E-COMMERCE SITE DESIGN, ONLINE CONSUMER BEHAVIOR AND BUSINESS VALUE
A Doctoral Dissertation Proposal

(DRAFT VERSION)

Jungpil Hahn
Information and Decision Sciences
Carlson School of Management
University of Minnesota
Email: jhahn@csom.umn.edu
Last Revised: March 27, 2002

ABSTRACT

This research consists of three studies that offer an in depth analysis of the impact of e-commerce web site design. The first study examines online consumer behavior – does online consumer behavior follow the process model of purchasing?, can we observe the various motivations for online shopping (e.g., directed buying, hedonic browsing, search and deliberation vs. knowledge building)?, how are these online behaviors instantiated?, does online consumer behavior change over time? If so, how? We develop a theoretical typology of online consumer behavior and test it by uncovering taxonomies of online consumer behavior from clickstream data. The second examines the measurement of the effectiveness e-commerce web site design – how can we conceptualize and measure the effectiveness of an e-commerce web site?, can we identify web site design inefficiencies? For this, we define the concept of web site design efficiency as the degree to which the intended usage of a web site matches a customer’s actual usage given her current shopping goal. We empirically evaluate the effectiveness of an e-commerce web site using data envelopment analysis (DEA) on clickstream data. Finally, the third study investigates the business value of e-commerce site design – does web site design efficiency translate into business value?, what are the sources of business value created? We empirically assess the business value of web site design efficiency on process level measures as well as on customer value using clickstream data.

Keywords: Electronic commerce, e-tailing, web site design, online consumer behavior, web site design efficiency, business value, customer value.
BACKGROUND AND MOTIVATION

Since the crash of the DotComs in the American stock market in May and June of 2000, the evaluation of e-commerce web sites in terms of business value has become increasingly important (Varianini and Vaturi, 2000). No longer are the venture capital firms willing to make portfolio investments in e-commerce properties that only present future return on investment (ROI) opportunities. Instead, they seek e-commerce firms with more immediate opportunities, especially firms that can demonstrate an increasingly well-developed discipline for evaluating investments in web-related software development and e-commerce business models, so that the risks and uncertainties of investing in this emerging market are better balanced with the rewards. Indeed, the industry has progressed to the second phase of e-commerce. During the first phase of e-commerce, the goal for most companies was to secure a share of the virtual market space through an online presence by attracting as many visitors as possible to their web site, whereas now, as e-commerce begins to mature, the ability to conduct online operations justified by ROI is the only way an e-business can survive.

Recent industry analyses, however, point out that e-commerce retailers are earning low scores on ROI, by failing to meet consumers’ purchase needs with the poor usability and errant designs of their web-based storefronts (Souza, Manning, Sonderegger, Roshan, and Dorsey, 2001). For example, a study by Zona Research reported that 60% of web-savvy users dropped out of the purchasing process because they could not find the products in the online retailers’ web sites (Zona Research, 1999). Another study conducted by A.T. Kearney showed that 80% of experienced online shoppers gave up shopping on e-commerce web site sites due to problems they encountered while interacting with the web site (Rizzuti and Dickinson, 2000). Yet another study conducted by Creative Good showed that 43% of purchase attempts ended in failure due to poor utility of the web sites (Rehman, 2000). This shortfall in realized value compared to the potential value that web-based selling approaches offer is dramatic. The Creative Good study points out that this level of failed purchase attempts is consistent with an estimated loss of $14 billion in sales for e-commerce retailers in the 2000 Christmas-New Year’s holiday shopping season alone. Recent academic research reinforces the picture that emerges. Apparently the quality of the online customer experience that effectively-design web sites create not only has a positive effect on the financial performance of a firm, but also possesses the potential to create unique and sustainable competitive advantage for Internet-based sellers and other e-commerce firms (Rajgopal, Venkatachalam, and Kotha, 2001).

Developing, launching and maintaining an e-commerce web site is a significant investment for e-commerce firms. Simple e-commerce web sites can cost $1-2 million per year for setup and maintenance, whereas more sophisticated web sites with dynamic capabilities require annual investments nearing $52
million (Rizzuti and Dickinson, 2000; Dalton, Hagen, and Drohan, 2001). Despite the importance of web site development requiring such significant investments, the process of designing high quality web sites for e-commerce is still more of an art than a science. E-commerce companies are still relying largely on intuition when it comes to designing their web sites (Hahn, Kauffman, and Park, 2002). To make matters worse, design changes and their impacts are not tracked, making it impossible to measure the benefits of site design (Wallach, 2001). This situation brings to the foreground the importance of value-driven evaluation and design of e-commerce web sites. However, e-businesses are facing difficulties due to the lack of proven tools and methods for accomplishing this. It is clear that e-commerce companies are in dire need of disciplined methods for justifying the business value of IT investments in systems design.

The primary objective of this research is to bring the art of e-commerce systems design closer to management science. A complete understanding the impact of e-commerce web site design on e-business management can be a daunting task. Hence, I propose a three-part investigation of the impact of e-commerce web site design. First, I investigate online consumer behavior – how consumers (i.e., users) use e-commerce web sites for online purchasing. Second, I examine the effectiveness e-commerce web site design – how to measure the effectiveness of web site design. And third, I investigate the business value of e-commerce site design – how the quality of the web site design is effectively translated into business value.

This proposal is outlined as follows. In the next section, we present a review of the relevant literature. In the following three sections, we outline in detail the specifics of the three studies with background and motivation, research questions, research approach and expected contributions. Finally, we conclude the proposal with expected contributions of the overall dissertation, anticipated pitfalls and challenges along with contingency plans, and an overview of the research timeline for completing the dissertation.

LITERATURE REVIEW

This research investigates the impact of e-commerce web site design on consumer behavior. We briefly review the relevant literature related to online consumer behavior and e-commerce web site evaluation and design.

Online Consumer Behavior

Consumer purchasing behavior is typically characterized as a five-stage process model consisting of 1) need arousal, 2) information search, 3) product evaluation, 4) purchase decision and 5) post-purchase
evaluation (Bettman, 1979; Howard and Sheth, 1969; O'Keefe and McEachern, 1998). The stages of the purchasing process need not be restricted to either all online or all offline – different stages may be accomplished online or offline. For example, the need for an item may be aroused offline while being exposed to a print advertising in a magazine. Information search and product evaluation may occur both online and offline. The actual purchase decision and transaction may be conducted online at an e-commerce site, and finally, post purchase evaluation may occur offline as the physical product is delivered to the customer’s home. It is important to empirically validate whether all stages of the consumer purchasing decision making process is also conducted online. If not, what stages of the purchasing decision making process are conducted online versus offline1.

The nature of consumer behavior is also affected by the consumers’ motivation for purchasing. The consumer behavior literature identifies two distinct motivations for purchasing – goal-directed (or utilitarian) vs. experiential (or hedonic). The two motivations for shopping bring about considerable differences in consumer behavior (Babin, Darden, and Griffin, 1994). Goal-directed shopping, which typically occurs with a near-immediate purchase horizon, entails a highly focused information search process whereby consumers are seeking information specific to products in their consideration set so that it can be used in the purchase decision making. The focus of goal-directed shopping is on efficiency of the purchasing process and the desired outcomes are increased product knowledge, better purchase decisions, and greater satisfaction with the purchase outcome. On the other hand, experiential shopping focuses on the recreational or hedonic motives of the consumer. Experiential shopping entails on-going search without a specific purchase horizon (Bloch, Sherrrell, and Ridgway, 1986). Wolfinbarger and Gilly (2001) argue that goal-directed shopping will be predominant in online contexts as compared to experiential shopping. The rationale behind their argument is that time-starved consumers are more likely to adopt online channels to minimize the costs associated with going to the physical stores (Bellman, Lohse, and Johnson, 1999). However, other authors argue that an element of fun needs to be present in online contexts. For example, Hoffman and Novak (1996) propose the concept of flow (Csikszentmihalyi, 1990) and argue that the general customer experience online is important for e-commerce success. They argue that an online shopping environment can bring about a state of flow, which in turn leads to more exploratory browsing behaviors, which is characteristic of experiential shopping behaviors. Recent research has investigated different online shopping strategies with clickstream data. Moe (2001) theoretically proposes and empirically validates a typology of online store visits consisting of four shopping strategies differentiated by whether the purchase horizon is immediate

---

1 We may not verify whether the unobserved behaviors actually occur offline. However, we may assume that whatever stage is not observable online would be conducted offline.
or future and whether the search behavior is directed or exploratory (see Table 1). Through cluster analyses of clickstream data, she finds that visits can be categorized as a buying browsing, searching or knowledge building visit based on observed in-store navigation patterns with respect to the content of the web pages viewed. Given the contradictory arguments and evidence in the literature, this research investigates the whether different motivations for shopping are also observable online. And if so, how are the different shopping strategies instantiated in terms of web site usage.

<table>
<thead>
<tr>
<th>Table 1 - Typology of Shopping Strategies²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase Horizon</td>
</tr>
<tr>
<td>Immediate</td>
</tr>
<tr>
<td>Future</td>
</tr>
</tbody>
</table>

The previous discussion paints a somewhat static view of online consumer behavior. However, we cannot expect online consumer behavior to be fixed over time. It is important to consider the dynamic aspects of online consumer behavior since retention rather than attraction or conversion has become one of the most important metrics for evaluating the performance of e-businesses (Agrawal, Arjona, and Lemmens, 2001). Hence, we not only need to understand how consumers shop online but also how their online behaviors evolve over time in order to meet their changing needs. As consumers gain experience at navigating a site, they become increasingly efficient at using the site. Johnson, Bellman and Lohse (2000) model customers decreasing site visit duration at online e-commerce retail sites³ with the power law of practice (Newell and Rosenbloom, 1981). Whereas the Johnson et al. study investigated the dynamic nature of site visit duration between visits (i.e., how long do customers stay at e-commerce sites as they revisit these sites?), other authors have also investigated the dynamic nature of visit rates (i.e., do customers revisit sites more (or less) frequently as they revisit these sites?). Moe and Fader (2001a) find that overall site revisit frequencies tend to decrease over time but that those customers with increasing visiting rates have a higher propensity to purchase. Finally, the dynamic nature of customers’ browsing behaviors within a site visit session has also been investigated. Bucklin and Sismeiro (2001) find that a customer’s propensity to continue browsing a site changes dynamically with respect to the depth of the current visit and the number of repeat visits to the site. A limitation of these studies stems from the fact that they do not consider what the customers actually do within the visits – customers are assumed to perform the same purchasing tasks for each revisit. However, as discussed above, consumers may exhibit different shopping strategies (e.g., buying, searching, browsing or knowledge building) as they continue

² Adapted from Moe (2001).
³ The Johnson et al. study (2000) investigated 36 e-commerce sites (combined) in the travel, book and music industry.
to use a site. For example, a customer may first focus on efficient purchasing (e.g., buying or searching),
but as she gains more experience with navigating the site, she may opt to use the site for more hedonic
purposes (e.g., browsing or knowledge building) with the extra time resulting from the increased
efficiency of the purchasing tasks. In this research, we investigate the dynamic nature of online consumer
behavior – does online consumer behavior change over time? and if so, how?

E-commerce Web Site Evaluation

Recent reports concerning the poor quality and usability of web sites have led researchers and
practitioners to express increasing interests in the methodologies and approaches that are used to conduct
evaluations of web sites. Traditional approaches to web site evaluation fall into three different categories:
(Ivory and Hearst, 2000).

- **Testing**: Users perform representative tasks with a given web site and usability problems are
determined based on the range of observed user interactions (e.g., Spool, Scanlon, Schroeder,
Synder, and DeAngelo, 1999).

- **Inspection**: Usability experts use a set of criteria (e.g., web usability heuristics, such as those
suggested by Nielsen (1994)) to identify potential usability problems in the web site design (e.g.,
Nielsen and Mack, 1994).

- **Inquiry**: Users provide feedback on the web site via structured interviews, participation in focus
groups, responding to structured surveys, etc. (e.g., Schubert and Selz, 1999).

The methods mentioned above have been adopted from the discipline of user interface (UI) testing
within the broader field of human-computer interaction (HCI). However, even though these approaches
have been successfully applied for the evaluation of user interfaces of traditional IS applications, they are
not perfectly-suited for web-based applications. In fact, the Internet and World Wide Web introduce
several critical issues that impede with the wide applicability and effectiveness of these methods. For
example, web sites are very frequently updated and redesigned, which makes the cost of recruiting test
users and experts for each redesign overly excessive for most organizations with limited labor and capital
resources. Furthermore, it is difficult to reconstruct a representative environment for user testing since
developers do not have control over technologies that users employ when they browse. For example, they
are unable to control for the various web browsers that users select, which often have the impact of
changing the appearance of web pages. Moreover, they are unable to pre-specify the connection speed
with which a user accesses web pages at a site, affecting screen presentation and download times, and the
overall quality of the user experience (Nielsen, 2000). It is also important to emphasize that users of web-
based applications are most often customers, which is untypical of traditional IS applications within a firm. As a result, greater constraints are placed on what a designer must do to create a desirable setting for system use by a user/customer.

The above review of the literature on general web site evaluation suggests that the major focus has been on evaluating the usability of web sites. Even though ease of use of web sites may be an important and necessary condition for the success of a web site (i.e., people should be able to perform necessary tasks), the restricted concentration on usability only limits the applicability of these methods in linking design to performance of web sites.

The methods for web site evaluation described above also can be employed for e-commerce web site evaluation. For example, in terms of user testing, several authors point out that it is possible to simulate the use of an Internet-based seller’s web site by engaging users who are given the task of finding and purchasing items from an e-commerce storefront (e.g., Rizzuti and Dickinson, 2000; Spool et al., 1999). In terms of inspection, Gomez Advisors (www.gomez.com) rates e-commerce web sites based on multiple criteria. The criteria typically include ease of use, customer confidence, on-site resources, and relationship services, among others, depending on the business category of the firm’s web site that is being evaluated. Finally, Schubert and Selz (1999) have proposed the Web Assessment Tool, a survey-based inquiry method that is applied to evaluate the effectiveness of different phases of market transactions for e-commerce web sites. The phases include customer information gathering on potential products and services, reaching agreement when negotiations between customers and suppliers take place, and settling on the payments and logistics when the products and services are actually delivered.

Even though some of these approaches to e-commerce web site evaluation are useful, the fundamental limitations of each approach still hold. User testing often is not cost-effective, and in many situations, it fails to properly measure the outcomes associated with a range of users’ or customers’ experience with a web site. Inspection and inquiry, on the other hand, may generate useful insights about where to focus the firm’s efforts for web site maintenance and future additions to the software functionality. But these approaches may not be as effective in showing the details of what actually needs to be done to improve the web site to increase ROI.

Furthermore, the narrow focus on usability only also limits their relevance. So, even though the quality of the online customer experience (especially in terms of usability) has been shown to positively affect the overall firm performance of e-commerce firms (i.e., financial performance) (Rajgopal et al., 2001), it still is unclear how the usable design of e-commerce web sites translates into increased ROI and firm performance. Thus, it is imperative for e-commerce firms to understand how their site is performing against business performance metrics (e.g., customer acquisition, customer satisfaction, customer loyalty, online sales and
revenue, merchandizing effectiveness etc.). However, usability evaluations for most e-commerce web sites fail to provide the missing link between usability and business performance, and as a result, do not provide management decision makers with actionable information through which they can leverage ROI.

E-commerce Web Site Design

Previous research has investigated the impact of e-commerce site design in a number of ways. Conceptually, Hoffman and Novak (1996) propose the concept of *flow* for understanding consumer behavior in computer-mediated environments. They argue that, in order for e-commerce sites to induce a state of flow to online consumers, the design of the computer-mediated environment should provide an appropriate balance between users’ skills and the challenges of the online interaction. However, how exactly e-commerce sites should be designed to provide an appropriate balance between skills and challenges was not discussed. Another stream of research related to the design of e-commerce sites argue that the design of the electronic storefront needs to support online consumer behavior. For example, various authors argue that the design of e-commerce systems should support the various stages of the customers’ purchasing decision process (Maes, Guttman, and Moukas, 1999; O'Keefe and McEachern, 1998). Similarly, Miles, Howes and Davies (2000) present a framework for e-commerce technologies. This stream of research basically provides a conceptualization of the design space for e-commerce web sites – what basic systems functionalities need to be provided and what are various technological instantiations of such functionalities. The actual impact of the design of e-commerce sites, however, has not been investigated. Finally, the literature in Human-Computer Interaction (HCI) has also dealt with the design of e-commerce sites. The central focus of this line of work lies in the usability of web sites (Anderson, 2002). For example, Jakob Nielsen, a guru in hypertext usability, proposed general usability heuristics for web sites (Nielsen, 1999, 2000). A limitation in this stream of research is that the recommendations are process-oriented and context-free – the basic argument is that the design process needs to be user-centered and incorporate usability testing; if the process is followed then a high quality web site will result (Cooper, 1999; Nielsen, 2000).

There has also been significant empirical research investigating the impact of e-commerce web site design. In one of the earliest empirical studies of e-commerce site design, Lohse and Spiller (1998a; 1998b) have explored the effect of interface design on e-commerce site traffic and sales. They find that various interface design features lead to increased traffic and/or sales. For example, having a FAQ (frequently asked questions) section on the site leads to increased traffic, improved product lists lead to increased sales, whereas having more links between product categories (e.g., links to related or complementary products) and featuring product promotions on the site homepage lead to increased both
traffic and sales. One limitation of their study is that the study was exploratory in nature and did not provide a theory as to why the various interface design features would impact traffic or sales. More recently, Liu and Arnett (2000) surveyed webmasters from Fortune 1000 companies to investigate the factors that lead to design quality of web sites. They find that information quality, learning capability, playfulness, system quality, system use and service quality lead to high quality web site designs, which in turn results in perceptions of attractiveness, dependability, reliability, trustworthiness and pleasure.

Kotha, Rajgopal and Venkatachalam (2001) conceptualize online customer experience as consisting of five dimensions (i.e., web site navigability, relationship services, onsite resources, price leadership, and customer confidence) and find that two dimensions (web site navigability and relationship services) lead to increased traffic to the web site, whereas all five dimensions moderate the relationship between traffic and sales (i.e., conversion from browsers to buyers). Other studies that do not focus directly on the design of e-commerce sites also include various site design variables. For example, Chen and Hitt (2001), in their measurement of switching costs in the online brokerage industry, investigate the impact of site design variables such as provision of personalization, and ease of use of the web site on switching and attrition. Xue, Harker and Heim (2000) propose the concept of website efficiency and investigate the impact of various e-commerce site features (e.g., navigation, shopping tool, order fulfillment support, customer decision support, community and customization) on customer satisfaction and customer loyalty. They find significant impact of web site navigation and customer support features on website efficiency, which in turn impacts customer satisfaction and loyalty. A limitation in this line of research is that most studies only investigate whether or not a web site has certain features and measure their impact on various dependent variables (e.g., switching, attrition, customer satisfaction, customer loyalty, etc.). However, we cannot know whether having a certain feature necessarily leads to users using it (Orlikowski and Iacono, 2000). In this research, clickstream data is used to investigate how various web site design features are actually used in online shopping to analyze their impact on consumer behavior as well as business performance.

**STUDY 1 – UNDERSTANDING ONLINE CONSUMER BEHAVIOR**

**Background and Motivation**

The success of e-commerce relies heavily on the online experience provided by the e-commerce web site to the customer (Novak, Hoffman, and Yung, 2000). Hence, the effective design of the e-commerce web site becomes critical. However, the effective design of e-commerce web sites is difficult due to the nature of the systems development environment for e-commerce. There are several key differences between the systems development environment for e-commerce compared to that of traditional systems.
development for organizational information systems. Among others, the nature of target users is fundamentally different – with e-commerce, users are customers and not plain users or employees (or managers) as was the case with traditional MIS (Keeney, 1999; Kauffman and Walden, 2001). With this shift in the target users of an information system, various assumptions embedded in traditional systems development no longer hold. Even though external customers are the actual users of the information system, it is practically impossible to comprehensively elicit the users’ information requirements. To make matters worse, the actual site usage behaviors of customers is invisible. In other words, we are required to build systems without a clear specification of what the information requirements are, or how the system is to be used.

Understanding detailed behavior as consumers navigate within a Web site has enormous potentials for e-commerce firms in creating value. In online retail, consumers not only interact with a web-based application (i.e., the web site), but also through it with the firm’s business processes, and product and service delivery capabilities (Hoffman and Novak, 1996). Research investigating the within-site navigation behaviors has shown that the general customer online experience has significant impact on customer satisfaction, which in turn leads to actual purchases. For example, Mandel and Johnson (2002) show that preferences, and hence purchasing decisions, are constructed online while interacting with the online storefront. In a similar vein, Moe (2001) shows that online shoppers with different shopping strategies (e.g., directed buying, search and deliberation, hedonic browsing and knowledge building) show significant heterogeneity in their navigational behavior, which ultimately leads to differences in actual purchasing propensities (i.e., conversion). Given that online consumer behavior is shaped by the customer’s shopping goals and the structure of the online environment (Pirolli and Card, 1999) (i.e., design of the e-commerce web site), it is important to understand how the design of the e-commerce web site supports consumers’ purchasing processes.

Objectives and Research Questions

The review of the literature on online consumer behavior suggests that our current understanding of online consumer behavior is still far from complete. Various models from marketing (relating to consumer purchasing behavior in the physical world) are generally assumed to apply in the online environment. However, this assumption may not necessarily be true. The purpose of the first study is to explore online consumer behavior in more detail by focusing on what consumers do and how they

4 Other important differences include pressures for faster cycle times for e-commerce systems development (Gallaugher, 1999) and the rapid development and evolution of web technologies and standards (Brajnik, 2000).
accomplish what they do. Specifically, we examine the following questions related to online consumer behavior;

- Does online consumer behavior follow the process model of purchasing?
- Can we observe the various motivations for online shopping (e.g., directed buying, hedonic browsing, search and deliberation vs. knowledge building)? How are these online behaviors instantiated?
- Does online consumer behavior change over time? If so, how?

**Research Approach**

*Research Approach Overview*

The goal of the first study is to understand how consumers shop online. Given the exploratory nature of this study, both a top-down theory-driven approach, as well as a bottom-up data-driven approach to understanding what is it that consumers actually do online will be conducted. The top-down theory-driven approach involves typology testing, whereas the bottom-up data-driven approach entails uncovering taxonomies\(^5\).

The overall research procedure is as follows. *First*, an initial theoretical typology of online consumer behavior is developed based on the relevant literature. This typology is used to theorize about what patterns of behavior are expected to occur in the clickstream data. *Second*, cluster analysis of the clickstream data is performed in order to uncover the different patterns of online consumer purchasing behaviors. Two different methods for cluster analysis will be employed in this phase. Traditional distance-based cluster analysis (e.g., k-means cluster analysis) is employed by converting the customer navigation sessions into a vector representation of page view metrics. Next, model-based clustering (Cadez, Heckerman, Meek, Smyth, and White, 2001) will also be employed in order to capture recurring navigation patterns. The purpose of the two different methods of cluster analyses is that, even though the first method (i.e., distance-based clustering) is easier to perform, important information may be lost during the conversion of the navigational sequence raw data into a vector representation of page view metrics (Cadez et al., 2001). The second method (model-based clustering) takes the navigation sequence as is and can successfully cluster the raw data based on probability mixture models. Furthermore, the results of both clustering methods can be combined in order to uncover behavioral templates of online

---

\(^5\) Bensaou and Venkatraman (1995) provide a good overview of testing typologies versus uncovering taxonomies, and the combination of these two approaches.
purchasing behavior. For example, the distance based cluster analysis is expected to uncover various online purchasing strategies such as directed-buying, search and deliberation, hedonic browsing or knowledge building, similar to Moe (2001), whereas the model based clustering is expected to show different patterns of navigation sequences within those purchasing strategies. In other words, we expect to identify not only different consumer purchasing strategies but also how consumers go about in fulfilling those strategies. 

Third, the descriptive validity of the uncovered online shopping scenarios as well as the patterns of navigation is assessed with respect to the original theoretical model of online consumer behavior. This phase allows us to identify the distinguishing parameters for different shopping scenarios to provide insights for the definition and interpretation of the shopping scenarios. Finally, the predictive validity of the taxonomical scheme of shopping scenarios is also assessed. Predictive validity tests illustrate the meaningfulness of the shopping scenario taxonomy through its ability to capture and explain differences in a set of dependent variables (e.g., duration of visit, completion of purchase transaction, etc.)

In order to examine the dynamic aspects of online consumer behavior, we attempt to identify dynamic patterns in the purchasing strategies. For example, one may expect cyclical patterns (e.g., cycles consisting of one or more sessions of knowledge building or hedonic browsing scenarios followed by directed purchasing scenarios) or evolutionary processes (e.g., consumers beginning with knowledge building and hedonic browsing scenarios and tend to directed purchasing as they gain more experience with the design of the web site).

Data

The major source of data will be clickstream data collected from web server logs at an e-commerce firm in the online grocery industry\(^6\). Typical data pre-processing procedures for using web server logs will be used to extract navigation path sequences from the clickstream data (Cooley, Mobasher, and Srivastava, 1999). The result of the data pre-processing will be a set of sessions characterized by navigation sequences.

Expected Results and Contributions

This study aims at a detailed understanding of consumer behavior in online shopping environments, which is an important prerequisite for the design of effective online shopping environments. The results

\(^6\) The same dataset will be used for all three studies of this dissertation.
of the study will provide insights into what actually happens online with implications for e-commerce web site design. Specifically, this study aims at uncovering patterns of online consumer behavior. A typology of online consumer purchasing scenarios will be developed both top down from consumer behavior theory and also bottom up from empirical data. Navigational patterns will also be identified for each of the purchasing scenarios so that actual behavior can be linked to motivations / strategies.

Pitfalls and Challenges

The major challenge for this study is in the application of the model-based clustering method (Cadez et al., 2001). The method proposed by Cadez et al. used clickstream data from MSNBC.com’s web site where the nature of the navigation path was characterized by content areas (e.g., news, travel, weather, business, sports etc.). However, in our e-commerce retail domain, we are not only interested in the contents of the page viewed but also on the web site design features that present those contents. For example, the same item may be reached via different navigational paths. The key is to clearly define a clear and comprehensive taxonomy of web site features that can be derived from clickstream data.

STUDY 2 – MEASURING THE EFFECTIVENESS OF E-COMMERCE WEB SITE DESIGN

Background and Motivation

E-business management involves an on-going process of understanding consumer needs and developing online solutions to meet those needs. E-commerce firms need to understand not only the needs of online consumers but also how effectively their web site is performing in meeting those needs. However, as discussed in the literature review on e-commerce web site evaluation, current evaluative methods for analyzing the effectiveness of e-commerce web sites are still quite limited. Indeed, one of the major problems that e-commerce firms face today is the lack of tools for analyzing the effectiveness of their web site.

Understanding the effectiveness of web site design is important for e-business management. A clear understanding of how the e-commerce site is performing will provide important implications for e-business management. For example, e-commerce managers would like to know how their web site is being used, whether their current design is effective, what areas would need radical redesign versus minor incremental changes. The focus of the second essay is to investigate the effectiveness of e-commerce web site design.
Objectives and Research Questions

The purpose of this study is to investigate the effectiveness of e-commerce web site design. In particular, we are interested in understanding whether the design of an e-commerce web site is effective in supporting customers’ shopping goals. In order to accomplish this research objective, we ask the following research questions:

- How can we conceptualize and measure the effectiveness of an e-commerce web site?
- How can we identify web site design inefficiencies?

Research Model and Approach

Different shopping goals will invariably entail different uses of the e-commerce web site with respect to what features are used. In other words, different web site features will be more valuable given different shopping strategies. Table 2 presents the expected patterns of site usage by shopping goals. Given that the effectiveness of a web site is determined by how well the design of the web site supports users goals, we define the concept of a customer’s web site design efficiency as the degree to which the intended usage of a web site matches the customer’s actual usage given the customer’s current shopping goal. Web site design efficiency is similar to the concept of task-technology fit that has been investigated extensively in the information systems literature in the areas of graphical information presentation (e.g., Benbasat, Dexter, and Todd, 1986), tables vs. graphs (e.g., Vessey, 1991), and also general information systems (e.g., Goodhue and Thompson, 1995). In the context of online shopping, the task is the shopping goals and decision processes whereas the technology is the design of the e-commerce web site. In other words, web site design efficiency measures the fit between the intended usage via system design and actual usage (i.e., how well the design of the web site actually supports a consumer’s goal for using the web site).
In order to examine the first research question, this study develops a method for estimating web site design efficiency with clickstream data. Data Envelopment Analysis (DEA) is employed in order to estimate the web site design efficiency. The DEA approach is a mathematical programming procedure for estimating the relative efficiencies of decision-making units (DMUs). One thing to note is that the DEA estimates are relative measures. In other words, an efficiency score for a particular DMU is relative to other DMUs. However, this poses the problem of estimating the efficiency frontier envelope – how do we characterize absolute efficiency? For example, if a DMU obtains a score of 1\(^7\) (i.e., most efficient), this means that this particular DMU is the most efficient relative to the other DMUs. However, it may be the case that all DMUs are in fact very inefficient and the particular DMU in question is the most efficient among the inefficient DMUs. In order to overcome this limitation of DEA, idealized DMUs are included so that the efficiency ratings of the actual empirical observations are better calibrated.

The overall procedure of the modified DEA approach is as follows. We start by characterizing the goal state of a customer’s shopping transaction. We assume that the goal state is the set of items in a customer’s electronic shopping cart at the time of checkout. Given this goal state, we identify the customer’s optimal navigation path in order to find all required items. This navigation path is then translated into navigation metrics, which will be used as benchmarks for comparing the efficiency of a customer’s actual navigation path. Then, for each customer’s transaction, we also compute the navigation metrics from her actual navigation path as portrayed in the clickstream data. Finally, both ideal and actual navigation metrics as input and the respective transaction’s goal state as the output of the production function are used in DEA analysis. This way, the navigation metrics for the ideal DMUs will

---

\(^7\) Efficiency ratings produced from DEA are typically between 0 (i.e., least efficient) and 1 (i.e., most efficient).
always attain maximum efficiency (i.e., an efficiency score of 1), whereas the efficiency ratings of the actual observations will be calibrated relative to the ideal counterpart.

In order to investigate the second research question (i.e., how we can identify web site design inefficiencies), we analyze the distribution of web site design efficiencies. Conceptually, there may be two sources of inefficiency – customer and web site inefficiency. First, the customers may be inefficient in that they do not use the web site design features that are optimal for task performance. Second, the design of the web site may be poor such that even efficient consumers cannot complete their tasks efficiently. Hence, the overall web site inefficiency can be inferred by investigating the mean of the empirical efficiency distribution – if the mean of the empirical distribution is significantly different from the ideal usage pattern, then we may infer that the design of the web site is poor. Consequently, customer inefficiency can be inferred by investigating the variance of the empirical efficiency distribution – the greater the variance, the greater number of inefficient customers.

Expected Results and Contributions

We apply the proposed method for evaluating the effectiveness of web site design to a currently operational e-commerce web site. We expect the analysis of web site design efficiency will yield interesting results. Analyzing the different sources of web site design inefficiencies provides an innovative approach to analyzing the design performance of e-commerce web sites. The results of the design inefficiency analysis will have different implications depending on how the results turn out. If the source of inefficiency is present for all customers, then it becomes a signal to the e-commerce manager that she should think about a radical new redesign. However, if the source of inefficiency is the customer, then radical redesign would not be necessary. Rather, various other remedial approaches will be more effective. For example, the e-commerce firm may target email messages to those less efficient customers to inform them about features that exist, which they do not currently use, or redesign the site to make these hidden areas more prominent and accessible.

---

8 This distinction is similar to program vs. managerial efficiency investigated by Charnes, Cooper and Rhodes (1981) in the context of program follow through (PFT) in public education.
9 An extreme view (e.g., Norman, 1990) would still put the blame of inefficiency on the design.
10 The discussion seems to imply that the distribution of the empirical efficiencies will be normally distributed with mean and variance to be estimated. However, this is not the case with efficiency measures. Efficiency has the characteristics that it may not surpass the ideal efficiency.
Pitfalls and Challenges

A critical issue in DEA relates to the identification of the input and output variables. The set of input/output variables that will be used have yet to be finalized. We are currently developing a scheme of navigation and performance metrics to use. Second, another constraint of DEA lies in the minimum number of data points required for efficiency estimation. The rule of thumb in DEA is that the number of DMUs needs to be greater than the number of input variables times output variables. Given that we are working with clickstream panel data with more than 10,000 active customers, we see no threat from this potential limitation. Third, the original DEA method had the limitation of over-estimating the efficiency ratings in situations where a DMU would be dominantly efficient in only one input (or output) measure. We can overcome this potential limitation by incorporating additional constraints\(^\text{11}\) in the mathematical programming. Finally, the concept of web site design efficiency may not apply to all customer shopping goals. It is clear that efficiency of shopping is important for the goal-oriented shopping strategies, but it is not clear whether efficiency is appropriate for more hedonic goals. In such case, we will limit our analysis to the goal-directed shopping strategies.

STUDY 3 – ANALYZING THE BUSINESS VALUE OF E-COMMERCE SITE DESIGN

Background and Motivation

Since the crash of the DotComs in the American stock market in 2000, the justification of business value of e-commerce activities has become of paramount importance for e-businesses. Ideally, e-commerce managers would like to know how their web site creates value for their customers. This study undertakes an empirical assessment of the business value of e-commerce web site design. Given that the web site is the interface between the consumers and the firm (i.e., the consumers interact with the e-commerce firm through the web site in performing economic transactions), the analysis of whether the design of the e-commerce web site creates value is important and needed. However, most of the academic research investigating the business value of e-commerce web sites has not explicitly examined the value impacts of the design of web sites.

\(^{11}\) The constraint would be that the weights assigned to the input (and output) measures are greater than zero, meaning that all input and output measures need to be considered in the final estimation.
Objectives and Research Questions

The purpose of this study is to examine the business value of e-commerce web site design. This study examines the following research questions:

- Does web site design efficiency translate into business value?
- What are the sources of business value created?

Research Model and Approach

A retail web site is the storefront for the e-commerce retailer. It is the place where consumers see products, get product information, make purchases, choose a shipping method and possibly monitor their shipment, and handle any complaints and returns. The design of the web site inevitably influences how consumers perform the necessary tasks to complete a transaction. Consistent with the second study, we may conceptualize the quality of the web site design with web site design efficiency and hypothesize that web site efficiency is positively associated with business value.

IT value research suggests that the assessment of value should be measured at the level of business processes that are impacted by IT (Barua, Kriebel, and Mukhopadhyay, 1995). Consistent with this approach, this study conceptualizes value creation at the level of transaction processes. In order to identify the source of business value creation we investigate the impact of web site design efficiency on different sources of business value. If we conceptualize the process as taking a set of inputs to generate some output, there are three ways in which value is created at the process level – 1) less input for the same output (e.g., less time to complete a transaction), 2) more output given the same input (e.g., more items purchased, more frequent transaction), and 3) a more efficient process given the same input and output (e.g., reduced use of resources such as customer service staff intervention for manually correcting orders).

In addition, we also investigate a customer value perspective to complement the process-level analysis of business value. The customer value perspective focuses on the lifetime value of the customer, which has been regarded as a firm’s solid foundation for long-run profitability. We investigate the impact of web site design efficiency on customer satisfaction / loyalty. The measurement of customer satisfaction / loyalty may be somewhat tricky given the data source limitations of this research. It is difficult to define and measure customer satisfaction, especially with objective clickstream data. Previous research primarily measured customer satisfaction with perceived measures in survey instruments, which we cannot gain access to due to the unwillingness of the target organization to conduct surveys to its
customers. We use store visit patterns as a proxy for customer satisfaction. Site revisit and purchase frequencies, and the rate of increase (or decrease) of site revisit and purchase has been used in prior research as proxy measures for customer satisfaction (Moe and Fader, 2001b).

**Expected Results and Contributions**

The empirical test of the hypothesis will yield interesting results. We expect to confirm the positive relationship between web site design efficiency and business value. The more interesting results will stem from analyzing the sources of business value creation. Previous research has yet to analyze how the design of e-commerce web sites materializes into business value. The results of the analysis should shed some light into this issue.

**Pitfalls and Challenges**

There are several methodological challenges for this study. The assessment of business value involves several levels of analysis. For example, measures such as reduced time per transaction, and greater items purchased are at the level of a transaction, whereas other measures such as revisit and purchase frequencies are at the level of the customer. Hence, care should be taken, not only when characterizing the empirical model but also when interpreting the results of the assessment.

**CONCLUSIONS**

**Expected Contribution**

Study 1 examines how consumers make use of e-commerce web sites. Study 2 develops a measurement scheme for analyzing the effectiveness of the e-commerce web site. Finally, study 3 analyzes how the quality of the web site design creates business value. Take together, these three studies contribute to a deeper understanding of the impact of e-commerce web site design.

This research brings several innovations that are worth mentioning. First, a distinctive advantage of this research is that it relies on clickstream data automatically logged by web servers. With such data, our proposed analytical methods can overcome the limitation of previous evaluative approaches (e.g., testing, inspection and inquiry). Second, our analysis does not assume technological determinism. Previous research have typically conceptualized and operationalized e-commerce web site design by the features that the web site offered. This research not only investigates what features are provided by also how various features are actually used.
Pitfalls and Challenges

Since this research relies on data on actual systems usage, the most significant challenge for the completion of this research involves access to such data in a real-world e-commerce setting. I have secured a research relationship with an e-commerce company in the online grocery industry. The target organization has shown significant interest in our proposed research and has already provided access to web server log data, website design source codes, marketing data (e.g., product, customer, and transaction databases), systems design related documentation (e.g., redesign schedules, cost estimates in terms of man-hours etc.). The company has agreed to provide access to these kinds of data on an ongoing basis.

Another potential challenge involves the quality of the data that will be used for the empirical analysis. Web server logs have been notorious for their lack of completeness and clarity (Cooley et al., 1999). For example, cached pages and the use of “Back” buttons on web browsers are typically not detected by the web server and hence are not logged. I have devised data pre-processing techniques to remedy this situation. The techniques are currently being tested for accuracy and initial results suggest that the increased quality of the data is promising.
REFERENCES


Cooper, A. (1999). The Inmates are Running the Asylum: Why High-Tech Products Drive Us Crazy and How to Restore the Sanity. Indianapolis, IN: SAMS.


