A Usability Evaluation Framework for B2C E-Commerce Websites

Sandeep Kumar Panda

School of Computer Engineering, Research scholar, KIIT University, Bhubaneswar, Odisha, India

Skpanda00007@gmail.com

Abstract

The research aims to findings of quantitative and qualitative analysis of several usability problems of B2C e-commerce online shopping websites. Here we presented a framework for the usability evaluation process of B2C e-commerce websites. This involved user testing (usability testing, inspection and inquiry) and open source automated tool like Camtasia. The framework was represented by the advantages and disadvantages of these methods in lieu of the specific areas of usability problems. The framework gives proper attention that user testing is good for identifying specific major usability problems related to four areas: navigation, design, the purchasing process, and customer service.

Keywords: Navigation, Design, Purchasing Process, B2C E-Commerce, Framework, Usability Problem

1. Introduction

Of late, there has been a proliferation of Business to consumer (B2C) websites due to the increased use of the Internet. Due to the vast reach and different interactive capabilities supported, these have become an important trading medium for many organizations. There is a phenomenal increase in many organizations that are using the web for trading, marketing, promoting, and transacting products and services with consumer. Apart from firms and organizations, there seems to be very large growth of the Internet by consumers for various purposes, including online shopping and information search. The consumer interest rise in online shopping is affecting the traditional retail sales in that the growth in offline sales over the next decade expected to slow down from 5% to 3% per year. The rise in business to consumer electronics commerce has made many organizations looks for new ways to understand online shopping behavior in order to attract and retain the consumers.

To be successful, websites need to have good usability. Usability is an overall measure of how easy the interface is to use. (Najjar, 2005; Nielsen, 2003). Nielsen (2003) stated that if users are unable to find a product, they would not buy it. User based usability evaluation methods usually involve users being observed undertaking pre defined tasks with the purpose of identifying usability problem (Brinck *et al.*, 2001). User based approaches have been frequently used to evaluate the usability of e-commerce websites (McKinney *et al.*, 2002). For example, McKinney *et al.* (2002) developed constructs and corresponding measurement scales with users for measuring web customer satisfaction and Tilson *et al.* (1998) asked sixteen users to complete tasks on four e-commerce sites and report what they liked and disliked. A user-based usability evaluation method includes a set of methods that involves users. These methods aim to record users' performance while interacting with an interface and/or users' preferences or satisfaction with the interface being tested. The most common method in this category relates to user testing.

User testing method is considered to be a promising approach since it provides direct information regarding how real users use the interface; and consider the problems users' encounter in their interaction (Nielsen and Mack 1994). Dumas and Redish (1999) defined the user testing method as "a Systematic way of observing actual users trying out a product and collecting information about the specific ways in which the product is easy or difficult for them". Several extensions have been suggested for use during a user testing session, such as making different types of observation (e.g. notes, audio, video or interaction log file) to capture users' performance; questionnaires and interviews have also been suggested as ways of collecting data concerning users' satisfaction (Nielsen 1993; Sharp *et al.* 2007; Dumas and Redish 1999; Rubin 1994).

In this paper, we developed a usability evaluation process and an evaluation framework considering some attributes. After that, we conducted an user testing approach to identify usability problem areas based on the case studies applied on three online shopping websites, than we identify the major usability problems related to four areas that is navigation, design, Purchasing process and customer service. By this, our study identified factors that are critical to the success of a business to consumer website.

The rest of this paper is organized as follows:

Section 2 discusses the literature survey. We review user-testing methodologies, which are used in evaluating usability of websites in different domain. Section 3 provides a proposed usability evaluation process and a

suggested framework for evaluating the B2C e-commerce websites. Section 4 presents our method, user testing. We first discuss data collection and evaluation procedure to evaluate usability of B2C e-commerce websites, which collects from 10 novice users and 10 expert users. Next, we describe user-testing results followed by statistical analysis. Section 5 describes our methods to evaluating the usefulness of suggested framework, usability evaluation, usefulness and applicability of the suggested framework, types of problem the companies were interested to identify. Section 6 concludes the paper with a critical analysis and interpretation of our work. Finally, we discuss the possible future extensions to our work.

2. Literature survey

Claudio and Antonio developed a model by adapting the Technology Acceptance Model (TAM) to identify the design characteristics of CD e-retailing websites that would influence a user's intention to buy from these sites. Other studies compiled a set of design issues and used them to investigate which were preferable for users. The websites that were investigated included chocolate websites (Oppenheim and Ward 2006), food and drink websites (White and Manning 1998), clothing and product websites (Tilson *et al*. 1998) and supermarket websites (Freeman and Hyland 2003).

Although the studies identified above investigated different types of e-commerce website, there were a number of common design features preferred by users for inclusion in the sites. Examples of the common features included: Ease of use, ease of navigation and finding products (Claudio and Antonio; Tilson *et al*. 1998; Freeman and Hyland 2003), Simple and successful search facilities (Tilson *et al*. 1998; Oppenheim and Ward 2006; Freeman and Hyland 2003), Customer service or help functions (Tilson *et al*. 1998; Oppenheim and Ward 2006),Secure sites (Tilson *et al*. 1998; Oppenheim and Ward 2006),Site support and personalization/customization (White and Manning 1998; Oppenheim and Ward 2006),interesting sites (Claudio and Antonio; Oppenheim and Ward 2006),Attractive/innovative sites (Claudio and Antonio; White and Manning 1998).

Furthermore, additional design issues were identified uniquely by each study. Some of these issues related to the ability to purchase without registering with the site (Tilson *et al*. 1998); the availability of multilingual options; the clear provision of error messages on pages providing feedback on users' input (Oppenheim and Ward 2006); and the need for a fun, useful, clear, concise and informative design (White and Manning 1998).

3. Usability Evaluation Process

Usability evaluation is a process framework that dealt with some of the activities projected in figure 1, depending on the methods used. This section discusses each of these activities.

```
1. Specify usability evaluation goals
2. Determine usability interface to evaluate
3. Identify target users
4. Select usability metrics
5. Select evaluation method
6. Select tasks
7. Data Collection
8. Capture usability data
9. Analyze and interpret usability data
10. Usability interface to suggest improvements
11. Present results
```

Figure 1: Activities that occur during the usability evaluation process

3.1 Specify usability Evaluation Goals

Usability evaluation is applicable at all stages like design, implementation, and redesign of usability interface life cycle. At these various stages, different usability evaluation goals are relevant. Some of the usability evaluation goals are specify usability interface requirements, evaluate design alternatives, identify specific usability problems, and improve usability interface performance. Here we are clearly specified the goals of the usability evaluation that is identify specific usability problems.

3.2 Determine Usability Interface Aspects to Evaluate

Some usability interface can be extremely large and complex and an evaluation of all aspects may not be economically possible. Therefore, the evaluator must determine specific usability interface aspects to evaluate.

Here, we are specified in our work, usability interfaces are navigation, design, purchasing process, and customer service based on B2C electronics commerce websites.

3.3 Identify target users

For a larger user community an interface may be used, but it is important to know user characteristics most relevant to the study and for the usability interface in particular. Here in our work the users are employed during the study, and they are the representative of larger user community. In order to find the number of users to perform the user testing, an investigation into the literature was found. Brinck *et al*. (2002) suggested, if the budget allowed, recruiting eight to ten users to perform user testing. Rubin (1994) also suggested testing with more than five users, suggesting at least eight participants. It is worth noting that, in order to obtain statistically valid results, enough participants should be tested to perform the appropriate analysis and to generalize to a target population (Rubin 1994). In this context, Nielsen (2006) recommended testing 20 users in quantitative studies that included collecting quantitative usability metrics such as learning time, efficiency of use, memorability, user errors, and subjective satisfaction. However, while performing the user testing, it is suggested that there is a need to balance acquiring participants with the practical constraints of time and resources so issues such as the availability of the type of participants required and the duration of the test session need to be considered. Based on the illustration above an advertisement was prepared, it was decided that twenty users would be recruited in this research, 10 females' and10 males were chosen.

3.4 Select Usability Metrics

Usability metrics are the most important component of the usability evaluation. Here we were selecting a minimal number of metrics that gives us the maximum number of usability detail for the usability interface under study. ISO 9241 suggests effectiveness, efficiency, and satisfaction measures that are satisfaction reflects users' freedom from discomfort and positive attitudes about use of an interface. The metrics include ratings for satisfaction, ease of learning and error handling. Efficiency metrics include the time to complete a task and learning time. Effectiveness metrics include percentage of goals achieved and errors corrected successfully.

3.5 Select Evaluation Method

Choosing one or more usability evaluation methods is an important step of the usability evaluation process. There are five types of usability evaluation methods: usability testing, inspection, inquiry, analytical modeling, and simulation. An inspection entails an evaluator using a set of criteria to identify usability problems in an interface, while testing involves an evaluator observing participants interacting with an interface to determine usability problems. Inquiry methods entail gathering subjective input from participants, typically through interviews, surveys, and questionnaires. Analytical modeling and simulation approaches to usability evaluation that enable evaluators to predict usability with user and interface models.

Usability evaluation methods uncover different usability problems; therefore, we were used multiple assessment methods that are during a usability test, participants may also complete questionnaires to provide subjective input; thus, enabling evaluators to gather quantitative and qualitative data.

3.6 Select Tasks

Tasks are the most important part of the usability evaluation. They must be appropriate for the usability interface aspects under study, the target users, and the evaluation method.

3.7 Data Collection

The evaluator may need to collect usability data. In particular, the evaluator needs to decide on the number of participants (users), the evaluation procedure as well as the environment and system setup. Here we were furnished all requirements for usability testing and inquiry. We are also conduct pilot runs during the study.

3.8 Capture Usability Data

In this phase, we employ the usability evaluation method to record previously specified usability metrics that is at the time of usability testing and inspection. Here we were used to capturing user performance can be automated using tools such as Camtasia. Camtasia is a screen capture software package, provided by TechSmith Company that has proved to be an effective tool for capturing website usability data (Goodwin 2005). Camtasia records users' activities on screen (i.e. users' actions and movements that take place on the computer screen); it also has the capability to record users' voices along with their actions if a microphone is used (Goodwin 2005). Camtasia files, which include videos of each recorded session, are saved in Audio Video Interleaved (AVI) format that can be then compressed and played again to review and interpret users' actions with the interface being tested. Goodwin (2005) stated that Camtasia software is the best method for acquiring usability data in

terms of minimizing data loss and avoiding the bias of human recorders. This therefore helps to reduce the workload of the observer during the user testing session.

3.9 Analyze and Interpret Data

The goal of usability data analysis is to summarize the results in a manner that informs interpretation. This summarization entails statistical techniques based on the goals of the usability evaluation.

3.10 Usability Interface to Suggest Improvements

Analysis and interpretation of usability data gave flaws in the usability interface design as well as ways to possibly improve the design. Regular analysis may be required to verify that suggested improvements actually improve interface usability.

3.11 Present Results

The final step of the usability evaluation process is to communicate the results and interpretation of these results to the stakeholders. The evaluator presents the results such that they can be easily understood and acted upon.

4. Case Study

4.1 Data collection

In this paper, we describe the method of collection of data for user testing. The method involved using various types of observation, and the observer taking notes and using Camtasia software to capture performance data while questionnaires were used to assess user's satisfaction with the tested sites.

4.2 Evaluation Procedure

An evaluation procedure was developed to welcome the users and to provide an introduction to the research. A consent form acknowledging the users' agreement to participate in the test and to be observed through the testing session was also developed. The consent form was required to be read and signed by users. A pre test questionnaire was developed and hard to be filled out by the users after they had signed the consent form. A task scenario was developed for each of the three studied websites. This included tasks for the three B2C e-commerce websites that represented their actual use. In order to collect preference information from the users regarding the tested websites, three post test questionnaires were developed. Each user responded to the appropriate post test questionnaire after interacting with each website. A post evaluation questionnaire was developed to be filled out by the users after performing all three evaluation tasks and after filling out the three post test questionnaires. A pilot test was conducted before the main test to test the user testing methods. This is an essential step which helps to practice the test and to discover and refine any bugs in the testing process, such as un-applicable tasks or ambiguous questionnaire (Rubin 1994).

After completing the tasks for the tested website, the user was given the post test questionnaire to fill out in order to get his/her feedback. Then, he/she was given the post evaluation questionnaire to fill out in order to get his/her feedback about the usability of three tested websites.

4.3 Analysis and Interpretation of User testing results

This section presents an overview of the users in term of their characteristics, and their perceptions and experience of online shopping. This finding from the performance data and observations, the quantitative and qualitative analysis of the post test questionnaires, and the post evaluation questionnaires. There were ten novice and ten expert participants, five female and five male. For full details of the user's characteristics and the frequency distribution, see Table 1.

Table 1. User's characteristics and the frequency distribution

No.	Characteristic	Range		Frequency Distribution	
				Novice Group	Expert Group
Personal Information					
1	Age	18-29		60%	70%
		30-39		30%	30%
		40-49		10%	0%
		Over 50		0%	0%
2	Gender	Male		50%	50%
		Female		50%	50%
3	Education	Postgraduate Degree		10%	20%
		Higher Diploma		0%	0%
		Bachelors Degree		40%	60%
		Diploma		40%	20%
		High School		10%	0%
Computer Experience					
4	Experience using Computer	Under 1 year		0%	0%
		1-3 years		30%	0%
		More than 3 years		70%	100%
5	Daily use Computer	Less than 2 years		20%	0%
		2-4 hours		10%	30%
		More than 4 hours		80%	90%
Internet experience					
6	Browser	Internet Explorer		90%	90%
		Nets cape Navigator		10%	10%
		Other		0%	0%
7	Experience using Internet	Less than 1 year		10%	0%
		1-3 years		90%	0%
		More than 3 years		0%	100%
8	Weekly use of internet	Less than 2 hours		0%	0%
		2-4 hours		20%	10%
		More than 4 hours		80%	90%
9	Have you browsed the following websites before?	Website 1	Yes	100%	100%
			No	0%	0%
		Website 2	Yes	100%	100%
			No	0%	0%
		Website 3	Yes	100%	100%
			No	0%	0%
10	Did the user used the internet for Purchasing	Yes		100%	100%
		No		0%	0%

The Mann-Whitney test showed that there were no statistically significant differences between novice and expert users in their ratings regarding their perceptions towards the online shopping statements, except for one. That statement related to users' interest in information about companies presented on the sites. Novices were not interested whilst experts were. The Likert scores for the other statements (Table 2) showed that novice and expert users considered the cost of using the Internet as generally unreasonable, Liked websites to be easy to navigate and to be well organized, considered compulsory registration frustrating when shopping online, and worried about the security of their financial information, the privacy of their personal information, and the absence of legal regulations that govern online transactions when shopping online.

Table 2. Likert scores of the pre-test questionnaire for novice and expert users and the result of Mann-Whitney test

No.	Question	Likert Score		Mann-Whitney Test
		Novice Group	Expert Group	
Q 36	The cost of using the internet is generally reasonable	3.40	3.80	No (U=33.500, N1=10, N2=10, p=.218, two tailed)
Q 37	I am not interested in information about companies that is presented on their websites	3.60	5.80	No (U=17.000, N1=10, N2=10, p=.011, two tailed)
Q 38	I like websites to be easy to navigate	6.10	7.00	No (U=25.000, N1=10, N2=10, p=.063, two tailed)
Q 39	I am interested in well organized websites	7.00	7.00	No (U=50.000, N1=10, N2=10, p=1.000, two tailed)
Q 40	Compulsory registration when shopping online is frustrating	5.60	5.40	No (U=28.000, N1=10, N2=10, p=.105, two tailed)
Q 41	I am worried about the security of my financial information while shopping online	7.00	7.00	No (U=50.000, N1=10, N2=10, p=1.000, two tailed)
Q 42	I would worried about the privacy of personal information when shopping online	7.00	7.00	No (U=50.000, N1=10, N2=10, p=1.000, two tailed)
Q 43	I am worried about the absence of legal regulations that govern online transaction	7.00	6.56	No (U=25.000, N1=10, N2=10, p=.063, two tailed)

All the ten expert users who had purchased from the Internet provided information about their experience of online shopping (Table 3), Two thirds used the Internet annually for purchases, whilst one participant indicated his/her usage was monthly. The first purchase from the Internet was made less than a year ago for two participants and between one and two years for the others. Two thirds used their credit card as the method of payment, whilst one used the cash on delivery method. The products bought in their last purchase were a mobile phone, a digital camera, books and dresses.

Table 3. Experience of online shopping of expert users

No.	Question	Range	Frequency Distribution
11	Frequently use of the internet for purchasing products	Weekly	0%
		Monthly	25%
		Yearly	75%
12	The first time a user purchased from the internet	Less than a year ago	50%
		One or two years ago	50%
		Over two years ago	0%
15	Method of payment a user used	Credit card	75%
		Cash on delivery	25%
		Cheque by post	0%
		Bank transfer	0%
		Other	0%
No.	Question	Answer	
13	What was your last purchase online?	Mobile Phone Digital Camera Books Dresses	
14	Which site did you use to make this purchase?	www.ebay.com www.flipkart.com www.snapdeal.com www.quikr.com	

The Likert scores for the online shopping experience (Table 4) showed that these ten users shopped online because it saved time and they were able to buy products at any time of day from any location. Preferred to shop

online from well known sites with a good reputation; sites that provided alternative methods of ordering/payment/delivery; and sites that did not have limited delivery areas. Found the website's search function useful when shopping online. A detailed description of the products was also important. They preferred to research products in detail before purchasing and were encouraged to shop online from sites with a clear return and refund policy. Received the products within the time period specified by the company and were satisfied with the goods received. The products were accurately represented by the websites to obtained good customer service from online companies. They felt more comfortable with sites which kept them informed about the status of their order and did not find delivery costs reasonable.

Table 4. Likert scores for online shopping experience of Expert users

No.	Question	Likert Score
		Expert Group
Q 16	I shop online because it saves time	7.0
Q 17	I prefer to shop online from well known websites with a good reputation	6.8
Q 18	I do not find the website's search function useful when shopping online	6.5
Q 19	Generally I find it cheaper to shop online than to go to shops	5.8
Q 20	In general a detailed description of the product is not important to me	5.5
Q 21	I shop online because I can buy products at lower prices	4.3
Q 22	I prefer to research products in detail before purchasing	6.8
Q 23	I shop online because I can buy products at any time of day	7.0
Q 24	I shop online because I can buy products from anywhere	6.8
Q 25	I find it difficult to remember my password when shopping online	4.3
Q 26	In general products are received within the time period specified by the company	6.0
Q 27	In general I am satisfied with what I receive from Internet shopping and that products are accurately represented by websites	6.5
Q 28	Delivery costs are unreasonable	6.5
Q 29	In general I get good customer service from online companies	5.5
Q 30	Prices online are generally lower than elsewhere	4.8
Q 31	I find it encouraging to shop online from sites which have a clear return & refund policy	5.5
Q 32	It is important for me if a shopping site has the ability to deliver the order to an address other than my own	3.8
Q 33	It makes me feel more confident when the site keeps me informed about my order status	7.0
Q 34	I prefer to shop online from sites that provide alternative methods of ordering/payment/delivery	6.5
Q 35	I find it frustrating that some sites have limited delivery areas	5.8

The performance data is presented in two tables (Table 5 and 6). Table 5 presents the mean time in seconds and the standard deviation for each task for novice and expert users. Table 6 presents the accuracy of the tasks for each task across the sites.

Table 5. Mean time (in Seconds) for each task across the three sites for novice and expert users

Task	Expert and Novice Groups	Site 1		Site 2		Site 3	
		Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Task 1	Novice Group	81.0000	36.72117	107.5000	42.28803	117.6000	58.10948
	Expert Group	53.4000	18.79835	83.0000	37.44032	99.0000	55.07409
	Total	**67.2000**	31.72679	**95.2500**	40.85388	**108.5000**	55.88758
Task 2	Novice Group	296.1000	79.11799	406.5000	28.76437	243.7000	111.85511
	Expert Group	247.3000	76.24529	326.9000	68.20793	169.3000	33.47653
	Total	**271.7000**	79.65855	**366.7000**	65.29214	**206.5000**	88.96096
Task 3	Novice Group	116.1000	46.29963	154.5000	44.28004	131.2000	42.12627
	Expert Group	71.3000	24.95796	97.8000	37.15373	72.8000	21.97878
	Total	**93.7000**	42.87939	**126.1500**	49.28144	**102.0000**	44.35028
Task 4	Novice Group	140.3000	32.23887	128.8000	50.74074	170.1000	15.37278
	Expert Group	123.4000	40.74092	127.1000	28.14822	168.3000	12.51710
	Total	**131.8500**	36.79284	**127.9500**	39.94532	**169.2000**	13.67518
Task 5	Novice Group	86.5000	39.39614	105.7000	14.56060	76.5000	35.31839
	Expert Group	80.7000	29.65749	73.5000	32.04944	82.1000	36.25052
	Total	**83.6000**	34.06866	**89.6000**	29.32288	**79.3000**	34.95124
Task 6	Novice Group	155.6000	43.06636	109.8000	54.05923	164.4000	21.59321
	Expert Group	112.8000	49.73664	80.5000	31.86865	159.4000	27.71762
	Total	**134.2000**	50.32275	**95.1500**	45.73065	**161.9000**	24.31785
Task 7	Novice Group	33.2000	21.82659	27.9000	29.08016	20.0000	11.84155
	Expert Group	31.2000	20.82093	17.6000	11.12754	17.2000	12.06280
	Total	**32.2000**	20.78613	**22.7500**	22.07136	**18.6000**	11.72222
Task 8	Novice Group	72.1000	45.01469	97.7000	33.25675	67.0000	45.69464
	Expert Group	62.2000	38.48752	68.8000	41.67013	50.2000	40.07992
	Total	**67.1500**	41.07666	**83.2500**	39.57521	**58.6000**	42.71127
Task 9	Novice Group	63.8000	40.26247	57.4000	23.33429	116.8000	8.50882
	Expert Group	33.2000	21.92310	35.9000	18.50195	99.9000	31.30655
	Total	**48.5000**	35.24127	**46.6500**	23.27473	**108.3500**	23.95231
Task 10	Novice Group	111.7000	70.16021	50.0000	62.85256	43.9000	36.80112
	Expert Group	74.6000	57.64296	34.1000	42.44591	18.0000	12.24745
	Total	**93.1500**	65.32854	**42.0500**	52.83188	**30.9500**	29.81782

Table 6. Tasks accuracy

Task	Expert and Novice Groups	Site 1	Site 2	Site 3
		Accuracy Score		
Task 1	Novice Group	100%	100%	60%
	Expert Group	100%	100%	80%
Task 2	Novice Group	100%	30%	80%
	Expert Group	100%	90%	100%
Task 3	Novice Group	70%	40%	70%
	Expert Group	90%	100%	100%
Task 4	Novice Group	100%	100%	50%
	Expert Group	100%	100%	80%
Task 5	Novice Group	70%	60%	80%
	Expert Group	90%	100%	80%
Task 6	Novice Group	10%	30%	0%
	Expert Group	60%	70%	0%
Task 7	Novice Group	100%	100%	100%
	Expert Group	100%	100%	100%
Task 8	Novice Group	50%	40%	60%
	Expert Group	70%	60%	80%
Task 9	Novice Group	100%	100%	20%
	Expert Group	100%	100%	40%
Task 10	Novice Group	60%	100%	80%
	Expert Group	80%	100%	90%

Consequently, sixteen common areas of usability problems were identified which suggested identifying sixteen problem sub-themes. These sixteen problem sub-themes suggested identifying seven problem themes based on

the types of the identified problems. The seven problem themes related to: navigation, content, design, architecture, internal search, purchasing process and accessibility and customer service. Table 7. Shows the sixteen problem subthemes, their themes and the description of each.

Table 7. Usability problem themes and sub themes that were identified by the performance data and observations, together with their descriptions

Problem theme	Problem Sub theme	Description of the problem
Navigation	Misleading Links	The destination page, which was opened by the link, was not expected by users because the link name did not match the content of the destination page.
	Links were not obvious	Link was not situated in an obvious location on a page for it to be recognized by the users.
	Weak navigation support	A page did not have a navigation menu or links to other pages in a site.
Content	Irrelevant content	The content of the page are not clear to users because the page displayed an unclear message or had repetitive content or had empty content(that is the page was under construction)
Design	Misleading images	An image did not function as users expected. For example, it did not have a link when it suggested to users that it had one.
	Inappropriate page design	A page did not clear represent its content or it had an inappropriate design such as being long and/or displaying large number of images, or had inappropriate headings.
Architecture	Poor Structures	The structure or architecture of a site was neither simple nor straightforward enough to find information or products.
Internal search	Inaccurate results	The results of the internal search were inaccurate.
Purchasing process	Difficulty in knowing what was required for some fields	The site had pages with some entry fields where the required information to be entered was not clear to users.
	Difficulty in distinguishing between required and non required fields	The site had pages with some entry fields where there was no clear distinction between required and non-required fields.
	Difficulty in knowing what links were needed to be clicked	The site had pages with information that could be updated. Links had to be clicked in order to confirm this update but the links did not reveal that users had to click them to update the information.
	Session problem	The site had a session problem in which it did not save users' information, so users had to enter their information for each transaction during the same session.
	Required fields were not logical	The site had pages with some entry fields where the required fields were not logical.
	Expected information not displayed after adding products to cart	The site did not display expected information (i.e. confirmation) after users had added products to their cart.
Accessibility and Customer Service	Not easy to find help/customer support information	The site did not display the help/customer service information in an obvious location to be noticed and accessed by users.
	Inappropriate information provided within a help section/customer service	Some pages that displayed help /customer information had inappropriate content that did not match users' needs or expectations.

Table 8 shows the common areas of usability problems, the tasks that identified each problem, and the location of each problem on each site. The location of the problems was named either "entire-site" or by the title of the page with the problem. Entire site problems were identified as problems users faced in any page on the site. Table 8 also shows that during some tasks more than one problem was identified.

Table 8. Usability problem themes and sub-themes identified by performance data and observation and their locations per task

Problem Theme	Problem Sub-Theme	Site 1		Site 2		Site 3	
		Tasks identified the problem	Location	Tasks identified the problem	Location	Tasks identified the problem	Location
Navigation	Misleading Links	Task 2	Any product's page('Check out' link)	Task 2	Shipping page('go' link)	Task 10	Home page of the site('Our services' link)
		Task 2	Add to Cart End page('buy now' link)				
		Task 5	Entire site – Top Menu('sign in' and 'register' links)	Task 5	My account page('address book' link)		Home page of the Mall('Our Services' link)
		Task 10					
		Task 10	Entire site('advanced search' link)				
	Link are not obvious	Task 3	Entire site('Shopping cart' link)	Task 3	Entire site('Shopping cart' link)	Task 2	Any product's page('complete product' and 'shopping basket' links)

				Task 4	Order preview page('home page' link)	Task 3	Entire site('Shopping basket' link)
						Task 6	Home page of the Mall('online catalog' link)
	Weak Navigation Support	Task 3 Task 4 Task 5 Task 6	Order preview page(did not have navigational menus or links to the home page or to other pages)	Not Exist	Not Exist	Task 3	Order page(did not have navigational menus or links to the home page or to other pages)
						Task 2 Task 3	Shopping cart page(did not have navigational menus or links to the home page or to other pages)
Content	**Irrelevant Content**	Task 2 Task 4	Shipping information page(confusing error message was displayed all the time)	Not Exist	Not Exist	Task 1 Task 4 Task 6	Online catalog sub section (displayed products which were not ready to selling)
						Task 1 Task 4 Task 9	Search Mal Page(under construction page)
Design	**Misleading Images**	Task 3	Order preview page(site's logo)	Not Exist	Not Exist	Task 3	Entire site(site's logo)
	Inappropriate page design	Not Exist	Not Exist	Task 1 Task 2 Task 4	Any product's page(inappropriate presentation of product's description)	Not Exist	Not Exist
				Task 2	Login page(;new and current customer fields) Address page('shipping and billing' fields)		
Architecture	**Poor Structure**	Not Exist	Not Exist	Not Exist	Not Exist	Task 1 Task 4 Task 9	Entire Site
Internal search	**Inaccurate results**	Task 6	Entire Site(product search)	Task 6	Entire Site(product search)	Not Exist	Not Exist
Purchasing Process	**Difficulty in knowing What was Required for some Fields**	Task 2	Free shipping coupon page('free shipping coupon' field)	Task 2	Shipping page('gift certificate code' field)	Not Exist	Not Exist
	Difficulty in Distinguishing between Required and Non-Required Fields	Not Exist	Not exist	Task 2	Login page('password' fiels)	Task 2	Personal information page (some field required)
					Address page(some required field)	Task 3 Task5	
	Difficulty in Knowing what links were required to be clicked	Task 3	Shopping cart page('update order' link)	Not Exist	Not Exist	Task 3	Shopping cart page('ok' link)
	Session problem	Not Exist	Not Exist	Not Exist	Not Exist	Task 4 Task 9	Personal Information page(did not keep the users information)
	Required fields were not logical	Task 2	Registration page('state/province' field)	Task 2	Address page('state/region' field)	Not Exist	Not Exist
	Expected Information was not displayed after adding products to cart	Task 2 Task 4	Add to cart end page	Not Exist	Not Exist	Task 2	Product page
Accessibility and customer service	**Not easy to find help/customer support information**	Task 8	Entire site	Task 8	Entire site	Task 8	Entire site
	Inappropriate information provide within a help section /customer service	Not Exist	Not Exist	Not Exist	Not Exist	Task 8	FAQ page

Analysis of the performance data and observations provided the following general findings regarding the overall usability of the sites; the observation summary showed that expert and novice users experienced many similar problems, obstacles or difficulties performing the different tasks across the sites. The difference between experts and novices is the fact that experts recover faster. This explains why novice users had a larger number of problematic tasks, as shown in Table 6. The total number of tasks successfully performed by all the users (experts and novices) was lowest in site 3 (Table 6). This indicates that sites 1 and 2 were noticeably better than site 3. As expected, the percentage of experts who successfully completed each task was higher than the percentage of novices. This was due to their higher level of knowledge. A one-way within-subject ANOVA test showed the time spent performing the majority (eight) of the ten tasks was significantly different for the three sites. Table 9 shows the results of the ANOVA test for each task.

Table 9. Result of One-Way within-Subjects ANOVA test for each task among the three sites

Task	ANOVA Test (One-Way within- Subjects)
	Was there a statistically significant difference among site 1, site 2 and Site 3
Task 1	**Yes** $F_{(2,38)} = 6.021$, $p=.005$
Task 2	**Yes** $F_{(2,38)} = 33.183$, $p=.000$
Task 3	**Yes** $F_{(2,38)} = 4.471$, $p=.018$
Task 4	**Yes** $F_{(2,38)} = 10.873$, $p=.000$
Task 5	**No** $F_{(2,38)} = .502$, $p=.609$
Task 6	**Yes** $F_{(2,38)} = 16.517$, $p=.000$
Task 7	**Yes** $F_{(2,38)} = 4.369$, $p=.020$
Task 8	**No** $F_{(2,38)} = 2.364$, $p=.108$
Task 9	**Yes** $F_{(2,38)} = 40.407$, $p=.000$
Task10	**Yes** $F_{(2,38)} = 8.814$, $p=.001$

A mixed ANOVA design test showed. Experts performed all the tasks significantly faster than novices; this was determined by assessing the effect of the Group factor: $f_{(1, 18)} = 13.644$, $p =.002$. The total time spent on each site to perform all the tasks was not significantly different, demonstrated by the assessment of the effect of Sites factor $f_{(2, 36)} = 2.010$, $p =.149$. The time spent on performing each of the ten tasks was significantly different for the three sites, determined by assessing the interaction between Sites and Tasks factors $f_{(18,324)} = 16.439$, $p =.000$. This result is consistent with the one-way within-subjects ANOVA analysis.

A list of usability problems were identified from the negative statements (statements with Likert score rating of 1 to 3) in the satisfaction questionnaires. Each problem in the list and the problem sub-themes which were identified by the performance data and observation method, were compared for agreement. Consequently, these Statements were mapped to the identified problem themes and sub-themes. Four statements identified three new problem sub-themes that were mapped to the navigation, design and purchasing process problem themes. These problems, as well as their description, are shown in Table 10.

Table 10. New problem themes and sub themes that were identified by the quantitative data of the post test questionnaires, together with their descriptions

Problem theme	Problem sub theme	Description of the problem
Navigation	Broken links	The site had pages with broken links.
Design	Unaesthetic design	The site did not have an aesthetically pleasing nor attractive interface.
Purchasing Process	Compulsory registration	The site requires users to register to the site to proceed in the checkout process

The negative statements, their Likert scores and the problem themes and sub-themes identified by these statements, are shown in Table 11.

Table 11. Usability problem themes and sub-themes identified by the post test questionnaire

Problem theme	Problem Sub Theme	Statement Number in the Post Test questionnaire	Likert Score		
			Site 1	Site 2	Site 3
Navigation	Weak Navigation Support	9			2.55
		10			3.70
	Broken Links	24			3.85
Content	Irrelevant Content	14			3.50
		27			3.25
Design	Unaesthetic Design	20			3.80
	Inappropriate Page Design	25			2.95
Architecture	Poor Structure	1			2.95
		2			2.60
		8			2.70
Purchasing Process	Compulsory Registration	15	3.25	2.75	
		16			2.25

The following points represent the general findings for the overall usability of the sites; The Mann-Whitney test showed there were no significant differences between novice and expert users for a large number of the post test statements (Table 12).

Table 12. Likert Scores of the post-test questionnaire for the three sites for novice and expert users and the result of Mann-Whitney test

No.	Site 1			Site 2			Site 3		
	Likert Score		Mann-Whitney test	Likert Score		Mann-Whitney test	Likert Score		Mann-Whitney test
	Novice Group	Expert Group	Was there statistically a significant difference between Novice and Expert Groups	Novice Group	Expert Group	Was there statistically a significant difference between Novice and Expert Groups	Novice Group	Expert Group	Was there statistically a significant difference between Novice and Expert Groups
Q 1	5.1	6.6	Yes (U=18.500,N1=10, N2=10,p=.015,two tailed)	4.5	5.8	Yes (U=22.500,N1=10, N2=10,p=.035,two tailed)	3.1	2.8	No (U=48.500,N1=10, N2=10,p=.912,two tailed)
Q 2	5.6	6.6	No (U=24.000,N1=10, N2=10,p=.052,two tailed)	5.7	6.3	No (U=34.500,N1=10, N2=10,p=.247,two tailed)	2.6	2.6	No (U=44.500,N1=10, N2=10,p=.684,two tailed)
Q 3	5.7	6.5	No (U=27.000,N1=10, N2=10,p=.089,two tailed)	5.0	6.2	No (U=26.500,N1=10, N2=10,p=.075,two tailed)	NA	NA	NA
Q 4	5.2	6.5	Yes (U=20.500,N1=10, N2=10,p=.023,two tailed)	4.9	6.1	No (U=28.500,N1=10, N2=10,p=.105,two tailed)	3.7	2.0	No (U=27.500,N1=10, N2=10,p=.089,two tailed)
Q 5	5.9	6.1	No (U=45.000,N1=10, N2=10,p=.739,two tailed)	5.4	5.8	No (U=43.500,N1=10, N2=10,p=.631,two tailed)	4.4	3.9	No (U=41.000,N1=10, N2=10,p=.529,two tailed)
Q 6	4.9	5.7	No (U=39.000,N1=10, N2=10,p=.436,two tailed)	4.6	5.8	No (U=33.500,N1=10, N2=10,p=.218,two tailed)	3.8	2.9	No (U=37.500,N1=10, N2=10,p=.353,two tailed)
Q 7	4.3	5.6	No (U=28.000,N1=10,	4.4	5.5	No (U=29.500,N1=10,	4.2	3.4	No (U=39.500,N1=10,

Q									
			N2=10,p=.105,two tailed)			N2=10,p=.123,two tailed)			N2=10,p=.436,two tailed)
Q 8	5.5	6.6	Yes (U=22.000,N1=10, N2=10,p=.035,two tailed)	5.1	5.9	No (U=32.500,N1=10, N2=10,p=.190,two tailed)	3.4	2.0	No (U=25.500,N1=10, N2=10,p=.063,two tailed)
Q 9	5.9	5.4	No (U=49.000,N1=10, N2=10,p=.971,two tailed)	3.5	4.8	No (U=31.500,N1=10, N2=10,p=.165,two tailed)	2.7	2.4	No (U=40.500,N1=10, N2=10,p=.481,two tailed)
Q 10	4.9	6.3	Yes (U=23.000,N1=10, N2=10,p=.043,two tailed)	5.1	6.3	No (U=29.000,N1=10, N2=10,p=.123,two tailed)	4.1	3.3	No (U=39.500,N1=10, N2=10,p=.436,two tailed)
Q 11	5.3	6.3	No (U=27.000,N1=10, N2=10,p=.089,two tailed)	5.4	6.1	No (U=35.500,N1=10, N2=10,p=.280,two tailed)	NA	NA	NA
Q 12	5.6	6.5	No (U=22.000,N1=10, N2=10,p=.035,two tailed)	5.4	6.1	No (U=37.500,N1=10, N2=10,p=.353,two tailed)	NA	NA	NA
Q 13	5.4	5.9	No (U=35.500,N1=10, N2=10,p=.280,two tailed)	5.8	5.6	No (U=50.000,N1=10, N2=10,p=1.000,two tailed)	NA	NA	NA
Q 14	4.5	5.8	No (U=25.000,N1=10, N2=10,p=.063,two tailed)	4.5	5.6	No (U=32.000,N1=10, N2=10,p=.190,two tailed)	3.9	3.1	No (U=37.500,N1=10, N2=10,p=.353,two tailed)
Q 15	3.4	3.1	No (U=46.500,N1=10, N2=10,p=.796,two tailed)	2.0	3.5	No (U=26.500,N1=10, N2=10,p=.075,two tailed)			
Q 16							2.0	2.5	No (U=40.500,N1=10, N2=10,p=.481,two tailed)
Q 17	4.7	6.2	No (U=26.500,N1=10, N2=10,p=.075,two tailed)	3.9	4.8	No (U=35.500,N1=10, N2=10,p=.280,two tailed)	2.5	2.0	No (U=40.500,N1=10, N2=10,p=.481,two tailed)
Q 19	5.8	6.0	No (U=49.500,N1=10, N2=10,p=.971,two tailed)	5.6	5.2	No (U=35.500,N1=10, N2=10,p=.280,two tailed)	4.1	4.1	No (U=49.500,N1=10, N2=10,p=.971,two tailed)
Q 20	5.9	6.0	No (U=44.500,N1=10, N2=10,p=.684,two tailed)	5.8	5.2	No (U=36.000,N1=10, N2=10,p=.315,two tailed)	4.1	3.5	No (U=40.500,N1=10, N2=10,p=.481,two tailed)
Q 21	4.4	4.1	No (U=47.500,N1=10, N2=10,p=.853,two tailed)	5.7	4.8	No (U=41.000,N1=10, N2=10,p=.529,two tailed)	4.6	4.1	No (U=46.000,N1=10, N2=10,p=.796,two tailed)
Q22	5.5	6.1	No (U=47.000,N1=10, N2=10,p=.089,two tailed)	4.4	5.6	Yes (U=27.000,N1=10, N2=10,p=.089,two tailed)	4.8	4.7	No (U=43.500,N1=10, N2=10,p=.631,two tailed)
Q23	5.8	5.0	No (U=40.000,N1=10, N2=10,p=.481,two tailed)	4.7	5.5	No (U=33.000,N1=10, N2=10,p=.218,two tailed)	4.7	4.5	No (U=49.500,N1=10, N2=10,p=.971,two tailed)
Q 24	5.5	4.7	No	4.5	4.6	No	4.8	2.9	No

			(U=45.500,N1=10, N2=10,p=.739,two tailed)			(U=47.500,N1=10, N2=10,p=.853,two tailed)			(U=25.500,N1=10, N2=10,p=.063,two tailed)
Q 25	5.1	5.8	No (U=34.500,N1=10, N2=10,p=.247,two tailed)	4.0	4.7	No (U=36.500,N1=10, N2=10,p=.325,two tailed)	3.2	2.7	No (U=41.500,N1=10, N2=10,p=.529,two tailed)
Q 26	4.3	6.2	Yes (U=18.000,N1=10, N2=10,p=.015,two tailed)	4.3	5.2	Yes (U=36.500,N1=10, N2=10,p=.315,two tailed)	2.6	2.5	No (U=50.000,N1=10, N2=10,p=1.000,two tailed)
Q 27	5.2	5.3	No (U=47.500,N1=10, N2=10,p=.853,two tailed)	4.4	5.5	Yes (U=36.500,N1=10, N2=10,p=.315,two tailed)	3.1	3.4	No (U=46.500,N1=10, N2=10,p=.796,two tailed)
Q 28	4.5	5.6	Yes (U=30.500,N1=10, N2=10,p=.143,two tailed)	4.8	4.9	No (U=49.500,N1=10, N2=10,p=.971,two tailed)	3.5	2.0	No (U=26.500,N1=10, N2=10,p=.075,two tailed)
Q 29	5.7	5.9	No (U=46.000,N1=10, N2=10,p=.796,two tailed)	5.7	5.3	Yes (U=43.500,N1=10, N2=10,p=.631,two tailed)	5.2	5.0	No (U=48.500,N1=10, N2=10,p=.912,two tailed)
Q 30	5.5	4.9	No (U=42.500,N1=10, N2=10,p=.579,two tailed)	5.0	5.2	No (U=47.500,N1=10, N2=10,p=.853,two tailed)	5.0	3.7	No (U=34.000,N1=10, N2=10,p=.247,two tailed)
Q 31	5.5	5.3	No (U=47.000,N1=10, N2=10,p=.853,two tailed)	5.3	5.1	No (U=45.500,N1=10, N2=10,p=.739,two tailed)	4.6	3.4	No (U=32.000,N1=10, N2=10,p=.190,two tailed)

Consequently, the ratings of novice and expert users were combined for each statement concerning the post test questionnaire. The Friedman test was used after combining the ratings of novice and expert users. This showed that there were statistically significant differences between users' ratings of the three sites for all the statements, as shown in Table 13.

Table 13. Likert scores of the post-test questionnaire and the result of Friedman test

No.	Question	Likert Score			Friedman Test
		Site 1	Site 2	Site 3	Was there a statistically significant difference among site 1,site 2 and site 3
Architecture and Navigation					
Q 1	Finding the information related to the tasks(was very easy)	5.85	5.15	2.95	**Yes** X^2 (2, *N*=20)= 30.714, *p*=.000
Q 2	Finding the products(was very easy)	6.10	6.00	2.60	**Yes** X^2 (2, *N*=20)= 34.125, *p*=.000
Q 3	Using the internal search facility(was very easy)	6.10	5.60	NA	NA
Q 8	The organization of information of the website was clear	6.05	5.50	2.70	**Yes** X^2 (2, *N*=20)= 35.273, *p*=.000
Q 9	Moving around the website without getting lost was difficult	5.65	4.15	2.55	**Yes** X^2 (2, *N*=20)= 25.016, *p*=.000
Q 10	The table of contents was helpful	5.60	5.70	3.70	**Yes** X^2 (2, *N*=20)= 31.356, *p*=.000
Q 11	The site's search function was quick enough	5.80	5.75	**NA**	**NA**
Q 12	Accuracy of internal search result was good	6.05	5.75	**NA**	**NA**
Q 13	Results if internal search were poor	5.65	5.70	**NA**	**NA**
Q 21	It was difficult to go to the home page from any sub page of the sites	4.25	5.25	4.35	**Yes** X^2 (2, *N*=20)= 17.644, *p*=.000
Q 24	There were few broken/not working links	5.10	4.55	3.85	**Yes** X^2 (2, *N*=20)= 15.796, *p*=.000
Content					
Q 14	The information of the website was effective in helping me complete the purchasing tasks	5.15	5.05	3.50	**Yes** X^2 (2, *N*=20)= 22.172,

					p=.000
Q 27	The terminology/terms use throughout these website were clear	5.25	5.05	3.25	**Yes** X^2 (2, N=20)= 22.116, p=.000
Design					
Q 19	I liked the interface of this website	5.90	5.40	4.10	**Yes** X^2 (2, N=20)= 31.115, p=.000
Q 20	The interface of this website was pleasant/attractive	5.95	5.50	3.80	**Yes** X^2 (2, N=20)= 33.323, p=.000
Q 22	The choice of colors was appropriate	5.80	5.00	4.75	**Yes** X^2 (2, N=20)= 18.473, p=.000
Q 23	The size of the text made the site easy to read	5.40	5.10	4.60	**Yes** X^2 (2, N=20)= 12.792, p=.000
Q 25	It was clear to know the position of any page of the site	5.45	5.25	2.95	**Yes** X^2 (2, N=20)= 29.284, p=.000
Purchasing Process					
Q 4	Registering on the site(was very easy)	5.85	5.50	NA	NA
Q 5	Purchasing a product (was very easy)	6.00	5.60	4.15	**Yes** X^2 (2, N=20)= 30.632, p=.000
Q 6	Changing customer information (was very easy)	4.95	4.95	3.35	**Yes** X^2 (2, N=20)= 9.033, p=.011
Q 7	Changing the control of shopping cart (was very easy)	5.30	5.20	3.80	**Yes** X^2 (2, N=20)= 24.824, p=.000
Q 15	Compulsory registration in order to purchase products was convenient	3.25	2.75	NA	NA
Q 16	I prefer to register before purchasing products	NA	NA	2.25	NA
Q 29	I trust that the company will not misuse my personal information	5.80	5.50	5.10	**Yes** X^2 (2, N=20)= 14.176, p=.001
Q 30	I feel that the security of my financial information is protected while purchasing from this website	5.20	5.10	4.35	**Yes** X^2 (2, N=20)= 14.245, p=.001
Q 31	I have confidence in purchasing from this website	5.40	5.20	4.00	**Yes** X^2 (2, N=20)= 26.655, p=.000
The Overall Evaluation of the Sites					
Q 17	This website had all the functions and capabilities that I expected it to have	5.45	4.35	2.25	**Yes** X^2 (2, N=20)= 37.014, p=.000
Q 26	I felt comfortable using this website	5.25	4.75	2.55	**Yes** X^2 (2, N=20)= 25.400, p=.000
Q 28	I would recommend this site to a friend	5.05	4.85	2.75	**Yes** X^2 (2, N=20)= 28.212, p=.000

In these statements, site 3 had the lowest ratings for all the following aspects: navigation and architecture, content, design and purchasing process. The Likert scores for the overall evaluation statements also showed that site 3 rated negatively with the lowest rating for all statements. Site 1 rated positively with the highest rating and site 2 rated neutral. However, the Friedman test was not used for seven statements. For these statements, site 3 had no ratings for six statements and sites 1 and 2 had no ratings for one statement. Site 3 had no rating for four statements (3, 11, 12, 13) concerning the internal search as it did not have such a facility and for two statements (4, 15) as it did not enable registration. Sites 1 and 2 had no ratings for one statement (16) as they did not have optional registration.

Analysis of the qualitative data from the post-test questionnaires showed novice and expert users experienced similar usability problems in the sites. For this reason (and since the results of the Mann-Whitney test showed no significant difference between novice and expert users for many of the post-test statements (Table 12), answers from novice and expert users for each question of the post test questionnaire were combined. However, usability problems identified only by expert users were highlighted by noting 'expert' next to these answers. These problems were compared and then mapped to the appropriate problem themes and sub-themes identified by the previous two methods (performance data and observation, and the quantitative data from the satisfaction questionnaires). No match was found between nine problems and the identified problem sub-themes. Therefore, two new problem sub-themes identified two new problem themes relating to an inconsistency problem and missing capabilities. Seven new subthemes were also identified. These sub-themes were mapped to six

appropriate problem themes (navigation, internal search, content, design, purchasing process and customer service). Table 14. Shows the new problem themes and sub-themes and their descriptions.

Table 14. New problem themes and sub themes that were identified by the qualitative data of the post test questionnaires, together with their descriptions

Problem theme	Problem sub theme	Description of the problem
Navigation	Orphan pages	The site had dead-end pages that did not have any link.
Internal Search	Limited options	The internal search facility had limited options to search the site.
Content	Inaccurate information	The site displayed inaccurate information. For example, it displayed out of stock products or gave an inaccurate description for some products.
	Missing information about the products	Adequate information about the products was not displayed, such as: availability/stock indication, fabric, representative (large) images, length and width of some products, size guide.
Design	Inappropriate choice of fonts and colors	The site used an inappropriate font size (i.e. small size) or inappropriate font style (i.e. bold font style for many sentences on the same page) or inappropriate combination of background and link colors.
Purchasing Process	Long ordering process	Ordering process pages included more than one page with similar content which increased the number of steps required to purchase from a site.
Accessibility and Customer Service	Not supporting more than one language	The site did not display its content in languages other than English.
Inconsistency	Inconsistent design/layout/content.	The site's design, layout or content was Inconsistent throughout the site. For example, the content on Arabic and English interfaces was inconsistent.
Missing Capabilities	Missing functions/information	The site did not have some functions or Capabilities (i.e. an internal search facility) or it did not display adequate information.

Table 15 summarizes all the usability problem themes and sub-themes identified by the qualitative data of the post test questionnaires and their location on the sites

Problem Theme	Problem Sub-Theme	Site 1	Site 2	Site 3
		Location	Location	Location
Navigation	Misleading Links	Entire site – Top Menu('sign in' and 'register' links)	Shipping page('go' link)	Home page of the site('Our services' link)
			Not Exist	Not Exist
	Link are not obvious	Not Exist	Login page(Home page link)	Not Exist
			Address page (Home page link)	
			Shipping and Payment Page(home page link)	
			Shipping cart problem (home page link)	
			Order preview page('home page' link)	
	Weak Navigation Support	Order preview page(did not have navigational menus or links to the home page or to other pages)	Not Exist	Not Exist
				Home page
	Broken links	Not Exist	Not Exist	Online Catalog Subsection-search results page
				Online Catalog subsection Banner
				Related Links pages
	Orphan pages	Not Exist	Not Exist	Product's Image page(Larger view)for any product's page
Content	Irrelevant Content	Not Exist	Not Exist	Entire site(most pages had repetitive/not concise content)
	Inaccurate Information	Any product's page(displayed out of stock products)	Any product's page(displayed out of stock products)	Not Exist
	Missing Information about the Products	Any product's page (availability)	Any product's page (availability)	Any product's page (availability)
Design	Inappropriate page design	All product category pages(long pages with large number of images)	Any product's page(inappropriate presentation of product's description)	Not Exist
		Best/Most seller page (products are displayed at the bottom)		
	Inappropriate choice of Fonts and colors	Not Exist	Entire site(small font size for menus and text, combination of background and link colors)	Not Exist
Architecture	Poor Structure	Not Exist	Not Exist	Entire Site
Internal search	Inaccurate results	Entire Site(product search)	Entire Site(product search)	Not Exist
	Limited option	Entire site(product and advanced search)	Entire site(product and advanced search)	Not Exist
Purchasing (Check out)Process	Difficulty in Distinguishing between Required and Non-Required Fields	Not exist	Login page('password' fields)	Personal information page (some field required)
			Address page(some required field)	
	Long Ordering Process	Add to Cart End page	Not Exist	Not Exist
		Checkout Page		
	Session problem	Not Exist	Not Exist	Personal information page (did not keep the users information)
Accessibility and customer service	Not supporting the more than one language	Entire site	Entire site	Not Exist
Inconsistency	Inconsistent Design/Layout/content	Not Exist	Not Exist	Not Exist
Missing Capabilities	Missing Information/Functions	Not Exist	Not Exist	Not Exist

Analysis of the seven open-ended questions on the post evaluation questionnaire (relating to the site with the best features from the users' point of view) did not explicitly identify specific usability problems. It only provided

information on the overall usability of the sites from the users' point of view in terms of six features of the sites, Navigation: The answers to two questions (2, 6) indicated that the navigation support of sites 1 and 2 enabled users to find products and information easily. The number of users who recommended site 1 was higher than the number who recommended site 2. Site 1 had the most obvious and simplest methods for finding products and was the easiest site to find information related to the tasks, Internal Search: Answers to two questions (2, 6) indicated that the internal searches of sites 1 and 2 enabled products and information to be easily located, Architecture: Answers to two questions (2, 6) on the post-evaluation questionnaire indicated that the simple, straightforward architecture of sites 1 and 2 enabled users to find products and information easily. More users recommended site 1 than site 2. A few users (two) preferred the architecture of the Arabic interface of site 3 to the architecture of the other two sites because it used their first language, Design: The answer to one question (1) on the post-evaluation questionnaire indicated that site 1 had the most professional appearance. Few users recommended site 2 and none recommended site 3, Purchasing Process: The answers to three questions (3, 4, 7) showed that most users recommended site 1 as the site with the most obvious method for ordering items. Most users recommended site 1 as having the best support for customers (to continue shopping) and to change the contents of their shopping cart. Most users recommended site 2 as the easiest for changing customer information. No user recommended site 3. Security and Privacy: The answers to question 5 (related to the site users trusted the most) recommended site 1. Few users recommended site 2 and none recommended site 3. Only two users indicated that their reason for trusting sites 1 and 2 related to the sites' use of the secure socket layer. All the users who recommended site 1 indicated other reasons for their recommendations which did not relate to the site's design issues. They mentioned that this site is a famous and well-known company with a good reputation.

5. Evaluating the usefulness of the suggested framework

An interview conducted with each site's manager. During the interviews, and after discussing the results, the usefulness of the framework was tested. The results obtained from the interviews with the managers concerning testing the usefulness of the framework are presented below.

5.1 The usefulness of usability evaluation

All the companies agreed that the usability evaluation of their websites was useful and was an important technique. All the companies were interested in gaining the knowledge regarding the usability method that is employed in this research, and in their ability to identify the large number of problems on their sites. They indicated that they did not have any knowledge regarding usability evaluation methods before taking part in this research. Two of the companies (companies one and two) indicated that they were using other methods to collect feedback from their customers regarding what they liked or disliked on their websites. They used survey by email which was sent to their customers more than once. The companies were interested in receiving useful information about the weaknesses of their websites by taking part in this research and by trying these new methods. After receiving the results, they said that there was no comparison between the results gained from employing the usability evaluation methods and the survey that they generally use. The usability evaluation method provided them with rich, useful and detailed information which was above their expectations. These companies were glad they had decided to take part in this research.

5.2 The usefulness and expectations of the results

All the companies indicated that the results were very useful, interesting and unexpected. None of them expected the number and types of problem that were identified on their sites. Two of the companies (companies one and two) indicated that once they received the results, they fixed certain problems on their websites which were easily implemented. The recommendations that were presented with each problem in the report that was sent to them encouraged them to correct these problems. The companies provided examples regarding the problems that were dealt with. For example, company one reported that they fixed eight problems: two relating to the navigation area and six related to the content area. Company two reported that they fixed four problems: two related to the navigation area and two related to the content area. Furthermore, these companies indicated that they are in the process of fixing the different types of problem priority being given to all the purchasing process problems. They stated that addressing the problems is now within their short-term plans.

Company three, however, did not indicate that they had fixed any problems on their website based on the outcomes of this research, in spite of having stated that they did not expect their site to have such a large number of problems. However, they did say that it was their intention to make major changes to the design of their site shortly. The large number of problems and the recommendations encouraged this company to take this decision. The companies' feedback regarding the problems that were fixed, and the decision these companies made regarding fixing the other types of problem, represent further evidence of the usefulness of the results.

5.3 Types of problem the companies were interested to identify

The companies, by referring to their results which were categorized in terms of the specific problems themes and sub-themes, reported the specific types of problem they were interested or not interested in identifying on their websites. They also indicated the methods which they would employ to identify these problems:

• Company one reported that they were interested in all the navigation, internal search, content and purchasing process problems that were identified by the user testing method. This company also was interested in one design problem that related to inappropriate page design; this was identified by user testing.

• Company two reported that they were interested in all the navigation, internal search, content, design and purchasing process problems that were identified by the user testing. This company was also interested in one accessibility and customer service problem which related to it not being easy to find help/customer support information; this problem was identified by the user testing.

• Company three reported that they were interested in all the problems that were identified on their website by both the user testing evaluation. They explained the reason for this by indicating that they were planning to make major changes in their website. Therefore they were interested in fixing all the problems which, from their perspective, were important and should be fixed. They will employ user testing.

6. Conclusion

A usability evaluation process framework was developed to evaluate the usability of e-commerce websites which is the strategic use of user testing method. It is based on the benefits and drawbacks of the method in term of the specific usability problems that could or could not identify on these types of websites. The suggested framework has managerial and academic implications. Regarding the managerial implication: E-commerce companies need to evaluate and improve their e-commerce websites in a way that will improve their success.

Regarding the academic implications: This paper presents an evaluation of three e-commerce sites in India as the basis for proposing a new approach for evaluating the usability of websites, specifically e-commerce sites. A particularly novel approach is the use user testing. This research has provided a detailed account of the use and evaluation of usability technique for e-commerce websites.

The aim of this research was to develop a methodological framework, which would comprehensively and effectively investigate usability problem areas of e-commerce websites. The development this framework was an attempt to raise awareness of usability and usability evaluation method in order to gain the benefits of e-commerce. This aim was achieved by meeting the specific objectives of this research.

The framework also offers a base for future research. In particular, the extent to which the application of a framework, which uses the method to reduce the time. Research should be undertaken after the three companies involved in this research have changed the design of their websites based on the recommendations offered by this research.

References

Barnard, L. & Wesson, J., 2003. Usability issues for e-commerce in South Africa: an empirical investigation. *In the Proceedings of SAICSIT 2003*, 258-267.

Brinck, T., Gergle, D. & Wood S.D., 2001. *Usability for the web: designing websites that work*. Morgan Kaufmann Publishers.

Claudio O. & Antonio, 2007.J. *A model for evaluating B2C e-commerce websites: Application in the CD e-retailing industry in Brazil*.

Dumas, J. S. & Redish, J. C., 1999. *A practical guide to usability testing*. Second Intellect Ltd, Rev Sub edition.

Rubin, J., 1994. *Handbook of usability testing: how to plan, design, and conduct effective tests*. Wiley.

Freeman, M.B. & Hyland, P., 2003. Australian online supermarket usability.

Goodwin, S., 2005. Using screen capture software for website usability and redesign buy-in. *Library Hi Tech*, 23(4).

Kuniavsky, M., 2003. *Observing the user experience: a practitioner's guide to user research*. San Francisco, Calif; London: Morgan Kaufmann.

Lewis, J.R., 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57-78.

McKinney, V., K. Yoon, F. Zahedi,2002. "The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach," Information System Research, 13(3), 296-315.

Najjar, L., 2005. Designing E-commerce User Interfaces. In: R. W. Proctor, K-P. L., Vu. (Eds.) *Handbook of Human Factors in Web Design*, pp. 514-527. Mahwah, NJ: Lawrence Erlbaum.

Nielsen, J. & Mack, R. L. (Eds.),1994. *Usability inspection methods*. John Wiley & Sons, New York.

Nielsen, J., 2003. *Usability 101: Introduction to usability*.

Oppenheim C. & Ward L., 2006. Evaluation of websites for B2C e-commerce. *AslibProceedings: New Information Perspectives*, 58(3).

Pearrow, M., 2000. *Website usability handbook*. Charles River Media.

Preece, J., Sharp, H. & Rogers, Y., 2002. *Interaction design: beyond human computer interaction*, John Wiley & Sons, Inc.

Sharp, H., Rogers, Y. & Preece, J., 2007. *Interaction Design: Beyond Human-Computer Interaction*. Wiley; Second Edition.

Tilson, R., Dong, J., Martin, S. & Kieke E., 1998. Factors and principles affecting the usability of four e-commerce sites. *In the Proceedings of the 4th Conference on Human Factors and the Web (CHFW).* AT&TLabs, USA.

White, G.K. & Manning B.J., 1998. Commercial WWW site appeal: how does it affect online food and drink consumers' purchasing behavior? *Internet Research: Electronic Networking Application and Policy*, 8(1), 32-38.*Technical report*, Decision Systems Lab, University of Wollongong.

Sandeep Kumar Panda has received his Bachelor degree from Utkal University, India in 2001 and his Master degree from the Biju Ptanaik University of Technology, India in 2009. He is currently a Ph.D. Scholar of the School of Computer Engineering at KIIT University, India. His academic interests lie in Human Factors, especially, usability engineering and human–computer interaction: constructing a conceptual framework for user preferences, developing a user, heuristic and tool based usability evaluation method, and analyzing human cognitive processes based on users' judgments for interactive applications.