An Experimental Study on User Satisfaction and Comparison Shopping Agents for Product Evaluation

Juan Manuel Gomez Reynoso  
Claremont Graduate University  
juan.reynoso@cgu.edu

Kamla Al-Busaidi  
Claremont Graduate University  
kamla.al-busaidi@cgu.edu

Bengisu Tulu  
Claremont Graduate University  
bengisu.tulu@cgu.edu

Terry Ryan  
Claremont Graduate University  
terry.ryan@cgu.edu

ABSTRACT

Comparison-shopping agents (CSA) are intelligent agents designed to find product information based on the search criteria specified by customers. This research evaluated the effects of CSA on customers’ purchasing decisions. The determinants of user satisfaction are analyzed by comparing two different technologies, CSA and search engines, in a particular task of finding relevant information for purchasing books. Information quality and system quality were the two dependent variables in the research model. The independent variable was information search technology: a search engine or a CSA. Two experimental groups were used. The first group used a search engine and the second group used comparison-shopping agents during the experiment to purchase books. First group was not allowed to use any comparison-shopping agent. The sample size was 57 and 53 respectively for each group. Sample was randomly selected among students and faculty members of a Mexican university. Data analysis are conducted using t-test.

Keywords

Intelligent Agents, Comparison-shopping agent, online shopping, consumer satisfaction, product evaluation

INTRODUCTION

The advent of the broadband communication and hence the soaring number of the Internet users provide a golden opportunity for e-retailing. As of January 2004, there are more than 358 million online users in the World and more than 10 million are from Mexico alone (AMIPCI, 2003). Moreover, the same report (AMIPCI, 2003) identifies that in terms of online users Mexico shows an estimated growth of 269.9% from 2002-2004. E-retailing is the major business-to-consumer (B2C) application and the wide spread of the Internet offers a promising market for e-retailing. Several intelligent agents have been developed to improve the online shopping experience. One intelligent-agent-based technology called comparison-shopping agent was specifically developed to improve the product evaluation phase of the online shopping experience.

Understanding the effects of these comparison-shopping agents on customer satisfaction and hence purchasing decision is important. Studies were conducted on this new subject for the last few years regarding the effects of comparison-shopping agents on brand loyalty (Smith and Brynjolfsson, 2001), on price sensitivity (Lynch and Ariely, 2000), on consumers’ search for product information in terms of the size and quality of their consideration sets, the quality of their purchase decision in an online shopping environment (Haubl and Trifts, 2000), and on consumer buying behavior (Pedersen, 2000). It is also important to understand the effects of such comparison mechanisms on user satisfaction and its determinants. Comparison-shopping agents enable users to make a comparison without visiting dozens of sites for a single product. This value-added way of shopping may increase user satisfaction with the online shopping experience.

This paper investigates the effects of using comparison-shopping agents for product evaluation by comparing them with regular search engine-based product evaluations within a book purchasing scenario. The next section will give a brief overview of e-retailing and online shoppers, comparison-shopping agents, and online-purchasing decision-making. The paper continues with the study framework and the methodology. Finally, data analysis and our conclusions are presented.
BACKGROUND

E-Retailing and Online Shoppers

Electronic commerce (EC) is “the carrying out of business activities that lead to an exchange of value, where the parties interact electronically, using network or telecommunications technologies (Jones, Wilikens, Morris and Masera, 2000)”, primarily Internet. E-commerce is divided into three major areas, B2C, Business-to-Business (B2B), and intra-business, whereas the major B2C application is e-retailing (Turban, McLean and Whetzebe, 2001). The Internet and the increasing number of the Internet users are providing a promising market for brick-and-mortar retailers to go online. The Internet provides an integrative infrastructure that allows organizations to have more information about their customers and to provide better service for customers by overcoming time and geographic barriers (Liang and Huang, 2000). Thus, EC is an attractive market for business and individuals shopping. The number of consumer shopping online and the spending by online buyers has been increasing. For example, according to the U.S. Department of Commerce, U.S. online retail sales jumped from $11.2 billion in the fourth quarter of 2001 to $14.3 billion in 2002’s final period, a gain of close to 30%. According to Ernst & Young, “eMarketer predicts that U.S. online retail sales will continue to grow rapidly from $45.5 billion in 2002 to $88 billion in 2005, almost double in three years. This growth is expected to be fueled by consumers’ adoption of broadband (making buying online faster), their growing familiarity with online shopping, and the great convenience of the Web.” According to an ITU1 estimation, in Mexico more than eight million Personal Computers exists (a ratio of 8.2 per 100 habitants) and 17% of the population has shopped online from home (ITU, 2002). One report projects sales of $25 million USD in 1999 to $1,542 million USD in 2005 (Opinamos, 2001).

There are two types of e-commerce (hence, e-retailing) customers: individual consumers and organizational buyers (Turban et al., 2001). This research focuses on the individual consumers. Depending on their motivations for searching, online individual consumers are classified as either experiential or goal-oriented (utilitarian) (Wolfinbarger and Gilly, 2001). The same study reported that 71% of the subjects said their most recent purchase had been previously planned, 29% said they had been browsing when they made the purchase. Thus, most of the online consumers are goal-oriented rather than experiential.

The profile of the online shopper is significantly changing. Previous research reported that the online population is younger, more educated, and wealthier than the average household population in the U.S. (Onemerchant.com, 2001, Dholakia and Dholakia, 2001). On the other hand, a recent AMIPCI2 study (AMIPCI, 2003) described online shopper as: 67% male and 33% female, 74% 25 years old or older, 70% well educated (at least with some college studies). According to AMIPCI study, the top selling consumer products online were books, flight tickets, CDs, computer equipment, and electronics. In addition, most online shoppers paid services (e.g. utilities) through the Internet. Finally, 47% of the online shopping in Mexico was with international vendors and 53% with Mexican vendors. One important issue affecting the number of Mexican online shoppers as well as the gained experience is the fact that not many Mexican organizations have B2C in place as of 2004. Therefore, online shoppers are depending on the international vendors. However, due to customs restrictions and other laws in place, the products that can be purchased from international vendors are limited and are likely to be taxed at customs.

Comparison-Shopping Agents

Several technologies are developed to support e-commerce applications and online shopping. Intelligent Agent (IA) is one of these technologies. IAs are programs, used extensively on the Web, that perform tasks such as retrieving and delivering information and automating repetitive tasks. IAs perform different tasks depending on their purpose. Some of them help users to reduce work and information overload (Maes, 1994), some of them help in buying and selling processes (Maes, Guttmann and Moukas, 1999), and some others help in business (Papazoglou, 2001). In general, “autonomous agents can take many different forms depending on the nature of the environment they inhabit” (Maes, 1995), and doing so helps in the specific task of their design. The main idea behind the design of IAs is to engage and help all types of end users (Riecken, 1994). IAs have been classified in a number of ways. The classifications have evolved over time as new advances are achieved in the field, and new types of agents are listed in new classifications. Turban et al. (Turban, Lee, King and Chung, 2000) broadly classified IAs into five basic categories: Intelligent Agents for Information Search and Filtering, Intelligent agents for Products and Vendor Finding, Negotiation Agents, Intelligent Agents for Customer service, and Learning Agents. A more online shopping specific classification was made based on the role agents play in consumer decision making (West,

---

1 International Telecommunication Union
2 The Mexican Internet Advertising Association (AMIPCI -Asociacion Mexicana de la Industria Publicitaria y Comercial en Internet)
Ariely, Bellman, Bradlow, Huber, Kahn, Little and Schkade, 1999). This research identified four basic roles for IAs: tutor, clerk, advisor, and banker.

Our research focuses on the IAs for product and vendor finding and they will be called comparison-shopping agents. Comparison-shopping agents provide “a ranked list of product/merchant information gathered from the Web” in order to make the purchasing decision easier and quicker for the customers (Yuan, 2003). Comparison-shopping agents are capable of making comparison and filtering based on different attributes (e.g. price) of a product a consumer is considering to buy. These IAs fall into the “clerk agents” (West et al., 1999) category since they assist consumers in performing tasks such as information search and product screening. Comparison-shopping agents have their own vendor database and they retrieve information by visiting the vendors’ web sites and hence they are not multi-agent systems. Comparison-shopping agents, by definition, only extract information from multiple sites about products that a user is willing to buy. Previous research on the effects of IAs on users approached the subject differently. One source conducted a study about online wine buyers and the effects of IAs on price sensitivity (Lynch and Ariely, 2000). This study focused on differentiated products like wines and concluded that lowering the cost of search on quality information has positive effects on consumers’ shopping experiences. Another study was conducted to test the effects of interactive decision aids, such as a recommendation agent and a comparison matrix, on purchasing decisions (Haubl and Trifts, 2000). They conclude that decision aids that provide an in-depth comparison among selected alternatives may have strong favorable effects on both the quality and the efficiency of purchase decisions. This study (Haubl and Trifts, 2000) defined comparison matrix as a decision aid that helps consumers make in-depth comparisons among selected alternatives, which is almost identical to the previous definition provided for comparison-shopping agents. Another study used a different terminology for product and merchant brokering agents, and referring to them as shopbots (Pedersen, 2000) while investigating their effect on reducing the consumer search cost and increasing the rationality behind buying behavior.

The previously mentioned research concluded that the effects of comparison-shopping agents on purchasing decisions regarding high-complexity products that can be differentiated is different than the products that are hard to differentiate such as books. Our study focused on books with a goal of understanding the effects of different search mechanisms on the determinants of user satisfaction.

**Online-Purchasing Decision Making**

Several variables that affect the decision-making process for online shopping were reported in a study (Ernst & Young LLP, 1999). It concluded that Internet stores attract consumers for several reasons such as cost savings, greater variety, and convenience. Also, well-designed and easy-to-use web sites are important factors for overall satisfaction in the online shopping experience. Retailers consider these features to be the most important factors in achieving success online. Having a strong company brand and well-known branded products are also highly rated factors in online success. On the other hand, security, privacy, reliability and trust are the required factors for e-commerce (Jones et al., 2000). The development of trust between the e-commerce stakeholders such as businesses, consumers, and service providers, is vital to the expansion of e-business markets and the full exploitation of technological developments in this area (Jones et al., 2000). The decision making process model (Figure 1) shows the variables that affect decision making in online shopping (Turban et al., 2000).
There are several stages involved in the purchasing decision making process. Each step in the consumer purchasing decision model (Turban et al., 2000), (Figure 2) is supported by different IA types. Comparison shopping agents support the “information search” step in this model by gathering information about products that a user is willing to buy from multiple sites. They also partially support the “alternative evaluation” step by providing information about the evaluation of merchant alternatives based on consumer-selected criteria (e.g., price, warranty, availability, delivery time, reputation, etc.). The combination of these two steps is called the Merchant Brokering stage, which is one of the six stages of the Consumer Buying Behavior (CBB) model (Guttman, Moukas and Maes., 1998) and this is an important part of the overall purchase decision-making process. This study investigated the determinants of online consumer satisfaction in using comparison-shopping agents for “information search” and “product-alternatives evaluation”. The final purchasing decision was not included in this study.

RESEARCH FRAMEWORK AND HYPOTHESES

The main research question of our study is: How do comparison-shopping agents affect the online customer satisfaction of product evaluation? This study experimentally investigated the effects of using comparison-shopping agents for online product evaluation by looking at the differences in system quality and information quality while manipulating the information search technology variable (comparison-shopping agent versus web-search engine). A previous study carried out this investigation by conducting a web-based survey (Al-Busaidi, Reynoso and Tulu, 2004). Because of the uncontrolled nature of their methodology, the study was unable to detect any significant differences in the system quality of the comparison-shopping agent and the web-search engine. This study was based on a laboratory experiment to have a better-controlled environment, which can also improve the internal validity.

Framework Development

Several frameworks have been introduced to measure user satisfaction of an information system (Bailey and Pearson, 1983, DeLone and McLean, 1992, Ives, Olson and Baroudi, 1983, Raymond, 1985). Delone and McLean’s framework (Figure 3), the most cited framework in the literature, indicates that the user satisfaction is a multidimensional attitude towards various aspects of the information systems such as information quality and system quality (DeLone and McLean, 1992, Raymond, 1985). Information quality is related to the characteristics of the output information produced by the system; it is generally defined as the characteristics of data quality such as accuracy, relevance, format, objectivity, timeliness, accessibility and completeness (Bailey and Pearson, 1983, Turban et al., 2000). System quality factors are related to the characteristics of the system. The common measures of system quality are response time, reliability, system flexibility, accessibility, and ease of use (Turban et al., 2000).

Based on DeLone and McLean framework, our study measured the determinants of customer satisfactions; information quality and system quality provided by comparison-shopping agents and web search engines. Information quality is measured by accuracy and format. Many different measurement items for system quality have been proposed in the literature.
Based on Srivasan’s (Srinivasan and K.M., 1987) constructs, two of them were identified as relevant to this research: time spent and system reliability. Accessibility is another construct used in the previous research. However, in the experiment the accessibility item was controlled, therefore it was not included as a construct for system quality. Figure 4 shows the framework for this study. The independent variable is the information search technology: a web search engine or a comparison-shopping agent. The four dependent variables are the users’ perception of the information accuracy, information format, time spent, and system reliability.

Figure 4. Customer Search Satisfaction Model

Hypotheses Development

*Information Quality*

The content quality that a search engine provides has been debated in the literature (Brin and Page, 1998). It is known that a web search engine can provide low quality matches when the automated search engine relies on keyword matches only. Also, intelligent agents are considered to add more value. Hence they provide a better information (TERENA, 2003). The hypotheses in the study also followed the same direction as explained below.

*Accuracy* means that the recorded value conforms to the real-world fact or value (Fisher and Kingma, 2001). Accuracy in the context of the information search focuses on the match between the real-world information and the information provided by the search mechanism. The information search provided by the search mechanism must reflects real-values. Otherwise, the customers will be affected by the inaccurate information. In terms of accuracy, the comparison-shopping agents provide more accurate information by reducing the bias caused by the search engines (Brin and Page, 1998, Lucas, Crostt and Schiano, 2001). This reduction is a result of a focused information search which eliminates the less relevant information captured by a keyword search via a search engine. Therefore, it is expected that:

\[ H1a. \text{The comparison-shopping agents provide more accurate information about the product than the web-search engines.} \]

In information search the *format* is the way that the search results are presented to the user. Data must be presented in a format that serves the user’s purpose and stated in terms familiar to the user (Fisher and Kingma, 2001). The information provided must be in a format that the user can read and use to compare products easily based on the entered search criteria. The argument from the previous paragraph suggests that intelligent agents present information in better format. Therefore, it is expected that:

\[ H1b. \text{The comparison-shopping agents provide information in a better format than the web-search engines.} \]
System Quality

*Time spent* represents the time spent to gather the necessary information to make a purchasing decision. Since the intelligent agents’ most promising feature is reducing information overload (Maes, 1994), this will result in reducing the time spent to make a decision or to reach the targeted information. Therefore, it is expected that:

H2a. Time spent to do online-product evaluation is shorter using comparison-shopping agents than web search engines.

The *reliability* of a system refers to the consistency and dependability of the output information under a variety of conditions (Bailey and Pearson, 1983). Comparison-shopping agents gather information from the original sources of data and present it to the online consumer in the requested format. Instead, search engine results are based on the metadata provided by the websites. Therefore, it is expected that:

H2b. The comparison-shopping agents provide more reliable information than web search engines.

**METHODOLOGY**

**Research Design**

Our study investigated the effects of using comparison-shopping agents on information quality and system quality compared versus a web search engine for product comparison. Since there were many extraneous independent variables such as technology ease-of-use and users familiarity with technology that might affect the findings of this study, a controlled environment was set to eliminate the effects of these extraneous factors on the study output. An experimental design is the only research design that allows to control for these extraneous factors while manipulating the independent variable (Kerlinger and Lee, 2000). The control of these variables allowed the measurement of the direct effects of the independent variable on the dependent variables without any interfering factors enhancing the internal validity of the study.

A two-group design was used in the experiment. The first group used a web search engine and the second used a comparison-shopping agent for an information search. To control for the extraneous factor of technology ease of use, both search technologies were selected from the most popular search engines (UC Berkeley - Teaching Library Internet Workshops, 2003) and comparison-shopping agents. It was hoped that the selected participants would be regular online shoppers in order to control the familiarity with the task extraneous factor. However, the results of the descriptive analysis showed that this goal was not achieved completely, since the sample was composed of mainly university students in their first or second year.

The task assigned to both groups was to search for three books and continue with the purchasing process until the final payment step. The first group was asked to complete the task using only Google\(^3\) as the web search engine. The second group was asked to use either AddALL\(^4\) or BookFinder\(^5\) comparison-shopping agents\(^6\). Each participant was assigned a computer. All computers had the same system configuration in terms of CPU speed, memory space, Internet connection speed and so on. Computer technicians, guided by one of the researchers, monitored the groups and assisted them with any technical problems encountered. At the end of each search task, the participants were asked to fill out an online questionnaire in response to the experience of the technology used.

**Sample**

The total sample size for this study was 110 participants, composed of faculty members and students from a Mexican University. The potential participants were invited to participate in the experiment either via email or verbally. The participants were randomly selected from the respondents and assigned to one of the two experimental groups. The sample size of the first, the comparison shopping agent group, was 53 people, while the sample size of the web search engine group was 57.

---

\(^3\) http://www.google.com

\(^4\) http://www.addall.com

\(^5\) http://www.bookfinder.com

\(^6\) As pointed by the anonymous reviewers of this paper, the use of two different agents could affect the results of this study. However, the information regarding the agent used was not captured during the experiment and this eliminates the option of a 3-group design.
DATA ANALYSIS

Demographics statistics

The preliminary analysis of data based on frequencies provided useful information about the sample characteristics. Only 31.8% of the participants were female compared to 68.2% male. 83.6% were single, 85.5% were undergraduate students in Mexico, and 14.5% were faculty members. 77.3% of the participants were less than 25 years old whereas only 9.1% were older than 35. Only 41.8% had shopped online and of those who shopped online, 79.1% had shopped online less than 3 times. Only 8.2% of the population described their computer skills as fair. The rest ranked their computer skills as either good (55.5%) or excellent (36.4%). 62.7% had been using Internet between 4 to 6 years, with only 4.5% using it less than 2 years and 20.9% more than 6 years.

Measurements Reliability

The measurement scale of the items was based on a 7-point Likert scale as used in (Bailey and Pearson, 1983), using 3 items to measure each construct. Figure 5 shows the items in the study. Constructs Accuracy, Format, and Reliability were measured using values ranging from strongly disagree=1, to strongly agree=7. The first two items for Time Spent have the same scale, the third item was measure with range of minutes (>=5,6-10,15-20,21-25,26-30,>30).

<table>
<thead>
<tr>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The output information of the search website is as accurate as the one found on the online store website</td>
</tr>
<tr>
<td>2. The output information of the search website is as sufficient as the one found on the online store website</td>
</tr>
<tr>
<td>3. The output information of the search website is not consistent with the one found on the online store website</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The display of the output of the information search content is simple</td>
</tr>
<tr>
<td>2. The display of the output of the information search content is unreadable</td>
</tr>
<tr>
<td>3. The display of the output of the information search content is good</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Spent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The time spent to finalize an order is very short</td>
</tr>
<tr>
<td>2. The time spent to reach the desired site is unreasonable</td>
</tr>
<tr>
<td>3. How many minutes did you take to finalize an order?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The output information of the search website is consistent with among different searches I conducted</td>
</tr>
<tr>
<td>2. The reliability of the output information is low</td>
</tr>
<tr>
<td>3. The reliability of the output information is sufficient</td>
</tr>
</tbody>
</table>

Figure 5. Items In The Study

Figure 6 reports Cronbach’s alpha values obtained for each construct. Only the Reliability construct alpha value was above .70 (the minimum acceptance value (Kline, 1998)). Accuracy and Format constructs alpha values were within the valid interval (.50 - .70) reported in (Kline, 1998). Time Spent was the only construct with a very low reliability value. One reason for this may be the last item of that construct, which asks users the exact time spent for the search while the other two items were questioning their perception about the time spent in general.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Coefficient Alpha</th>
<th>Average Item Intercorrelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>.65</td>
<td>.39</td>
</tr>
<tr>
<td>Format</td>
<td>.56</td>
<td>.31</td>
</tr>
<tr>
<td>Time Spent</td>
<td>.49</td>
<td>.24</td>
</tr>
<tr>
<td>Reliability</td>
<td>.76</td>
<td>.50</td>
</tr>
</tbody>
</table>

Figure 6. Reliability Test
Hypothesis Testing

The results of t-test, which were used for hypothesis testing (Figure 7), are as follows: Comparison-shopping agents provided better Format than regular web-search engines, and comparison-shopping agents were perceived to have better Reliability than web-search engines. Hence, the second (H2) and forth (H4) hypotheses were supported. However, the t-test for the first hypothesis (H1) was not statistically significant, which means that web search engines and comparison-shopping agents were perceived to have about the same Accuracy. Results for the third hypothesis (H3) were not statistically significant. Hence, users’ perceived time spent for a search using both web search engines and comparison-shopping agents about the same.

<table>
<thead>
<tr>
<th>Construct</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>-1.898</td>
<td>.060</td>
</tr>
<tr>
<td>Format</td>
<td>-2.269</td>
<td>.025</td>
</tr>
<tr>
<td>Time Spent</td>
<td>-1.716</td>
<td>.089</td>
</tr>
<tr>
<td>Reliability</td>
<td>-3.713</td>
<td>.000</td>
</tr>
</tbody>
</table>

* p < .001

Figure 7. t-Test Results

CONCLUSIONS

There are some limitations to the study. Since this is an experimental study, the findings of it will be difficult to generalize. The population that was used to select the sample may be limited, since it was selected based on accessibility. A broader sampling frame might increase the strength. The items selected for the model were based on the items that already exist in the early IS literature. More recent studies (Rai, Lang and Welker, 2002, Seddon, 1997) should be utilized for identifying items that are related to this study other than those included for future research. Since the selected items are from an old source, it might not apply well to the Internet era, which can explain some of the results presented. Only first generation intelligent agents in our study are included. However, the second generation of intelligent agents is different in terms of the features that they provide to the consumers. Therefore, this study cannot be generalized for all intelligent agents. The Time Spent construct has a low reliability coefficient (0.49). If the measurement is unreliable, it may not produce an accurate measurement and consequently it may be unable to provide significant results. Therefore, the results may not be representative and cannot be generalized. Adding more items to measure the construct might improve reliability.

Considering the above limitations, the following further research options are suggested. The same study can be conducted with different samples in order to strengthen external validity. Further research is necessary to identify the items that are not included in this study but might have an effect on the results. The model can be tested by adding new items. To measure how user satisfaction is affected by the use of different algorithms and different management schemes for intelligent agents can be a further research area. The same study can be conducted by using second generation intelligent agents. In this research, it is planned to go up to the point where the purchasing decision is made, which is not included at the present research. A further study can expand the model including the purchasing decision and how the intelligent agents affect this decision.

ACKNOWLEDGMENTS

Our special thanks to Dr. Magid Igbaria for the advice he gave us during the initial stage of the paper.

REFERENCES

2. AMIPCI (2003) Asociacion Mexicana de la Industria Publicitaria y Comercial en Internet, Mexico, D.F.
7. Ernst & Young LLP (1999).