input side of control was frequently cited as particularly applicable. This in turn was reflected in several themes of what enhances resilience with respect to effective transition between modes:

- Leadership centralization on-site
- Decentralized execution
- Use of informal agreements when formal policies become a hindrance
- Recognition and universal acceptance of common goals
- Increased communications
- Flexibility in operations
- Experience in operating in various modes

Themes of limitations on operational resilience included:

- Lack of understanding about various modes of operations
- Lack of awareness of the current mode
- Limited risk taking
- Adherence to policies versus intent
- Sub-optimization of functional metrics versus enterprise-level objectives

Thus the modes of operation as developed by Simpson and Hancock (2012) were validated with this initial sample. The four modes were found to be distinct, requiring unique preparatory management and in-process management techniques to allow the timely and effective movement between modes. For the application of this sample within the U.S. Air Force, the taxonomy was adapted to represent the variety of mission requirements, as shown in Figure 6.

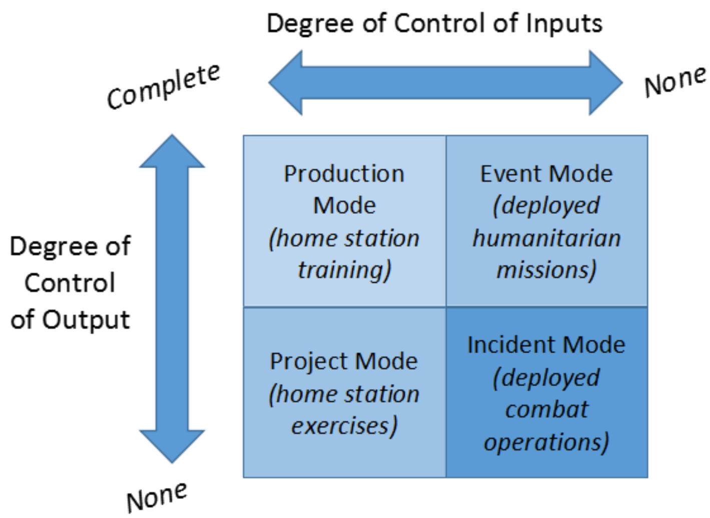


Figure 6: Modes of Operations as Translated in User Terminology (adapted from Simpson and Hancock, 2012)

CONCLUSIONS

Results of this analysis validated the need to operate in various modes depending on the situation, as defined by the level of control over inputs and outputs. Organizations concerned with operational resilience should identify potential states as related to the modes of operation (Figure 5), then create organizational structures and internal policies to not only optimize each mode but also to expedite the transition between modes when necessary. For organizational structure, a centralized control with decentralized execution was found to be more effective. This blended use of two otherwise seemingly mutually exclusive organizational structures resonates strongly with a stream of published findings from other fields, including rapid software development (Austin and Devin, 2003; Thomas and Bendoly, 2009), municipal firefighting (Bigley and Roberts, 2001), and service recovery sub-systems (Smith et al, 2012). For internal policies, a flexible set of guidelines combined with investment in advanced technical training and education in operations management will show resilience enhancements. As experience at the individual and organizational level was shown to be critical, units without recent experience in operating in each mode should invest in practice exercises regularly. Each exercise should be designed with the forethought of reinforcing operational changes required to move from Production to Project Mode, from Production to Event Mode, or from Production to Incident Mode. All throughout, increased communications via technological enhancements (hands-free radios, shared data, common situational awareness using augmented reality, etc.) and strengthening of personal relationships will not only speed the transition of mode change but also make the new mode simultaneously more efficient and effective. Clearly additional exploration should be conducted, with focus on each specific dyad of modes. Analysis of disruption case studies may provide additional insight, as well as evaluation of pre-planned exercises and drills common to those conducted by the military and emergency responders. Generalization will eventually follow into manufacturing and service industries as operational resilience evolves into our ever-changing world.

REFERENCES

Altay, N. & Green, W. (2006). “OR/MS research in disaster operations management.” *European Journal of Operational Research*, Volume 175, Number 1, 475-493.

Ambulkar, S., Blackhurst, J. & Grawe, S. (2015). “Firm's resilience to supply chain disruptions: Scale development and empirical examination.” *Journal of Operations Management*, Volumes 33–34 (January 2015), 111–122.

Astrom, K. & Murray, R. (2012). *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton, NJ: Princeton University Press.

Austin, R. & Devin, L., (2003). *Artful Making: What Managers Need to Know About How Artists Work.* Upper Saddle River, NJ: FT Prentice Hall.

Beamon, B. & Balcik, B. (2008). “Performance measurement in humanitarian relief chains.” *International Journal of Public Sector Management*, Volume 21, Number 1, 4-25.

Bigley, G. & Roberts, K. (2001). “The incident command system: High-reliability organizing for complex and volatile task environments.” *Academy of Management Journal*, Volume 44, Number 6, 1281-1299.

Blackhurst, J., Dunn, K. & Craighead, C. (2011). “An Empirically Derived Framework of Global Supply Resiliency.” *Journal of Business Logistics*, Volume 32, Number 4, 374–391.

Boin, A. (2009). “The new world of crises and crisis management: Implications for policymaking and research.” *Review of Policy Research*, Volume 26, Number 4, 367-377.

Boin, A., Kelle, P. & Whybark, C. (2010). “Resilient supply chains for extreme situations: Outlining a new field of study.” *International Journal of Production Economics*, Volume 126, Number 1, 1-6.

Brandon-Jones, E., Squire, B., Autry, C. W. & Petersen, K. J. (2014). “A Contingent Resource- Based Perspective of Supply Chain Resilience and Robustness.” *Journal of Supply Chain Management*, Volume 50, Number 3, 55-73.

Charles, A., Lauras, M. & Van Wassenhove, L. (2010). “A model to define and assess the agility of supply chains: Building on humanitarian experience.” *International Journal of Physical Distribution and Logistics Management*, Volume 40, Number 8/9, 722-741.

Christopher, M. & Peck, H. (2004). “The Five Principles of Supply Chain Resilience.” *Logistics Europe*, Volume 12, Number 1, 16-21.

Fiksel, J. (2006). “Sustainability and Resilience: Toward a Systems Approach.” *Sustainability: Science, Practice, & Policy*, Volume 2, Number 2, 1-8.

Fiksel, J., Polyviou, M., Croxton, K. & Pettit, T. (2015). “From risk to resilience: Learning to deal with disruption.” *MIT Sloan Management Review*, Volume 56, Number 2, 79-86.

Fowler, A. (1999). “Feedback and feedforward as systemic frameworks for operations control.” *International Journal of Operations and Production Management*, Volume 19, Number 2, 182-204.

Galindo, G. & Batta, R. (2013). “Review of recent developments in OR/MS research in disaster operations management.” *European Journal of Operational Research*, Volume 230, 201-211.

Green, L. & Kolesar, P. (2004). “Improving emergency responsiveness with management science.” *Management Science*, Volume 50, Number 8, 1001-1014.

Gunasekaran, A. & Ngai, E. (2012). “The future of operations management: An outlook and analysis.” *International Journal of Production Economics*, Volume 135, 687-701.

Habermann, M., Blackhurst, J. & Metcalf, A. (2015). “Keep Your Friends Close? Supply Chain Design and Disruption Risk.” *Decision Sciences,* Volume 46, Number 3, 491–526.

Hayes, R. (2002). “Challenges posed to operations management by the `New Economy'.” *Production and Operations Management*, Volume 11, Number 1, 21-32.

Holguin-Veras, J., Taniguchi, E., Jaller, M. & Aros-Vera, F. (2014). “The Tohoku disasters: Chief lessons concerning the post disaster humanitarian logistics response and policy implications.” *Transportation Research Part A*, Number 69, 86-104.

Janssen, M., Lee, J., Bharosa, N. & Cresswell, A. (2010). “Advances in multi-agency disaster management: Key elements in disaster research.” *Information Systems Frontiers*, Volume 12, Number 1, 1-7.

Kovacs, G. & Spens, K. (2009). “Identifying challenges in humanitarian logistics.” *International Journal of Physical Distribution and Logistics Management*, Volume 39, Number 6, 506-528.

Leipold, B. & Greve, W. (2009). “Resilience: A conceptual bridge between coping and development.” *European Psychologist*, Volume 14, Number 1, 40–50.

Mitzenberg, H. (1980). “Structure in 5'S: A synthesis of research on organization design.” *Management Science*, Volume 26, Number 3, 322-341.

Oltra, M., Maroto, C. & Segura, B. (2005). “Operations strategy configurations in project process firms.” *International Journal of Operations and Production Management*, Volume 25, Number 5, 429-448.

Pedraza, A., Martinez, O. & Van Wassenhove, L. (2011). “Field vehicle fleet management in humanitarian operations: A case-based approach.” *Journal of Operations Management*, Volume 29, Number 5, 404-421.

Pettit, S. & Beresford, A. (2009). “Critical success factors in the context of humanitarian aid supply chains.” *International Journal of Physical Distribution and Logistics Management*, Volume 39, Number 6, 450-468.

Pettit, T., Croxton, K. & Fiksel, J. (2013). “Ensuring supply chain resilience: Development and implementation of an assessment tool.” *Journal of Business Logistics*, Volume 34, Number 1, 46-76.

Pettit, T., Fiksel, J. & Croxton, K. (2010). “Ensuring supply chain resilience: Development of a conceptual framework.” *Journal of Business Logistics*, Volume 31, Number 1, 1-21.

Schmenner, R. (1986). "How can service businesses survive and prosper?" *Sloan Management Review*, Volume 27, Number 3, 21-32.

Sheffi, Y. (2005) *The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage*. Cambridge, MA: MIT Press.

Simpson, N. & Hancock, P. (2009). "Fifty years of operational research and emergency response." *Journal of the Operational Research Society*, Volume 60, S126-S139.

Smith, J., Nagy, P., Karwan, K.. & Ramirez, E. (2012). "The contingent nature of service recovery system structures." *International Journal of Operations and Production Management*, Volume 32, Number 7, 877-903.

Tatham, P. & Pettit, S. (2010). "Transforming humanitarian logistics: the journey to supply network management." *International Journal of Physical Distribution and Logistics Management*, Volume 40, Number 8/9, 609-622.

Thomas, A. & Kopczak, L. (2005). *From Logistics to Supply Chain Management: The Path Forward in the Humanitarian Sector*. San Francisco, CA: Fritz Institute.

Thomas, D. & Bendoly, E., (2009). "Limits to effective leadership style and tactics in critical incident intervention." *Project Management Journal*, Volume 40, Number 2, 70-80.

Yin, Robert K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

U.S. Air Force. (2015). *B-1B and CV-22 Fact Sheets*, www.af.mil/AboutUs/FactSheets/, (Sep 29, 2015 and Feb 6, 2015).

Weick, K. & Sutcliffe, K. (2001). *Managing the Unexpected: Assuring High Performance in an Age of Complexity*. San Francisco, CA: Jossey-Bass.

Journal of Business and Behavioral Sciences
Vol 28, No. 2; Fall 2016

METHODOLOGICAL CONSIDERATIONS IN CROSS-CULTURAL RESEARCH: A DISCUSSION OF THE TRANSLATION ISSUE

Mahesh N. Rajan
Bobbi Makani
San Jose State University

ABSTRACT

It is generally believed that cross-cultural studies signal the maturation of an academic discipline. Yet, comparative research in the international business field (especially, international marketing and management) is still marked by uncertainty and misconception as concerns about issues of methodological validity and or appropriateness persist. Furthermore, some scholars have pointed out that in addition to the fact Americans in general are monolingual, the most common type of cross-cultural management studies are either parochial or ethnocentric. Therefore, it is absolutely imperative that (American) international marketing and management scholars ensure methodological correctness and or appropriateness in their research efforts in foreign countries and cultures. While other methodological issues are equally important, this paper will focus on the translation issue. Though there is a limitless body of anecdotal evidence on the pitfalls of translation errors written mainly for the practitioner audience, there is a paucity of discussion on the theoretical aspects of translation in cross-cultural studies. By undertaking a systematic and comprehensive review of the translation issue from a theoretical standpoint, this paper attempts to fill the above void in the literature. Guidelines, in the form of a couple of frameworks, are also provided to aid scholars' deal with translation of materials in their cross-cultural research.

Key Words: Cross-cultural research, methodology, translation

INTRODUCTION

A major area of concern in cross-national or cross-cultural (hereafter the term comparative research will be used to denote both terms) research is that of methodological validity and or appropriateness (Sekaran, 1983; Wheeler, 1988; Adler and Graham, 1989; Temple, 1997; Olohan, 2000; Munday, 2002; Harkness *et al*, 2004; Maclean, 2007; Pëna, 2007; Hurtado de Mendoza, 2008; Choi *et al*, 2012; Nurjannah *et al*, 2014). While other issues such as sample design, research instrument choice and design, construct measures to be utilized, etc. are equally important, this paper focuses on the translation issue.

As pointed out by Brislin (1970) comparative research involves the translation of instructions and or data collection instruments from one language to another. While researchers need not necessarily be so, it is strongly believed that bilingual or multilingual competencies are definitely advantageous in comparative research. Yet, in Beamish and Calof's (1989) survey of (Canadian) corporate managers on international business education, foreign language skills/aptitude was listed as the least important in terms of selection for international assignments! Professor Richard Lambert, founder and director of the National Foreign Language Center at the Johns Hopkins University best sums up the monolinguistic orientation of the American society thus:

> The greatest barrier to the expansion of foreign language competences in the United States is the low value placed on such competences by American society as a whole. For instance, a recent Gallup survey commissioned by the National Geographic Society ranked foreign language at the bottom of a list of school subjects ranked in order of importance to adults. …Survey after survey...informs us that American companies, even those with numerous foreign transactions, place foreign language competence close to the bottom in their list of desiderata in hiring new employees...The generally low evaluation of foreign language skills and the limited occupational demand for them are accompanied by a generally low level of competence among adults in our society. (1990, p.48)

Similarly, the Head of the Defense Technical Information Center's (DTIC) Information Systems and Standards Division pointed out that due to the costs of translating documents and the lack of a sufficient number of people with both the technical and language skills, much foreign technological information does not get translated contiguously (United States General Accounting Office, 1990).

Relatedly, in a survey of over 900,000 source items from the 6100 journals in 31 languages in the Institute of Scientific Information (ISI) database, Garfield and Welljams-Dorof (1990) found the English language to be the predominant language of research communication. Source items in English accounted for 85% of the total and were cited four times greater than that of the next leading languages - German and French respectively. Moreover, citations in American publications were almost all drawn from English (99.6%), and primarily from American (72.4%) sources. From the above figures it is quite apparent that American scholars basically restrict their literature reviews to publications in the English language, and to a large extent to those by fellow Americans (in fact, the United States had the highest self-citing rate among all nations in the database). Interestingly, West Germany, France, and Italy indicated their bilingual competences in that they published in both their native language and English.

While the above data establish the fact that English is the internationally accepted language of research communication, Garfield and Welljams-Dorof (1990, p.10) warn of an inherent danger in the sense that "native English-speaking researchers risk being ignorant of significant findings reported in foreign languages." Hence, they conclude that "conversational fluency in more than one language remains a valuable professional asset for researchers [as it enables them] to appreciate more deeply the expression of other nations and cultures - their art as well as their science." (1990, p.24).

Moreover, familiarity of the language does not necessarily guarantee an understanding of the cultural phenomena as even within the same language words and expressions could have different meanings and or connotations in different regions and or countries. Consider another example of Rabassa (1989, p.10) that "a Puerto Rican in Buenos Aires who innocently announces that he is going to catch a bus (*Voy a coger la guagua*) would be arrested as a child molester. *Bicha* in Portugal is a line; in Brazil, it is a drag queen." More familiar examples might be terms such as *flat*, *automobile bonnet*, and *lift* used by the British to denote what in America are considered as apartment, automobile hood, and elevator respectively. These examples emphasize the need for and importance of judicious language usage in cross-cultural studies. Hence, the purpose of this paper is to undertake a systematic, empirical investigation of the translation issue. Specifically, this paper will focus on the theoretical aspects of translation in cross-cultural research. The Literature Review (on translation in management and marketing) section is followed by a detailed discussion of the translation process including definitions, theories, and constructs. Some practical frameworks are provided in section four followed by conclusions.

LITERATURE REVIEW

In addition to the fact that Americans in general are monolingual, Adler (1984) points out that the most common type of cross-cultural management studies are parochial studies and ethnocentric studies, in that order. In other words, parochial studies are "research projects originally designed and conducted in one culture by researchers from that culture," (p.33), whereas ethnocentric studies are "studies which attempt to replicate American management research in foreign countries." (p.32). Therefore, it is absolutely imperative that management scholars ensure sociolinguistic correctness and or appropriateness in their research efforts in foreign countries and or cultures. While there is almost a limitless body of anecdotal evidence or recounts on the pitfalls of translation errors written especially for the practitioner audience (Ricks, 1988), very little information exists on the theoretical issues of translation in cross-cultural studies. Several authors have lamented on this fact that there is no comprehensive review or discussion of the translation issue (Werner and Campbell, 1970; Green and White, 1976, Temple and Young, 2004). Our review of the literature led to somewhat similar results.

THE TRANSLATION PROCESS: DEFINITIONS, THEORIES AND CONSTRUCTS

According to one author, the origins of translation can be traced as early as 3000 B.C. when inscriptions in two languages during the Egyptian Old Kingdom were discovered. The dubbing of the twentieth century as the "age of translation" (Newmark, 1981, p.3) reflects the importance and volume of translation accompanying greater political, economic, cultural, and scientific cooperation between nations of the world. However, little has been written about the topic of translation.

Biguenet and Schulte (1989) explain that the act of translation does not entail the mere reduction of texts into simplistic interpretative statements. Instead, translators attempt to discover relationships from the text - "relationships between the words, relationships between the word and its philological and etymological background, relationships between the word and its cultural ambiance, relationships between the word and its historical tradition, relationships between the word and its context within a text." (p. xii).

While many versions exist, an authoritative source, (Brislin, 1976, p.1) defines translation as thus:

> Translation is the general term referring to the transfer of thoughts and ideas from one language (source) to another (target), whether the languages have established orthographies or do not have such standardization; or whether one or both languages is based on signs, as with sign languages of the deaf. Interpretation is one type of translation, and it refers to oral communication situations in which one person speaks in the source language, an interpreter processes this input and produces output in a second language, and a third person listens to the source language version

In other words, translation involves the competent use of dictionary, grammar, and other extra-linguistic criteria to transfer meaning contained in one set of language signs into another set of language signs (Bassnett-McGuire, 1980). As pointed out by Douglas and Craig (1983) while the need to translate verbal stimuli is well recognized, both verbal and non-verbal stimuli need to be translated for use in different linguistic and cultural contexts.

THEORIES OF TRANSLATION

There are three different approaches to the problems of translation. While the principle and procedures of each approach have been formulated and defended, in reality there is a considerable degree of overlap among them.

<u>The Philological Theories</u> - These theories of translation, predominant until World War II, are primarily concerned with so-called literary texts - that is, the focus of attention is on particular texts. Such texts "are regarded as sufficiently important to warrant special study and as having a sufficiently complex history (both in their development and in their subsequent interpretation) to require attention to the circumstances of production and transmission, if the form and content are to be adequately understood and appreciated, and if their translation is to be satisfactorily carried out." (Nida, 1976, p.67).

Philological theories of translation, while based on the philological approach to literary analysis, focus on the corresponding structures of the source and target languages and their equivalences, rather than just treating the form in which the original text was initially composed. Proponents of the philological approach are generally concerned about kinds of stylistic features and rhetorical devices while at the same time becoming increasingly interested in many of the larger units of structural discourse with regards to hierarchical structures, dependency relationships, and semantic interpretations. Philological theories began to lose their popularity when scholars started calling for a more linguistic orientation to translation theory and practice.

Linguistic Theories - These theories are based on a comparison of the linguistic structures of the source and target texts rather than on literary genres and stylistic features. Two factors that were primarily responsible for the development of linguistic theories were: (i) the application of the rapidly expanding science of linguistics to various areas of intellectual exercises; and (ii) machine (or mechanical) translation. The key issue debated by the various linguistics theories is the extent of focus on surface structures or corresponding deep structures.

From a linguistic perspective, the notion of underlying structures is important for translation theory and practice because: "(i) on the kernel or subkernel levels the syntactic structures of various languages are much more alike (basic underlying sentence structures can probably be reduced to ten or so, whereas the variety of surface structures is relatively numerous); (ii) one can more readily identify the semantic structures of subsurface levels and thus be in a position to determine more accurately the extent of equivalence and the need for supplementation or redistribution of semantic components; and (iii) on the deeper levels of structure one can more easily determine the symbolic relations and their hermeneutical implications. It is, of course, the surface structures that must point to the deep-structure relations, but it is the analysis of the deep structures that greatly facilitates a comparison of structures and relations." (Nida, 1976, pp.72-73).

However, linguistic theories do have some important drawbacks in that propositional equivalence is often considered as synonymy (and thereby meaning), whereas in reality propositionally equivalent expressions need not have the same meaning (or be synonymous). Moreover, linguistic theories assume that all individuals involved in the communication process have no limitations in terms of activities such as speaking and hearing (an assumption which translators cannot obviously assume). Finally, linguistic theories attempt to interpret texts in isolation, far removed from the actual contexts - discuss language as verbal communication occurring in a cultural vacuum.

Sociolinguistic theories - Communication is the basic model for these theories in order to reflect or accommodate the crucial role of the receptors in the translation process. The communication model illustrates the interlingual communicative nature of the translation process, with such communication being based on the use of a code - i.e. language. Sociolinguistic theories incorporate the principles of both the linguistic and philological approaches in that linguistic structures are elevated to a higher level of relevance thus enabling them to be viewed in terms of their function in communication. At the same time, sociolinguistic theory is concerned about all of the elements involved in the social setting of communication such as the author, the historical backgrounds of the text, the circumstances involved in producing the text, and the history of the interpretation of the text.

Since sociolinguistic theories recognize the fact that there are no ideal speaker-hearers among the receptors of the language, it does not expect nor attribute normative responses (that is, what are the actual responses rather than how people ought to respond). Moreover, because human behavior exhibits changes in terms of likes, dislikes, preferences, etc. no permanent criteria are set by the sociolinguistic theories to judge the quality of the translation.

Nida (1976) also points out that translators' seldom base their work on the principles or procedures of one underlying theory of translation. Instead, he concludes that most translators are highly eclectic in practice and are hence, fully prepared to meet the obstacles or challenges in the complex task of translation.

CONSTRUCT MEASURES

Construct measures in cross-cultural literature are broadly divided on they basis of whether they are universally applicable (etic measures) or culturally unique (emic measures). Etic measures are generally concerned with those aspects of human behavior such as certain needs, goals, etc. that transcend cultural nuances. Emic measures on the other hand are measures developed specifically to study the unique elements of each culture. Clearly both approaches have drawbacks as etic measures could lead to overgeneralizations while emic measures offer limited opportunities for comparative research. Researchers need to be aware of these limitations and choose the appropriate measures based on their research agenda (Richins and Verhage, 1985).

Though fundamentally suspect, most management cross-cultural studies because of their ethnocentric nature entail the translation of *emic* measures into the languages of the other cultures or nations studied. As the main methodological goal is standardization (rather than equivalence), "the rule guiding the translation of ethnocentric research is not that the two versions have the same meaning in each of the target populations." (Adler, 1984, p.39). Hence, differences occur between the two groups during (due to) the translation process, which are then incorrectly interpreted as cultural differences." The magnitude of this erroneous research design (i.e. standardization rather equivalence) of ethnocentric studies is brought to light by Rabassa's (1989) assertion that a translation can never equal the original text, but instead approach the latter. Therefore the accuracy of the translation in terms of its proximity to the original serves as measure of the quality of translation. Similarly, Biguenet and Schulte (1989) state that even though equivalence is the guiding rule of translation, exact equivalences across languages are impossible. Instead they contend that translation is "an act of `transformation' that adopts a new metaphor to the original metaphor." (p. viii).

TYPES OF EQUIVALENCE

Sechrest, Fay and Zaidi (1972) point out that a translation process usually entails different kinds of equivalence with differing effects and implications. The different kinds of equivalence are:

Vocabulary equivalence - Perhaps the most obvious, this equivalence is concerned about the words used in two or more translations. While some pairs of languages easily lend to vocabulary equivalence, often it is difficult to find comparable terms for words such as semantic differential words (e.g., strong-weak, rich-poor, fast-slow, etc.). Also, certain terms in one language may really not have an equivalent in another. For example, there are really no ready English equivalents for words such as *moonis*, *hundun*, and *kabeeb* in the Urdu language denoting differing degrees of friendship and closeness (Sechrest, Fay and Zaidi, 1972). One possible solution to this problem might be a good dictionary and some researchers suggest the usage of several synonyms in the target language to convey equivalent meaning of the word in the source language.

Idiomatic equivalence - Idiomatic speech, which is so deeply ingrained in our speech patterns and an integral part of any language poses the next problem. Idioms can never be properly translated and so the best solution is to make sure that the meaning and the level (i.e. scholarly, more stilted, etc.) of the idiomatic speech in both languages are more or less equivalent. For example, while the literal translation of the Tagalog (a Philippino dialect) idiom *galit-bulkan* is "angry volcano", it is interpreted as the sudden expression of anger, with the English equivalent being "blow up". Similarly, Rabassa's (1989, p.11) discussion illustrates an incorrect translation from the idiomatic equivalence perspective. He notes that the title of Faulkner's novel Light in August was a "country expression used for a cow who is comin in expectant, `heavy in June, light in August'". The Spanish translation *Luz en Agosto*, while preserving the mysterious lyricism of the English counterpart, and perfect in terms of a direct matching of words is far from the English [idiomatic] meaning.

Grammatical-Syntactical equivalence - This equivalence problem is caused by the grammatical and syntactical differences among languages which are often critical to the meanings in various translations. Though more of a problem with longer paragraphs, shorter passages and even single words pose difficulties in terms of grammatical-syntactical equivalence. Moreover, other parts of the speech such as adjectives, and adverbs may not be present in other languages and hence, pose additional problems. Also since problems of verb mood, voice, or tense are not dealt in the same way it is difficult to find expressions which have the same verb form or meaning in English. For example, the Tagalog does not have any subjunctive moods. Hence the English sentence, "If I had the money, I would have bought the dress," will be literally translated into Tagalog as "If I have the money (understood I have not), I bought the dress (understood I did not) - *Kung mayroon sana akong pera, nabili ko sana ang barro*. Similarly, Seidensticker (1989) notes that the Japanese language has no relative pronouns; that adjectival clauses, no matter how long, always precede the nouns; that nouns are not declined and are without gender; that verbs are conjugated with no specific references to person or number; and that adjectives, in the absence of relative pronouns, directly modify nouns and are conjugated with no reference to person, number, or gender.

Experiential equivalence - Though not involving purely linguistic considerations, this form of equivalence (called "cultural translation") is concerned about the utilization of terms referring to real things and real experiences which are familiar (even if not exactly equal in terms of familiarity) in both cultures. Thus, when the differences in the nature of the objects, or the social arrangements, or the overall ways of life between two cultures are such that objects or experiences familiar in one culture and unfamiliar in the other, then obtaining experiential equivalence from linguistic statements is difficult no matter how carefully the translation is carried out from the standpoint of the target language. For example, most Americans might understand the statement, "*I would like to be a florist*". However, because flower shops are not found in most parts of the world and the idea of it being a place where flowers are arranged in fancy designs are not commonly known to people in many countries, herein arises a problem of experiential equivalence. Hence, if the work of a florist is considered as a feminine activity in America, one solution to achieve experiential equivalence is to look for a similar kind of activity primarily identified with females in that country (such as a market vendor in the Philippines where most vendors are women). The other alternative is to eliminate the term or experience completely if no equivalent or counterpart exists in the other culture, or if the experiences vary so much that it is difficult to find an equivalent (e.g., since "department stores" are an unknown in Pakistan, the closest translation in Urdu for the same term would be a "large shop").

Rather than seek equivalence of stimuli across cultures, some authors suggest that experiential equivalence can be achieved through equivalence of inferences within cultures or social systems. Hence, "an instrument is equivalent across systems to the extent that the results provided by the instrument reliably describe with (nearly) the same validity a particular phenomenon in different social systems" (Sechrest, Fay, and Zaidi, 1972, p.48).

<u>Conceptual equivalence</u> - This problem is somewhat different than the others in the sense that there may be certain words, which, when translated mutually yields high agreement and yet the concepts implied by the two words are neither identical nor particularly close in nature. For example, the word "love" as used in the English language has far different implications from the concept which might be implicit in other languages. Moreover, even in the English language the word "love" is used to mean several different things or at least connote varying degrees of a positive feeling towards some objects. A better example is Rabassa's (1989, p.1) observation that "the word `dog' and the word *perro* may conjure up a like image in the mind of the Englishman and the mind of the Spaniard, but other subliminal images may accompany the two versions and thereby give the two words further differences beyond sound; the Portuguese word *cao* is closer to the Latin root that gives us `canine' opening up our minds to broader connotations."

Another form of the problem of conceptual equivalence is when a concept well understood and frequently used in one culture is altogether lacking in the other culture, or when it appears in such varied fragmented forms that it is difficult to pull together terms that make the true concepts equivalent. For example, it is very difficult to find even a concept, let alone words, in the Philippines that have the same connotations as the common American concept of homosexual. "The available Philippine words are used in a variety of ways which suggests that the concept simply does not exist in the same highly developed form as it does in the American culture." (Sechrest, Fay, and Zaidi, 1972, p.49).

Rabassa's (1989, p.3) discussion on the problems of translating curses and oaths once again provides a good illustration of this problem:

> In English, when we insult some one's maternal descent, we call the person a "son of a bitch", while in Spanish he is an hijo de puta, "son of a whore". The closest in English to this latter is the archaic "whoreson", which, even if understood, would not arouse much more than a ripple of indignation. Portuguese leaves it up to the listener's imagination to deduce the impact by simply saying filho da mae, "son of your mother", which is patent and obvious, but leaves the door open to all manner of vile conjectures regarding one's dam. A most common insult in Spanish is cabron, "cuckold". There is really no exact equivalent in English, for the word itself would be about as effective as a "whoreson". Indeed, many people would not know what you were talking about.

While great care and much attention to detail is required to ensure all of the different equivalences in the translation process, too much emphasis on the same could also become dysfunctional and affect the quality of research adversely. This "paradox of equivalence" is defined by Sechrest, Fay, and Zaidi (1972, p.49) as follows:

> The paradox of equivalence is that if one demands that a form of a test or other measure yield comparable results in two cultures in order to demonstrate equivalence, then the more equivalent two forms become the less the probability of finding cultural differences. On the other hand, if one looks predominantly for cultural differences and ignores the problem of

> equivalence, then the less attention that is paid to the problem of equivalence the greater the probability of finding cultural differences.

While difficult to resolve, the authors suggest some method of triangulation whereby measures are subject to increasingly rigorous tests of equivalence thus enabling the researchers to determine the extent of convergence in the responses of two (or more) cultures.

TYPES OF TRANSLATION

Out of the five methods or procedures of translation discussed below, the first two are the popular ones in comparative research.

Direct translation - The most common procedure of translation is direct translation, wherein bilingual translator(s) attempt to translate as best as they can from one (source) language to another (target) language. Except for brief topics such as orientation, instructions, objective check lists, etc. direct translation procedures are likely to be most inadequate for capturing the cultural differences or nuances in cross-cultural studies. While being a relatively quick and low-cost procedure, there are several problems with direct translation. A major area of concern is the translator induced bias in the research instrument (and process). Such bias is likely to occur if the translator is (a) more skilled in one language than the other, (b) not a cultural representative of the sample groups in the study - i.e. unfamiliar with idiomatic and colloquial expressions, and (c) based on his/her peculiar experiences, interprets words or phrases in such a way that is not shared by others, especially those in the research population. Popular anecdotal examples such as "the translation into French of a `full' airplane as a `pregnant' airplane, and in German, of `Body by Fisher' as `Corpse by Fisher'" (Douglas and Craig, 1983, pp.186-187) illustrate errors in direct translation.

As pointed out by Adler (1984), ethnocentric studies are primarily concerned about discovering similarities across cultures. Since replication is the most common approach to such studies, standardization becomes the main methodological goal. This results in a literal translation of the research instruments, instructions, etc. from the source language to the target language, with the two versions not having the same meaning(s) in each of the target populations. However, while similarities are interpreted as proof of universality of (certain aspects) of the theory being tested, differences are interpreted as cultural rather than as methodologically (translation) driven. Needless to say, both interpretations are fundamentally flawed and pose substantial implications not only in research design issues, but also in theory building and testing.

Back-translation - Though not the most ideal alternative, this procedure is clearly superior to direct translation. Basically, the procedure starts with a questionnaire being translated from the source or base language by a bilingual who is a native speaker of the target language or language into which the translation is to be made. A bilingual who is a native speaker of the first or source language then retranslates the translated version back into the original language. The two versions of the original language are then checked for discrepancies - in other words, the adequacy of the translations. If need be (and if properly carried out), the process is repeated until the researchers are of the opinion that (near) equivalency has been achieved. A critical feature of this iterative procedure that warrants attention is that discrepancies, if any, are corrected only in the target language versions. In other words, even back-translation assumes the source language as the base language (not unlike the starting assumption of the direct translation procedure) and hence, is considered as a "*psuedo-etic*" approach (Douglas and Craig, 1983).

The biggest advantage of this procedure is that it serves as a filter through which non-equivalent terms are not readily acceptable or easily added to the research instrument (conversely, the opposite would be true for equivalent terms). With the exception of a couple of studies, the method of back translation was the most commonly employed method of translation in studies reviewed by some authors (Adler, Campbell, and Laurent, 1989).

The utility and value of (back) translation is illustrated by Adler *et al*'s (1989) narration that:

> Prior to its use in the P.R.C., the questionnaire was translated into Chinese and then back-translated into English...A number of mistranslations were detected, including some fairly funny ones, such as one item that states in English that "Most conflicts in a company can be productive" while in Chinese it became, "Much hitting and physical violence in a company can be productive." In the process of translation, conflict had gone from verbal disagreement to physical violence. Items continued to be translated and back-translated until the English and Chinese versions substantially agreed with each other. (p.63)

However, Brislin (1970) notes that even back-translation does not always represent an effective procedure. Douglas and Craig (1983) point out some of the disadvantages of back-translation. From a translator's standpoint, any of the factors discussed in direct translation that adversely affect equivalence in translation are equally a threat in this procedure also. Specifically, translators' ignorance of the idiomatic expressions of the research populations, unequal language competences, non-representativeness, etc. all introduce translator induced bias or non-equivalence in the translation procedure. Yet another source of difficulty in achieving equivalent translations in two or more cultures might be due to the fact that no satisfactory equivalent concepts exist. For example, the word "homosexual" when translated into Tagalog usually becomes *bakla* which when literally translated into English reads as "sissy" or the like. Usually, the only way to resolve such dilemmas is by consulting a larger and more representative group of bilinguals with the hope of finding a (near) equivalent.

The associative problems with greater numbers (of individuals) need to be considered by the researchers before making a decision in this regard.

De-centering - An ultimate solution to the problem of translation, de-centering is focused on the development of genuinely bilingual versions of the research instruments. Since most research instruments are developed in one country (culture) and then translated into another language for use in that country (culture), the narrow cultural base or scope of these research instruments leads to obvious disadvantages. In other words, researchers following the above strategy attempt (or hope) to somehow transform emic measures to etic measures through the translation process! Hence, Werner and Campbell (1970) suggest the joint development of research instruments, construct measures, and other related items by research partners in both cultures. While the process of de-centering is similar to that of back-translation, the key difference between the two is that both languages are iteratively modified until satisfactory equivalence levels have been reached (or until the dominant bias of one language is reduced to approximate those of the other language). Terms such as source and target languages are not applicable in this situation, since, by definition, de-centering (if properly carried out) involves the translation of two languages - i.e. a dialectic process.

De-centering's obvious advantages definitely enhance its stature as a relatively more appropriate translation procedure than all of the procedures discussed thus far. The first and foremost advantage is that de-centering allows cross-cultural researchers to incorporate concepts that are more relevant to the cultures being studied, while concurrently reducing the prevalent and often incorrect usage of Western (especially, American) concepts in such endeavors. The significance of this advancement is evident from the fact that researchers can now study foreign language concepts such as *nak uchi* (Pakistani concept related to pride and literally translated as "nose high"), *utang na loob* (Philippino concept for reciprocal obligation), and on (the Japanese concept for filial loyalty and obligation) that are seldom studied in the Western cultures. Additionally, de-centering increases the likelihood of greater equivalence being achieved by enabling longer, more elaborate versions of the research instruments.

However, de-centering is not without drawbacks either. Depending upon the type of stimuli, characteristics of respondents, etc. the longer, more elaborate versions of the data collection instruments discussed above may become a significant disadvantage in comparative research. Hence, while the relatively more appropriate or logically sound research instrument enhances the quality of the research on the one hand, it may suffer on the other hand because of fewer respondents, meaningful responses, etc. The second major drawback of the de-centering procedure is that because there are not too many people with multilingual competences, the number of cultures a researcher may study is severely restricted. Moreover, there might be certain risks of overgeneralization by attempting to classify several cultures on a least common denominator(s).

<u>Carrier language</u> - This procedure calls for the identification of a common language (perhaps English) in which the basic concepts for cross-cultural research are carried from one culture to each of the others. An example is that involving the Urdu language which has the potential to be a carrier language from which translations can be carried out into Bengali, Sindhi, and Pashtun without ever resorting to the English language. However, practical constraints and uncertainty seem to diminish the appeal and or adoption of this procedure.

<u>Parallel translation</u> - Douglas and Craig (1983) mention that a committee of translators with at least bilingual competences in the languages being studied compare the various translated versions. Agreement on a final version is reached through a discussion of the various alternatives. Though some critics have expressed concern over the lack of formal controls, and the possibility of personality clashes Douglas and Craig deem the parallel translation procedure useful in international marketing research, especially in those tolerant of less exact and rigid translations.

However, as noted earlier, an increasing number of cross-cultural studies in marketing do seem to make a concerted effort to discuss the translation process (albeit, sparsely). However, since the present article is even more comprehensive and theoretically grounded/based, we feel that this as the main contribution(s) to the field.

EVALUATION CRITERIA OR GUIDELINES

Finally, Biguenet and Schulte's (1989, p. xv) analysis of six different translations of the same German poem led to the conclusion that no two translations were alike and that each of the translators had taken a different (in their intensity of perception) approach in the interpretation of the opening line of the poem. This is neither unusual nor unacceptable, according to the above scholars who contend that "we know that two different translators will never come up with exactly the same translation, since their initial way of seeing a work varies according to the presuppositions they bring to a text. No two translators perceive the every moment in a text with similar awareness of intensity, which leads to varying value judgements within a text about what elements should be chiseled out for the act of translation from the source-language situation into the target language." Moreover, as even an individual translator does not ("will never") approach the same text similarly on more than one occasion, variation in repetitive individual interpretations are also not anomalous. "On one day the translator might say `maybe' and on the next day `perhaps' for the same word in the source-language text." Biguenet and Schulte's comments raise an interesting issue in that how does one control for (significant) variance in the translation field? In other words, how can we ensure consistency (in and or equivalence) and yet at the same time not infringe upon the creative and intellectual freedoms of the translators? What constitutes a successful translation?

Several scholars have pointed out the fact that since a translation can never equal the original, the true measure of its quality (success) is judged by the accuracy of its interpretation of the original text (Biguenet and Schulte, 1989; Rabassa, 1989). If that is the case, what kind(s) of evaluative criteria do we have, or are to be used in the evaluation of translation?

PRACTICAL FRAMEWORKS/GUIDELINES FOR TRANSLATION

Two frameworks have been suggested in this regard. The first framework consists of the following guidelines for writing translatable English and was put forth by Brislin, Lonner, and Thorndike (1973 - as cited by Brislin, 1976):

1. Use short, simple sentences of less than 16 words.
2. Employ the active rather than the passive words
3. Repeat nouns instead of using pronouns.
4. Avoid metaphor and colloquialisms. Such phrases are least likely to have equivalents in the target language.
5. Avoid the subjunctive mood (e.g., verb forms with could, would).
6. Add sentences that provide context for key ideas. Reword key phrases to provide redundancy. This rule suggests that longer items and questions be used than in single-country research.
7. Avoid adverbs and prepositions telling where or when (e.g., frequent, beyond, upper).
8. Avoid possessive forms wherever possible.
9. Use specific rather than general terms (e.g., the specific animal such as cows, chickens, pigs, rather than the general term, livestock).
10. Avoid words indicating vagueness regarding some event or thing (e.g., probably, frequently).
11. Use wording familiar to the translator wherever possible...
12. Avoid sentences with two different verbs if the verbs suggest two different actions.

Another framework more useful for evaluating the translations was also elucidated by Brislin (1970), and incorporated by Jones *et al* (2001). The basic guidelines suggested by Brislin are:

1. Evaluation of the initial version and the translated version of the research instrument in the source language by monolinguals for clarity and comprehension.
2. Evaluation of the source and target language versions by bilinguals for potential meaning errors.
3. Testing of the subjects' ability to answer questions about the content matter or perform tasks contained in the questionnaire.
4. Testing of both original and translated versions of the research instrument on bilingual subjects.

There have also been some suggestions involving the use of quantitative techniques to asses the accuracy or equivalence of translation such as ANOVA tests (Brislin, 1970), psychometric equivalence tests (Hulin and Mayer, 1986), statistical tests based on item response theoretical models (Ellis, 1989), item response theory tests (Ellis, Minsel, and Becker, 1989), systemic models (Munday, 2002), etc. However, a detailed discussion of such quantitative models and tests of translation is beyond the scope of this paper, and interested readers are asked to refer to the original articles for the same.

CONCLUSIONS

Research on comparative management (has grown) and is growing by leaps and bounds. However, concerns of methodological validity and or appropriateness doggedly persist scholarly discussions and forums. One such major issue of concern is the translation. The literature is marked by a paucity of systematic or comprehensive discussion of the theoretical aspects of the translation issue in cross-cultural studies. This study is an attempt to fill such a void in the literature. In addition to a detailed discussion of the translation process (including definitions, theories, and constructs), some practical frameworks/guidelines are provided. If this study does nothing more than raise awareness among international business scholars of the nuances and practicalities of translation while conducting research in foreign countries/cultures, we would consider its purpose fulfilled.

REFERENCES

Adler, Nancy J. (1984). Understanding the Ways of Understanding: Cross-Cultural Management Methodology Reviewed. Advances in International Comparative Management, Vol. 1, pp. 31-67, Greenwich, Conn.: JAI Press.

Adler, N. J., N. Campbell, and A. Laurent (1989). In Search of Appropriate Methodology: From Outside The People's Republic of China Looking In. Journal of International Business Studies, Vol. 20, Spring: 61-75.

Adler, N. J., and J.L. Graham (1989). Cross-Cultural Interaction: The International Comparison Fallacy. Journal of International Business Studies, Vol. 20, Fall: 515-538.

Basnett-McGuire, S. (1980). Translation Studies. London: Metheun & Co.

Beamish, P.W., and J.L. Calof (1989). International Business Education: A Corporate View. Journal of International Business Studies, Vol. 20, Fall: 553-564.

Biguenet, J. and R. Schulte (1989). Introduction, in John Biguenet and Rainer Schulte (Eds.), The Craft of Translation, Chicago: The University of Chicago Press.

Brislin, R.W. (1970). Back-Translation For Cross-Cultural Research. Journal of Cross-Cultural Psychology, Vol. 1, No. 3, September: 185-216.

Brislin, R.W. (1976). Introduction, in Richard W. Brislin (Ed.), Translation: Applications and Research, New York: Gardner Press Inc.

Choi, J., K.E. Kushner, J. Mill, and D.W. L. Lai (2012). Understanding the Language, the Culture, and the Experience: Translation in Cross-Cultural Research. International Journal of Qualitative Methods, Vol. 11, No. 5, 652-665.

Douglas, Susan P., and C. S. Craig (1983). International Marketing Research. Englewood-Cliffs, New Jersey: Prentice-Hall International Series in Management.

Ellis, B.B. (1989). Differential Item Functioning: Implications for Test Translations. Journal of Applied Psychology, Vol. 74, 6: 912-921.

Ellis, B.B., B. Minsel, and P. Becker (1989). Evaluation of attitude survey translations: An investigation using item response theory. International Journal of Psychology, Vol. 24, 6: 665-684.

Garfield, E., and A. Welljams-Dorof (1990). Language Use in International Research: A Citation Analysis, in Richard D. Lambert and Sarah Jane Moore (Eds.), Foreign Language Use in The Workplace, The Annals of The American Academy of Political and Social Science, Vol. 51, September: 10-24.

Green, R.T., and P.White (1976). Methodological Considerations in Cross-National Consumer Research. Journal of International Business Studies, Vol. 7, Fall/Winter: 81-86.

Harkness J, Pennell BE, Schoua-Glusberg A. (2004). Survey questionnaire translation and assessment, in Presser S, Rothgeb J, Couper M, Lessler J, Martin E, Martin J, et al., editors. Methods for testing and evaluating survey questionnaires. Hoboken, NJ: John Wiley; 453–473.

Hulin, C.L., and L.J. Mayer (1986). Psychometric Equivalence of a Translation of the Job Descriptive Index into Hebrew. Journal of Applied Psychology, Vol. 71, February: 83-94.

Hurtado de Mendoza, A. (2008). The Problem of Translation in Cross-Cultural Research on Emotion Concepts (Commentary on Choi & Han). International Journal for Dialogical Science, Fall 2008, Vol. 3, No. 1, 241-248.

Jones, P.S., J.W. Lee, L.R. Philips, X.E. Zhang, and K.B. Jaceldo (2001). An Adaptation of Brislin's translation model for cross-cultural research. Nursing Research, Sep-Oct, 50(5), 300-304.

Lambert, R.D. (1990). Foreign Language Use Among International Business Graduates, in Richard D. Lambert and Sarah Jane Moore, (Eds.), Foreign Language In The Workplace, The Annals of The American Academy of Political and Social Science, Vol. 51, September: 47-59.

Maclean, K. (2007). Translation in Cross-Cultural Research: An Example from Bolivia. Development In Practice, Vol. 17, No. 6, November, 784-790.

Munday, J. (2002). Systems in Translation: A Systemic Model for Descriptive Translation Studies, in Theo Hermans (Ed.), Cross-cultural Transgressions Research Models in Translation Studies II: Historical and Ideological Issues, London: St. Jerome Publishing.

Newmark, P. (1981). Approaches To Translation. New York: Pergamon Press.

Nida, E.A. (1976). A Framework for the Analysis and Evaluation of Theories of Translation, in Richard W. Brislin (Ed.), Translation: Applications and Research, New York: Gardner Press Inc.

Nurjannah, I, J. Mills, T. Park, and K. Usher (2014). Conducting a Grounded Theory Study in a Language Other Than English: Procedures for Ensuring the Integrity of Translation. Sage Open, January-March, 1-10.

Oches, N. (1989). Cross-Cultural Presentations - How to Make Them More Effective. Sales and Marketing Management, Vol. 141, September: 82-84.

Olohan, M. (2000). Shifts, But Not as We Know Them? Research Models and Methods in Translation Studies, in Maeve Olohan (Ed.), Intercultural Faultiness Research Models in Translation Studies I: Textual and Cognitive Aspects, London: St. Jerome Publishing.

Pëna, E. (2007). Lost in translation: methodological considerations in cross-cultural research. Child Development, Vol. 78, N0. 4, Jul-Aug, 1255-1264.

Rabassa, G. (1989). No Two Snowflakes Are Alike: Translation Metaphor, in John Biguenet and Rainer Schulte (Eds.), The Craft of Translation, Chicago: University of Chicago Press.

Richins, M.L., and B.J. Verhage (1985). Cross-cultural differences in consumer attitudes and their implications for complaint management. International Journal of Research in Marketing, 2: 197-206.

Ricks, D.A. (1988). International Business Blunders: An Update. Business and Economic Review, Vol. 34, 2: 11-14.

Sechrest, L., T.L. Fay, and S.M.H. Zaidi (1972). Problems of Translation in Cross-Cultural Research. Journal of Cross-Cultural Psychology, Vol. 3, No. 1, March: 41-56.

Seidensticker, E. (2011). Tokyo from Edo to Showa 1867-1989: The Emergence of the World's Greatest City. Tuttle Publishing.

Sekaran, U. (1983). Methodological and Theoretical Issues and Advancements in Cross-Cultural Research. Journal of International Business Studies, Vol. 14, Fall: 61-73.

Temple, B. (1997). Watch Your Tongue: Issues in Translation in Cross-Cultural Research. Sociology, Vol. 31, No. 3, 607-618.

Temple, B., and A. Young (2004). Qualitative Research and Translation Dilemmas. Qualitative Research, 4(2), 161-178.

United States General Accounting Office (1990). Foreign Technology: U.S. Monitoring and Dissemination of the Results of Foreign Research. Washington D.C.

Werner, O., and D.T. Campbell (1970). Translating, Working Through Interpreters and the Problems of Decentering, in Raoul Naroll and Ronald Cohen (Eds.), A Handbook of Method in Cultural Anthropology, New York: Columbia University Press.

Wheeler, D.R. (1988). Content Analysis: An Analytical Technique for International Marketing Research. International Marketing Review, Vol. 5, Winter: 34-40.

Journal of Business and Behavioral Sciences
Vol 28, No. 2; Fall 2016

AN ANALYSIS OF THAI LISTED HOTELS: FINANCIAL AND OPERATIONAL PERFORMANCE

Slisa Bhamornsathit
Punthumadee Katawandee
Chulalongkorn University

ABSTRACT: The advent of the ASEAN Economic Community (AEC) is expected to bring tremendous opportunities for tourism as a surge in investment in the region is anticipated. The hospitality industry in ASEAN will also compete more aggressively. The region's changing business landscape will open the prospect for freer flow of capital to Thailand's hotel sector. The country's listed hotel companies stand a good chance to benefit. Appropriate measures are needed to gauge hotel performance in terms of financial and operational dimensions.

The study examines ratio analysis of Thai listed hotels during 2010 - 2014. The financial ratios comprise liquidity, solvency, activity, and profitability. The measures for operational performance include occupancy rate, ADR, RevPAR, and operating costs. Selected descriptive statistics are then used for the analysis and interpretation. The computed ratios are measured against industry averages for internal comparison. Further, an external comparison is made with hotels in ASEAN.

The strength of Thai listed hotels is identified in terms of operational performance. To further enhance competitiveness, management needs to increase profitability and the ability to meet financial obligations. The results make a contribution to academicians and practitioners in assessing hotel business. The study is an initiative to produce comparative data for hotel industry in view of both financial and operational performance. Additionally, the study provides a comprehensive method for ratio analysis that can be applied to measure hotel performance.

Key Words: Ratio analysis, hotel performance, Thai listed hotels, hotel industry comparison

INTRODUCTION

International tourist arrivals to Thailand grew from 14.58 million in 2008 to 24.81 million in 2014 with an average growth rate of 9.80%. The hotel industry had also seen a rise in hotel supply to accommodate an increase in visitor arrivals. The number of hotels increased from 5,502 to 6,996 during the same period. The occupancy rate of hotels in Thailand averaged at 56.36% (Table 1).

From 2015 onwards, inbound tourism demand in Thailand and the whole region has risen to meet up with trade opportunities brought about by the ASEAN Economic Community (AEC). The World Tourism Organization (UNWTO) has forecast that in 2020 approximately 400 million tourists will travel to Asia-Pacific and around 160-200 million to ASEAN countries (World Tourism Organization n.d.). The hotel industry in Thailand has to prepare for the emerging opportunities and challenges. The surge in investment to the region is expected, and tourism businesses in ASEAN will become more rigorously competitive.

Table 1: Number of international tourists, hotels, and hotel rooms in Thailand (2008-2014)

Year	No. of international tourists[1]	Δ (%)	No. of hotels[2]	Δ (%)	Occupancy rate[4]	Δ (%)
2008	14,584,220	-	5,502	-	56.16	-
2009	14,149,841	-2.98	5,775	4.96	49.22	-12.36
2010	15,936,400	12.63	6,056	4.87	50.18	1.95
2011	19,230,470	20.67	6,344	4.76	57.73	15.05
2012	22,353,903	16.24	6,566	3.50	60.81	5.34
2013	26,546,725	18.76	6,784	3.32	64.86	6.66
2014	24,809,683	-6.54	6,996	3.13	55.58	-14.31
Avg.		9.80		4.09	56.36	0.39

Source: Modified data [1,4]Bank of Thailand (2008-2014), *Tourism Indicators*. [2,3]Euromonitor International (2015), *Lodging in Thailand* and Euromonitor International (2014) *Travel Accommodation in Thailand*.

The study addresses Thailand's hotel sector for two reasons. First, the hotel industry has high potentials as the country boasts high service quality, relatively low transportation costs, and developed supporting industries (i.e., restaurants and retail). Secondly, in the face of the AEC, Thailand's hotel sector needs to find ways to sustain its competitiveness in the changing business landscape. The scope of the study concerns hotels listed on the Stock Market of Thailand (SET) as they stand a good chance to benefit from freer flow of capital. The purpose of this study is to assess relevant ratios of the Thai listed hotels during 2010 – 2014. Among many assessment tools, the study employs a set of financial and operational performance measures commonly used in the hotel industry. Additionally, the study aims to produce comparative data in hotel industry.

LITERATURE REVIEW

A firm's various stakeholders, for instance, managers and investors, use various types of indicators to evaluate the firms' performance and highlight areas of improvement. The meta-analysis undertaken by Sainaghi (2010) reviewed 101 academic papers on hospitality management and identified 3 dimensions of hotel performance; namely, financial, operational, and organizational.

Several kinds of financial and operational ratios have been developed and widely used by practitioners and scholars alike. The important measures of hotel financial performance are liquidity, solvency, activity, profitability, and operating ratios (Jagels, 2007; Andrew and Schmidgall, 1993). Liquidity ratios measure the firm's ability to meet its short-term obligations. The key liquidity indicators are current ratio, quick or acid-test ratio, accounts receivable turnover, and average collection period. Solvency ratios measure the ability of the firms to cover long-term obligations. The common solvency indicators are debt ratio, debt to equity ratio (D/E), and time interest earned. Activity ratios measure the firm's efficiency in managing its assets. Activity ratios used in the hotel industry are inventory turnover, inventory holding period, fixed asset turnover, and total asset turnover. Profitability ratios measure overall performance of the firm's ability to generate income and the return on revenue and investment. These ratios include gross profit margin, operating profit margin, net profit margin, return on assets (ROA), and return on equity (ROE). Listed companies usually take into consideration additional ratios such as price - earnings ratio (P/E), market value to book value ratio (P/BV), and dividend yield.

Various measures have been suggested to evaluate hotel operational performance. The most common ones are occupancy rate, average daily room rate (ADR) and revenue per available room (RevPAR). All of these are revenue related. Operating ratios that are cost related include food cost percentage, beverage cost percentage, and labor cost percentage. They assist managers in analyzing and controlling hotel's operations. Some ratios are deemed to be more important than others. According to Singh and Schmidgall's study (2002) financial executives in the U.S. lodging industry considered profitability and operating ratios the most important ones.

Using ratios to compare financial characteristics of firms within the same and across different industries has been a research methodology in a number of financial literature (Kim and Ayoun, 2005). For instance, in his study using financial ratio to measure financial performance of three listed companies in the education sector in the Philippines, Tugas (2012) produced benchmark figures against which other firms in the same sector can compare. Dong Jin Kim (2006) compared 15 financial ratios in companies in similar industries, which were hotel and restaurant businesses. The findings suggested that liquidity and activity ratios of hotels were higher than those of restaurants. Also, based on the solvency ratios, restaurants were more capable to meet long-term obligations, while there was no significant difference in terms of profitability ratios between these two sectors. A cross-sectoral comparison is found in the work of Kim and Ayoun (2005) who conducted a comparison of 13 financial ratios in 4 hospitality sectors: lodging, restaurant, airline, and amusement sectors. The results indicated that eight out of thirteen financials ratios were statistically different across the four sectors. However, caution is often made for cross industry comparisons, suggesting that a meaningful analysis should be limited to comparisons with industry averages or companies in similar industries (Boundless, 2015).

Further, ratios analysis over a long period of time provide a more significant evaluation. A series of ratios presents a trend that provides additional information important to managers (Andrew and Schmidgall, 1993). A number of studies on financial analysis consider a time frame of five years appropriate for comparing financial ratios (Kim, 2006).

In Thailand, a number of researches have been conducted in the field of hospitality. Nevertheless, none academically relates to ratio analysis. Thai hotel industry itself has not thus far produced national comparative data for industry comparison. Besides, financial information related to the industry is rarely publicly available. The major source of financial data that can be obtained publicly is the Stock Exchange of Thailand (SET). Those data are only limited to listed firms.

METHODOLOGY

Data collection: The Stock Exchange of Thailand (SET) (2014) listed 12 hotels under the hotel and leisure subsector. Only 10 firms were assigned with Thailand Standard Industrial Classification TSIC 55101 (hotel). Of these 10 firms, 2 were excluded because their main lines of business were in other areas. Some items in their financial statements therefore differed from those of the hotel-focused business. Finally, the 8 listed hotels were selected for the ratio analysis in this study.

The financial data of the selected hotels were mainly obtained from the Stock Exchange of Thailand (SET) 2010-2014. Proprietary data were collected from Business Online Public Company Limited (*BOL)* which compiled a database of hotels registered in Thailand (BOL, 2015). The major source of data on operational performance were the hotels' annual reports and official reports, Form 56-1. This form is known as a disclosure report submitted to Thailand's supervisory agencies on additional information on company operations. However, some data from these sources were relatively limited. For instance, not all hotels adequately disclose data pertaining to hotel operational performance.

Calculation: Although the literature review yielded a large set of ratios, the numbers of ratios used in this study were compromised due to data availability. Relying on publicly available data deemed important, the total of 27 ratios were selected and classified under financial and operational dimensions. The indicators for financial performance comprise liquidity, solvency, activity, and profitability ratios. The indicators for operational performance include occupancy rate, average daily room rate (ADR), RevPAR, and operating costs. Selected descriptive statistics, mean and standard deviation, were then used for the analysis and interpretation. In order to get normal comparative figures, some outliners were deleted from the calculations. The average value of each ratio was calculated for each year. The five-year-averages were computed to yield the overall comparative figures. The computed ratios were measured against industry averages for internal comparison. An external comparison was made with hotels in ASEAN using selected indicators.

ANALYSIS AND DISCUSSION

In this section, ratio analysis is discussed followed by hotel industry comparison. Table 2 shows the computed ratios of financial performance of Thai listed hotels between 2010 and 2014.

Table 2: The financial ratios of Thai listed hotels (2010-2014)

Ratios	2010	2011	2012	2013	2014	Avg.	S.D.
Liquidity Ratios							
Current ratio (times)	0.34	0.33	0.70	0.55	0.61	0.51	0.46
Quick or Acid-test ratio (times)	0.31	0.31	0.67	0.52	0.59	0.48	0.46
Accounts receivable turnover (times)	19.54	18.35	11.09	13.01	10.70	14.54	7.77
Average collection period (days)	18.68	19.90	32.93	28.05	34.12	25.11	20.26
Operating cash flows to current liabilities (%)	63.69	13.42	47.98	37.82	38.19	40.22	61.41
Solvency Ratios							
Debt ratio (times)	0.44	0.42	0.41	0.43	0.44	0.43	0.18
Debt to equity (times)	1.06	1.01	0.79	0.86	0.89	0.92	0.61
Times interest earned (times)	-0.93	2.37	9.88	4.70	3.09	3.82	7.49
Operating cash flows to total liabilities (%)	46.16	4.59	24.53	28.61	20.49	24.88	49.25
Activity Ratios							
Inventory turnover (times)	45.31	59.83	65.48	64.65	70.58	61.17	54.67
Inventory holding period (days)	8.06	6.10	5.57	5.65	5.17	5.97	10.99
Fixed asset turnover (times)	0.61	0.68	0.74	0.84	0.82	0.74	0.57
Total asset turnover (times)	0.26	0.30	0.32	0.37	0.35	0.32	0.29
Profitability Ratios							
Gross profit margin (%)	41.30	39.19	48.21	47.95	42.04	43.74	21.89
Operating profit margin (%)	-14.20	-11.81	0.97	4.73	-5.66	-5.19	20.38
Net profit margin (%)	3.11	-2.20	26.25	19.78	7.07	10.80	25.51
Return on assets (ROA) (%)	-0.58	-0.70	9.25	5.30	1.43	2.94	8.56
Return on equity (ROE) (%)	-2.15	-2.29	16.53	10.63	3.56	5.26	17.45
Earnings per share (EPS) (Baht)	0.79	1.05	4.95	2.91	1.99	2.34	5.23
Price/Earnings ratio (P/E) (times)	48.31	14.77	39.66	15.34	32.68	30.15	18.32
Market value to Book value (P/BV) (times)	1.82	1.79	2.51	2.74	3.80	2.53	3.09
Dividend yield (%)	1.58	2.03	1.46	1.79	2.89	1.95	1.40

Liquidity Ratios

Current ratio: This ratio compares current assets to current liabilities. The ideal level of a current ratio varies depending on types of business. In hospitality industry, hotels can operate with a current ratio of 1.2 - 1.5 (Andrew and Schmidgall, 1993; Jagels, 2007). During the studied period, the average current ratio was below 1.2. Thai listed hotels should therefore focus more on the policy to meet short-term obligations.

Quick or acid-test ratio: This ratio indicates the firm's ability to pay its liabilities with liquid assets in a timely manner. The truly liquid current assets are, for example, cash, marketable securities, and accounts receivable. Generally, the quick ratio should be at least 1.0. With the average of 0.48, Thai listed hotels should manage their liquid assets with caution.

Accounts receivable turnover: Normally, the turnover of 10 to 30 times per year is acceptable (Jagels, 2007). Thai listed hotels' accounts receivable turnover ratios were found to be greater than 10, with an average of 14.54 times. This indicated the efficiency of the firms in collecting accounts receivable. However, the ratios showed a downward trend from 19.54 times in 2010 to 10.70 times in 2014. The firm's managers should tighten their policy on credit extension and collection process.

Average collection period: In practice, an average collection period of 30-35 days is reasonable for a firm that allows 30-day credit terms for its accounts receivable (Bragg, 2010). The accounts receivable collection period of Thai listed hotels in each year showed a good sign of performance, and was in line with accounts receivable turnover ratios calculated above. However, the collection period indicated an upward trend. Management should keep an eye on their credit policy and receivable collection procedures to bring the ratio to its acceptable figure.

Operating cash flows to current liabilities: This ratio considers operating cash flows generated from its business operation for a period of time. It is an indicator in addition to current ratio and quick ratio which have limitations as they use static number at a specific point in time. It is suggested that the higher the ratio, the more financially healthy organization (Andrew and Schmidgall, 1993). The computed operating cash flows to current liabilities averaged at 40.22%. Consequently, the listed hotels should improve their operations to generate more cash flows to meet current liabilities.

Solvency Ratios

Debt ratio: The ratio indicates the proportion of total assets that are financed with debts. Generally, the lower the ratio, the higher security the creditors will earn. Traditionally, the hospitality industry is financed with debt ranging from of 0.60 - 0.90 (Jagels, 2007). The debt ratio in each of the 5 years was below the traditional range with an average of 0.43. This implied that the listed hotels were less exposed to financial risk.

Debt to equity ratio (D/E)**:** This ratio presents the relative proportion of debt and equity. A firm may choose to use debt instead of additional equity to increase the return on the equity already invested. A low debt to equity is preferred by creditors. Usually, the ratio should not exceed 1 to 1 as it suggests that the firms' assets are equally funded by creditors and investors. The computed debt-to-equity ratios between 2010 and 2014 were within the normal range, with an average of 0.92. Improvement was seen during 2012 - 2014 when the ratios were brought under 1.00.

Time interest earned**:** This ratio refers to the number of times interest expenses can be covered by the earnings before taxes and interest expenses (EBIT). Creditors favor higher ratios as it means that the company can afford to make interest payments. Jagels (2007) suggested the times interest earned ratio of 2.0 or more. All, except the calculated ratio of 2010, were greater than 2.0. Apparently, Thai listed hotels had sufficient earnings to pay for annual interest expenses. Nonetheless, the decreasing ratios from 9.88 in 2012 to 3.09 in 2014 could be a cause for concern.

Operating cash flows to total liabilities**:** The concept behind this ratio is the same as the concept behind operating cash flows to current liabilities. In general, the cash flows from operations should be high relative to total liabilities. The calculated average showed that the operating cash flows of the listed hotels covered only 24.88% of total liabilities. Operations are in need of improvement to generate higher cash flows.

Activity ratios

In hospitality industry, the activity ratios determine management's ability to control certain classes of assets such as inventory and long-term assets (Jagels, 2007). These ratios measure management efficiency in using and controlling these assets.

Inventory turnover: This ratio measures the firm's ability in managing its inventory. Inventory turnover varies from industry to industry. In hotel operations, the inventory supplies are typically food, beverages, and operating supplies. Most inventories are perishable stocks and need quick turnover. Theoretically, inventory turnover should be calculated separately for food and beverage supplies. Normally, food turnover rate is 2-4 times per month while beverage turnover rate varies from 1-4 times a month (Jagels, 2007). However, the financial statements of Thai firms did not classify types of inventory. The calculation in this study was therefore based on the overall inventory. A substantial increase was found from 2010 to 2014. An average of 61.17 times annually or a turnover of 5.10 times a month showed an effective management of hotel inventory.

Inventory holding period**:** Inventory holding period measures the average number of days to hold the inventory one time. The computed holding period decreased substantially over the years, with an average of 5.97 days. A low holding period signaled a good performance of Thai listed hotels in inventory management.

Fixed asset turnover**:** This ratio evaluates the firm's efficiency in utilizing its property and equipment. It indicates the number of times a firm can turn its fixed assets into total revenue. In hotel business, both fixed asset turnover and total asset turnover tend to be relatively low as the business typically has high capital intensity. Fixed assets represent the majority of total assets. The acceptable range for fixed asset turnover ratio in hotel operations lies between 1.5 and 2.0 per year. The higher the ratio, the more capable the firm is of managing assets (Jagels, 2007). The ratios of Thai listed hotels were lower than the normal range, albeit an increasing trend. An average of 0.74 annually indicated that the efficiency in fixed asset management was rather low. The situation called for a better utilization of fixed assets.

Total asset turnover: This ratio determines the use of total assets in relation to total revenues. The calculated total asset turnover was quite low over the studied period - all below 0.50. The figures presented an increasing trend until a drop in 2014. Thai listed hotels seriously needed to manage and use all of their assets more efficiently.

Profitability Ratios

Gross profit margin: This indicates the margin available to cover operating expenses. The calculated gross profit margin of the listed hotels was around 39 – 48%, with an average of 43.74%. Thus, the cost of sales of these hotels was approximately between 52 – 61%.

Operating profit margin**:** This measures operating income relative to total revenue. The calculated figures showed that, despite positive gross profit margin, the operating profit margin was negative, averaged at -5.19%. Management needed to handle selling and administrative expenses more effectively as they proved too high.

Net profit margin**:** This is a key indicator of firm's profitability. It represents a firm's ability to generate revenue and control expenses. This margin shows the net return on sales after tax deduction. All, except the calculated ratio of 2011, were positive, rendering an average of 10.80%. The reason for positive net profit margin, albeit negative operating profit margin, lies in hotel firms' other revenue sources. Such revenues as interest income, dividend income, and gain on disposal of fixed assets and properties help compensate for the loss of profit margin.

Return on assets (ROA): Generally, a high ROA ratio is preferred because it shows management's efficiency in its use of total assets. From the calculation, an average ROA in all five years was only 2.94%. The ROA ratio peaked in 2012 before declining again in the following years. The listed hotels needed more control on asset utilization to produce greater profits.

Return on equity (ROE): This measures the rate of return on common stock of shareholders' investment in the firm. ROE should be at least 15% (McClure, 2014). The trend in ROE over the studied period was up and down. The calculated ratios were below the recommended percentage. This should draw management's closer attention to profit generation.

Generally, profitability ratios of any listed firm also include market value ratios. Traders in the stock market usually concern themselves with such indicators as earnings per share (EPS), price-earnings ratio (P/E), market value to book value ratio (P/BV), and dividend yield.

Earnings per share (EPS): This is the ratio of the net income divided by the average common shares outstanding. The calculated earnings per share ratios showed a fluctuating trend with an average of 2.34 Baht. Management could strive more for a higher EPS, as this is an indicator of a company's profitability on a shareholder basis.

Price-earnings ratio (P/E): This is the ratio of stock price per share divided by earnings per share (EPS). Financial analysts often use the P/E ratio in analyzing investment opportunities of the firms, i.e., what the market is willing to pay for a stock based on its current earnings. It is noted that the P/E ratios of Thai listed hotels fluctuated from year to year, ranging from 14.77 to 48.31 times. Overall, an average P/E ratio of 30.15 was higher than that of the SET of 16.30 (Sawangpak, 2015).

Market value to book value (P/BV): This ratio compares the market value to book value per share of stock. Investors expect to see high P/BV. Since 2011, the calculated P/BV increased every year, with an average of 2.53. This indicated that the market value was greater than the book value at an average of 2.53 times per share. However, investors are advised to also use other stock valuation methods, as many factors could affect the market value of stock.

Dividend yield: This is the ratio of dividend per share divided by price per share. This shows the return that shareholders actually achieve on their investment. The computed dividend yield of Thai listed hotels fluctuated with an average of 1.95%. This was considered relatively low comparing to investments in other businesses. During 2010 - 2014, the interest rates in Thailand's money market outperformed dividend yield of the listed hotels for most of the studied period (Bank of Thailand, 2015). The lower returns should warrant management attention.

Operating ratios

Operating ratios include occupancy rate, average daily room rate (ADR), revenue per available room (RevPAR), food and beverage (F&B) cost, and labor cost (Table 3).

Table 3: The operating ratios of Thai listed hotels (2010-2014)

Operating ratios	**2010**	**2011**	**2012**	**2013**	**2014**	Avg.	**S.D.**
Occupancy rate (%)	47.43	55.05	64.17	67.47	57.17	58.26	23.62
ADR (Baht)	4,861.45	4,857.54	5,814.04	5,328.79	6,738.32	5,520.03	3,547.75
RevPAR (Baht)	2,467.69	2,680.96	3,591.02	3,403.25	3,638.96	3,156.38	1,804.47
F&B cost (%)	14.94	14.45	14.30	12.92	14.01	14.12	7.28
Labor cost (%)	36.72	34.72	35.90	27.04	34.61	33.80	18.05

Occupancy rate**:** This rate is calculated by dividing the number of rooms sold by the number of rooms available for sale, and multiplying it by 100 to reflect the percentage. Andrew and Schmidgall (1993) identify "the paid occupancy percentage" as an activity ratio as it is a key measurement of facility utilization. This study, however, classified occupancy rate as an operating ratio in the same way as Jagels (2007). The listed hotels' occupancy rate went up for 3 consecutive years before a decline in 2014 as a result of the country's political situation, averaging 58.26% annually.

Average daily room rate (ADR): ADR is an average rate paid for rooms sold in a given period. The price per room is typically different depending on types of room, market segment, and seasonality. Most hotels employ an ADR as a critical indicator of operations. The ADR of Thai listed hotels peaked in 2014 at 6,738.32 Baht, despite a declining occupancy rate.

Revenue per available room (RevPAR): RevPAR is calculated by dividing total room revenue by the number of available rooms, taking into account the amount of unoccupied available rooms. The use of RevPAR therefore provides more comprehensive information than using the ADR or the occupancy rate alone. From the calculation, the RevPAR fluctuated year-to-year with an average of 3,156.38 Baht.

Food and beverage cost percentage: The ratio shows food and beverage costs relative to food and beverage revenue. During the studied period, the range of this cost was between 12.92% and 14.94%, with an average of 14.12%. The figures should be compared with the industry standard or the firms' predetermined cost target to determine if proper operational control is needed.

Labor cost percentage: Labor cost includes salaries, wages, and employee benefits. Labor cost percentage is the ratio of total labor costs relative to the revenue. It is generally the largest proportion of hotel operating costs. The labor cost percentages of the listed hotels showed a sign of fluctuating trend with an average of 33.80%.

Industry Comparison

Internal comparison: It is beneficial to compare the ratios over a period of time for the same company (or groups of companies) to see if the performance has improved or deteriorated over the years. Likewise, the comparison to industry standards is crucial to get a sense of how well the companies have performed within the industry. Industry averages used in this study derived from many sources. Financial ratios were obtained from BOL, occupancy rate and ADR, from the Bank of Thailand. RevPAR was calculated by the authors.

Table 4 presents a comparison between the industry averages and the calculated ratios of Thai listed hotels in 2013. Mostly, the hotels performed well in terms of using and controlling assets (activity ratios), generating profits (profitability ratios), and operational efficiency (operating ratios). On the other hand, most liquidity ratios were below industry averages. With regards to solvency ratios, Thai listed hotels showed mixed performance. Overall, the financial ratios of the listed hotels signaled a good performance.

External comparison: The computed ratios can also be compared with different benchmarks (i.e., those of other countries) to get a broader picture of Thai listed hotels' performance. Table 5 shows an example of the selected

Table 4: The comparison of Thai listed hotels' ratios and industry averages (2013)

Ratios	Thai Listed Hotels	Industry Averages	Difference	Performance of Thai listed hotels	
				Out-perform	Under-perform
Liquidity Ratios					
Current ratio (times)	0.55	4.56	-4.01		√
Quick or Acid-test ratio (times)	0.52	4.41	-3.89		√
Accounts receivable turnover (times)	13.01	24.43	-11.42		√
Average collection period (days)	28.05	14.82	13.23		√
Solvency Ratios					
Debt ratio (times)	0.43	0.64	-0.21	√	
Debt to equity (times)	0.86	0.21	0.65		√
Times interest earned (times)	4.70	0.05	4.65	√	
Activity Ratios					
Inventory turnover (times)	64.65	38.19	26.46	√	
Inventory holding period (days)	5.65	14.82	-9.17	√	
Fixed asset turnover (times)	0.84	0.80	0.04	√	
Total asset turnover (times)	0.37	0.42	-0.05		√
Profitability Ratios					
Gross profit margin (%)	47.95	48.40	-0.45		√
Operating profit margin (%)	4.73	-2.81	7.54	√	
Net profit margin (%)	19.78	-7.54	27.32	√	
Return on assets (ROA) (%)	5.30	-1.03	6.33	√	
Return on equity (ROE) (%)	10.63	-2.28	12.91	√	
Operating Ratios					
Occupancy rate (%)	67.47	64.86	2.61	√	
ADR (Baht)	5,328.79	1,948.96	3,379.83	√	
RevPAR (Baht)	3,403.25	1,264.10	2,139.15	√	

Remarks: Ratios comparisons are made when industry averages are available.
This study used the latest industry averages of 5,573 registered hotels in 2013 as the comparative figures.

measure, which is hotel occupancy rate, being compared across ASEAN during 2010-2013. The results indicated that the average occupancy rate of Thai listed hotels was close to that of Thailand's hotel industry as a whole (Bank of Thailand, 2010-2013). However, comparing the averages among ASEAN countries, Thai hotels' occupancy rate was much lower than that of Singapore and the Philippines.

Table 5: The ASEAN hotels' occupancy rate (%) (2010 – 2013)

Rank	Country	2010	2011	2012	2013	Avg.
1	Singapore	85.1	86.5	87.8	90.6	87.5
2	Philippines	67.6	67.3	66.9	67.4	67.3
3	Malaysia	59.3	60.6	59.3	58.9	59.5
4a	**Thai listed hotels**	47.4	55.1	64.2	67.5	58.5
4b	**Thailand (All hotels)**	50.2	57.7	60.8	64.9	58.4
5	Laos	56.0	56.4	56.6	57.1	56.5
6	Vietnam	51.9	51.9	51.9	51.9	51.9
7	Indonesia	48.9	51.1	51.8	52.6	51.1
8	Brunei	41.3	41.4	41.4	41.5	41.4
9	Myanmar	34.4	36.4	37.1	38.2	36.5
10	Cambodia	N/A	N/A	N/A	N/A	N/A

Sources: 1-3, 5-9 data taken from Euromonitor International (2014) *Hotel Bed Occupancy Rates*; 4a calculated by the authors; 4b denotes Thailand's hotel industry (Bank of Thailand, 2014)

Based on the selected measure, occupancy rate, Singapore was the best-in-class. Singapore's listed hotels therefore became a benchmark for Thai listed hotels to compare themselves against. It was found that Thai listed hotels showed a good sign of performance in terms of activity ratios. On the other hand, issues that called for management attention were ensuring profitability and the ability to meet financial obligations (Table 6).

Table 6: Ratio comparison between the listed hotels in Thailand and Singapore (2014)

Ratios	Thai Listed Hotels	Singapore Listed Hotels[1]	Difference	Performance of Thai listed hotels	
				Out-perform	Under-perform
Liquidity Ratios					
Current ratio (times)	0.61	1.86	-1.25		√
Quick or Acid-test ratio (times)	0.59	1.17	-0.58		√
Solvency Ratios					
Debt to equity ratio (times)	0.89	0.55	0.34		√
Times interest earned (times)	3.09	5.78	-2.69		√
Activity Ratios					
Inventory turnover (times)	70.58	43.67	26.91	√	
Inventory holding period (days)	5.17	8.36	-3.19	√	
Total asset turnover (times)	0.35	0.22	0.60	√	
Profitability Ratios					
Gross profit margin (%)	42.04	51.35	-9.31		√
Operating profit margin (%)	-5.66	29.00	-34.66		√
Net profit margin (%)	7.07	36.98	-29.91		√
Return on assets (ROA) (%)	1.43	1.83	-0.40		√
Return on equity (ROE) (%)	3.56	6.62	-3.06		√
EPS (Baht)	1.99	3.86*	0.41		√
P/E (times)	32.68	29.34	3.34	√	
P/BV (times)	3.80	0.87	2.93	√	
Dividend yield (%)	2.89	2.65	0.24	√	

Remarks: Ratios comparisons are made when data are available. *Converted to Thai Baht
Source: [1]Calculated from the data of 12 listed hotels obtained from the Stock Exchange of Singapore (SGX)

CONCLUSION

This study discussed ratio analysis of the hotels listed on the Stock Exchange of Thailand (SET) during 2010 - 2014. The study analyzed 2 dimensions of hotel performance; namely, financial and operational. The ratios evaluating financial performance were liquidity, solvency, activity, and profitability ratios. The measures for operational performance included occupancy rate, average daily room rate (ADR), RevPAR, and operating costs.

The study found that Thai listed hotels should improve the ability to meet financial obligations. With regards to the activity ratios, inventory management proved efficient. However, a more effective use and control of assets, especially the fixed ones should be emphasized. Most profitability ratios showed inconsistent year-to-year fluctuation during the studied period. The operating profit margin, ROA, and ROE needed to be increased, along with asset utilization and profit generation. All market value ratios fluctuated from year to year, with relatively low dividend yield. The occupancy rate, ADR, and RevPAR revealed good performance. Also, operating costs were efficiently controlled, especially food and beverage cost and labor cost.

The computed ratios were then measured against industry averages for internal comparison. Overall, the listed hotels performed better than the industry ones. An external comparison was made with hotels in ASEAN using the selected indicator. Thai listed hotels ranked fourth in occupancy rate after Singapore, the Philippines, and Malaysia. Benchmarking with the listed hotels in Singapore, the Thai ones showed a good performance in activity ratios. On the other hand, the ability to increase profitability and meet financial obligations should be a concern for management to ensure competitiveness.

The results of the study make a contribution to academicians and practitioners in evaluating hotel business. The calculated ratios can serve as benchmarks for the hotel industry. Additionally, the method employed in this study can be applied as a tool to measure hotel performance. The suggested future research should involve comparing Thai listed hotels' ratios before and after the establishment of the ASEAN Economic Community (AEC) to assess its impact. Looking outward, a comparative study of ratios among hotels in the AEC will be beneficial to hotel industry in the region.

REFERENCES

Andrew, William P. and Schmidgall Raymond S. (1993). *Financial Management for the hospitality Industry*. MI: Education Institute of the American and Hotel & Motel Association.

Bank of Thailand (2015). *Tourism Indicators*, Retrieved from https://www.bot.or.th/statistics, June 15, 2015.

Bank of Thailand (2015). *Interest rate*, Retrieved from https://www.bot.or.th/statistics, June 15, 2015.

Boundless (2015). "Comparisons within an Industry." *Boundless*, July 21, 2015, Retrieved from https://www.boundless.com, November 7, 2015.

Bragg, Steven M. (2010) *Business Ratios and Formula: A Comprehensive Guide.* John Wiley & Sons.

Business Online Public Company Limited (2015). *BINGO Report.* Retrieved from http://corpus.bol.co.th, October 12, 2015.

Euromonitor International (2015) *Lodging in Thailand*. Retrieved from Euromonitor Passport Database, November 14, 2015.

Euromonitor International (2014). *Travel Accommodation in Thailand*. Retrieved from Euromonitor Passport database, May 19, 2015.

Euromonitor International (n.d.). *Hotel Bed Occupancy Rates*. Retrieved from Euromonitor Passport database, May 19, 2015.

Garrison, Noreen and Brewer (2012). *Managerial Accounting*, 14th edition, McGraw-Hill Companies.

Jagels, Martin (2007), *Hospitality Management Accounting*, 9th Edition, John Wiley and Sons.

Kim, Woo Gon and Ayoun, Baker (2005). "Ratio Analysis for the Hospitality Industry: A Cross Sector Comparison of Financial Trends in the Lodging, Restaurant, Airline, and Amusement Sectors." *Journal of Hospitality Financial Management*, Vol. 13, Issue 1. Article 26.

Kim, Dong J. (2006). "A Comparative Study of Financial Ratios between Hotels and Restaurants." *International Journal of Tourism Sciences*, Vol. 6, No. 1, 95-106.

McClure, Ben (2014). "ROA and ROE Give Clear Picture of Corporate Health" Retrieved from www.investopedia.com, July 20, 2015.

Parkan, Celik (2005). "Benchmarking Operational Performance: the Case of Two Hotels." *International Journal of Productivity and Performance Management*, Vol. 54, No. 8. 679-696.

Sainaghi, Ruggero (2010). "A Meta-Analysis of Hotel Performance. Continental or Worldwide Style," *Tourism Review*, Vol. 65, No. 3, 46-69.

Singh, A.J. and Schmidgall, Raymond S. (2002). "Analysis of Financial Ratios Commonly used by US Lodging Financial Executives", *Journal of Retail and Leisure Property*, Vol. 2, No.3, 201-213.

Sawangpak, Arparporn (2015). The Stock Exchange of Thailand. *Year 2015: Highly Fluctuate…Buyer Beware, Stock Information*. Retrieved from www.set.or.th/yourfirststock/research/DBSV_01_58pdf, June 22, 2015

Tugas, Florences C. (2012). "A Comparative analysis of the Financial ratios of Listed Firms belonging to the Education Subsectors in The Philippines for the Year 2009-2011." *International Journal of Business and Social Sciences*, Vol. 3. No. 21, 173-190.

Department of Employment, Ministry of Labor. (2009). *Thailand Standard Industrial Classification*. Bangkok.

World Tourism Organization: UNWTO (n.d.). *Tourism 2020 Vision*. East Asia and Pacific. Vol. 3. Madrid. Retrieved from https://pub.unwto.org/WebRoot/Store/Shops/Infoshop /Products/1189/1189-1.pdf, January 28, 2015.

Journal of Business and Behavioral Sciences
Vol 28, No. 2; Fall 2016

EPA RISK AVERSION UPON AN INNOVATIVE EN-PRODUCT ADOPTION

Tom Suraphol Apaiwongse
Clark Atlanta University

ABSTRACT: Grounded by Contingency Theory, the study focuses on the impact of Environmental Protection Agency (EPA) regulatory action on market innovation. Based upon detailed interview from eighty one marketing executives representing both the industry and non-industry, the study identifies fourteen different uncertainty categories relating to EPA regulations. The results indicate a very high level of perception within the government regulatory action posing a significant risk to ecologically related innovation among industrial firms.
Key Words: Environmental Marketing, Adoption of EN-Product Innovation, Regulatory Uncertainty, and EPA Market Based “Bubble” Policy.

INTRODUCTION

The 2000s have become the "Decade of the Environment," and both business and government are coming to the same conclusion: environmental protection is not an option; it is an essential part of the complex process of doing business. Global concern about protecting the environment fostered regulatory reforms and a growing variety of specific legislation favoring environmental friendly marketing. More recently, the Obama Administration has favored the replacement of government regulatory programs with an innovative marketing-based policy known as the "Bubble" voluntary program. Such a change in the EPA orientation, which is already somewhat in evidence, has gone on to make a transition from a "command-and-control" orientation to a "marketing" orientation. The Bubble is an approach that attempts to elicit voluntary action based on a realistic understanding of industry's needs, attitudes and behavior. The voluntary policy of regulatory alternatives for industry is compatible with the concept of ecological marketing, a term used to identify marketing that takes into consideration the environmental consequences of its actions. Unfortunately, U.S. environmental policy has been dominated by the "market failure" paradigm - the belief among businesses that environment problems are caused by the failure of the market to provide the right signals. Surprisingly, many businesses are not aware of the Bubble policy, and, thus, only a relatively small proportion of industrial firms are engaged in the policy. Despite the increasing importance of voluntary reaction to green policies, scholarly inquiry on the topic has been hampered in two ways. First, the literature suggests that only a relatively small proportion of industrial firms react to the green voluntary policies (Apaiwongse, 2013). Second, very little academic research has attempted to practically explain the interactive role played between government and business in an ecological as well as political setting (Apaiwongse, 2007). Another question is of major concern for regulatory agencies and industrial firms alike: whether or not this marketing-based policy may produce some barriers to the acceptance among businesses. Yet, the issue remains one about which marketers know relatively little. The scant attention that has been given to studying this issue is in no way commensurate with its importance. To investigate this issue, a survey was conducted to identify how industries adopt the regulatory policy. After describing the concept of [1] contingency theory in marketing, and [2] the concept of uncertainty, the paper identifies different types of regulatory uncertainty.

CONCEPTUAL FRAMEWORK

Prior Marketing Research on Contingency Theory: Early contingency theorists studied the relation between the environment and structure at a macro level. Lawrence and Lorsch (1967) studied firms representing three levels of environmental uncertainty (as measured by clarity of information, uncertainty of causal relationships, and the time span required for definitive feedback). The demands of the external environment lead to differences in decision-making structures, not only across firms but also between functional units within the same organization. Managers facing dynamic and complex environments sensed the need to implement organic structures whereas units operating in simple and more stable environment were more likely to develop short-term, formalized outlooks and implement mechanistic structures.

Galbraith (1973) provided significant theoretical support for the contingency perspective by demonstrating the important link between uncertainty and information. The greater the task uncertainty is, the greater the amount of information that must be processed among decision makers during task execution. Thus, variations in organizational structure are often simply variations in an organization's strategy for processing such information. For the mechanistic organization, tasks are usually routine and predictable and behavior can be coordinated among units by rules and programs. As the level of uncertainty increases, a few situations arise for which no rules are available and these exceptions are referred to a level in the hierarchy where information is adequate to address them. The real dilemma for an organization occurs, however, when environmental uncertainty reaches a level at which the exceptions are so frequent that none of these more simple integrating devices is adequate to cope with them. In contrast to stable environments in which bureaucratic structures are the preferred design strategy, uncertainty environments involve so many exceptions that information must be processed.

One view, rooted in the contingency theory and applied in the context of marketing centers is that increasing levels of task uncertainty lead to a marketing decision process characterized by increasing participation by lower level members of the organizational hierarchy and less subject to the control of formalized rules and procedures (Apaiwongse, 2009). In contrast to this traditional contingency model, extensive support is found in the marketing literature for the view that at high levels of uncertainty, organizational decision-making processes are characterized by a constriction of authority and an increase in rule-governed behavior as decision units act to minimize the errors often associated with decision making in uncertain situations (Apaiwongse, 2011).

A major difficulty to integrate different perspectives on the uncertainty-structure relation in marketing decision units is the construct of uncertainty. Uncertainty refers to events that the organization can't forecast. It is not mere change or the rate of change, but the unpredictable change which affects critical organizational decision-making. Lawrence and Lorsch (1967) operationalized the concept of uncertainty by describing organizational uncertainty as: [1] lack of information clarity; [2] general uncertainty of causal relationships; and [3] a long intervening time span between events and feedback of results. These factors, in turn, lead to an individual's perception of organizational or group uncertainty. Contingency theorists generally have used a view of uncertainty centering on the perceived uncertainties of the marketing task. Some evidence reflects the application of the perceptual-measure approach to the industrial marketing area (Apaiwongse, 2008).

Duncan's measure of perceived uncertainty was developed from a semantic analysis of individual verbalizations of the concept of uncertainty. The three dimensions included in Duncan's measure of uncertainty include: [1] lack of information regarding environmental factors associated with a given decision-making situation; [2] lack of knowledge about the outcome of a specific decision in terms of how much the organization would lose if the decision were incorrect; and [3] ability or inability to assign probabilities as to the effect of a given factor on the success or failure of a decision unit in performing its function.

RESEARCH OPERATIONALIZATION

Measuring Instrument for Uncertainty: In an effort to be consistent with the logical and methodological contributions made by Lawrence and Lorsch (1967) and Duncan (1972), measure of uncertainty are modified and applied to the present study. To operationalize the concept of regulatory uncertainty for this research, the specific dimensions of the EPA regulation are specifically required. Viewpoints expressed by company executives and in some leading environmental journal articles are also presented. The operationalization of this variable involved marketing-center members asked, by a means of a set of ten, five-point scale, Likert-type items, whether they believe or disbelieve each statement.

STUDY 1

To determine whether regulatory uncertainty, as perceived by marketing executives, has an impact on new product innovations, two focus group interviews in one group of four and another group of five were conducted. Interviews were used to determine the uncertainty perceived by marketing executives. The findings from the interview were used as a part of the final questionnaire construction.

STUDY RESULTS

All marketing executives found the concept of uncertainty (or barrier) meaningful and generally expected to see a variation in uncertainty across firms. The executives suggested some novel or practical approaches for measuring uncertainties. All executives counted the subject matter of firm's adaptive response to the EPA policy as crucial. Findings are presented by including supportive verbatim quotations from executives and from viewpoints expressed consistently by those conveyed in leading environmental journal concerning the policy uncertainty.

Finding 1: An executive's perceived degree of uncertainty regarding EPA policy will affect the company adoption of the EN-Product innovations.

"Uncertainty is the primary problem. Uncertainty, in turn, leads to inactivity."

"EPA established so many constraints that no reasonable case could succeed."

Finding 2: EPA's inability to provide clear information on the revision of the EPA policy will affect the company adoption of the policy.

"EPA's inability to modify environmental regulations definitely works against EN-Product innovations. The EPA has been talking about amendments to the new policy since late 1980, but it has been unable to promulgate the amendments because of the internal problems."

"[Company name] has been an innovator in the pollution control products. The company has done the homework, but EPA adds a wrinkle to the assignment. In announcing its policy, the agency will view each application as a revision to a particular State Implementation Plan (SIP) under the Clean Air Act. This requires public notification and a hearing at the state level, followed by similar proceedings at the federal level."

Finding 3: Air pollution control regulations are ambiguous to the point that they affect the company adoption of the EPA policy.

"Secondary ambient air quality standard for particulate cannot be universally met in maintaining overall program uncertainty."

"Whether EPA intends to monitor each innovative plan or whether the agency will opt for year-end audits of state practices is not clear."

Finding 4: The air pollution control regulations are not flexible enough to encourage the company to adopt the EPA policy.

"The long period (15 months) which has elapsed since beginning discussions on the innovative project and the numerous "red tape" procedures have been very frustrating and make it questionable whether it is worth the effort. Part of the SIP revision is still not finalized and flexible."

"The EPA policy was so restrictive that it didn't help much."

Finding 5: The firm's ability to determine whether a proposal will be approved or what action should be taken to meet expectations of air pollution control regulation are crucial for the adoption of the EPA policy.

"The policy EPA referred to was an issue only discussed verbally in the presence of the state people. The EPA indicated willingness to explore it, but the state vetoed the concept. EPA admitted the state's right to do this. The company never formally submitted an application."

Finding 6: The information about conflicts and inconsistencies between air pollution regulations is not adequate for the company to adopt the EPA policy.

"The federal-state relations may be most important. It takes states eight months to one year to adopt regulations. It takes EPA one federal regulatory notice of change in strategy to make a state plan unreasonable for bubbling. Many state regulations have inconsistent EPA regulations with federal programs."

Finding 7: Delays by regulators in clarifying regulations affect the adoption of the EPA policy.

"Explicit guidance is mandated by EPA to get this program rolling as soon as possible. It is the greatest program in Air the EPA has, and they are destroying their creditability and image by not pushing it to the hilt."

STUDY 2

The study used a descriptive research design to address the following issues raised in focus group interviews and the literature review.

1. What are opinions of marketing executives toward the EPA policy?
2. Can firms specify barriers that adversely affect their response to the EPA policy upon EN-product innovation?

The study focuses on the opinions among 81 marketing executives about EPA regulatory barriers on innovation in industry. The study, based in detailed interview from marketing executives, identifies 14 risk (barrier) categories relating to EPA regulations with the order of importance.

STUDY RESULTS

The results indicate a very high level of opinions among regulatory agencies regarding the regulatory action posing a significant risk to ecologically related innovation among industrial firms. The findings from regulatory agencies indicate the following impact of uncertainty on the EPA policy on industrial innovation:

1. EPA regulatory time pressures leading to non-optimal innovations;
2. Prohibitively high costs of complying with EPA regulations;
3. Unclear scope or implications of EPA regulations;
4. Delays by the agency in promulgating guidelines required by the environmental law;
5. Inability or unwillingness of agency to modify EPA regulations;
6. Disagreement within the agency about the application of EPA regulations;
7. Inconsistency over time in the agencies application of EPA regulations;
8. Inability of firm to allocate the resources to comply with EPA regulation;
9. Conflicts and inconsistencies between EPA regulations;
10. Inability of firm to meet prescribed deadlines in EPA regulations;
11. Lack of mechanism within the agency for explaining EPA regulations;
12. Lack of effective appeal procedure;
13. Differential treatment by the agency of the entities affected by the EPA regulations; and

14. Unwillingness of the agency to explain EPA regulations.

STUDY 3

The study compares uncertainty attributes, as viewed by both marketing executives and environmental officers. The basis for constructing the questionnaire to measure uncertainty was theoretically grounded in the work of some organizational theorists. Factor analysis is employed to identify the key dimensions of the perceived policy uncertainty. Three orthogonal factors were generated. Those factors (or attributes) are: [1] Lack of Knowledge Concerning Outcomes, Predictions and Expectations; [2] Ambiguity of the Regulation and Delayed Feedback; and [3] Lack of Enough Information for Decision Making.

The three interactions of the above uncertainty attributes with the adoption decision for new EN-Product innovations were all shown to be substantively and statistically significant. The attribute which has the greatest impact on innovation decision is "Lack of Knowledge Concerning Outcomes, Predictions and Expectations," followed by "Lack of Enough Information for Decision Making." The uncertainty attribute "Ambiguity of the Regulation and Delayed Feedback" seemed to have the least impact on innovation adoption.

STUDY RESULTS

The findings from the personal interview indicate the following impact of uncertainty on the EPA policy on industrial innovation:

1. As a constraint, a marketing executive's perceived degree of uncertainty regarding the environmental policy will affect the marketing decision to innovate new environmentally-related products [EN-Products].
2. EPA's inability to provide clear information on the revision of the environmental policy affects the decision on new EN-Product innovations.
3. Environmental regulations are ambiguous to the point that they affect the decision on new EN-Product innovations.
4. The environmental regulations are not flexible enough to promote new EN-Product innovations.
5. The firm's ability to determine whether a new EN-Product will be approved or what action should be taken to meet expectations of environmental control regulation is a crucial factor which affect the decision to innovate new EN- Products.
6. The available information about conflicts and inconsistencies among environmental control regulations affect the decision on new EN-Product innovations.

The findings do indicate that the EPA regulations impose considerable direct and indirect barriers to the decision to adopt new EN-Product innovations. There appear to be two main areas of impact through regulation on industrial firms. The first is that members of many marketing centers perceived the unclear and ambiguous content of the environmental regulation as posing a significant barrier to its adoption of new innovation. The second is related to such problems of uncertainty as lack of information, inability to predict the result from the decision to innovate a new environmentally related products, and confusion concerning how to meet the expectation of the government regulation and regulators.

STUDY 4

The views expressed in both the marketing executives, government agencies, and environmental marketing journals are crucial to provide insight into issues explored in the questionnaire. The questionnaire was constructed by modifying these viewpoints to take the form of structured queries.

METHOD

Population: From information obtained from state and regional offices of the EPA, the list of industrial firms was originally reported to be 63. Names and addresses of participating firms, as well as the names of certain key officials, were obtained from the EPA. Of 42 firms contacted, 33 (about 79%) responded with complete questionnaires. A research design included contact among national, regional, and state EPAs, as well as company environmental officials; a pretest, and a mail survey. Of the 88 questionnaires returned, all but five responses were complete enough to be included in the analysis. An average of three responses was obtained from each of the 33 firms involved in the marketing centers. In each of these firms there were four to six potential respondents. Thus, a response rate of 52 percent was obtained.

HYPOTHESES

As mentioned earlier, the dependent variable in the present study is the innovation decision toward the EN-Products. The measures of three uncertainty dimensions of EPA regulation from the previous study are treated as separate predictor variables. The preceding discussion and studies provide the conceptual foundations for the four propositions examined in this research:

Proposition 1: The greater the uncertainty of the EPA policy as perceived by marketing executives, the less favorable the executives' decision to new EN-Product innovations.

Proposition 2: The lesser knowledge concerning outcomes, predictions and expectations of the EPA policy, the less positive decision to adopt new EN-Product innovations.

Proposition 3: The lesser ambiguity of the EPA regulation, the more positive decision to adopt new EN-Product innovations.

Proposition 4: The lesser amount of information for decision making, the less positive the decision to adopt new EN-Product innovations.

ANALYSIS

Separate factor analyses were performed on the items comprising the regulatory uncertainty scale to insure that the measures retained their construct validity. Factor analysis explains more than 70% of the variance. The dimension of uncertainty from regulation is represented by a set of three variables. Three factors which represent a different dimension of the regulatory uncertainty scale are Lack of Knowledge of Outcome (LACKNOW), Ambiguity of Regulation (AMBIREG), and Lack of Enough Information (LACKINF). The analysis produces a clean factor structure with items loading on the appropriate factors.

The results of a multiple regression of the factor scores for the three dimensions of regulatory uncertainty on the dependent variable indicate a strong support for the proposed relationship between uncertainty from lack of knowledge and the adoption attitude (beta of 0.62, and significant at 0.0001 level) as well as weak support for the relationship between uncertainty about information and the adoption attitude (beta of -0.30, and significant at 0.001 level). In addition, it also appears that the uncertainty about the ambiguity of regulation is not a determinant of the innovation decision. The standardized beta weight is small and statistically insignificant.

Thus, the Propositions met with mixed results. The overall regression equation explains about 50.0% of the total variance at the 0.001 statistically significant levels. In addition, the signs on the coefficients are in the proposed directions, except for the coefficient of the ambiguity dimension, which are insignificant. Therefore, this finding partially supports Proposition (1), (2), and (4).

CONCLUSIONS, IMPLICATIONS AND DIRECTIONS

The results of this study point to some interesting findings as well as important implications for public policy decision makers. The study compared uncertainty attributes. The attribute which has the greatest impact on the adoption attitude is "Lack of Knowledge Concerning Outcomes, Predictions and Expectations," followed by "Lack of Enough Information for Decision Making." The uncertainty attribute "Ambiguity of the Policy and Delayed Feedback" seems to have the least impact on the adoption attitude. The findings indicate that the EPA regulations impose substantial direct and indirect barriers to the decision toward the adoption of EN-Product innovations. There appears to be two main areas of impact through regulation on firms. The first is that when marketing executives perceive the EPA policy as unclear and ambiguous, this creates a significant unfavorable attitude toward innovation decision. The second is related to such uncertainties as lack of information, inability to predict the result from the decision to innovate EN-Products, and confusion concerning how to meet the expectation of the government regulation and regulators.

The findings have important practical implications. The trade-off between costs and benefits of each regulatory alternative could be emphasized for the best regulatory selection. The market perceptions about the comparative benefits and costs of different EPA regulations, both regulatory alternatives and the traditional command-and-control regulation, could present an opportunity favoring the innovation decision. The relative advantages of a specific EPA regulation compared to other regulatory alternatives could be highlighted.

Additionally, if government regulatory agencies are involved in the information-yielding process, they, too, can enhance business's ability to make the more rational choice (as in the EPA regulatory concept). The more rational choice is one that reduces non-monetary as well as monetary costs. These include frustrations, unpredictable results from proposal submission, government paperwork and bureaucracy, and needless delay. When such information exchange exists, business executives would realize a reduction in uncertainty. Market transaction can be made more efficient when both the business firms (or buyers) and government agencies (or sellers) understand the full costs and benefits associated with the transactions they undertake.

Moreover, the diffusion theory postulates that adoption of a new innovation requires continuous guidance to industrial organizations during the innovation adoption process. This two-step process consists of an initiation phase and an implementation phase. In the marketing of the business and government alternative for the EPA policy, the diffusion theory would involve a firm's awareness of the regulatory alternatives and its advantages, as well as the development of favorable executives' attitudes toward the innovation. This would be followed by the firm's search for necessary regulation information prior to the submission of a policy proposal for EN-Product innovations to the USEPA. The success of the process is partly influenced by the individual member's regulation-information processing activities and the amount of, as well as the quality of, regulation information available. As the present study indicates, marketing executives generally recognize the conflicts, confusions, and inconsistencies between and/or lack of practicality in federal and state laws concerning environmental standards and measures. These uncertainties cause both inaction in the adoption of new innovative products and increased operating and investment risks. Therefore, industry must request clear information from the government agencies about comparative costs and benefits, risks and social effects, and other data connected with the command-control and voluntary EPA alternative options. When the agency provides such material, and industry then utilizes it for deciding the advantages and disadvantages of EPA policy choice, industry would then become convinced that the voluntary alternative is the attractive option. Industry could then communicate much better with the agency, and the result would be the development of mutual credibility and trust.

Generally, then, the value of the information-exchange process may be measured by the capacity of the regulation-related information to reduce perceived uncertainty among firms. The credibility and the reputation of the government agency and its representatives may grow if the information-exchange takes place. The problems of government intervention in private enterprise could be alleviated by encouraging information flow among federal, state, and local government agencies and business. Representatives of the government agencies and marketing executives would realize that open channels of communication and mutual trust and understanding between them are essential and possible. Without such understanding and trust, innovative public policies for adoption decision are less likely to be responsive to the real issues at hand and would be exceedingly difficult to implement.

An incidental uncertainty (or cost) and benefit of this entire process of adopting the innovative products, therefore, would be the encouragement of continuing and productive interchange between representatives of government agencies and business in order to arrive at mutually reinforcing policies and actions.

Finally, at a minimum, the findings imply that marketing executives must be flexible to cope with varying levels of complexity and uncertainty associated with regulatory choices. The rapid development of regulatory alternatives left a communication void not only between government regulators and business firms (Apaiwongse 2007), but also among marketing executives as well. All marketing executives could then be made aware of the critical need to accept voluntary alternatives. In order to improve attitudes toward and awareness of the new innovation decision, formal communication networks between business and government as well as among marketing members must be flexible and informal as well as open networks of communication must be established which encourage coordination as a team unit among other things.

Some major directions can be suggested for future research. These directions consist of theoretical and methodological considerations. First, this research suggests that interpersonal relationships within marketing decision units should have a strong impact on innovation decision. The nature and frequency of interpersonal communications among marketing members is the main variable which may be explored in future research. The study of interpersonal communication networks would require more rigorous methods of data collection than mail surveys. Additionally, the communication networks among company executives and government regulators should be emphasized in future studies, especially in the area of negotiating marketing. The interaction processes between the two dyads may then be considered.

Second, the company process of adopting government regulations should be viewed as a complex process rather than as a dichotomous state. Longitudinal surveys of different regulatory alternatives would allow valid analysis of causal processes involved in measuring adoption attitude and innovation decision. As indicated earlier, the innovation process may be focused on the studies of initiation and implementation stages. In addition to cross-time design, there is a need, as well, for replication studies to verify the results of this research. Such longitudinal data are needed to establish cause and effect relationships. Third, to ascertain the relative validity of the several constructs and to strengthen the degree of generalization of the research results, the multitrait - multimethod analysis can then be employed. This analysis can offer new constructs and new operationalizations that may shed light on regulatory uncertainty and acceptance attitude towards innovation. For instance, the results of objective and protocol measures could be compared with the results of subjective (or perceptual) measurers.

Fourth, the presence of several interaction effects suggests that the desired hypotheses may need to be modified by bringing in other relevant variables directly into the formulations. For example, demographic variables attached to marketing decision units (such as size, type of industry, and so on) may improve the ability to predict speed of innovation process. Fifth, predictive ability could be improved using multiple responses within a marketing executives as well as marketing decision units together with consensus measures. However, the complexity of the consensus measures and the current stage of scientific knowledge do not allow any single investigation to suffice. Sixth, empirical research regarding the impact of regulation on innovation requires the evaluation of many factors, and is difficult to observe because the innovation process involves many parties and takes time. For instance, not only government agencies and businesses, but also the third parties (i.e., public citizen groups, environmental groups, local citizens in the affected communities, etc.) are relevant actors in this process. Then, one future research direction may focus on the answer to the question of what role, if any, should third parties or the public play to influence this innovation process for EN-Products.

Finally, there is a need, in general, for more research into different uncertainty types of government regulations (or regulatory alternatives) in order to increase understanding of the specific factors that affect attitudes toward new innovations. Research findings could be compared and constructed to determine the amount and type of generalization across government regulations and alternatives that are justified.

REFERENCES

Apaiwongse, Tom Suraphol (2013), "Voluntary Reaction to Green Policies Among Market Mavens: An Application of the Parallel Political Marketplace Conceptualization" in **Environmental Marketing: Strategies, Practice, Theory, and Research,** Alma Mintu and Michael Polonsky eds., Haworth Press, PP: 307-339.

Apaiwongse, Tom Suraphol (2011), "Green Policy Uncertainty upon Global Market Adaptation between Business and Government," **Journal of Business and Behavioral Sciences**, Vol. 25, No. 1, 5-17.

Apaiwongse, Tom Suraphol (2009), "Ecological Market Exchange for Current Green Policy: The Parallel-Political Marketplace Approach," **Journal of Business and Behavioral Sciences**, Vol. 21, No. 1, 5-17.

Apaiwongse, Tom Suraphol (2008), "Barriers of EPA Voluntary Policy upon Market Innovations: An Exploratory Qualitative Approach," **Journal of Business and Behavioral Sciences**, Vol. 19, No. 2, 11-21.

Apaiwongse, Tom Suraphol (2007), "Environmental Policy and Market Adaptation," **Journal of Business and Behavioral Sciences,** Vol. 15, No. 1, (Spring), 23-36.

Duncan, Robert B. (1972), "Characteristics of Organizational Environments and Perceived Environmental Uncertainty," **Administrative Science Quarterly**, 17, 313-327.

Galbraith, J.R. (1973), **Designing Complex Organizations.** Reading, MA: Addison Wesley Publishing Company.

Lawrence, P. and J. Lorsch (1967), **Organization and Environment, Managing Differentiating and Integration** (Boston, MA: Harvard University, Division of Research, and Graduate School of Business Administration).

Journal of Business and Behavioral Sciences
Vol 28, No. 2; Fall 2016

VALUE CO-CREATION IN RETAILING IN THE U.S. AND JAPAN: A CROSS-CULTURAL COMPARISON

Miwa Y. Merz
Michael A. Merz
Kenneth C. Gehrt
San José State University
Ikuo Takahashi
Keio University

ABSTRACT: This research examines value co-creation in retailing in the U.S. and Japan. Specifically, it examines the extent to which different levels of customer participation (i.e., firm-creation, joint-creation, customer-creation) are valued in both cultures in the pre-purchase and post-purchases phases. The results of an experimental study show that customers in the U.S. and Japan are most favorable toward joint-creation in the pre-purchase phase but toward firm-creation in the post-purchase phase. These effects are significantly more pronounced for U.S. customers. Unexpectedly, in the pre-purchase phase Japanese (versus U.S.) customers are more favorable toward customer-creation whereas U.S. (versus Japanese) customers are more favorable toward firm-creation.

Key Words: Value co-creation, customer participation, service-dominant logic, retailing, value dimensions, cross-cultural comparison

INTRODUCTION

WeBook provides aspiring writers a platform to easily publish their books with the help of other writers, editors, and reviewers. Nike+ constitutes a platform that brings together millions of runners helping Nike co-design Nike shoes. Through MyStarbucksIdea, Starbucks takes advantage of crowd creativity to provide customers with the opportunity to share their ideas and vote on ideas. Finally, Apple's iPhone has been successful, largely because of the many applications available that were co-developed with outside programmers.

These examples illustrate the value of integrating customers into the value creation process. In fact, one of the key foundational premises of the evolving service-dominant (S-D) logic in marketing is that value is always co-created between the firm and its customers (Vargo and Lusch, 2004). S-D logic is philosophically grounded in a commitment to collaborative processes between firms and their customers, partners, and employees (Lusch, Vargo, and O'Brien, 2007). *Co-creation* is considered an important manifestation of customer engagement behavior (van Doorn, 2010; Hoyer, Chandy, Dorotic, Krafft, and Singh, 2010).

Despite its importance, research on co-creation in consumer settings is limited (Hoyer et al., 2010). Most of the existing work on consumer co-creation has been conducted in the context of *business-to-business markets* (Baumann, Le Meunier-FitzHugh, 2015; Chan, Yim, and Lam, 2010; Hoyer et al., 2010; Spann, Ernst, Skiera, and Soll, 2009). In addition, most of the existing work on consumer co-creation has focused on two levels of customer participation (i.e., firm-creation, joint-creation), neglecting the possibility of the customer doing most of the work (i.e., customer-creation; Bendapudi and Leone, 2003). Furthermore, prior research has focused on examining co-creation in pre-purchase service encounters, but has neglected the examination of co-creation in post-purchase service encounters (Edvardsson, Enquist, and Johnston, 2005). Moreover, most of the existing work on consumer co-creation has been conducted in the U.S. with little knowledge about its applicability in other cultures and potential cross-cultural differences (Chan, et al., 2010; Merz and Takahashi, 2011). To fill these gaps, this study aims to advance existing research by examining value co-creation (1) in the *business-to-consumer context*, (2) across three levels of co-creation (i.e., firm-creation, joint-creation, customer-creation), (3) in pre-purchase and post-purchase service encounters, and (4) in a cross-cultural context (i.e., U.S. and Japan).

Specifically, the purpose of this research is threefold. First, it examines whether joint-creation (versus firm-creation or customer-creation; Bendapudi and Leone, 2003) affects customers' perceived service quality and satisfaction with their retail experience. Lusch, Vargo, and O'Brien (2007) have pointed out the importance for retailers to use joint-creation as a competitive resource. Second, this research examines the effects of different levels of co-creation on customers' perceived service quality and satisfaction with their shopping experience in a pre-purchase and post-purchase retail service encounter. Edvardsson et al. (2005) have pointed out the importance of not only examining co-creation in the pre-purchase service encounter, but also the post-purchase service encounter. Finally, this research examines in two very different national cultures – the U.S. and Japan – whether cross-cultural differences regarding customers' preferred level of co-creation exist.

THE CONCEPT OF CO-CREATION IN RETAILING

Marketers are increasingly realizing the importance of learning from and collaborating with customers to create value that meets their individual and dynamic needs (Chan et al., 2010; Prahalad and Ramaswamy, 2000). Involving customers in the value creation process is discussed as customer participation in the marketing literature (e.g., Bendapudi and Leone, 2003). Value co-creation is the main premise of customer participation (Vargo and Lusch, 2004). Encouraging customer participation in the service delivery process benefits both firms and customers. It helps firms increase customer satisfaction and productivity (Chan et al., 2010; Mills and Morris, 1986). It helps customers have a better service experience, which includes improved service quality, more customization, and better service control (Xie, Bagozzi, and Troye, 2008).

However, Chan et al. (2010) point out that customer participation might not necessarily create positive value. For example, if customers' increased participation shifts power from service employees (i.e., the firm) to customers, it might result in an increased workload for customers (Gehrt, O'Brien, and Sakano, 2009). As the beneficiary always determines the value (Vargo and Lusch, 2004), this increased workload might not always be perceived as favorable.

This suggests that customers might have different preferences with regard to their desired level of participation (Bendapudi and Leone, 2003). Therefore, the concept of customer participation is one that employs many forms and degrees, ranging from firm-creation to joint-creation to customer creation (Bendapudi and Leone, 2003; Meuter and Bitner, 1998). In line with prior research, therefore, we define customer participation as a behavioral construct that measures the extent to which customers provide or share information, make suggestions, and become involved in decision-making during the service creation and delivery processes (Auh, Bell, McLeod, and Shih, 2007; Bolton and Saxena-Iyer, 2009; Chan et al., 2010; Hsieh, Yen, and Chin, 2004). Dabholkar (1990, p. 484) defines customer participation similarly as "the degree to which the customer is involved in producing and delivering the service" (see also Bendapudi and Leone, 2003).

Given this definition of customer participation, we adopt Meuter and Bitner's (1990) model of co-creation to distinguish between different levels of co-creation, in line with previous research (e.g., Bendapudi and Leone, 2003). The authors distinguish between three types of co-creation: firm creation, joint creation, and customer creation. *Firm creation* is a situation in which the service is provided predominantly by the firm and its employees, without much participation by the customer. *Joint-creation* is a situation in which the customer and the firm's contact employees interact and participate, in similar measure, in the service encounter. *Customer creation* is a situation in which the service is provided predominantly by the customer, without much participation by the firm or its employees. As a result, the level of customer participation is highest in a customer creation condition, whereas it is the lowest in a firm creation condition.

The existing research on co-creation in the context of retailing includes conceptual or empirical work. Regarding the existing conceptual work, Arnould (2005) substantiates the importance of co-creation in retailing by arguing that a Consumer Cultural Theory (CCT)-based approach to retailing is important as it accounts for the concept of co-creation. In the retail context, CCT provides a basis to explain how consumers deploy their own cultural resources, aided by retailer-provided resources. Verhoef et al. (2009) in their conceptual model of customer experience creation in retailing identified co-creation as one of the key service interfaces. Regarding existing empirical work, Grissemann and Stokburger-Sauer (2012) investigated the degree of co-creation in the tourism industry in Austria. The authors find that co-creation positively affects customer satisfaction with the service company (i.e., travel agency), customer loyalty, and service expenditures. Furthermore, Cheung and To's (2011) experimental research suggests a positive relationship between customer participation and perceived service performance. Auh, Bell, McLeod, and Shih (2007) examined co-creation in the financial services and medical services industries. The authors find that co-creation directly affects attitudinal loyalty and indirectly behavioral loyalty. Baumann and Le Meunier-FitzHugh (2015) examined value co-creation between sales personnel and end-customers. Their exploratory interviews suggest that co-creation leads to episode value in discrete transactions, and to mutual episode and relationship value in relational exchanges.

Finally, Andreu, Sánchez, and Mele (2010) examined value co-creation among retailers and consumers. On the basis of in-depth interviews and observations of retailers and customers, the authors propose a framework of value co-creation that differentiates between three types of co-creation processes: customer value-creating processes, supplier value-creating processes, and encounter processes. The authors also identify different phases of value co-creation: before visiting the store, in the store, preparation, delivery-installation, and follow-up. The latter two are particularly interesting as they suggest that value co-creation is not only important during the service encounter (i.e., prior to the purchase), but also after the actual encounter (i.e., post purchase). Given this, we distinguish in our examination of co-creation between two phases of customer co-creation: the pre-purchase phase and the post-purchase phase. Table 1 provides an overview of our conceptual framework.

Table 1: Conceptual Co-Creation Framework in Retailing

	Firm-Creation	**Joint-Creation**	**Customer-Creation**
Pre-Purchase Phase	Service contact employees participate predominantly in the decision-making process	Customers and service contact employees participate equally in the decision-making process	Customers participate predominantly in the decision-making process
Post-Purchase Phase	Service contact employees participate predominantly in the order fulfillment process	Customers and service contact employees participate equally in the order fulfillment process	Customers participate predominantly in the order fulfillment process

CONCEPTUAL DEVELOPMENT

Cultural Value Dimensions: The question arises whether there exist cultural differences with regard to customers' perceived service quality and their shopping experience satisfaction across the different levels of co-creation and different co-creation phases. Prior research has argued that cultures differ in terms of various value dimensions (Hofstede, 2012; Schwartz, 1992). According to Hofstede (2012), national cultures differ along five main dimensions: power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, and long-term versus short-term orientation. The former three appear particularly relevant to our framework. Therefore, we will focus on those three in the following.

Hofstede's (2012) *individualism versus collectivism* distinction describes the degree to which a culture values independence, freedom, and personal and individual time. Thus, it focuses on the extent to which culture reinforces individual or collective achievement and interpersonal relationships. Customers in individualistic cultures stress self-esteem, self-centeredness, independence from others, personal goals, and self-expression. Their focus is their own self-image ("I") and the forming of a larger number of looser relationships. In contrast, customers in collectivistic cultures are more group-focused and interdependent. Their focus is their in-group's self-image ("We") and their saving face in front of their in-groups. Conformity to group norms and values are highly valued in collectivistic cultures. The U.S. ranks relatively high on individualism (91 of 100), whereas Japan ranks relatively low on individualism (or high on collectivism; 46 of 100).

Power distance describes the extent of cultural inequality that underlies the functioning of each particular society. Its focus is on how a culture shapes its individuals and groups with regard to authority, prestige, power, status, wealth, and material possessions (Javidan and House, 2001). Cultures high in power distance emphasize social status, class affiliation (rich versus poor), wealth, and hierarchic values. Customers in such cultures are motivated to maintain, increase, and show their "power" or social status as a source of satisfaction (Roth, 1995). Members of high power distance cultures tend to be reluctant to trust one another. In contrast, cultures low in power distance do not place an emphasis on hierarchic values thereby believing in equal rights for all members of the society. They de-emphasize the differences between their citizens' wealth and power and embrace consultation, participation, cooperation, and practicality (Hoftede, 2012). The U.S. ranks relatively low on power distance (40 of 100), whereas Japan ranks relatively high on power distance (54 of 100).

Uncertainty avoidance describes the extent to which members of a culture feel either comfortable or uncomfortable in unstructured, ambiguous, and uncertain situations (Hofstede, 2012). Its focus is on the extent to which cultures value novelty, surprise, predictability, stability, and low stress rather than change. Javidan and House (2001) discuss that uncertainty avoidance is the extent to which people see orderliness, consistency, structure, and laws. Customers in high uncertainty-avoidance cultures prefer consistency, structured lifestyles, certainty, and clearly articulated expectations. They feel threatened by ambiguous situations and are less tolerant of ambiguity, more risk averse, and more resistant to change (Roth, 1995). In contrast, customers in low uncertainty-avoidance cultures are more tolerant of a variety of opinions and less concerned about ambiguity and uncertainty. They take more and greater risks and more readily accept change. They accept new ideas, develop innovative products, and are willing to take risks to try something new or different (Hosftede, 2012). The U.S. ranks relatively low on uncertainty avoidance (46 of 100), whereas Japan ranks relatively high on uncertainty avoidance (92 of 100).

Level of Co-Creation: The evolving S-D logic suggests that value creation is always a dynamic and interactive process between service recipients and service providers (Vargo and Lusch, 2004). This suggests for all customers irrespective of cultural background that joint-creation is likely to result in more favorable outcomes than firm-creation or customer-creation where such dynamic and interactive processes between service recipient and provider are limited. In support of this view, Bendapudi and Leone (2003) found in the context of retailing that participating customers are more satisfied than non-participating customers. Similarly, Ennew and Binks (1999) found that customer participation in the context of retailing positively affects service quality and satisfaction.

To the extent that customer participation is embodied most obviously in joint-creation, we expect that joint-creation in retailing will positively affect U.S. and Japanese customers' perceived service quality and satisfaction with their shopping experience in both purchase phases. Because prior research has shown a positive relationship between perceived service quality and satisfaction with the shopping experience, we hypothesize the following relationship for both U.S. and Japanese customers:

> **H1:** Customers will (a) have a higher perceived service quality and (b) be more satisfied with their shopping experience across both purchase phases when the context is joint-creation versus firm-creation or customer-creation.

Cultural Differences Across Levels of Co-Creation During the Pre-Purchase Phase: As mentioned, service provider and recipient are likely to engage in more dynamic and interactive processes during the pre-purchase (versus post-purchase) phase. Because a service encounter constitutes a social exchange, it is likely that culture influences customers' and service employees' norms, roles, and expectations (Chan et al., 2010; Patterson, Cowley, and Prasongsukarn, 2006) in particular during the pre-purchase phase. Therefore, it is likely that culture moderates the relationship between level of co-creation and outcome variables. This seems likely especially during the pre-purchase phase where a more intensive interaction between service provider and customer is possible. We will look at each type of co-creation separately in the following.

The first type of co-creation is *firm-creation.* In the pre-purchase phase it is likely that Japanese (versus U.S.) customers feel more comfortable with a shopping experience where the customer is less involved in the decision-making process and where the recommendations come from the firm (i.e., the service employee). To illustrate, the more collectivist Japanese customers try to fit in with their in-group. Thus, they might desire to have the advice of a service employee to help them avoid negative consequences associated with the selection of a "wrong" product. They are high in uncertainty avoidance. Having a knowledgeable salesperson help them make their decision might be beneficial. Firm creation might effectively reduce their perceived uncertainty and fear that their selection does not conform to group norms and values that are highly valued in their in-group. In contrast, the more individualist U.S. customers might try to at least partially get involved in the pre-purchase decision-making process as part of their desire for self-expression. As a result, we hypothesize the following interaction between level of co-creation and culture:

> **H2:** In the pre-purchase phase, Japanese customers will (a) have a higher perceived service quality and (b) be more satisfied with their

shopping experience than U.S. customers when the context is firm-creation.

The second type of co-creation is *joint-creation*. Members of cultures low in power distance do not place an emphasis on hierarchies and embrace consultation, participation, cooperation, and practicality (Hofstede, 2012). Consequently, the lower power distance customers in the U.S. are likely to embrace a joint-creation context. Furthermore, U.S. customers might feel accepting and even be looking for new insights and a diversity of opinions prior to making a decision. For U.S. customers, variety is good and they are tolerant toward other people's opinions. In contrast, the higher power distance customers in Japan are likely to place less emphasis on consultation, participation, and cooperation (Hofstede, 2012). It is likely, therefore, that U.S. customers are more favorable toward joint-creation than their Japanese counterparts. Hence, we hypothesize the following interaction between level of co-creation and culture:

H3: In the pre-purchase phase, U.S. customers will (a) have a higher perceived service quality and (b) be more satisfied with their shopping experience than Japanese customers when the context is joint-creation.

The third type of co-creation is *customer-creation*. The more individualist U.S. customers might feel more comfortable than their Japanese counterparts in service encounters where they have to do all of the work (i.e., customer creation). This is because they like to make independent decisions. They take such situations as opportunities for enhancing their self-esteem and finding something that matches their personal goals. Their lower level of uncertainty avoidance suggests that they are likely to feel more comfortable with the many options they might face than their Japanese counterparts. In addition, they are likely to be more willing to take risk in their product selection when not being helped and without a third opinion. As a result, we hypothesize the following interaction between level of co-creation and culture:

H4: In the pre-purchase phase, U.S. customers will (a) have a higher perceived service quality and (b) be more satisfied with their shopping experience than Japanese customers when the context is customer creation.

Pre-Purchase Phase versus Post-Purchase Phase: In addition to culture, we expect the purchase phase to moderate the relationship between level of co-creation and the dependent variables. Previous research has examined joint-creation predominantly in the pre-purchase phase. While the S-D logic and the above-cited literature suggest that joint-creation is important, we hypothesize that it is less important after the purchase transaction has been made than prior to the purchase transaction. Due to the nature of order fulfillment, customers might not feel that they need to get involved to the same extent in the post-purchase phase as they do during the selection and decision-making processes in the pre-purchase phase (Edvardsson et al., 2005). Therefore, we predict for the joint-creation condition and for both cultures that customers will have higher service quality perceptions and be more satisfied with the shopping experience in the pre-purchase versus post-purchase condition. Consequently, we expect the following purchase phase type main effects for both U.S. and Japanese customers:

> **H5:** In the joint-creation condition, customers will (a) have a higher perceived service quality and (b) be more satisfied with their shopping experience in the pre-purchase phase than the post-purchase phase.

METHOD

295 (U.S.) and 300 (Japan) undergraduate students participated in this study in exchange for course credit. We assigned subjects randomly to one of three pre-purchase scenarios and one of three post-purchase scenarios that were each followed by the dependent and independent measures. The scenarios varied in level of co-creation (firm-creation, joint-creation, customer-creation).

The stimuli were refined through a pre-test involving 62 U.S. subjects who were randomly assigned to one of the three pre-purchase and post-purchase conditions. The ultimate stimuli were scenarios about a person visiting a department store to purchase a bookshelf. The pre-purchase scenarios described the purchase experience prior to the financial transaction and in the department store. The post-purchase scenarios described the experience after the financial transaction in both the department store and the person's house. We generated a total of three pre-purchase and three post-purchase scenarios varying the extent to which co-creation between the customer and the department store personnel takes place. The scenarios were back-translated by two native speakers for the Japanese sample (Brislin, 1980).

Subjects filled out an online survey at their own pace. The survey contained three parts. First, we randomly assigned subjects to the pre-purchase scenario and asked them to complete several measures related to the scenario (in this order): satisfaction with shopping experience, service quality, co-creation manipulation check, scenario and claim believability measures, and scenario clarification measures. Second, we randomly assigned subjects to the post-purchase scenario and asked them to complete these same measures related to the second scenario. Finally, we asked subjects for their demographics as well as their interest in and importance of shopping in general. We also assessed subjects' perceptions of the study's purpose. The questionnaire took between 10-15 minutes to complete.

We used existing scales for all measures. Customers' satisfaction with the shopping experience was measured by averaging four seven-point scales (very dissatisfied/very satisfied, very pleased/very displeased, frustrated/contented, delighted/terrible; $\alpha > .82$ for both the U.S. and Japanese pre-purchase and post-purchase measures), in line with Bendapudi and Leone (2003) and Spreng, MacKenzie, and Olshavsky (1996). We measured customers' perceived service quality by averaging three seven-point scales (extremely poor/extremely good, awful/excellent, very low/very high; $\alpha > .94$ for both the

U.S. and Japanese pre-purchase and post-purchase measures), based on Spreng and Mackroy (1996).

Subjects indicated the perceived level of participation during and after the shopping experience on two seven-point scales (low work/high work, low effort/high effort; $\alpha > .85$, for both the U.S. and Japanese pre-purchase and post-purchase measures), which were taken from Bendapudi and Leone (2003). This index constituted the manipulation check for type of co-creation (firm-creation, joint-creation, customer-creation). Subjects furthermore indicated on five seven-point scales (not believable/believable, not credible/credible, not relevant/relevant, realistic/not realistic, could happen/could not happen) how believable, relevant, and credible the scenarios were (Bendapudi and Leone, 2003; Kent and Allen, 1994; Jain and Posavac, 2004). We averaged the five scales to form one index of claim believability ($\alpha > .82$, for both the U.S. and Japanese pre-purchase and post-purchase measures). This measure ensured that subjects perceived the scenarios as believable, relevant, and credible. Additional measures assessed subjects' understanding of the scenarios and their awareness of the pre-purchase versus post-purchase situations, interest in and importance of shopping in general, demographics, and demand characteristics (Yi, 1990).

RESULTS

Sample Information and Manipulation Checks: Of the completed and submitted online questionnaires, we obtained 284 and 295 usable questionnaires for data analysis for the U.S. and Japan, respectively. 46% of the U.S. sample and 69% of the Japanese sample were male. 84% of the U.S. sample and 96% of the Japanese sample were single. The majority of the U.S. sample was Asian (48%), White (22%), or Hispanic (10%), whereas the majority of the Japanese sample was Asian (96%). The average age of the U.S. subjects was 24 and of the Japanese subjects 19. Subjects in the U.S. and Japan were generally interested in shopping (M= 5.5 in both countries) and indicated that shopping in general is important to them (M= 5.5. and 5.6, respectively).

To ensure the effectiveness of the co-creation manipulations, we submitted subjects' responses to a 3 (co-creation type) x 2 (Country) between-subjects ANOVA, separately for the pre-purchase and post-purchase scenarios. The perceived level of co-creation was the dependent variable. The analysis yielded a main effect for the co-creation factors (pre-purchase scenario: $F(2, 553) = 26.2$, $p < .05$; post-purchase scenario: $F(2, 554) = 125.3$, $p < .05$). Pairwise comparisons for the *pre-purchase* scenarios – using Tukey's post-hoc test – revealed that firm-creation was perceived as significantly less co-creative than joint-creation (U.S.: M= 4.4 vs. 4.8, $p<.056$; Japan: M= 3.6 vs. 4.1, $p< .05$). In addition, firm-creation was perceived as significantly less co-creative than customer-creation (U.S.: M=4.4 vs. 5.5; Japan: M=3.6 vs. 4.6; all $p< .05$). Furthermore, subjects perceived the joint-creation condition as significantly less co-creative than the customer-creation condition (U.S.: M=4.8 vs. 5.5; Japan: M=4.1 vs. 4.6; all $p< .05$).

Turkey's post-hoc pairwise comparisons for the *post-purchase* scenarios revealed that the firm-creation condition was perceived as significantly less co-creative than joint-creation (U.S.: M= 3.6 vs. 5.6; Japan: M= 2.7 vs. 5.1; all p< .05) and customer-creation (U.S.: M= 3.6 vs. 6.0; Japan: M= 2.7 vs. 5.2; all p< .05). Furthermore, subjects perceived the joint-creation condition as less co-creative than the customer-creation condition (U.S.: M= 5.6 vs. 6.0, p < .098; Japan: M= 5.1 vs. 5.2, p>.05). While this was not significant for the Japanese sample at the 0.1 significance level, it was directionally consistent with expectations. As a result, the co-creation conditions varied as intended and indicated that the co-creation types in our study were appropriate for hypothesis testing.

In addition, subjects perceived the scenarios as believable, realistic, and credible (U.S.: all M > 5.4; Japan: all M > 4.6). Finally, our scenarios were well understood such that subjects remembered the content of the scenarios, the general level of customer-sales person interaction, and whether the scenario focused on the pre-purchase versus post-purchase phase thereby providing further support for the reliability of our scenario co-creation manipulations.

Main Effect Results: H1 predicts for both the pre-purchase and post-purchase phases that U.S. and Japanese customers will prefer a shopping experience that is characterized by joint-creation rather than firm-creation or customer-creation. To test H1, we first submitted U.S. subjects' responses to simple between-subjects MANOVAs, separately for the pre-purchase and post-purchase scenarios. Customers' perceived service quality and their satisfaction with the shopping experience were the dependent variables. For the *pre-purchase* scenarios, as expected, the analysis revealed a main effect for co-creation for the service quality (F(2, 248)=129.3, p<.05) and the shopping experience satisfaction (F(2, 248)=21.9, p<.05) dependent variables. To further examine these main effects, we used independent-samples t-tests. Analyses revealed that the joint-creation condition led to significantly more favorable service quality perceptions (M=5.9 vs. 3.2, t(158)=12.9, p<.05) and shopping experience satisfaction (M=5.7 vs. 4.7, t(154)=5.1, p<.05) than the customer-creation condition. However, the differences were not significant for the joint-creation versus the firm-creation conditions. For the *post-purchase scenarios*, as expected, the analysis similarly revealed a main effect for co-creation type for the service quality (F(2, 256)=54.6, p<.05) and shopping experience satisfaction (F(2,256)=42.9, p<.05) dependent variables. Post hoc mean comparisons revealed that the joint-creation condition led to significantly more favorable service quality perceptions (M=4.5 vs. 3.8, t(171)=2.6, p<.05) and shopping experience satisfaction (M=4.8 vs. 4.2, t(171)=2.4, p<.05) than the customer-creation condition. However, opposite of what was expected, joint-creation led to significantly less favorable service quality perceptions (M=4.5 vs. 6.1, t(168)=-8.0, p<.05) and shopping experience satisfaction (M=4.8 vs. 6.2, t(168)=-7.0, p<.05) than the firm-creation condition. Overall, H1 was only partially supported for the U.S. sample.

Similarly, we submitted Japanese subjects' responses to simple between-subjects MANOVAs, separately for the pre-purchase and post-purchase scenarios. Customers' perceived service quality and their satisfaction with the service experience were the dependent variables. For the *pre-purchase* scenarios, as expected, the analysis revealed a main effect for co-creation for the service quality (F(2, 283)=38.3, p<.05) and the shopping experience satisfaction (F(2, 283)=3.9, p<.05) dependent variables. To further examine these main effects, we used independent-samples t-tests. Analyses revealed that the co-creation condition led to more favorable service quality perceptions (M=5.0 vs. 3.9, t(192)=7.2, p<.05) and shopping experience satisfaction (M=4.9 vs. 4.7, t(197)=1.1, p>.05) than the customer-creation condition. However, the differences were only significant for the perceived service quality dependent variable. Furthermore, there were no significant differences for the co-creation versus the firm-creation conditions. For the *post-purchase scenarios*, as expected, the analysis similarly revealed a main effect for co-creation for the service quality (F(2, 281)=79.9, p<.05) and shopping experience satisfaction (F(2,281)=54.5, p<.05) dependent variables. Post hoc mean comparisons revealed that the co-creation condition led to significantly more favorable service quality perceptions (M=4.0 vs. 3.4, t(194)=3.4, p<.05) and shopping experience satisfaction (M=4.4 vs. 3.6, t(192)=4.2, p<.05) than the customer-creation condition. However, opposite of what was expected, co-creation led to significantly less favorable evaluations of the perceived service quality (M=4.0 vs. 5.5, t(187)=-9.1, p<.05) and the shopping experience (M=4.4 vs. 5.5, t(185)=-6.3, p<.05) than the firm-creation condition. Overall, H1 was partially supported for the Japanese sample.

Interaction Results: H2-H4 predict interaction effects between level of co-creation and culture. Therefore, we submitted subjects' responses to a 3 (level of co-creation) x 2 (country) between-subjects MANOVAs with service quality and customer satisfaction as the dependent variables. As expected, there was a significant interaction effect for both the service quality (F(2, 531)=6.9, p<.05) and customer satisfaction (F(2, 531)=26.2, p<.05) dependent variables. To further examine these interaction effects and test H2-H4, we conducted independent-samples t-tests.

Specifically, H2 predicts for the pre-purchase phase that Japanese versus U.S. customers will be more favorable toward firm-creation. Independent-samples t-tests revealed—opposite of what was expected—that U.S. customers had higher service quality perceptions (M=5.8 vs. 5.1, t(195)=5.2, p<.05) and were more satisfied with the shopping experience (M=5.8 vs. 5.1, t(190)=4.2, p<.05) than Japanese customers when exposed to the firm-creation condition. As a result, H2 was not supported and was opposite of what was expected.

H3 predicts for the pre-purchase phase that U.S. versus Japanese customers will be more favorable toward joint creation. Independent-samples t-tests revealed that U.S. customers showed significantly higher service quality perceptions (M=5.9 vs. 5.0, t(162)=5.6, p<.05) and were significantly more satisfied with their shopping experience (M=5.7 vs. 4.9, t(163)=5.3, p<.05) than Japanese customers. Hence, H3 was supported.

H4 predicts for the pre-purchase phase that U.S. versus Japanese customers will be more favorable toward customer creation. The results showed that U.S. customers had less favorable service quality perceptions than Japanese customers when the context was customer creation ($M=3.2$ vs. 3.9, $t(188)=-3.6$, $p<.05$). This was opposite from what was expected. There was no significant difference for the satisfaction measure. Overall, therefore, H4 was not supported.

H5 predicts that customers from both cultures value joint-creation more in the pre-purchase phase versus post-purchase phase. To test this hypothesis, we submitted subjects' responses to the dependent variables to paired-samples t-tests, separately for both cultures. For this analysis, we focused on subjects who were exposed to the co-creation condition in both the pre-purchase and post-purchase treatment conditions to avoid potential confounds. The results for the U.S. sample revealed that subjects had significantly higher service quality perceptions ($M=6.3$ vs. 5.3, $t(17)=2.3$, $p<.05$) and were significantly more satisfied with their shopping experience ($M=6.1$ vs. 5.3, $t(19)=3.2$, $p<.05$) in the pre-purchase versus the post-purchase phase when the scenario was joint-creation. The results for the Japanese sample revealed that subjects had significantly higher service quality perceptions ($M=5.0$ vs. 3.9, $t(30)=4.4$, $p<.05$) and were significantly more satisfied with their shopping experience ($M=5.0$ vs. 4.3, $t(32)=2.8$, $p<.05$) in the pre-purchase versus post-purchase phase when the scenario was joint-creation. Overall, H5 was supported.

DISCUSSION

This research investigated the effect of co-creation on U.S. and Japanese customers' service quality perceptions and their satisfaction with the shopping experience. It did so in the context of retailing and for two purchase phases, the pre-purchase phase and the post-purchase phase.

We found for the pre-purchase phase that U.S. and Japanese customers had higher service quality perceptions and were more satisfied with their shopping experience when joint-creation took place than when they had to do the majority of work (i.e., customer-creation). This finding is in line with the premises of the service-dominant logic and the extant literature highlighting the importance of co-creation.

In the post-purchase phase, however, customers from both cultures preferred it when the firm created most of the value for them (i.e., firm-creation) rather than when they participated in the order fulfillment creation process (i.e., joint-creation). This finding was unexpected. It might suggest that joint-creation is more important in some of the parts of a multi-phase process that can potentially include pre-purchase, purchase, consumption, and disposal phases. In retailing, it might be that joint-creation particularly matters during the pre-purchase/search phase but not so much during the consumption/service-provision phase.

Our research also suggests that there exist significant cross-cultural differences with regard to co-creation. Specifically, we found that U.S. (versus Japanese) customers had higher service quality perceptions and were more satisfied with their shopping experience when joint-creation took place during the pre-purchase and post-purchase phases. Furthermore, unexpectedly, we found for the pre-purchase scenarios that Japanese (versus U.S.) customers were more favorable toward the customer-creation condition, whereas U.S. (versus Japanese) customers were more favorable toward the firm-creation condition.

Our findings suggest that retail managers across cultures might benefit from implementing a servicescape that invites customers to participate in the service provision co-creation process. This could for example be done by simply involving customers in open conversations to guide them during their decision-making process. Our findings also suggest that retail managers across cultures might benefit less from involving customers in the co-creation process during the consumption/service-provision phase. Customers had higher service quality perceptions and were more satisfied with their shopping experience when the firm took care of the service-provision phase than when they were part of it. Consequently, retail managers might want to invest into creating physical store space that invites customers to participate in the co-creation process, but create a service-provision phase that is less focused on co-creation.

Our findings suggest further that co-creation has different meaning across cultures. Consequently, global retail marketers that are cross-culturally active might benefit from knowing that co-creation seems to play a bigger role in more individualistic cultures than collectivistic cultures. Therefore, creating retail space that invites customers to co-create value together with the retailer might be more important in individualistic cultures (such as the U.S.) than collectivistic cultures (such as Japan). This finding may also be the result of the fact that the U.S. retail infrastructure has moved further along the self-service continuum than has the Japanese retail infrastructure (Meuter and Bitner, 1998). Japanese consumers, consequently, are less habituated to co-creation.

SELECTED REFERENCES

Auh, S., Bell, S. J., McLeod, C. S., & Shih, E. (2007). Co-production and customer loyalty in financial services. Journal of Retailing, 83(3), 359-370.

Baumann, J., Le Meunier-FitzHugh, K. (2015). Making value co-creation a reality – exploring the co-creative value processes in customer-salesperson interaction. Journal of Marketing Management, 31(3-4), 289-316.

Bendapudi, N., & Leone, R. P. (2003). Psychological implications of customer participation in co-production. Journal of Marketing, 67(1), 14-28.

Hofstede, G. (2014). “National Culture – Countries.” Retrieved on September 15, 2014 from http://geert-hofstede.com/united-states.html.

Hoyer, W. D., Chandy, R., Dorotic, M., Krafft, M., & Singh, S. (2010). Consumer Cocreation and New Product Development. Journal of Service Research, 13(3), 283-296.

Lusch, R. F., Vargo, S. L., & O’Brien, M. (2007). Competing through service: Insights from service-dominant logic. Journal of Retailing, 83(1), 5-18.

Prahalad, C. K., & Ramaswamy, V. (2000). Co-opting customer competence. Harvard Business Review, *78*(1), 79-90.

Journal of Business and Behavioral Sciences
Vol 28, No. 2; Fall 2016

THE CONSISTENCY AND VALIDITY OF ONLINE USER RATINGS OF MOVIE AND DVD QUALITY

Dan Baugher
Chris Ramos
Alan Eisner
Pace University

ABSTRACT: Internet word-of-mouth is an increasingly discussed topic, but comparisons of user ratings across different platforms are uncommon. Mean user ratings were obtained for a broad mix of 500 movies and their corresponding DVD using the rating systems in place on Netflix and Amazon. Mean user ratings on Netflix averaged 3.57 (SD = .37). Mean ratings for DVDs on Amazon averaged 4.29 (SD = 1.83). Ratings showed a significant, though modest correlation of .153 ($p \leq 001$) indicating some similarity in user views, but far from substantial similarity. In ratings of DVDs, users consider both movie quality and DVD technical quality which may account for the weak correlation across the sites. Rater volume across the sites showed a correlation of .418 ($p \leq 001$) suggesting movie visibility played a similar role in generating more ratings across the sites. Netflix and Amazon user ratings showed criterion validity correlations for DVD sales of -.503 ($p \leq 001$) and -.293 ($p \leq 001$), respectively, when Amazon's Best Seller's Rank in Movies and TVs was used. Netflix ratings had the greater validity for DVD sales as shown by their higher correlation with the Amazon sales index.

Key words: Electronic Word of Mouth, Online Consumer Reviews

INTRODUCTION

Consumer buying behavior has changed considerably with the substantial growth of the Internet. Word-of-mouth now often takes the form of online ratings made by users of a product. Online reviews of products are available for many products from hotels to books (Zhang et al., 2010) and a major source of information for product purchases (Liu, 2006). With this increasing availability of ratings, researchers have begun to use rating information, especially average user ratings, typically referred to as valence, in many ways. For movies, valence has often been related to sales with some finding valence predicts sales (Dellarocas & Zhang, 2007) and others finding no relationship between the two (Chintagunta, Gopinath, & Venkataraman, 2010).

While it is possible this difference in results is due to differences in the movies considered (Purnawirawan et al., 2015), it is also possible investigators should not so quickly assume user ratings are reliable and valid sources of information or sites presumably measuring quality are valid for the same purposes. Further, as consistency across sites decreases, it can suggest low reliability which can lower the correlation between such ratings and criteria under investigation because the power to discover relationships is weakened (Nunnally, 1978, Traub, 1994) or it can suggest the validity of the ratings is not strong or the measures reflect different quality dimensions (Kline, 1993).

These problems are, in part, due to the nature of online ratings. They are not collected in controlled settings and the rating scales used can sometimes be weak. None have been developed for the purpose of academic research. Though the focus of this study is on movies, these weaknesses can foster random measurement error and result in rating inconsistency and validity problems across database sources for books, computers, or any other product. Generally, there is very little research on the cross-platform rating consistency and validity of user ratings.

PRIOR RESEARCH

The reliability of movie ratings has been a concern for quite some time and the results have varied. Cosley et al. (2003) found test-retest reliability to be relatively strong at .70 for 40 randomly selected movies rated by viewers over time, suggesting at least that users agree with their own ratings. Similarly, Amatriain, Pujol, and Oliver (2009) found strong test-retest reliability of .88 for 100 movies rated a second time at least 15 days from when first rated.

Thus, it appears users agree with themselves over time, but this is not the same as two users agreeing across two different rating platforms. It is possible for individual ratings to be reliable in a test-retest sense, but for users to differ. For example, agreement between movie critics has been found to be quite weak by some researchers (Agresti & Winner, 1997). Assessing the reliability of individual ratings is a good way for better understanding average ratings, but it does not assure the average ratings will be reliable or valid. And ultimately, they are the predictor of concern (Tett, Jackson, & Rothstein, 1991)

One method for assessing the reliability of average ratings is to check consistency across platforms. This type of research has been infrequent and shown varying results for movies. Plucker et al. (2009) found a low-moderate correlation of .43 between mean ratings by students and critics. However, they found a moderate correlation of .65 between mean student ratings and mean user ratings on both the IMDb and boxofficemojo.com sites.

One factor that seems to affect the consistency of mean ratings across users is user experience. For students who rarely saw movies, their average rating showed a low correlation in the Plucker et al. study (2009) of .22 with mean critic ratings and .30 and .29 for mean ratings on the IMDb and boxofficemojo.com platforms, respectively. However, the correlation between mean student ratings and mean critic ratings increased to .44 for experienced student raters, those who viewed more movies than most. Likewise, the correlations with IMDb and boxofficemojo.com mean ratings increased to .58 and .59, respectively, for experienced students.

In the Plucker et al. study (2009), those who used IMDb and boxofficemojo.com were defined as novice critics because they were more than casual raters of movies. These more experienced raters showed a high correlation of .86 for mean ratings across the IMDb and boxofficemojo.com platforms. Hon (2014) conducted a similar study of critics and users from the Yahoo! and metacritic.com sites for 243 popular movies released between 2008 and 2009 and found the correlation between mean critic and mean user ratings from Yahoo! was .413, but the correlation between mean critic ratings was quite high at .92, similarly suggesting experience may contribute to consistency across sites.

Summarizing, test-retest reliability for move ratings by individual users is typically acceptable. However, user consistency is much more variable, showing very good to not so good correlations between ratings. One possible factor accounting for consistency in user ratings is movie viewing experience though users can vary on other unknown qualities. Since the mean ratings across platforms can often be based on different users, it is quite possible that consistency across platforms will also vary considerably.

STUDY PURPOSE

This study seeks to determine if mean ratings of movie quality are consistent and valid across two sites. It assesses consistency of mean ratings (valence) across the Netflix and Amazon sites. The Netflix site was selected because it is one of the most popular benchmarks used in the Recommender Systems (RS) literature (Amatriain et al., 2009) and is heavily used by researchers. No other site provides the high volume of ratings of movies as Netflix where volume is frequently in the millions. The Amazon site was selected because Amazon user ratings are readily available and often used.

Consistency across two sites can provide evidence for validity since consistency (or reliability) is often a first step in determining validity (Nunnally, 1978). However, another method for assessing validity is to find the correlation between a predictor such as mean ratings of quality (valence) and a criterion, such as sales, often referred to as criterion validity (Goodwin, 2010). Comparative studies of criterion validity for user ratings across different platforms for DVD sales have not been conducted. To assess criterion validity, this study compares the correlation between mean ratings on the two sites with DVD sales on Amazon.

This study also considers the correlation between rater volume across the two sites. It does this to provide a degree of assurance that the sample of movies used had similar visibility to users on both sites.

HYPOTHESES

While the results vary, the research on reliability and consistency of movie ratings discussed above suggests mean ratings across platforms should be correlated. In prior research, the correlation of mean movie ratings only met the criterion of .80 for rating reliability (Nunnally, 1978) when the raters were expert critics or more than casual critics. As rater experience declined, so did the correlation between mean ratings (Hon, 2015; Plucker et al., 2009).

Since Netflix raters tend to make quick judgments after viewing a movie more akin to inexperienced student raters and Amazon ratings of DVDs involve both movie quality and DVD quality which do not always correspond, it is not expected that the correlation between the two will be any greater than moderate. For example, a high quality movie can be issued on a low quality DVD, resulting in a lower rating on Amazon than on Netflix. Likewise, a Netflix user who makes a quick rating could lack care in their judgment of movie quality, fostering instability of Netflix user ratings or less validity in the ratings.

With regard to criterion validity, it is expected ratings from both sites will have validity for DVD sales. Both involve user ratings of movies though Amazon ratings also include considerations of DVD technical quality. The sales criterion used is Amazon's Best Seller's Rank in Movies and TV. This index is updated hourly and provides an indication of how a product is selling overall in its category. It reflects recent and historical sales in an item's category when a sub-category is used (http://www.amazon.com/gp/help/customer/display.html?nodeId=525376). Amazon does not have a rank solely for movies. It combines movie and TV DVDs. A low value for this index shows strong sales activity. Thus, a negative correlation is expected with high quality ratings associated with low index values, or better sales, and vice-versa.

Rating volume is considered because it can serve as an indicator of movie visibility (Dellarocas & Zhang, 2007). Given that visibility should relate to the number of ratings made on both sites, rating volume should correlate across the sites with movies rated most often on one site corresponding to those most often rated on the other site and vice-versa.

Three hypotheses were tested. In Hypothesis 1, it was predicted that a positive correlation would exist between mean user ratings of movie quality (valence) on the Netflix and Amazon sites. In Hypothesis 2, it was predicted that valence on both sites would negatively correlate with Amazon's Best Seller's Rank in Movies and TV showing both have criterion validity. In Hypothesis 3, it was predicted that a positive correlation would exist between rater volume across the two sites. The following formal hypotheses were tested:

H1: Mean user ratings of movie and DVD quality (valence) will be positively correlated across the sites.

H2: Mean user ratings of movie and DVD quality (valence) will negatively correlate with sales as measured by Amazon's Best Seller's Rank for Movies and TV.

H3: Rater volume will be positively correlated across the sites.

METHODOLOGY

Mean ratings of movie quality were obtained from the Netflix site (http://movies.netflix.com/WiHome) and mean ratings of DVD quality were obtained from the U.S. Amazon site (http://www.amazon.com). A five-point rating scale for quality is used on both sites with a low of one and a high of five. Amazon's Best Seller's Rank for Movies and TV was captured for each DVD when valence from the Netflix and Amazon sites was ascertained. It can range from one to the maximum number of DVDs on sale in any given day. Volume of user ratings was also input for each movie/DVD. It was obtained from each respective site.

A sample of 500 movies was selected for rating comparison to assure a wide range of movie quality, movie type, and visibility. The process began with a search of DVD listings on the U.S. Amazon site from its decade by decade listing. Blockbusters and movies from all traditional movie genres (Mckenzie, 2010) including drama, horror, and musicals were included. Foreign films were also included, most of which were drama. As an example, blockbusters included such movies as *The Shawshank Redemption*, *Titanic*, and *The Terminator*. On the other hand, art films with far less marketplace visibility were also included such as *Autumn Sonata* by Bergman, musicals such as *Kiss Me Kate* and *Moulin Rouge*, some horror movies considered to have very low quality but a following such as *The Incredible Melting Man*, and plays, like *Dragon Country*, by Tennessee Williams.

RESULTS

The movies comprised eight genres: drama, horror/sci-fi, comedy, mystery/thriller, action/adventure, musicals, foreign films, and gay/lesbian. Drama dominated the list followed by horror/sci-fi, foreign films and comedy. Most foreign films were drama. A one-way-ANOVA showed no significant difference in valence by genre for Netflix ratings (F = 1.42, df = 7, 492, ns) or Amazon ratings (F = .362, df = 7, 492, ns). Movie release date ranged from 1931 to 2015.

The average valence for Netflix ratings was 3.57 with a standard deviation of .37 and range of 1.90 to 4.50. The average valence for Amazon ratings was 4.29 with a standard deviation of 1.83 and range of 1.80 to 5.00. The Amazon Best Seller's Rank for Movies and TV for the DVDs considered had a mean of 64,561 with a standard deviation of 60,696. The index ranged from 15 to 304,270 across the 500 DVDs showing a very broad range of sales activity.

As expected, volume was considerably different across the Netflix and Amazon platforms. The average volume for Netflix was 1,480,931 while that for Amazon was 306. The mix of movies also resulted in considerable variability in rating volume for both sites with a standard deviation higher than the mean on both. The standard deviation for volume was 2,636,388 and 454 for Netflix and Amazon, respectively.

H1 was tested by finding the correlation between mean user ratings on the Netflix and Amazon sites. The zero-order product moment correlation is the preferred statistic for studying reliability and validity (Tett et al., 1991). The results show support for H1. The Pearson correlation between the two sets of ratings was .153 ($p \leq .001$). The results suggest a modest degree of consistency, in keeping with the proposed hypothesis.

H2 was tested by finding the correlation between mean user ratings on the two sites with Amazon's Best Seller's Rank for Movies and TV. The results support H2. The Pearson correlation was -.503 ($p \leq 001$) and -.293 ($p \leq 001$) for the Netflix and Amazon ratings, respectively, showing criterion validity for both.

To assess the relative contribution of the two user ratings for predicting sales, a multiple regression was conducted with Netflix and Amazon mean ratings as the independent variables and Amazon's Best Seller Rating for Movies and TV as the dependent variable. The regression resulted in a multiple R of .503 with standardized betas -.04 (t = .673, ns) and -.482 (t = -8.188, $p \leq 001$) for Amazon and Netflix ratings, respectively, showing Amazon ratings added nothing incremental in prediction to the prediction provided by Netflix ratings.

H3 was tested by finding the correlation between rater volume on the Netflix and Amazon sites. The results support H3. The Pearson correlation between volume on the Netflix and IMDb sites was .418 ($p \leq 001$), despite huge differences in volume across the two sites.

DISCUSSION

This study focused on determining the consistency of mean user ratings (or valence) for movies available on DVDs across the Netflix and Amazon platforms. While user ratings were found to correlate across the sites, the association was not strong. This was not totally unexpected as ratings of DVDs involve a more complex task than ratings of movies. Since reliability places an upper bound on the correlation of a measure with criterion variables, examining the consistency of mean user ratings across platforms is one approach for determining their reliability and also understanding the extent to which their validity might be lower than expected. In this regard, the modest correlation between user ratings on the two sites suggests reliability and validity may be issues of concern for the ratings.

In making this assessment, a very diverse range of movies was considered with movies ranging in release date from as early as 1931 to 2014. While the mix was not a random selection of all movies and thus the results are limited by the movies considered, the mix has the key benefit of including movies with staying power in the marketplace and likely to be sold as DVDs for a long time after this study. They reflect a list of movies that would easily be recognized by anyone with even a passing interest in the genres and includes a full range of movies from blockbusters to art films.

While user ratings from both sites showed a significant correlation with sales on the Amazon site and thus, criterion validity, the weak relationship between Amazon ratings and those on Netflix suggests the two measures are not measuring the same thing. If rating unreliability was the major source of this low consistency, the correlation of user ratings with sales should have been closer to the correlation between mean user ratings, that is .153, because it should have set a ballpark maximum on the correlation with the sales criterion since low reliability attenuates correlations with criteria (Nunnally, 1978; Traub, 1994).

Rather, the correlation of user ratings with sales was considerably higher, at -.503 and -.293, for Netflix and Amazon ratings, respectively, suggesting user unreliability is probably less of a problem than differences in views on what constitutes quality. This is not surprising in that ratings of DVD quality and movie quality are not quite the same. Technical quality of DVDs may not always align with movie quality thereby resulting in a low correlation across the sites. For example, an excellent movie could be crippled by poor resolution or sound on a DVD.

Of course, there is always the possibility of self-selection in online reviews which can bias their representativeness (Li and Hitt, 2008). This is less likely for the Netflix site where ratings are based upon many thousands of raters, but the Amazon ratings certainly have this potential. They may reflect users who are much more concerned over technical DVD issues than the typical buyer of DVDs. In some cases, user ratings on Amazon may also be impacted by other concerns such as delivery and packaging. This study sheds light on this possibility by showing that the correlation between user ratings and DVD sales was strongest for Netflix ratings which are far less impacted by technical issues.

Also, Amazon ratings which clearly considered technical issues added no additional predictive power for sales to the Netflix ratings. This suggests that despite issues with technical quality that may be reducing the correlation between Netflix and Amazon ratings, ratings that only involve movie quality are more related to DVD sales. It appears buyers acquire DVDs more on the basis of movie quality than technical quality. Further, raters on both sites were impacted in like ways with regard to movie visibility as volume of ratings was correlated across the sites even though large differences in the number of users making ratings on the sites was apparent. Thus, it is not likely that differences in movie visibility were a factor causing the low correlation between ratings on the two sites

It may be that raters on the Amazon site do not reflect the broader marketplace because they are too few in number or have biased views, possibly even a greater sensitivity to such issues as DVD quality. Bias due to insufficient raters has been shown to affect ratings of movies and their correlation with outcomes such as sales by others (Zhou & Lange, 2009). There is no simple solution to this problem other than to be on the lookout for multiple sources of online user ratings when available and to recognize the findings of this study can occur for any product or site, but especially those with few ratings relative to number of users or products sold. For example, if only Amazon ratings were used in this study, the relationship of user views to DVD sales would have been underestimated.

More research on the reliability and validity of user ratings across platforms for varying products is needed. Such research can help investigators select the best predictor for their purpose. It can also provide data useful for better understanding how subsets of users who dominate the ratings on a site may not reflect the marketplace. In an era where such ratings can be easily obtained and reliability and validity may not be well known, research on these properties is important for better understanding the nature of Internet word-of-mouth.

SUMMARY AND CONCLUSIONS

We conclude that mean user ratings (or valence) for movie and DVD quality have a modest relationship across the Netflix and Amazon sites. While it might at first seem that sale of DVDs would be more highly correlated with Amazon ratings which include views of movie quality, as well as DVD quality, analysis showed this is not to be the case. In the broad marketplace, Netflix ratings of movie quality appear to be the greater driver of DVD sales on Amazon not ratings on Amazon where the technical quality of the DVD is mixed with movie quality. It could be Amazon DVD ratings reflect a small subset of DVD buyers with greater concerns over technical quality than matters to the average consumer or their views on movie quality are divergent from what drives DVD sales

While the availability of Internet word-of-mouth ratings makes them a tempting source of information, they continue to have the same issues of problematic rating systems, biased subjects, and a potential lack of reliability and validity that existed for such ratings prior to the Internet (Miller, 2001). For example, Amazon DVD ratings mix movie quality with technical DVD quality and concerns over packaging and delivery. As a composite measure, they include a number of potentially unrelated factors. Also, Amazon ratings often include ratings of all versions of a DVD issued over time and in some cases technical quality may have changed. Thus, the problem could be less the users and more what they rate.

With multiple online platforms, not all can be expected to offer the same criterion validity or the best subset of users for reflecting marketplace views of quality. There are few studies to guide investigators on which site may offer the best predictor for sales of a given product. For DVD sales, it appears Netflix ratings, though they are on another platform, have a stronger relationship with sales than ratings on the platform where the sales are made, Amazon. This is an interesting and possibility unexpected result and why studies of this sort are valuable.

REFERENCES

Agresti, A., & Winner, L. (1997). Evaluating agreement and disagreement among movie reviewers. Chance, 10 (2), 10-14.

Amatriain, X., Pujol, J. & Oliver, N. (2009). I like it. I like it not: Evaluating user ratings noise in recommender system. User Modeling, Adaptation & Personalization. In Lecture Notes in Computer Science, ed. G. J. Houben et al. Vol. 5535, 247-258. Berlin: Springer-Verlag.

Cosley, D., Lam, S. K., Albert, J., Konstan, J. A. & Riedl, J. (2003). Is seeing believing? How recommender system interfaces affect users' opinions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 585-592. April 5-10, Ft. Lauderdale.

Chintagunta, P. K., Gopinath, S. & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. Marketing Science, 29 (5), 944-957.

Dellarocas, C., & Zhang, X. M. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. Journal of Interactive Marketing, 21 (4), 23-45.

Goodwin, C. J. (2010). Research in Psychology: Methods and Design, 6th edition. Hoboken, NJ: John Wiley and Sons.

Hon, L.Y. (2014). Expert versus audience's opinions at the movies: Evidence from the North-American box office. Marketing Bulletin, 25, 1-22.

Kline, P. (1993). The handbook of psychological testing. New York, NY: Routledge

Li, X., & Hitt, L. (2008). Self-selection and information role of online product reviews. Information Systems Research, 19 (4), 456-474.

Lui, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenues. Journal of Marketing, 70 (3), 74-89.

Mckenzie, J. (2010). How do theatrical box office revenues affect DVD retail sales? Australian empirical evidence. Journal of Cultural Economics, 3, 159-179.

Miller, T. W. (2001). Can we trust the data of online research? Marketing Research, 13 (2), 26-32.

Nunnally, J. C. (1978). Psychometric theory. 2nd ed. New York, NY: McGraw-Hill.

Plucker, J. A., Kaufman, J. C., Temple, J. S. & Qian, M. (2009). Do experts and novices evaluate movies the same way? Psychology & Marketing, 26 (5), 470-478.

Purnawirawan, N., Eisend, M., De Pelsmacker, P. & Dens, N. (2015). A meta-analytic investigation of the role of valence in online reviews. Journal of Interactive Marketing, 31, 17-27.

Tett, R., Jackson, D. & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. Personnel Psychology, 44(4), 730-742.

Traub, R. (1994). Reliability for the Social Sciences: Theory and Applications (Vol. 3). Thousand Oaks, CA: Sage.

Zhang, Z., Quiang, Y., Law, R. & Yijun, L. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. International Journal of Hospitality Management, 29, 694-700.

Zhou, H. & Lange, K. (2009). Rating movies and rating the raters who rate them. American Statistician, 63 (4), 297-307.

Journal of Business and Behavioral Sciences
Vol 28, No 2; Fall 2016

CORPORATE FRAUD AND EMPLOYEE THEFT: IMPACTS AND COSTS ON BUSINESS

Stephanie Peters
Balasundram Maniam
Sam Houston State University

ABSTRACT: Corporate fraud and employee theft can affect almost any type of business in today's business environment. There are many different forms of fraud and theft that can impact the business results and reputation. Business culture, ethical strategies, and business activities can all have direct relations to how an employee or a group of employees have the opportunity of creating scenarios that result in negative business impacts. Impact costs to the organization can be both direct and indirect costs to the business, its employees, and its shareholders. This research also looks at ways that businesses can improve their structure, strategies, and policies to reduce these fraudulent and theft incidents.

Keywords: Fraud, employee theft, segregation of duties, internal controls, prevention, culture, impact costs, ethics, ethical behavior.

INTRODUCTION

Unfortunately, corporate fraud and employee theft within businesses today are not uncommon circumstances. These activities can create far reaching impacts and effects on the business and have substantial impacts on the cost performance metrics of the business and the overall morale of the employee workforce. These affected businesses can carry the weight of the "bad decisions" and unethical practices of a few associates for many years. A person can barely go a day without hearing or reading a news article that describes some type of activity that occurred by a business or its associate that questions one's ethical behavior practices. There are varying statistics of the costs that business bear due to employee theft and fraud, but in 2014, the amount ranged from $20 to $50 billion dollars (Foley, 2014). The cost, however, goes well beyond the total dollar value of the fraud or theft activity. The impact is also felt on overall company morale, especially the team morale of the involved associate. Activities such as these also cause other business and suppliers to question the practices of the "injured" company as well. Questions are raised about the ethical behaviors of the business as a whole. Only time and consistent performance and results can improve and remove the tarnish from the reputation of the affected company or business.

There are many types of corporate and employee fraud that affect business of all sizes and all types. The fraud and theft can come in many forms – actual theft of cash, property, or other assets; procurement or contractual fraud; payroll theft; and misrepresentations of financial statements are just a few. No business is one hundred percent immune to having a circumstance occur in their business. Even businesses that have the most "impactful" and detailed ethics statements and detailed procedures and segregations can still be a victim. Companies send employees through training sessions that discuss ethical decision making and practices that employees should not engage in while conducting business; however, these occurrences still happen.

A business must ensure that it has practices in place that can help prevent, or at the least, identify the issue very quickly to prevent larger impact to the organization. Segregation of duties where there are financial transactions and audit checks to verify processes are just a few ways that an organization can help to prevent fraud and theft. Also, simple employee behaviors and actions can also be indicators of concern – living well above their means and wanting access to management systems or details that do not relate directly to their role are a few examples. The instances that allow a person to consider and actually participate in fraudulent and theft practices are sometimes so innocent in the beginning. This requires that employers and businesses keep a watchful eye on areas where there is opportunity for one to be involved in such activity. This report will provide more examples of fraud and employee theft activities within the business world over the last several years and will look into the business culture and structure that allows for the employee behaviors. The paper will provide the impact costs to business – both direct and indirect costs. The paper will also provide actions or changes that a business can take to prevent fraud and employee theft.

LITERATURE REVIEW

Corporate fraud and employee theft are not new to this day and age. However, the rates seem to have become more staggering over the last several years. Back in 2007, it was estimated that theft and fraud by a business's own employees cost businesses $100 billion dollars worldwide each year (Sauser, 2007). Another article shows that in 2010 the magnitude of fraud and theft on businesses worldwide could be as large as $2.9 trillion dollars (Davis, 2013). In 2013, the number of cases of employee fraud and theft that exceeded $100,000 for each case grew at a record pace of almost five percent (2013: Highest Rate). In five years of case studies from the Marquet *Report on Embezzlement*, there were more than 2100 cases of employee theft that exceeded $100,000 and nearly two thirds of perpetrators were women (Powell, 2014). These numbers are all staggering no matter how one looks at them. From the research conducted by Free, employee misconduct behavior had begun with small incidents and then gets larger as the associate realizes no one is aware of the issues (2015). Activities generally do start out very small and almost somewhat innocent like; then, as the individual, or individuals in many cases, raises their confidence level, the activities become larger and wider spread across the organization.

Johnson describes the far-reaching impacts of corporate fraud on a business in terms of "reputational sanction"; in other terms, customers being weary of dealing with a firm that has dishonest management, which causes reduction in demand for their services or products (2014). Also mentioned in this research is the results that in the fraud cases studied, nearly six percent of customers stop doing business with a company that has been affected by a fraud event. The research from Marcel shows that not only are customer relationships affected by fraud and theft but that employee and board of director relationships are also affected. In the article written by Marcel, it states that board of directors who separate themselves from the affected business, normally come out ahead in terms of reputation (2014, p. 929). By vacating their board post early, it is possible that their name will not be mentioned in news and press release articles.

Another impact of fraudulent and theft activities is increased government regulation and policies to address or help to alleviate concerns from the business world, investment markets and consumers around the world. Such a regulation would be the Sarbanes – Oxley Act approved in 2002 (also known as the Public Company Accounting Reform and Investor Protection act of 2002). This act's purpose is to focus on internal controls, accounting reports, auditing processes and review and structure of boards and committees (Milliar, 2009). This Act was a response to one of the most notorious cases of corporate fraud, Enron. In a study of the cases of fraud by the Securities and Exchange Commission (SEC) between 1998 and 2007 found that ninety eight percent of cases had involvement of the CEO or CFO, according to Venkatesan (2015).

Businesses must take actions and implement control procedures with departments and policies that will help prevent fraud and theft attempts by associates. The first step is strategic to business function – simply care about your associates and the work they are performing (Timofeyev, 2014). Internal controls and segregation of duties are always considered top methods of preventing and eliminating the opportunities for associates to create fraudulent details or commit an act of theft against the company. However, it is not as simple as putting controls and policies in place, a company must "reinforcing and vigilantly adhering to the existing controls" (Cheney, 2008, p. 20). In looking at various instances of fraud and theft activities found in businesses today, each business affected must be able to develop practices that will prevent a reoccurrence.

FRAUDUENT & EMPLOYEE THEFT ACTIVITIES FOUND IN UNITED STATES BUSINESS

Business examples of fraud and employee theft can come from a business that is large or small, public or private, and in any shape or form. According to Curtis Verschoor, a simple search of database in the month of August 2014, uncovered 200 stories of some type of business related fraud, theft or embezzlement activity (2014). As shown in other research as well, small to midsize companies actually have the greatest risk of having a fraud or theft issue within the business. These instances are also easier for the local new sources to bring details quickly to the public.

Bernie Madoff's Ponzi scheme is a well know example of securities fraud over the last few decades that affected a great deal of people and their personal retirement future. This is a situation that had far reaching impacts than just another's retirement account. Individuals took their own lives based upon Madoff's scheme that affected their whole life's plan. Another example of securities fraud would be "pump and dump scheme" in which a party pumps up the hype about a particular business to get investors to put their money into the business to raise the market prices, and then when the price reaches the point that the fraudulent party desires, they dump the stock for a profit for themselves and causes a loss for the other investors (Galletta, 2015, p. 55).

An example of accounting fraud would be Health South inflated earnings in the early 2000's to exceed Wall Street expectations of the business and create inflated market prices for the business. This inflation of the books was the result of the CEO's instructions to the accounting and finance team. Another example would be the Enron debacle from the early 2000's as well, that caused both substantial business investment tidal waves and large losses to the employees of their business through their life savings. These type of fraudulent activities often involve the CEO or CFO position of the business. In 347 cases of fraud, 89% of the cases reviewed by the "Fraudulent Financial Reporting: 1998-2007" involved the CEO and CFO of the corporation; CEO are implicated in over 72% of the cases, and CFO are implicated is over 65% (Boyle, 2012).

Employee theft can occur in many shapes, forms, and at any level of associate within the business. Seventy five percent of employees have taken items from employers according to the survey conducted by the Association of Certified Fraud Examiners (Galletta, 2015). This can be an associate that regularly takes advantage of entering excess hours on their payroll documentation than what was physically worked; or an associate that takes merchandise from the business without paying or being approved to take the items and benefits from this in a positive manner; or could be an instance where the employee is taking funds from accounts without anyone's knowledge and depositing them into their personal accounts. Employee theft could also be an associate that is stealing research ideas and products from the company and selling them to the competition. In naming off these few examples of practices that could be considered employee theft, it helps to document the problems encountered when trying to determine an estimated amount or value and the scale of employee theft from a business (Sauser, 2007). It is very difficult to put a single amount to what these type of activities cost the business world each year.

Other types of fraud and theft that impacts business of today are ones that do not even initially start because of the business or its employees. Corporate identity theft and database breaches are also large scale issues that affecting businesses today in terms of their business results and reputations. A recent example of this would the Target database breach that affected millions of credit cardholders that had shopped at Target. This type of activity is yet another focus area for businesses to ensure they have the correct internal controls and processes in place to protect not just their business data but also the data of the customers that have expectations of the company to protect them as well. In addition to employee behaviors, business culture and practices that are allowed impact what an employee is capable of committing in terms of fraud and theft.

BUSINESS CULTURE ASPECTS THAT ALLOW FOR FRAUD & THEFT ACTIVITIES

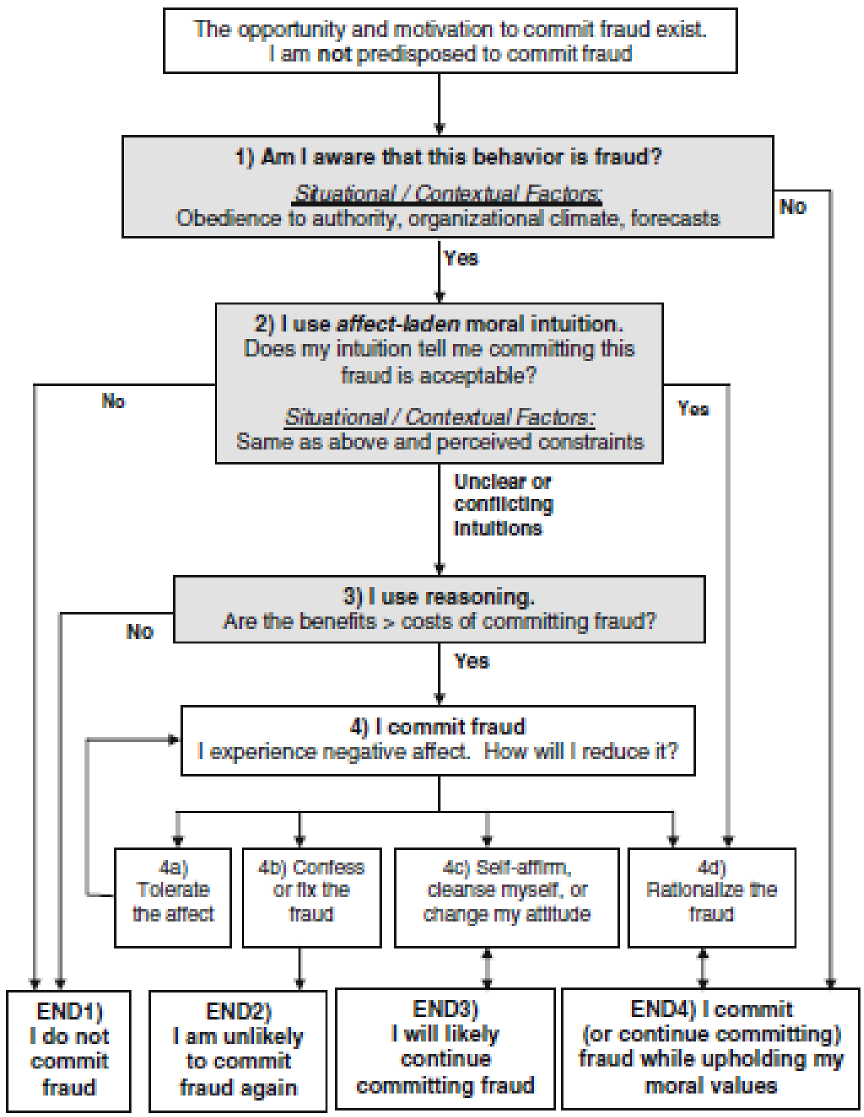
In many cases of fraud and theft, "the ends justifies the means" to the associates that end up caught up in the tangles of deceit (Campbell, 2014, p. 301). An employee or group of employees will rationalize their decision to make changes to statements, thinking that it does not affect anyone else outside of their business, when in fact it has far reaching negative effects on many. In Campbell's study, because the organizations connect the employee's success to the metrics of the business, therefore making it a "win-win situation", corrupt business practices can occur without someone raising the awareness flag. Employees that start committing theft or fraud against a company have the following components: financial need, opportunity, and rationalization (Rood, 2014).

The business overtone that is set by the CEO is one that will drive its way through an organization. Research shows that a "suspect" CEO arrival into an organization can completely drive existing culture out; this comes about through replacement of key employees with ones that agree with CEO's focus and requirements and one that is more than willing to follow the directions provided (Biggerstaff, 2015). When comparing the cases of fraud and employee theft, there are commonalities amongst the cases – inadequate controls, no segregation of duties, follow up and auditing, and poorly completely background checks on employees (Verschoor, 2014).

Another aspect that comes into the business or employee culture that brings about theft and fraud is what defines "theft". This is a question that when asked of associates the answers vary. Some feel it is not theft if they take time paid for lunch break where they were not working, others feel like it is not stealing if items are taken home that have been purchased by the company. These are all part of business theft, in much smaller scale than the cases that are being discussed in this research but it is what begins the definition of theft and fraud. An employee may have issues at home that start the process of trying to figure out how to survive; the solution may sometimes include utilizing business funds or payment methods. These situations are ones that employees must be aware of and must consider in terms of whether or not an employee would have a need or desire to "steal" from the company.

From Murphy's research on the prevention of fraud comes the psychological pathways to fraud diagram shown below. This all starts with an associate that is not predisposed to fraud activities and has not knowingly committed fraud (2011). This diagraph walks through the thought process an associate would go through when considering an activity out of their normal behavior zone. Along this diagram, there are many points where an organization or business can provide training, policies, procedures and reminders that would make an associate chose the path away from the fraudulent or theft behavior choice. When an employee gets to the point in the diagram that they rationalize the fraud, and determination is that it is okay, the employee continues to commit the fraudulent act at a much larger scale than first intended.

Psychological Pathways to Fraud



The grayed boxes indicate key decision or process points. Interventions can be designed to prevent fraud at each of these points.

Note: Diagram source - "Psychological Pathways to Fraud: Understanding and Preventing Fraud" in Organizations in *Journal of Business Ethics*, 2011.

Employees feeling like they are not being rewarded properly for the work that they provide a business is sometimes the starting point for behavior that leads into a theft or fraud situation. According to Hrncir and Metts, these employees may have a feeling the business owes them when fraud or theft is being contemplated; this feeling sometimes starts because of lack of compensation (2012, p. 64). The chart (table 1) below from Hrncir and Metts' article shows some red flags from researched theft cases that businesses can look for within their employee workforce. From this chart, it is seen that living above one's means is a top warning sign, along with employee financial difficulties. When reading various case studies, these two areas appear to be at the top of many research listings for reasons an employee would venture into a theft or fraudulent activity. Many times these acts will start out innocently with a small dollar amount; then, as the employee realizes it is being successful, the scale goes larger and sometimes for years before it is uncovered.

Table 1: Behavioral Red Flags of Perpetrators*

Living Beyond Means	43.0%
Financial Difficulties	36.4%
Control Issues, unwillingness to share duties	22.6%
Unusually close association with vendor/customer	22.1%
Wheeler-dealer attitude	19.2%
Divorce/family problems	17.6%
Irritability, suspiciousness or defensiveness	14.1%
Addiction problems	11.9%
Refusal to take vacations	10.2%
Past employment-related problems	9.3%
Complained about inadequate pay	7.9%
Excessive pressure from within organization	7.5%
Past legal problems	6.3%
Instability in life circumstances	5.6%
Excessive family/peer pressure for success	5.1%
Complained about lack of authority	4.6%

*The sum of percentages exceeds 100% because in many cases perpetrators displayed more than one behavioral red flags.
Source: 2010 ACFE Report to the nation on occupational fraud and abuse, p. 70.

Note: Table was retrieved from the article "Why Small Businesses Fall Victim to Fraud: Size and Trust Issues" by Hrncir and Metts in 2012 from the Business *Studies Journal*.

IMPACTS AND COSTS ASSOCIATED WITH FRAUDUENT & THEFT PRACTICES

Business loss of sales is a direct cost impact from fraudulent and employee theft incidents far outweigh the government fines or class action lawsuits that can be filed against them (Johnson, 2014). With the reduction in production due to lost sales, the business also faces an increase in the SG&A costs associated with producing their products. Another direct cost to business in terms of fraud and theft would be the marketing and promotion costs associated with trying to improve the business image after such an event occurs. The statistics below come from the Association of Certified Fraud Examiners research dated 2015. The median case incident is $175,000 impact to business impacted by fraud or theft. And over 33% of bankruptcies are caused by employee theft according to the statistics provided in this review (Association, 2015).

Employee Theft Statistics	**Data**
Amount stolen annually from US businesses by employees	$50,000,000.00
Percent of annual revenues lost to theft or fraud	7%
Percent of employees who have stolen at least once from their employer	75%
Percent of employees who have stolen at least twice from their employer	37.50%
Percent of all business bankruptcies caused by employee theft	33%
Amount Stolen	**Percent**
$1,000,000 and up	25.30%
$500,000 to $999,999	9.60%
$100,000 to $499,999	28.20%
$10,000 to $99,999	16.80%
$1,000 to $9,999	7%
$1,000 or less	19%
Median Amount Stolen	**$175,000**

Note: Table above was adapted from the article and survey from the Association of Certified Fraud Examiners in 2015 for Employee Theft Statistics. http://www.statisticbrain.com/employee-theft-statistics/

The environment or culture is important, not only for success business results, but it is also important in terms of preventing fraud and theft from occurring each and every day. The negative effects of a tarnished reputation will not be felt immediately by companies that have been a victim of employee theft or fraud. (Guiso, 2015). But over time and additional focus on the situation, companies do not want to be associated with other companies that have blemishes on their record for fraud and employee theft, as these areas can be the starting points of other unethical practices that management has allowed or turned their heads too.

From Timofeyev's study, the charts below show the median loss from fraud of $200,000 in 2014 and theft that and the frequency rate rose to 36.8 percent reported in the Certified Fraud Examiners study (2015). This data shows the losses related to incidents are actually dropping; however, the number of cases per year are rising by approximately 3% each year.

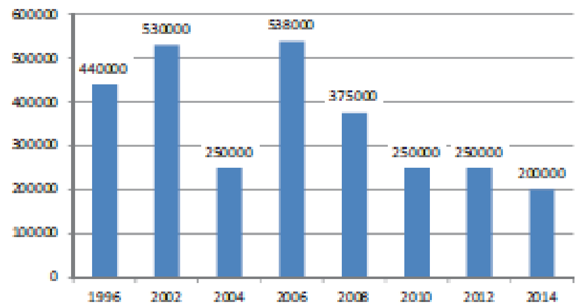


Fig. 1. Median loss due to occupational corruption, USD.

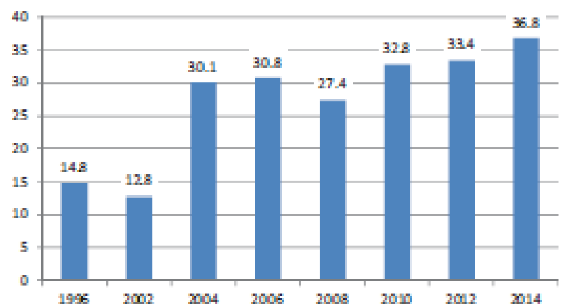


Fig. 2. Frequency of occupational corruption, percent of all reported cases.
Source: ACFE (2014).

Business must be willing to make changes within their environments,

Note: Figure 1 and 2 were retrieved from the article "Analysis of Predictors of Organizational Losses Due to Occupational Corruption" by Timofeyev in 2015 in the *International Business Review*.

practices, policies and procedures in order to be able to prevent fraud and theft activities from impact the business. The amount of change required or needed to improve the culture and environment that allows for employees to create a situation where fraud and theft can occur can vary greatly between the affected businesses.

CHANGES TO OCCUR WITHIN BUSINESS TO PREVENT FRAUD AND THEFT

Incidents of fraud and theft within an organization or business can be reduced and prevented with multiple aspects of control. Team awareness, management support, and cross functional team audits are just a few ways that could impact the business in a positive way in terms of fraud and theft. In the article written by Boyle (2012), a business must set a strong ethical tone from the top leaders of the business to set the culture of theft and fraud being an unacceptable practice (p. 65).

One way to prevent these activities from occurring is to provide the leadership structure that makes another employee feel safe to be able to share concerns and knowledge concerning another associate's unethical behaviors. Together, the leadership and the personal safety aspect give the associate the feeling that the business will accept the information being shared and protect the source of the knowledge within the process of investigation. (Liu, 2015, p. 115). The fear of retaliation is most likely reason some will not come forward with details about other associates. This fear comes about through thoughts of losing a job or creating a hostile work environment with other team mates. Also, when it comes to the discussion on whistleblowing on teammates, some associates feel it is someone else's place to step up to share the information so they remain quiet (Fredin, 2012). The work environment must be one that is positive enough to bring about a more supportive and encouraging environment that also would not allow for the unethical practices to continue to a point to where it would hurt the morale and success of the business. And it must be one that encourages team members to share concerns.

Internal auditors are "eyes and ears of management" and also should have a high level of professionalism based upon their professional standards (Xu, 2008). These individuals are definitely in the middle ground between management behavior that drives the environment of the business and then also the ethical standards that have been established for one in their profession. An auditor must be open minded and look at every detail and process with care to ensure that a business does not have substantial risk in a process or reporting mechanism. A company does not have to have an "auditor" role to get the benefit that an internal look at processes and employee behaviors. An associate with an eye for details can be trained to review another department's process to verify that the controls and policies are enforced within the department. Through this process, findings can be reviewed for improvement opportunities to reduce the possibility of theft or employee fraud.

Other ways that a business can reduce the opportunity an employee to create a situation of fraud or theft include:

- Background checks on new hires and current employees if they are handling large funds

Separation of duties for financial department position. Independent outside review of financial statements to ensure validity

- Protecting documentation that contains sensitive information or details that could be used in fraud or theft situation
- Audits of departments where there are potential risks – Human Resources and Payroll, Purchasing, and Credit
- Pay close attention to employees who have business credit cards and have a process in place
- According to the article "School Your Clients to Stop Theft", employees that will not take vacations (2007). These are employees that could be concern to the business if the role is one where there is financial reporting or payment processing. The concern would come about because of not wanting another associate to look into the processes that could covered up.

In the statistics below from the ACFE 2015 study, tips of other employees is the largest percentage of fraud and theft detection. Referencing back to the strong leadership culture and environment where employees feel safe to sharing details and knowledge about another associates' behavior is valuable to business owners and leaders when trying to prevent fraud and theft practices. Accidental discovery and internal audits are the next methods of finding about an employee's theft or fraudulent practices, with nearly 38% of the cases being discovered through these methods.

Reported Fraud Statistics	Percent
Tip from employee	26.3%
Accidental discovery	18.8%
Internal audit	18.8%
Internal control	15.4%
External audit	11.8%
Tip from customer	8.8%
Anonymous tip	6.2%
Tip from vendor	5.1%
Notification from law enforcement	1.7%

Note: Table was retrieved from the research information from the Association of Certified Fraud Examiners in 2015 on Employee Theft Statistics.

SUMMARY AND CONCLUSION

Employee theft and corporate fraud is costly impact to businesses in today's world. The managers and financial departments today must be on top of internal controls, procedures and policies, and employee business practices to ensure the business is not affected by these types of incidents. According to Ravi

Venkatesan, strong leadership can dramatically change the atmosphere and environment that allows for misconduct like fraud and theft to occur; policies, procedures, practices and audits are all needed as well but the leadership will make the biggest impact to prevent incidents (2014). With varying degrees of controls, audit, segregation of duties, and employee background checks to ensure that associates do not have issues in the past that would be an indicator of such activity again. There is not one solid method to prevent fraud or employee theft from occurring with a business today, but managers and leadership teams must have practices in place that will reduce the opportunity for an employee to choose to steal from the company and they must be able to reduce the severity of the incident. Ronald Reagan said many times while he was in office, "Trust but Verify" so that is always a good adage for businesses to remember.

REFERENCES

Association of Certified Fraud Examiners, (2015). Employee Theft Statistics. http://www.statisticbrain.com/employee-theft-statistics/

Barra, A. R. (2010). The Impact of internal controls and penalties on fraud. *Journal of Information Systems*, 24(1), 1-21. doi:10.2308/jis.2010.24.1.1

Biggerstaff, L., Cicero, D. C., & Puckett, A. (2015). Suspect CEOs, unethical culture, and corporate misbehavior. *Journal of Financial Economics*, 117(1), 98-121. doi:10.1016/j.jfineco.2014.12.001

Boyle, D. N., Carpenter, B. W., & Hermanson, D. R. (2012). CEOs, CFOs, and accounting fraud. *CPA Journal*, 82(1), 62-65.

Campbell, J., & Göritz, A. (2014). Culture corrupts! A qualitative study of organizational culture in corrupt organizations. *Journal of Business Ethics*, 120(3), 291-311. doi:10.1007/s10551-013-1665-7

Cheney, G. (2008). Are your internal controls sufficient? *Financial Executive*, 24(4), 20-23.

Davis, J. S., & Pesch, H. L. (2013). Fraud dynamics and controls in organizations. *Accounting, Organizations & Society*, 38(6/7), 469-483. doi:10.1016/j.aos.2012.07.005

Foley, K. (2014). How Big of a Problem is Employee Theft and Fraud? *Incorp.* http://www.incorp.com/employee-theft-and-fraud-part1.aspx

Fredin, A. (2012). The unexpected cost of staying silent. *Strategic Finance*, 93(10), 53-59.

Free, C., & Murphy, P. R. (2015). The ties that bind: the decision to co-offend in fraud. *Contemporary Accounting Research*, 32(1), 18-54. doi:10.1111/1911-3846.12063

Guiso, L., Sapienza, P., & Zingales, L. (2015). The value of corporate culture. *Journal of Financial Economics*, 117(1), 60-76. doi:10.1016/j.jfineco.2014.05.010

Hrncir, T., & Metts, S. (2012). Why small businesses fall victim to fraud: size and trust issues. *Business Studies Journal*, 4(1), 61-71.

Johnson, W. C., Xie, W., & Yi, S. (2014). Corporate fraud and the value of reputations in the product market. *Journal of Corporate Finance*, 2516-39. doi:10.1016/j.jcorpfin.2013.10.005

Liu, S., Liao, J., & Wei, H. (2015). Authentic leadership and whistleblowing: mediating roles of psychological safety and personal identification. *Journal of Business Ethics*, 131(1), 107-119. doi:10.1007/s10551-014-2271-z

Marcel, J. J., & Cowen, A. P. (2014). Cleaning house or jumping ship? Understanding board upheaval following financial fraud. *Strategic Management Journal*, 35(6), 926-937. doi:10.1002/smj.2126

Milliar, J. & Bowen, W. (2011). Small and large firm regulatory costs: the case of the Sarbanes-Oxley Act. *Corporate Governance: The International Journal of Effective Board Performance*, 11(2), 161-170.

Murphy, P., & Dacin, M. (2011). Psychological pathways to fraud: understanding and preventing fraud in organizations. *Journal of Business Ethics*, 101(4), 601-618. doi:10.1007/s10551-011-0741-0

Rood, D. K., & McCole, G. (2014). All in a dishonest day's work. *Journal of Accountancy*, 218(2), 20-21.

Sauser Jr., W. I. (2007). Employee theft: who, how, why, and what can be done. *SAM Advanced Management Journal* (07497075), 72(3), 13-25.

Timofeyev, Y. (2015). Analysis of predictors of organizational losses due to occupational corruption. *International Business Review*, 24(4), 630-641. doi:10.1016/j.ibusrev.2014.11.0071(1), 77-103.

Venkatesan, R. (2014). Confronting corruption. *McKinsey Quarterly*, (4), 92-101.

Verschoor, C. C. (2014). A Month of embezzlement stories. *Strategic Finance,* 96(10), 11-61.

Xu, Y., & Ziegenfuss, D. (2008). Reward systems, moral reasoning, and internal auditors' reporting wrongdoing. *Journal of Business & Psychology*, 22(4), 323-331. doi:10.1007/s10869-008-9072-2

Journal of Business and Behavioral Sciences
Vol 28, No 2; Fall 2016

HOW GEEK CHIC CAN LEAD THE WAY TO BETTER MARKETING MODELS: THE CALL TO RESEARCH FASHION ADOPTION

Elizabeth F. Purinton
Marist College

ABSTRACT: How can one explain the reversal of meaning in terms such as "geek" or "nerd"? How did the oxymoron "Geek Chic" become a good thing? There still exists no theory or model to explain fashion adoption and acceptance. Can the marketing models such as Engel Kollat Blackwell (1968) or the Diffusion of Innovation Model (EKB) (Rogers 1962) be used to explain Geek Chic in particular and fashion in general, in a similar way to the adoption of high tech products? Is this a matter of "trickle down", "trickle up" or re-cycle?
This paper seeks to shed light on the lack of fashion adoption theories and highlight the need for research. By exploring fashion and marketing models, suggestions are made for tweaking the EKB model and pursuing future research.

INTRODUCTION

Sometime in the twentieth century, the word "geek" went from meaning "a circus performer" (American Heritage Dictionary 1976), in particular one who bit the heads off live chickens, to an unfashionable and socially inept but technologically savvy person. This, in turn, has come to mean "a knowledgeable and obsessive enthusiast" (Oxford Dictionary). A similar semiotic shift has been seen with "nerd".

Fashionability is the degree to which a style is accepted or adopted. A measure of whether something is "cool" would be its acceptance by innovators, early adaptors, and market mavens. What seems counterintuitive is the acceptance of items which *should be* seen as "uncool." In marketing, it is important to be able to understand consumer meanings, to properly portray products but perhaps, more importantly, to predict consumer behavior.

A look at the profusion of current television shows such as *The Big Bang Theory* and *King of the Nerds* will show that geeks are in fashion. "Geek Chic" including flash drives commonly worn as a pendant and heavy, plastic eyeglass frames, is all the rage.

What can explain a complete reversal of fashion and why does this concern marketers? Awareness of what is currently fashionable has a role in marketing, both from the standpoint of developing marketing communications and of consumer behavior.

Geek Chic represents several things in marketing and fashion. It is at once the revival of previous fashions (large plastic eyeglass frames, bow ties, suspenders, wearable technology), adoption of a fashion by multiple target markets and adopter categories, a tie in between fashion and technology, and a change in the definitions (and social judgment) of words such as "geek" and "nerd". Geek Chic is currently in late Growth to early Majority.

The goal of this research is to further both marketing and fashion knowledge and theory. This paper uses the phenomenon of Geek Chic to examine some of the existing models in marketing and fashion to evaluate their ability to explain adoption of new fashion trends.

In order to explore this, one must note that Fashion does not contain models of adoption as marketing does. It has only theories, few empirically tested. Any progress toward the development of a model of fashion adoption (or finding a model in another discipline that is suitable for application) would be an improvement over the lack in this area.

The study of fashion is highly relevant to marketing. Fashion is a culturally-embedded product. In recent years, researchers (Belk 1988, Hirschman and LaBarbera 1990 and others) have advocated closer examination of the enduring, emotional ties that consumers cultivate with certain possessions. Secondly, Fashion is an important aspect of consumer behavior. It has both an economic component (Veblen 1894, 1899) and a cultural one (McCracken 1981, Hansen 2004).

> As a lens, culture determines how the world is seen. As a blueprint, it determines how the world will be fashioned by human effort. In short, culture constitutes the world by supplying it with meaning (McCracken 1981, p 72)
> The fashion system is less frequently observed, studied, and understood as an instrument of meaning movement, yet this system also serves as a means by which goods are systematically invested and divested of meaningful properties. (McCracken 1981, p 76)

This paper will examine fashion change theory and diffusion of innovation. Research propositions will be drawn and a future research study proposed.

LITERATURE REVIEW AND RESEARCH QUESTIONS

While fashion has few models and fewer theories, there is an opportunity to selectively apply marketing models and theories. The discipline of fashion belongs to the field of art. While there is practitioner-based literature and textbooks, empirical research is rare.

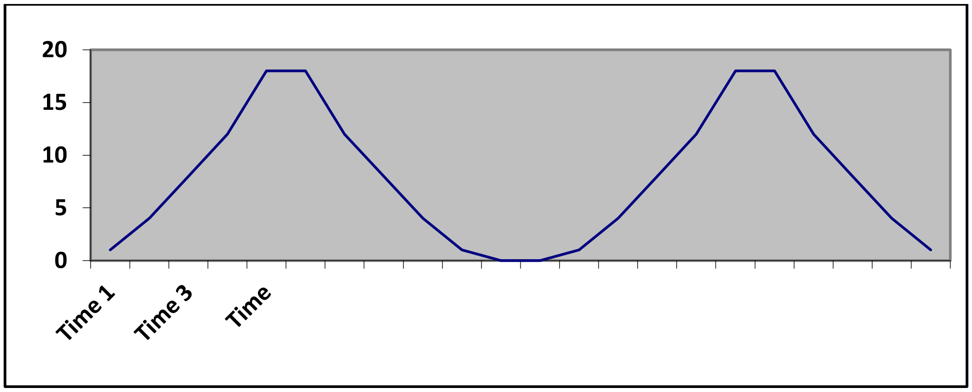
In general, fashion offers two models to explain the movement of fashion through society and time. The first is the concept of fashion cycles, similar to marketing's product life cycle.

The Fashion Cycle is one of the most common models used to explain fashion adoption. Essentially, the model states that fashions will come into being, become adopted, go mainstream, fall out of favor, lay dormant and then reemerge (see Sproles 1979, 1981 for a thorough explanation). These individual styles were often 5-10 years in popularity (Nystrom 1928) and the time between the popularity was approximately 40 years. This has changed to a 20 year cycle.

Geek Chic is an example of a fashion style, something that has come into fashion (been adopted). It is also accompanied by a social contradiction. Words, images, and styles that were considered unfashionable/ unacceptable are now lauded. Geek Chic, however, represents a 50 year cycle, another of its anomalies.

Perhaps Fashion is not as linear (as the trickle theories will suggest) or even curvilinear (as we will see in the Diffusion of Innovation). Fashion is generally considered to be cyclic, not linear, returning to previous fashions or ideals as seen in Figure 1. For example, after the upheaval of the French Revolution came the Greek Revival with women's hair and skirts becoming shorter. It was considered a return to more civilized times. (One of the most quoted fashion phenomena is that when economic times become more difficult, women's hem lengths become longer [more conservative, secure, and nostalgic] and rise again in optimistic times).

FIGURE 1: FASHION CYCLES



The cyclic nature of fashion may assist the Geek Chic movement by adding a bit of nostalgia to the mix. That is, the heavy plastic eyeglass frames of today remind us of the engineers, mathematicians, and even Clark Kent of yesteryear.

Research questions suggested by the cyclic nature of fashion include:

- Have here been previous styles that incorporate the technology of the time?
- How many styles within Geek Chic can be related to fashions from previous decades? How many of these have evolved simultaneously?

This framework falls short for two reasons. First, while it embraces the concept of the retro or revival fashion, it does nothing to address the changing meaning of the words. Secondly, it is purely descriptive and does not provide any suggestions for increasing spreading adoption. At least the product life cycle provide alternative strategies for product in marketing.

"Laver's Law" (Laver 1945) , though linear rather than cyclic, helps to explain the changing fashionability wherein he said the same costume will be:

Indecent	10 years before its time
Shameless	5 years before its time
Outré	1 year before its time
Smart	---
Dowdy	1 year after its time
Hideous	10 years after its time
Ridiculous	20 years after its time
Amusing	30 years after its time
Quaint	50 years after its time
Charming	70 years after its time
Romantic	100 years after its time
Beautiful	150 years after its time"

The reader will note that this does not fit Geek Chic at all. Here we are, fifty to sixty years after popularity of large plastic eyeglass frames, cat eye glasses, bow ties and pocket protectors and these items are not seen as "quaint" but instead are re-popular.

Fashion change theory is immersed in the movement ("trickle") of fashion throughout social classes and subcultures. Trickle theories help explain how a style move through society. They track the flow but offer few reasons why. There are actually three different trickle theories: trickle down, trickle up and trickle across. *Trickle down* theory became common vernacular during President Reagan's term as he discussed economic theory and how wealth would move from one higher socioeconomic class down to the next eventually benefitting everyone. This phenomenon, as applied to fashion, was vividly described by William Hazlitt in 1818:

> "…Strike the attention and excite the envy or admiration of the beholder, and which are no sooner made known and exposed to public view for this purpose, then they are successfully copied by the multitude, the slavish herd of imitators, who do not wish to be behind-hand with their belters in outward show and pretensions, and which then sink without any further notice, into disrepute and contempt" p. 73

Trickle up theory (put forth by Sproles 1979 and acknowledged by Polhemus 1994) reverses trickle down and explains how some fashions move from street wear to everyday wear. The hip-hop movement, in particular, has gifted the

American public with such things as baggy trousers, hooded sweatshirts, excessive *bling*, neck tattoos, and precious metal tooth covers known as *grills*.

The *trickle across* effect follows the adoption of fashion across all levels of society simultaneously. Georg Simmel (1904) wrote of fashion as a process of imitation. In each of these trickle theories, one group is imitating another. Emulation leads to imitation leads to equalization. This might be the most explanative of Geek Chic especially when one considers the emulation of high-tech moguls such as Bill Gates. Further, mass communication plays a strong role in adoption of fashion through trickle across and the explosion of computer accessed information has facilitated mass communication in de dissemination of styles.

Research questions suggested by these theories include:

- Did Geek Chic "begin" at a particular socioeconomic level and travel up or down?
- Did Geek Chic "begin" with a particular subculture (market mavens of fashion, opinion leaders within technology (i.e., Bill Gates), or those working in the high tech industry)?
- Is there emulation beginning with, or encouraged by, famous people or celebrities?

Having exhausted frameworks in fashion, we must now turn to marketing models to investigate whether they will add to our understanding.

Fashion is, after all, a "product". Items may be purchased privately or publicly, and consumed privately or publicly. Fashion carries performance risk, social risk and economic risk. Consumers vary in their involvement level, expertise, reference group and/or opinion leader influence, and subculture.

The Diffusion of Innovation Model, developed by Everett Rogers in 1962, is seminal in discussion of product adoptions. The change in behavior of the individual is known as adoption. It explains how a product or idea gains momentum and spreads (or diffuses) throughout society. The end result is that, over time, a new idea, social construct, or product becomes adopted or accepted as "normal". The idea or product begins as new, innovative, or novel and becomes accepted, mainstream or "old". Blumer (1969) inferred adoption as "collective selection" wherein taste derives from a group responding to the zeitgeist or "spirit of the times".

The Diffusion of Innovation Model, seen in Figure 2, describes adoption of new products but, as we know from both the product life cycle and fashion cycles, an "old" product can act in the market as a new one. Do adjustments need to be made? In particular, how do Innovators differ between the first adoption of a fashion and its reintroduction. Does motivation play a role?

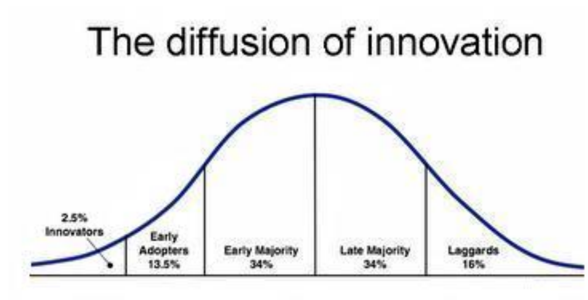
Solomon (2009) suggested that the Diffusion of Innovation Model requires adjustment to apply to fashion adoption. He categorized six stages: innovation risk acceleration general acceptance, decline and out-of-fashion.

Research questions suggested by the Diffusion of Innovation Model include:

- Who were the Innovators and Early Adopters of Geek Chic?

- Has the Geek Chic movement reached the critical mass that Rogers claimed was necessary for mainstream adoption?
- How many adopters' categories have adopted Geek Chic so far?
- Have other fashion phenomena followed the Diffusion of Innovation Model? What can be learned of their final fates?

FIGURE 2



Rogers 1962

One of the most prominent and enduring model in marketing is Engel, Kollat and Blackwell's (EKB) Model of Consumer Behavior (1968). Adopted by every Marketing Principles and Consumer Behavior textbook (i.e., Kotler and Armstrong 2016), this seminal work in marketing represents a comprehensive model of consumer behavior. Inputs into the model include Marketing Stimuli (the marketing mix: product, place, price, promotion) and Other Stimuli (the external environments: economic, technological, sociocultural, political/legal and competitive). Consumer Characteristics are categorized into external (demographic, psychographic, etc.) usually common to members of the target market, and internal (learning, memory, motivation, etc.) which are specific to the individual. This is followed by the Buyer' Decision Process and Choice. (Rather than presenting the entire EKB Model, we will use the simplified version, seen in Figure 3 and standard to Marketing textbooks (i.e., Kotler and Armstrong 2016).

FIGURE 3
MODEL OF CONSUMER BEHAVIOR

Marketing Stimuli	Other Stimuli	Buyer’s Black Box		Buyer’s Choice
Product Price Place Promo	Economic Techno- logical Socio- Cultural Political Legal Compet.	Buyer’s Characteristics: Internal External	Buyer Decision Process	Product Choice Brand Choice Dealer Choice Purchase Timing Purchase Amount

Adapted from Kotler 2016

Fashion fits into the EKB model in multiple places. Marketing Stimuli influence and are also part of fashion. The fashion literature cites the influence of designers and the promotion of their designs (runway shows, magazine articles, reviews). Designers and the celebrities who wear their designs, in turn, become opinion leaders (and part of the Buyer’s Decision Process). Word of mouth is consumer-generated promotion. In the case of Geek Chic, this is usually social media and online discussions.

Perhaps even more obvious is the impact of fashion on the socio cultural environment. Current “fashions” whether for clothing, accessories, automobiles, diets, prevalence of technology use, or other lifestyle factors, are inherent in the socio cultural environment.

In the case of Geek Chic, technology plays a major role. One influence is wearable technology itself. What was once a slide rule tie bars, or belt-clipped calculator is now a sequin covered flash drive clipped to one’s bedazzled iPhone. The combination of everyday technology adoption and the disaggregation of mainstream have coalesced into an acceptance of what would have been the pinnacle of “uncoolness” in earlier generations.

Another influence of technology is much richer and speaks of a power shift among subcultures. Previously, those people with intense interest in science and technology were generally considered to be overly studious or socially inept. “Mainstream” would have been the last word associated with them.

But, as technology advanced and became more and more prevalent in our lives, its contribution became more visible and appreciated. What was once esoteric, specialized and obscure became common, necessary and valued.

Those people with specialized knowledge to create, maintain and facilitate the workings of this technology found themselves in demand, first in the workplace and then, in social settings as well. It was no longer “uncool” to be able to study or practice science or technology. The reader will consider the popularity of such "science geeks" as Neil DeGrassi Tyson, the late Jack Horkheimer and Stephen Hawking. At the same time, as more people became involved "critical mass" (see Diffusion of Innovation, Rogers 1962) was reached. As this reevaluation of social judgment (values) took place, people with other interests felt free to express themselves. This is evidenced in the popularity of phenomena such as ComicCon, Dragon Con, and Cosplay.

THEORETICAL AND PRACTICAL IMPLICATIONS AND CONCLUSIONS

Adjustments to the EKB model to better explain fashion in general, and geek chic in particular, include relational loops within the effects of the external environments and between the external environments and the marketing mix. These, in turn, will affect both the buyer's characteristics (especially opinion leadership and reference groups) and even the buyer's decision process (especially information search). Such as these relationships are presented in the Table.

The contribution of studying these inter-relational factors helps us to better understand what facilitates or inhibits the adoption of fashion. Particularly, as technology becomes more intertwined with fashion and style, our simpler models of effects of external environments are no longer adequate. Instead, the interconnected nature of these influences must be appreciated.

While it is true that Geek Chic appeals to multiple target markets, it also means multiple things to multiple sub cultures. What is a visible affirmation of one group's interests is simply a fashion trend to another and is nostalgic fashion to yet another. This multiplicity, while difficult to study, adds to the richness of the phenomenon.

The contribution of studying these inter-relational factors helps us to better understand what facilitates or inhibits the adoption of fashion. Particularly, as technology becomes more intertwined with fashion and style, our simpler models of effects of external environments are no longer adequate. Instead, the interconnected nature of these influences must be appreciated.

While it is true that Geek Chic appeals to multiple target markets, it also means multiple things to multiple sub-cultures. What is a visible affirmation of one group's interests is simply a fashion trend to another and is nostalgic fashion to yet another. This multiplicity, while difficult to study, adds to the richness of the phenomenon.

TABLE 1
RELATIONSHIPS AMONG EKB ELEMENTS

EKB Element	**Fashion**	**Geek Chic**
Marketing Mix: Promotion	Designers' influence, runway shows, advertising, awareness of new fashions and trends	Role of mass communication in trickle across adoption
Marketing Mix: Product Technology	New fibers Wearable technology	Smartphones as accessories Flash drives as fashion accessories Google glasses Clothing with laptop pockets
Technology Socio Cultural Environment		Adoption of Technology and change in social power makes geek chic acceptable
Buyers' Characteristics	Adopter categories from Diffusion of Innovation model are often seen (trend makers/setter, followers, etc.)	Technophiles become the trend setters
Buyers' Characteristics Socio Cultural Environment	Effect of trend setters' behavior on social "norms"	Previously unacceptable geekhood becomes fashionable. "It's hip to be square."

There are two disparate approaches to studying consumer phenomena, generally known as "inside out" and "outside in". The inside out perspective uses primary research surveying consumers to derive cultural meaning behind purchases. The challenge to this approach is that the adoption of fashion takes place over time and surveys study only the current point in time. Further, consumers are rarely aware of all of the myriad influences on their own behavior.

Semiotic analysis employs an "outside in" approach of semiotic analysis and attempts to learn about a phenomenon by studying it in place. One source would be narrative texts devoted to the topic. What can be learned from motion pictures, TV programs, fiction, and TV and print ads? These ads encode many cultural beliefs about the meaning of fashions and communicate diverse norms about their purchase and wearing. A systematic analysis of these texts drawn from multiple time periods would likely go a long way toward revealing the evolution of consumer attitudes and adoption. For example, did the rise in status of the "nerd" coincide with the adoption of personal computers? Did Geek Chic gain major thrust with the advent of wearable technology?

Another approach would be in the study of other anachronisms in fashion, specifically steampunk which combines elements from the late 19th century with science fiction.

The research questions raised within this paper have value, not just in studying Geek Chic, per se, but also in attempting to approach the development of adoption models in Fashion. Understanding diffusion and adoption in Fashion can further our understanding of product positioning, target marketing, identification of trends and opportunities as well as a host of practitioner suggestions for marketing.

REFERENCES

Bartels, Robert, (1988), *The History of Marketing Thought (3rd. ed.),* Columbus: Publishing Horizons.

Belk, Russell W., (1985), "Materialism: Trait Aspects of Living in the Material World", *Journal of Consumer Research (Pre-1986), 12*(3), 265.

Blumer, Herbert, (1964), "Fashion: From Class Differentiation to Collective Selection" *The Sociological Quarterly*, 275-

Engel, James F., Kollat, David T. and Roger D. Blackwell, (1968), *Consumer Behavior,* Holt, Rinehart &Winston, New York.

Hansen, Karen Tranberg, (2004), "The World in Dress: Anthropological Perspectives on Clothing, Fashion, and Culture", *Annual Review of Anthropology,* 33 (2004): 369-392.

Hirschman, Elizabeth C., & Priscilla A. LaBarbera, (1990), "Dimensions of Possession Importance", *Psychology & Marketing (1986-1998), 7*(3), 215.

Kotler Philip T. and Gary Armstrong, (2016), *Principles of Marketing* (16th ed.), Prentice Hall

Laver, James, (1945), *Taste and Fashion: From the French Revolution to the Present Day,* London: George G. Harrap & Co. Ltd.

Nystrom, Paul H., (1928), *Economics of Fashion*, New York: Ronald Press.

Polhemus Ted, (1994), Streetstyle: From Sidewalk to Catwalk, New York: Thames &Hudson

Richins, Marsha L., & Dawson, Scott, (1992), "A Consumer Values Orientation for Materialism and its Measurement: Scale Development and Validation", *Journal of Consumer Research, 19*(3), 303.

Rogers, Everett M., (1962), *Diffusion of Innovations (first edition)*, Glencoe: Free Press.

Simmel, Georg, (1904), “Fashion.” *International Quarterly*, 10 (1904): 130-155.

Solomon , Michael R., (1983), “The Role of Products as Social Stimuli: A Symbolic Interactionism Perspective”, *Journal of Consumer Research*, 10(3), 319-329.

Sproles, George B., (1979), *Fashion*, Minneapolis, MN: Burgess Publishing Company.

Sproles, George B., (1981), “Analyzing Fashion Life Cycles-Principles and Perspectives”, *Journal of Marketing*, 45.4 (1981): 116.

Tait, Hugh (ed.), (1986), *Jewelry: 7000 Years*. Harry N. Abrams, Inc.: New York.

Veblen, Thorstein B., (1894), “The Economic Theory of Woman’s Dress”, *The Popular Science Monthly*, Vol. 46, New York 1894, 198-205

Veblen, Thorstein B., (1899), *The Theory of the Leisure Class: An Economic Study of Institutions, New York: The Macmillan Company.*

Wejnert, Barbara, (2002), “Integrating Models of Diffusion of Innovations: A Conceptual Framework”, *Annual Review of Sociology (Annual Reviews), 28: 297–306*

Journal of Business and Behavioral Sciences
Vol 28, No 2; Fall 2016

PROTECTED CLASSES, CREDIT HISTORIES AND CRIMINAL BACKGROUND CHECKS: A NEW TWIST TO OLD-FASHIONED DISPARATE IMPACT CASES?

Roy W. Copeland
Mel Schnake
Valdosta State University

ABSTRACT: While some industries employ the use of credit checks and criminal background checks to decrease the likelihood that those hired will engage in behavior, which might have a deleterious impact upon the business, others might use such checks for legally impermissible purposes. Credit checks and criminal background histories are neutral practices employers use in making employment decisions. These neutral practices or requirements sometimes result in an adverse impact on some protected classes. The Supreme Court, addressed in a broad sense, the application of Title VII of the 1964 Civil Rights Act to adverse impact cases in 1971. If a requirement, practice, test or procedure functions to preserve discriminating employment practices it might contravene Title VII. Employers, however, are able to rebut a prima facia case of adverse impact discrimination by showing the practice or requirement in question is job-related and consistent with business necessity. In general, Title VII does not prevent the appropriate use of credit and criminal history information. Those who challenge the use of credit and criminal background information must be prepared to present experts who are competent, reliable and base their opinions on complete and reliable data sets. Courts are intolerant of aggrieved employees who are only able to raise statistical inferences of disparate impact. This article will examine the use of statistics by experts and the scrutiny given to both by courts as well as present day use of credit checks and criminal histories by employers.

Keywords: credit checks, criminal background checks, Title VII, neutral requirement

INTRODUCTION

Credit reports are generally viewed as precursors to determine one's creditworthiness. Borrowers are accustomed to lenders mandating the disclosure of credit history as a condition to determining whether or not credit should be extended. Credit assessments are increasingly used to assist employers making hiring and promotion decisions. Few would doubt that in industries such as banking, finance and law enforcement, credit and criminal background checks may be necessary to ensure the integrity of those who are hired. The nexus between routine credit history and criminal background checks should be determined based upon a standard of job relatedness. However, employers who conduct routine credit history and criminal background checks and are unable to determine a connection between such checks and the type of employment may encounter legal challenges. One such challenge is that the use of credit and criminal background checks in making hiring and promotion decisions may have a discriminatory impact on the basis of protected class status. It is axiomatic that Title VII of the 1964 Civil Rights Act prohibits hiring and promotion practices which have a disparate impact on the basis of race, color, religion, national origin or sex unless the employer shows the practice is job related and consistent with business necessity. This article will examine employers' use of credit reports and criminal background checks in making hiring and promotion decisions. This article will also examine the use of statistics and statistical analysis by plaintiffs in establishing a *prima facie* case and by defendants in attacking the validity of such data.

TITLE VII of the 1964 Civil Rights Act

Title VII of the 1964 Civil Rights Act provides in pertinent part that

> It shall be an unlawful employment practice for an employer to: (1) fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privilege of employment, because of such individual's race, color, religion, sex, or national origin; or
>
> (2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin (Title VII Civil Rights Act of 1964).

Since the passage of the 1964 Civil Rights Act, racial discrimination in the employment arena has undergone a metamorphosis from blatant, to the smoking gun, to the sublime and subdued. The forms of blatant and overt discrimination were pervasive during the pre-Title VII period, and its likely roots are traceable to the appalling era of the Black Codes. After the passage of the Thirteenth Amendment which abolished slavery, "nothing in the Constitution prevented a state from enacting laws – Black Codes, as they came to be known – that were blatantly discriminatory . . ." (Goldstein, 2011).

The smoking gun is a form of direct evidence which generally does not involve a written policy of discrimination. Indicia of discriminatory intent is demonstrated by an employer's tolerating or encouraging an environment that condones discriminatory conduct and activity that thwarts job opportunities for members of protected classes or discourages them from seeking employment opportunities. In such circumstances, it is not likely that the punitive employer will admit involvement in promoting discriminatory activity by its workers. The "smoking gun" usually involves racially or sexually insensitive remarks, jokes or slurs.

The sublime and subtle form of discrimination is exemplified in cases involving "neutral" selection criteria which have a discriminatory impact on protected class members. Discrimination, in the general sense, is the denial of an employment opportunity based upon an individual's protected class status or characteristics. Adverse impact discrimination arises when the denial of an employment opportunity is the result of a "neutral" job requirement that is "not job related and consistent with business necessity." The origin of the disparate impact theory in Title VII cases is *Griggs v. Duke Power Co* (Griggs v. Duke Power, 1971). In *Griggs*, a unanimous Court held that using employment tests which have a discriminatory effect violates Title VII except in instances when the employer meets "the burden of showing that any given requirement [has] . . . a manifest relationship to the employment in question" (Title VII, Civil Rights Act of 1964). The *Griggs* court was unequivocal and firm in its pronouncement of the prohibitions of Title VII. "The Act proscribes not only overt discrimination but also practices that are fair in form, but discriminating in operation" (Title VII, Civil Rights Act of 1964). The court added "the touchstone is business necessity" (Title VII, Civil Rights Act of 1964). The employer's burden does not arise until the plaintiff meets its burden of making out a *prima facie* case of discrimination.

> "The employer's burden of proving job-relatedness to rebut a claim of disparate impact is greater than its burden of merely showing a legitimate nondiscriminatory reason in response to a claim of discriminating treatment. The hard, cold statistical record of impact provides a stronger circumstantial case of discrimination than a subjective claim of improper motivation. Despite some evidence of weakness in the statistics, where they disclose a glaring absence of minority representation in the jobs at issue, the burden on the employer increases since 'fine tuning' of the statistics will not rebut an inference of discrimination derived 'not from a misuse of statistics but from the inexorable zero'" (Grant v. Bethlehem Steel Corp., 1980).

Once a plaintiff meets the initial burden of persuasion that an employment policy or practice has a disparate impact due to his protected class status, the employer may defend the practice by presenting evidence of business necessity for the practice or policy. If the employer produces evidence of business necessity, a plaintiff must be allowed to show that a feasible alternative is available with less discriminatory effect or a reduced discriminatory effect that would also carry out the employer's "legitimate interest" in efficiency and trustworthiness. "Such a showing would be evidence that the employer was using its test merely as a 'pretext' for discrimination" (Albermarle Paper Co. v. Moody, 1975).

CREDIT REPORTS

Credit reports are generally associated with credit institutions, such as a bank or credit card issuer, rendering a decision as to an individual's creditworthiness. In addition to assessing one's ability to qualify for credit, credit reports are also used to determine one's employability and qualification for certain levels of security clearance in high level governmental positions or credit sensitive positions in private industry. The definition of a credit report extends well beyond simply serving as an evaluation tool used by those in the financial services industry.

The Fair Credit Reporting Act (FCRA) governs the obtaining and use of credit information gathered by those entities seeking to use the reports (Fair Credit Reporting Act, 2012). Credit reports are defined by the FCRA as follows:

> The term "consumer report" means any written, oral, or other communication of any information by a consumer reporting agency bearing on a consumer's creditworthiness, credit standing, credit capacity, character, general reputation, personal characteristics, or mode of living which is used or expected to be used or collected in whole or part for the purpose of serving as a factor in establishing the consumer's eligibility for
>
> (A) Credit or insurance to be used primarily for personal, family or household purposes;
> (B) Employment purposes; or
> (C) Any other purposes authorized under section 604 [1681(b)] (Fair Credit Reporting Act, 2012).

To gain a clear understanding of the breadth of what constitutes a consumer report, it is necessary to read and understand sections 1681(a)(d) and 1681(b). Under Section 1681(b), any consumer reporting agency may divulge credit information to an array of persons and/or entities, including, but not limited to, those who may evaluate ". . . the consumer's eligibility for a license or other benefit granted by a governmental instrumentality required by law to consider an applicant's financial responsibility or status. . . ." (Fair Credit Reporting Act, (2012) This statute consensually includes all professions and jobs which require any type of license or certification. The range of coverage would include certified nurse assistants to cosmetologists and neurosurgeons.

Criminal Background Checks

Criminal convictions have been used as a disqualifying factor by employers in making hiring decisions for decades. For example, Missouri Pacific Railroad "since 1948 . . . has followed the policy of disqualifying for employment any applicant with a conviction for any crime other than a minor traffic offense" (Green v. Missouri Pacific Railroad Co., 1975). Criminal background checks are viewed as one mechanism to insulate employers from civil liability based on negligent hiring and retention, to prevent inventory losses due to theft, and to curtail fraud, identity theft and other nefarious activities associated with prior criminal convictions. The adoption of blanket exclusions of those with criminal convictions from employment opportunities may lead to a disparate impact claim.

In *Green v. Missouri Pacific Railroad Co.* (MoPac), an African American job applicant brought an action under Title VII of the 1964 Civil Rights Act contending he was denied employment due to the railroad's policy of refusing to consider applicants convicted of a crime other than a minor traffic offense (Green

v. Missouri Pacific Railroad Co., 1975). Buck Green applied for a position as a clerk at MoPac's headquarters in St. Louis, Missouri. Green disclosed, on his application, that he had a criminal conviction for refusing military induction. Upon reviewing Green's application for employment, MoPac determined he was disqualified because of his criminal conviction. Green's suit was dismissed by the district court, and on appeal he asserted the following claims:

> "(1) MoPac's policy of not hiring any person convicted of a criminal offense has a racially discriminating effect and violates Title VII;
> (2) This policy is not justified by any business necessity; and
> (3) The district court erred in restricting the class only to black persons denied employment consideration because of a conviction record" (Green v. Missouri Pacific Railroad Co., 1975).

In finding that Green ". . . established a *prima facie* case of discrimination," the court validated the usefulness of statistics (Green v. Missouri Pacific Railroad Co., 1975). "The statistics established that MoPac's employment practice under consideration disqualifies black applicants or potential black applicants for employment at a substantially higher rate than whites" (Green v. Missouri Pacific Railroad Co., 1975). The circuit court noted the trial court failed to conduct the appropriate statistical comparison in ". . . comparing the number of black applicants rejected because of a conviction record to the total number of applicants does not reflect a disparity of impact separately against each race" (Green v. Missouri Pacific Railroad Co., 1975). The circuit court also found fault with the trial court's ". . . comparing the resulting percentage of (blacks rejected) against the percent of blacks in the relevant population is of no assistance for the issue in Title VII cases . . ." (Green v. Missouri Pacific Railroad Co., 1975). This kind of comparison sheds no light on discrimination due to protected class status. The court stated that statistical focus must highlight ". . . whether the questioned employment practice operates in a disparate manner upon a minority group . . ." and not whether the protected group impacted by the prohibited practice is statistically large (Green v. Missouri Pacific Railroad Co., 1975).

MoPac's business necessity defense was rejected despite proffering several reasons for its exclusionary policy. Employing language from *McDonnell Douglas vs. Green*, wherein Justice Powell postulated the petitioner in that case did not endeavor to exclude ". . . through some sweeping disqualification of all those with any past record of unlawful behavior, however remote, insubstantial, or unrelated to the applicant's personal qualifications as an employee" (McDonnell Douglas Corp. v. Green, 1973). The Eighth Circuit adopted Justice Powell's language from *McDonnell Douglas* because ". . . we perceive this comment [by Justice Powell] to suggest that a sweeping disqualification for employment resting solely on past behavior can violate Title VII where that employment practice has disproportionate impact and rests upon a tenuous or insubstantial basis" (Green v. Missouri Pacific Railroad Co., 1975). The circuit court refused to overturn the lower court's ruling limiting the class of plaintiffs, holding that the employment practice under consideration involved questions ". . . relevant only to the limited class of blacks with conviction records" (Green v. Missouri Pacific Railroad Co., 1975). Before obtaining a consumer report, the employer must provide the applicant with written notice, and the disclosure must be "clear and conspicuous" (Fair Credit Reporting Act, 1970). A post-report disclosure is required if the employer makes an adverse employment decision based upon the contents of the report. The affected employee must be provided with a copy of the report and an explanation of his or her rights before the adverse employment action is taken (Fair Credit Reporting Act, 1970).

Two years before its decision in *Green*, the Eighth Circuit held that applicants for the position of municipal firefighter should not be routinely disqualified for employment based upon prior felony or misdemeanor convictions (Carter v. Gallagher, 1971). The court was ". . . persuaded . . . that at least in the case was of aggravated offenses and multiple convictions, may have a bearing on the suitability of an applicant for a fire department position both from the standpoint of protecting fellow firemen and the public" (Carter v. Gallagher, 1971) The court ruled, however, that any rejection of an applicant because of a conviction must be based upon evidence that the applicant's behavior resulting in the conviction would prevent the applicant from performing the job (Carter v. Gallagher, 1971).

DISPARATE IMPACT

Disparate impact happens when a member of a protected class is denied an employment opportunity through the use of a neutral requirement or practice. "The disparate impact standard provides if an employer has a facially neutral hiring standard or policy . . . that has a disparate impact on a particular group, the court can strike down the test as violating Title VII unless the standard is directly related to future job performance" (Severin, 2013). Use of disparate impact analysis serves as a legal theory to thwart "artificial, arbitrary and unnecessary barriers to employment when . . . [those barriers] operate invidiously to discriminate on the basis of [protected class status]" (Griggs v. Duke Power Co., 1971). Title VII prohibits "not only overt discrimination, but also practices that are fair in form, but discriminatory in practice" (Griggs v. Duke Power Co., 1971). To prevail in a disparate impact case, a plaintiff must identify a specific practice or policy that has a disparate impact on a protected class. The Supreme Court has noted that the statistical proof required in disparate impact cases must meet an exacting standard. "[T]he plaintiff's burden in establishing a *prima facie* case goes beyond the need to show there are statistical disparities in the employer's work force" (Watson v. Fort Worth Bank & Trust, 1988). The Court in *Watson* observed that in disparate impact cases, a plaintiff must isolate and identify ". . . the specific employment practices that are allegedly responsible for any observed statistical disparities" (Watson v. Fort Worth Bank & Trust, 1988).

To discern whether a selection process disproportionately impacts a protected class, the Equal Employment Opportunity Commission (EEOC) promulgated a guideline known as the "four-fifths rule" (Uniform Guidelines on Employee Selection Procedures). The "four-fifths rule" provides that

> "A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact."

Under the EEOC guidelines, less significant differences in selection rates might lead to adverse impact "where they are significant in both statistical and practical terms" or where applicants have been disproportionately discouraged from seeking employment opportunities on the basis of protected class status (Uniform Guidelines on Employee Selection Procedures). Conversely, under the guidelines, "[g]reater differences in selection rate may not constitute adverse impact where the differences are based on small numbers and are not statistically significant . . . " (Uniform Guidelines on Employee Selection Procedures). The four-fifths rule is merely "a rule of thumb" and is generally not accepted by courts as dispositive of disparate impact inquiries. "It can be misleading when the number of applicants, and thus the number of test takers on which selection rates are based, is small" (Uniform Guidelines on Employee Selection Procedures). The EEOC guidelines allow for statistical analysis which compares the selection rate for one group versus another and an evaluation of "the individual components of the selection process" to show adverse impact (Walsh, 2016).

Two types of statistical measuring tools are generally engaged to determine the differences in selection rates in employment discrimination cases (Biddle, 1995). Those measuring tools are binomial distribution and hypergeometric distribution. "To calculate standard deviations, courts have applied the hypergeometric distribution test or the binominal distribution formula" (Employment Discrimination, Annual Survey, 1984). The hypergeometric test is employed by statisticians when the formula involves small samples. "The hypergeometric distribution test assumes that a sample population is finite and that each selection from that pool is not replaced in the pool once selected" (Employment Discrimination, Annual Survey, 1984). The binomial distribution test generally applies to sample populations in excess of 30 (Baldus & Cole, 1983). Binomial distribution tests apply to sample populations that have only two observed values. The binomial distribution, therefore, quantifies the probability that an employer will select a protected group member and the probability that the employer will select a non-protected group member (Baldus & Cole, 1983). Hypergeometric standard deviations will typically be higher than binomial distribution standard deviations. Thus, plaintiffs generally prefer the hypergeometric distribution, while defendants generally prefer the binomial distribution. The binomial statistical distribution approach was first employed by the U. S. Supreme Court in *Hazelwood School District v. United States (*Hazelwood School District v. United States, 1977*)*. In *Hazelwood*, an action was brought on behalf of African American teachers against the Hazelwood, Missouri School District located in St. Louis, alleging the school board engaged in a pattern or practice of under-employing African American teachers. The central issue was whether it was proper to compare the number of African American teachers in the district with the number of African American teachers in the relevant labor market (which included the City of St. Louis). The Supreme Court noted, "The Court of Appeals was correct in the view that a proper comparison was between the racial composition of Hazelwood's teaching staff and the racial composition of the qualified public school teacher population in the relevant labor market" (Hazelwood School District v. United States, 1977). Hazelwood represents an early use by the U. S. Supreme Court of binomial statistical comparison as it allowed for the comparison of black teachers employed by the district and qualified black teachers in the relevant labor pool. Statistical analysis, while useful, is rarely dispositive of the issue of disparate impact. However, "[a] plaintiff in a Title VII suit need not prove discrimination with scientific certainty, rather his or her burden is to prove discrimination by a preponderance of the evidence" (Bazemore v. Friday, 1986). "The United States Supreme Court has concluded that the EEO field will consider 2-3 standard deviations as the range for statistical significance in disparate impact cases. Therefore, as a general rule, if the difference between an expected value and the observed number is greater than 2-3 standard deviations, the difference is greater than that which can be attributed to chance. Standard deviations of five or more have been used in 'treatment' cases to infer intentional discriminatory treatment" (Hazelwood School District v. United States, 1977). The role of statistics in disparate impact cases has been succinctly stated by the U. S. Supreme Court. "We have emphasized the useful role that statistical methods can have in Title VII cases, but we have not suggested that any particular number of 'standard deviations' can determine whether a plaintiff has made out a *prima facie* case in the complex area of employment discrimination" (Hazelwood School District v. United States, 1977). While there is no strict formulaic approach adopted by the Court, at a minimum, courts may require a plaintiff to show that ". . . statistical disparities must be sufficiently substantial that they raise . . . an inference of causation" (Watson v. Fort Worth Bank & Trust, 1988). It should be noted that while a substantial disparity might raise an inference of causation it does not a *fortiori* lead to a conclusion of disparate treatment. Courts and statistical experts have determined that when "the number of standard deviations of the difference is less than approximately two (technically ± 1.96); or the probability of chance occurrence is greater than 5%; the actual number is 'close enough' to the expected number to be consistent with the result of a nondiscriminatory selection process" (Baker, 2011).

Plaintiff's burden of proof in a disparate impact case is to make ". . . a comparison between the composition of the qualified . . . population in the relevant labor market. It is such a comparison . . . between the racial composition of the qualified persons in the labor market and the persons holding at-issue jobs . . . that generally forms the proper basis for the initial inability in disparate impact cases" (Ward's Cove Packing Co., Inc. v. Antonio, 1989). Simply showing a racial imbalance in the number of whites versus non-whites in certain jobs is a misconstruction of the role of statistics in disparate impact cases. "Such comparison(s) fundamentally misconceiv(es) the role of statistics in employment discrimination cases" (Hazelwood School Disctrict v. United States, 1977). In situations where labor market statistics may not be ascertainable, courts ". . . have recognized that certain other statistics . . . such as measures indicating the racial composition of 'otherwise-qualified applicants' for at-issue jobs . . . are equally probative for this purpose" (Ward's Cove Packing Co., Inc. v. Antonio, 1989). One of the first steps in scrutinizing statistical analysis in disparate impact cases is to determine whether or not the analysis conducted is accurate and reliable. In *Watson*, the Supreme Court noted that identifying "statistical disparities in the employer's work force" alone is insufficient; a plaintiff must tender "statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion of applicants for jobs or promotions because of their membership in a protected group" (Watson v. Fort Worth Bank & Trust, 1988). Credibility of the statistical evidence and the analysis of such evidence are subject to challenges, and experts presenting statistical analysis are subject to impeachment. Neither the ". . . Courts nor defendants are obliged to assume that plaintiffs' statistical evidence is reliable" (Watson v. Fort Worth Bank & Trust, 1988). The Court in *Watson* elaborated on the means and manner in which statistical evidence may be challenged in disparate impact.

"If the defendants in a Title VII suit believe there to be a reason to discredit plaintiff's statistics that does not appear on their face, the opportunity to challenge them is available to the defendants just as in any other lawsuit. They may endeavor to impeach the reliability of the statistical evidence, they may offer rebutting evidence, or they may disparage . . . the probative weight which the plaintiffs' evidence should be accorded" (Dothard v. Rawlinson, 1977). The Court, in *Watson*, further noted that while it not endeavor to enumerate all the weaknesses statistical evidence is susceptible, ". . . typical examples include small or incomplete data sets and inadequate statistical techniques" (Dothard v. Rawlinson, 1977).

THE BUSINESS NECESSITY DEFENSE

If the plaintiff puts forth a *prima facie* case, the defendant is required to justify the employment practice by demonstrating it arises from "business necessity" and is "job related." The U. S. Supreme Court pronounced the business necessity defense in *Griggs v. Duke Power Company* (Griggs v. Duke Power Co., 1971). In *Griggs*, the Court addressed the question of whether Title VII prohibits an employer ". . . from requiring a high school education or passing of a standardized general intelligence test as a condition of employment in or transfer to jobs . . ." when the test is not ". . . significantly related to successful job performance . . ." In concluding that the testing practice violated Title VII, the Court reasoned, "[t]he Act proscribes not only overt discrimination but also practices that are fair in form, but discriminating in operation. The touchstone is business necessity" (Griggs v. Duke Power Co., 1971). A concise definition of business necessity was articulated in *Robinson v. Lorillard Corporation* (Robinson v. Lorillard Corp., 1971). "The test is whether there exists an overriding legitimate business purpose such that the practice is necessary to the safe and efficient operation of the business. Thus, the business purpose must be sufficiently compelling to override any racial impact" (Robinson v. Lorillard Corp., 1971). The *Griggs* court rejected the assertion that Duke Power lacked discriminatory intent because of the company's effort in providing financial assistance to the employees to overcome the education barrier. The focus of Congress in enacting Title VII, the Court noted, was ". . . the consequences of the employment practices, not simply the motivation" (Griggs v. Duke Power Co., 1971). The second component in adverse impact cases requires a step beyond business necessity. "More than that, Congress has placed on the employer the burden of showing that any given requirement must have a manifest relationship to the employment in question" (Griggs v. Duke Power Co., 1971).

The Supreme Court issued decisions in two other cases which solidified the law of business necessity in disparate impact cases. In 1975, the Court decided *Albermarle Paper Company v. Moody.* The Court stated

> "If an employer does then meet the burden of proving that its tests are "job related," it remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable racial effect would also serve the employer's legitimate interest in 'efficient and trustworthy workmanship'" (Albermarle Paper Co. v. Moody, 1975).

The Court in *Albermarle* retreated from the stringent business necessity standard set forth by a unanimous Court in *Griggs*. Justice Blackmun in a concurring opinion would require that pre-employment tests be ". . . fairly related to the job skills or work characteristics desired" (Albermarle Paper Co. v. Moody, 1971). Justice Berger seemingly accepted a less stringent standard (Albermarle Paper Co. v. Moody, 1971).

The third case in the Supreme Court's trilogy firmly establishing business necessity as a defense in disparate impact cases was *Dothard v. Rawlingson*. *Dothard* involved the use of Alabama's height and weight requirement, which had a disparate impact on women who sought work as prison guards. A majority of the Supreme Court found that the height and weight requirement violated Title VII because the "appellants produced no evidence correlating the height and weight requirements with the requisite amount of strength thought essential to good job performance" (Dothard v. Rawlinson, 1977). The Court held, however, that the State of Alabama had carved out a permissible *bona fide* occupational qualification exception due to the rampant violence and security related issues at the maximum security prison (Dothard v. Rawlinson, 1977).

CREDIBILITY OF EXPERTS

Two recent cases in the federal courts have dealt with experts and data analysis. In each case, the courts employed the standard set forth in Federal Rule 702 to determine the admissibility of the expert's testimony. In *EEOC v. Freeman*, the plaintiff filed a complaint in the U. S. District Court for the District of Maryland alleging violations of Section 706 and 707 of Title VII of the 1964 Civil Rights Act. The plaintiff also contended that the defendant engaged in a "pattern or practice" of discriminating against African American job applicants by using adverse credit history as a hiring requirement and against African American, Hispanic and male job applicants by using criminal history as a criterion. Such criteria, claimed the EEOC, have a disparate impact on identifiable protected classes and are not job-related or consistent with business necessity. The EEOC sought injunctive relief prohibiting the defendant from the use of credit checks and criminal history when hiring employees. The agency also sought compensatory relief to make the plaintiffs whole, including back pay, prejudgment interest and other monetary relief. The district court entertained several motions filed by the defendant, including a motion to dismiss claims related to "hiring decisions made . . . 300 days before the January 17, 2008 charge of discrimination that served as the legal predicate for the EEOC's claim" (EEOC v. Freeman, 2013). Partial summary judgment was granted on hiring decisions based on criminal history information made ". . . 300 days before the EEOC notified the defendant that its investigation would include race as a basis for the use of criminal history in making employment decisions" (EEOC v. Freeman, 2013). The parties stipulated that the claim filed against the defendants alleging discrimination against Hispanics would be dismissed (EEOC v. Freeman, 2013).

The defendant, *Freeman*, is a family-owned company which employs over 3,500 full-time and 25,000 part-time and seasonal workers. The company has offices throughout the United States. Due to problems of theft, embezzlement, substance use and workplace violence, *Freeman* began conducting background checks. The employee background checks were implemented for the following purposes:

(1) "avoid exposure to negligent hiring/retention suits;
(2) increase security of Defendant's assets and employees;
(3) reduce liability from inconsistent hiring or screening practices;
(4) proactively reduce the risk of employee-related loss; and
(5) mitigate the likelihood of an adverse incident occurring on company property that could jeopardize customer or employee confidence" (EEOC v. Freeman, 2013).

At the time of the district court's decision ". . . the EEOC's current 'credit class' consist[ed] of 51 African-Americans who allegedly were unlawfully excluded from employment between March 23, 2007 and August 11, 2011" and ". . . a current 'criminal class' consist[ing] of 83 African-American males who allegedly were unlawfully excluded from employment between November 30, 2007 and July 12, 2012" (EEOC v. Freeman, 2013). After a distillation of the issues and facts and a resolution of the remaining classes whose claims survived the various dispository motions, the focal point of *Freeman* turned on the credibility of the experts, the data and the analysis of the data proffered as evidence. The gravamen of the case was whether the analysis of the data presented by the EEOC's expert witnesses credibly demonstrated that African American applicants failed the defendant's credit and criminal background checks at a significantly higher rate than females and non-African Americans.

The court began its discussion by stating the burden of proof imposed on plaintiffs by citing *Ward's Cove Packing v. Antonio*. "The Plaintiff bears the burden of proving discriminatory impact by showing statistical disparities between the number of protected class members in the qualified applicant group and those in the relevant work-force" (EEOC v. Freeman, 2013). The EEOC offered an expert who ". . . analyzed data produced by the Defendant" which allegedly demonstrated that African Americans failed to meet the defendant's required credit score "at a significantly higher rate than other races" and that males and African Americans failed criminal history checks "at a significantly higher rate than females and non-blacks." (EEOC v. Freeman, 2013). The defendants asserted ". . . that experts' conclusions are based on non-reliable data and are rife with analytical errors" as well as inadmissible (Federal EEOC v. Freeman, 2013). The district court turned to Federal Rule of Evidence 702 as its baseline to begin its dismantling of the EEOC's expert testimony. Rule 702 provides that expert testimony is admissible if

(1) "the testimony is based on sufficient facts or data
(2) the testimony is the product of reliable principles and methods, and
(3) the witness has applied the principles and methods reliably to the facts of the case" (Federal Rules of Evidence, Rule 702).

A plethora of analytical, sampling and statistical errors were uncovered by the court. Judge Titus noted the experts' ". . . analysis suffers primarily because it was not based on a random sample of accurate data from the relevant applicant pool and time period" (EEOC v. Freeman, 2013). As noted above, the U. S. Supreme Court in *Ward's Cove Packing* defined the parameters of relevant labor pool data in disparate impact cases. The "proper comparison [is] between the racial composition of [the at-issue jobs] and the racial composition of the qualified . . . population in the relevant labor market" (Ward's Cove Packing Co., Inc. v. Antonia, 1989). The district court found there was no such comparison in *Freeman*. Citing *Ward's Cove*, the district court did not equivocate. "The general population pool 'cannot be used as a surrogate for the class of qualified job applicants because it contains many persons who have not (and would not be)' applying for a job with the Defendant" (EEOC v. Freeman, 2013). The district court issued a stinging rebuke of the EEOC's expert, opining that he "cherry picked" data from outside the relevant time period, failed to "perform a mathematical extrapolation to the time periods" outside the scope of his analysis, "deliberately ignored the data available from the relevant time period" and showed performance for data outside the time period (EEOC v. Freeman, 2013). Judge Titus wrote that the number of errors found in the expert's database was "mind-boggling." Moreover, the errors were so pervasive, they alone could ". . . render his disparate impact conclusions worthless" (EEOC v. Freeman, 2013).

The second case, *EEOC v. Kaplan Higher Education Corp. (EEOC v. Kaplan Higher Education Corp., 2014),* also involved the employment of an expert in an attempt by the EEOC to prove disparate impact. Kaplan is an educational institution which provides undergraduate and graduate degrees. Its students apply for and receive federal financial aid; therefore, Kaplan's employees are privy to students' financial information. Due to concerns related to breaches of confidentiality and fraud perpetrated by Kaplan's employees, the institution implemented measures to curtail the abuses. Among the measures initiated were credit checks on senior level executives, those employed in accounting and others with access to student financial aid information. The credit checks were conducted by a third party ". . . which reports . . . whether the applicant has ever filed for bankruptcy, is delinquent on child support payments, has any garnishments on earnings, has outstanding civil judgments exceeding $2,000 . . ." or has a fraudulent Social Security number *(EEOC v. Kaplan Higher Education Corp., 2014).* If the applicant's credit history reveals any of the designated problems, the applicant file is reviewed by Kaplan *(EEOC v. Kaplan Higher Education Corp., 2014).* The race of applicants is not disclosed during this process. The EEOC, however, subpoenaed motor vehicle records and was able to identify the race of many of the applicants. An expert for the EEOC devised a process called "race rating" whereby he ". . . purported to identify the race of each person in these drivers' license photos" *(EEOC v. Kaplan Higher Education Corp., 2014).* While according to the records the EEOC's expert had no ". . . expertise in constructing methodologies to identify race by visual means," he employed raters to view the job applicants' driver's license photographs to ". . . classify the person's race in one of five ways: 'African American,' 'Asian,' 'Hispanic,' 'White' or 'other'" *(EEOC v. Kaplan Higher Education Corp., 2014).* The question presented to the circuit court was whether the district court erred in concluding that the EEOC's expert's methodology and process were reliable under the standards set forth in *Daubert v. Merrell Dow Pharmaceuticals, Inc*. (Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993). The circuit court noted that the district court applied the *Daubert* standard and determined the EEOC's expert's "methodology flanked them all" *(EEOC v. Kaplan Higher Education Corp., 2014).* The Sixth Circuit Court of Appeals, "unimpressed" with the analysis and methodology employed by the expert, rejected the EEOC's disparate impact claim. The court found that the agency failed to meet the standards for admissibility articulated by the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc*. (Daubert v. Merrell Dow Pharmaceuticals, Inc. 1993). The circuit court signals its conclusion in its opening paragraph ". . . [T]he EEOC sued the defendants (Kaplan) for using the same type of background check that the EEOC itself uses" *(EEOC v. Kaplan Higher Education Corp., 2014).* Because of concerns associated with "overdue just debts" and "unethical acts . . . to meet financial obligations . . . the EEOC runs credit checks on applicants for 84 of the agency's 97 positions" *(EEOC v. Kaplan Higher Education Corp., 2014).* The court held "[t]he EEOC brought this case on the basis of a homemade methodology, crafted by a witness with no particular expertise . . . administered by persons with no particular expertise . . ., and tested by . . . only the witness himself" *(EEOC v. Kaplan Higher Education Corp., 2014).* The decision was reached largely because the theories employed by the expert had not been subjected to peer review, and the rating system was devoid of any standard by which to classify the applicants other than simply "eyeball[ing] the DMV photos" *(EEOC v. Kaplan Higher Education Corp., 2014).*

CONCLUSION

Employers who desire to conduct credit and criminal background checks on prospective employees must ensure that a sound business reason justifies such inquiries. Such checks raise Title VII and privacy issues. These types of pre-employment checks are not novel, as they have existed for decades. Past and recent cases suggest that employers recognize a disparate impact claim will likely result when a significant number of a protected class is disqualified from employment opportunities due to the use of such checks. Employers must also realize that the Federal Rules of Civil Procedure allow for the collection of employment data. The EEOC also requires that employers collect and maintain data of applicants. Title VII, however, does not prevent the appropriate use of credit and criminal history information as a part of the employment process. This is especially true when the employer is able to show a relationship between security-sensitive jobs and the need to screen certain employees as a means to protect company assets and other employees.

Employers should not develop general exclusionary policies or practices which automatically exclude from employment consideration any applicant who has been convicted of a crime, even a felony. Employers should develop policies pertaining to convictions that consider the seriousness and nature of the crime, the time that has elapsed since the conviction, the nature of the crime in relation to the job sought, and the applicant's rehabilitation since the conviction. Employers should exclude applicants from employment consideration based on a conviction only to the extent they are able to document that the behavior leading to the arrest and conviction would prevent the applicant from successfully performing the particular job sought by the applicant.

Credit checks and criminal background checks are likely to become generally accepted as part of the screening process in hiring decisions unless those who challenge these processes are able to present experts capable of offering credible analysis and data based upon reliable scientific principles and method. Moreover, one must question the propriety of the EEOC's filing lawsuits against employers who use background checks for essentially the same reason as does the EEOC.

Experts who are employed to analyze and disaggregate data must be cognizant of the requirements of Federal Rule of Evidence 702. The testimony of experts must be reliable and based upon statistics, data and principles that are credible. Claims of disparate impact can only be supported by a reliable expert who bases his opinion on complete and reliable data sets. Plaintiffs are not allowed to simply raise a statistical inference of disparate impact. Experts must provide analysis and complete data sets that are free of analytical and methodological errors and that encompass the appropriate relevant labor market comparisons for the applicable time period.

REFERENCES

Albermarle Paper Co. v. Moody, 1975 (422 U.S. 405, 1975).

Baker, Mary Dunn (2011) Statistical Analysis of the Outcomes of Employment Selection Process – It's All (or at Least Mostly) About the Model. National Conference on EEO Law, New Orleans, LA, April 6.

Baldus, David C. & James W.L. Cole (1983) *Statistics Proof of Discrimination*, 308 N. 36 (1980 & Supp. 1983).

Bazemore v. Friday, 1986 (478 U.S. 385, 1986).

Biddle, Richard E. (1995) Disparate Impact Reference Trilogy for Statistics, *Labor Law Journal*, 46(11): 651-660.

Carter v. Gallagher, 1971 (452 D.2d 315, 8th Cir. 1971).

Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993 (509 U.S. 5779, 1993).

Dothard v. Rawlingson, 1977 (433 U.S. 321, 1977).

EEOC v. Freeman, 2013 (961 F. Supp. 2d, 2013).

EEOC v. Kaplan Higher Education Corp., (748 F.3d 749, 6th Cir. 2014).

Fair Credit Reporting Act 2012 (15 U.S.C. et seq, 2012).

Federal Rules of Evidence, Rule 702.

Goldstone, Lawrence (2011) *Inherently Unequal: The Betrayal of Equal Rights by the Supreme Court, 1965-1903.* London: Walker Books.

Grant v. Bethlehem Steel Corp., 1980 (635 F2d 1007 (2d Cir. 1980).

Green v. Missouri Pacific Railroad Co., 1975 (523 F2d 1290 8th Cir. 1975).

Griggs v. Duke Power, 1971, 401 U.S.C 424 1971).

Hazelwood School District v. United States, 1977 (433 U.S. 299, 1977).

McDonnell Douglas Corp. v. Green, 1973, 411 U.S. 792 1973.

Robinson v. Lorillard Corp., 1971 (444 F.2d 791, 4th Cir. 1971).

Severin, Chade (2013) We Built This City: The Legality of Community Benefit Agreements for Big Box Construction Under Title VII and the Equal Protection Clause, *Columbia Journal of Race and Law,* 3(2): 215-252.

Title VII, Civil Rights Act of 1964, 42 U.S.C. 2000e-2(a) 1964).

Uniform Guidelines on Employee Selection Procedures, 1978 (29 CFR Part 1607).

Walsh, David J. (2016) Employment Law for Human Resource Practice, 5th Edition, Cengage Learning.

Ward’s Cove Packing Co., Inc. v. Antonio, 1989 (496 U.S. 542, 1989).

Watson v. Fort Worth Bank & Trust, 1988 (48 U.S. 977).