# USABILITY EVALUATION FRAMEWORK FOR E-COMMERCE WEBSITES

Layla Hasan

*Department of Management Information Systems, Zarqa Private University, Zarqa, Jordan*

Anne Morris, Steve Probets

*Department of Information Science, Loughborough University, Loughborough, U.K.*

Abstract:     The importance of evaluating the usability of e-commerce websites is well recognised and several studies have evaluated the usability of e-commerce websites using either user- or evaluator-based usability evaluation methods. No research, however, has employed a software-based method in the evaluation of such sites. Furthermore, the studies which employed user testing and/or heuristic evaluation methods in the evaluation of the usability of e-commerce websites did not offer detail about the benefits and drawbacks of these methods with respect to the identification of specific types of usability problem. This research has developed a methodological framework for the usability evaluation of e-commerce websites which involves employing user testing and heuristic evaluation methods together with Google Analytics software.

## 1 INTRODUCTION

Usability is one of the most important characteristics of any user interface and is a measure of how easy the interface is to use (Nielsen, 2003). Researchers have stressed the importance of making e-commerce sites usable and have stated that good usability is not a luxury but an essential characteristic if a site is to survive (Nielsen and Norman, 2000).

Usability evaluation methods can be categorised by how the usability problems are identified, for example by users, evaluators or tools.

- User-based usability evaluation methods**:** This category includes a set of methods that involves users in the process of identifying usability problems. The user testing method is the most common approach in this category.
- Evaluator-based usability evaluation methods: This category includes usability methods that involve evaluators in the process of identifying usability problems. The most common method in this category is heuristic evaluation.
- Software-based usability evaluation methods: This category involves software tools in the process of identifying usability problems. An example of this approach is web analytics. Web analytics is an approach that involves collecting, measuring, monitoring, analysing and reporting web usage data to understand visitors' experiences (McFadden, 2005).

User- and evaluator-based approaches have been frequently used to evaluate the usability of e-commerce websites. However, little research has employed web analytic tools in the evaluation of such sites. The research described here aims to address this gap and presents a methodological framework which outlines how each of three methods could be used in the most effective manner for evaluating the usability of e-commerce sites.

This paper is organised as follows: Section 2 reviews related work, Section 3 describes web metrics and provides an example of a web analytics tool, Section 4 presents the aims and objectives of this research, Section 5 describes the methods used, Section 6 presents the main results, Section 7 illustrates the framework and finally, Section 8 presents some conclusions.

## 2 USABILITY EVALUATION OF E-COMMERCE WEBSITES

Only a few studies were found in the literature that evaluated the usability of e-commerce sites. Tilson et al.'s study (1998) is one that involved users in evaluating the usability of e-commerce websites. The researchers asked sixteen users to complete tasks on four e-commerce websites and report what they liked and disliked. Another study, conducted by Freeman and Hyland (2003), also involved users in evaluating the usability of e-commerce sites, in this case three supermarket sites. These studies proved the usefulness of user-based methods in identifying major design problems which prevent users from interacting with the sites successfully.

Chen and Macredie (2005) involved evaluators using the heuristic method to evaluate the usability of four online supermarkets. The results demonstrated the usefulness of the heuristic evaluation method regarding its ability to identify a large number of usability problems on the sites.

Barnard and Wesson (2004) employed both heuristic evaluation and user testing methods together to identify usability problems on e-commerce sites in South Africa. Significant usability problems were identified based only on the common usability problems that were identified by both the user testing and heuristic evaluation methods.

## 3 WEB METRICS AND GOOGLE ANALYTICS

Web metrics are employed to give meaning to web traffic data collected by web analytics tools. Web metrics can be placed into two categories: basic and advanced. Basic metrics are raw data which are usually expressed in raw numbers (i.e. visits). Advanced metrics are metrics which are expressed in rates, ratios, percentages or averages instead of raw numbers, and are designed to guide actions to optimise online business. Inan (2006) and Phippen et al. (2004) criticised the use of basic metrics to measure the traffic of websites. Instead, they suggest using advanced metrics.

An example of a web analytics tool is Google Analytics. Google Analytics (GA) was released to the public in August 2006 as a free analytics tool. At least two studies have recognised the appearance of GA software and used this tool to evaluate and improve the design of web sites (a library web site and an archival services web site) (Fang, 2007;

Prom, 2007). However, these studies used the standard reports from GA (i.e., content by titles, landing pages) without deriving specific metrics. These studies suggested that the GA's reports enable problems to be identified quickly (Fang, 2007; Prom, 2007).

The literature outlined above indicates that there has been a lack of research that evaluates the usability of e-commerce websites by employing user-based, evaluator-based and software-based (GA) usability evaluation methods together. Studies by Fang (2007) and Prom (2007) have illustrated the potential usefulness of using GA to evaluate websites with the intention of improving their usability. However, there is a lack of research to illustrate the value of using GA for evaluating the usability of e-commerce websites by employing advanced web metrics. Furthermore, it is clear from the literature that there is a lack of research that compares user testing and heuristic evaluation methods for identifying detailed types of specific usability problems found on e-commerce websites.

## 4 AIMS AND OBJECTIVES

The aim of the research described here was to develop a methodological framework to investigate the usability of e-commerce websites.
The specific objectives for the research were:
- To use user testing, heuristic evaluation and GA to evaluate a selection of e-commerce websites.
- To identify the main usability problem areas.
- To determine which methods were the most effective in evaluating each usability problem area.
- To create a framework to identify how to evaluate e-commerce sites in relation to specific areas.

## 5 METHODOLOGY

The research involved three e-commerce case studies. It compared the usability findings indicated by GA software to the usability problems identified by user testing and heuristic evaluation methods.

In order to use GA software to track the usage of the e-commerce sites it was necessary to install the required script on the companies' web sites. The usage of the websites was then monitored for three months. In order to employ the user testing method, a task scenario was developed for each of the three

websites. Twenty users were recruited. In addition, a set of comprehensive heuristics, specific to e-commerce websites, was devised based on a thorough review of the HCI literature. A total of five web experts evaluated the sites using the heuristic guidelines.

The data were analysed to determine which methods identified each usability problem area. The analysis was undertaken in three stages. The first stage involved analysing each usability method for each case and identifying the usability problems obtained from each method within each case. The web usage of the three sites, tracked using GA, was measured using a trial matrix of 20 advanced web metrics (see Table 1). The second stage involved performing a comparison of each usability evaluation method across the three cases. The third stage was undertaken in order to generate a list of standardised usability problem themes and sub-themes to facilitate comparison among the various methods. Ten problem themes and 44 problem sub-themes were identified from an analysis of the methods (see Appendix).

Table 1: Trial matrix of web metrics.

| Metrics Category | Metrics |
|---|---|
| General usability metrics | Average time on site, average page views per visit, percentage of time spent visits, percentage of click depth visits, bounce rate. |
| Internal search metrics | Average searches per visit, percentage of visits using search, search results to site exits ratio. |
| Top landing pages metrics | Bounce rate, entrance sources, entrance keywords. |
| Finding customer support information metrics | Information find conversion rate, feedback form conversion rate. |
| Purchasing process metrics | Order conversion rate, cart start rate, cart completion rate, checkout start rate, checkout completion rate, ratio of checkout starts to cart starts, funnel report. |

# 6  RESULTS

This section reviews the usability problems identified by the three usability methods employed in this research.

## 6.1  Google Analytics Method

The results obtained from the trial matrix of web metrics (Table 1) were investigated. The intention was to determine the most appropriate web metrics that could then be used to investigate usability problems in an e-commerce site.

Specific metrics were devised to identify potential usability problems in six areas: navigation, internal search, architecture, content/design, customer service and the purchasing process. Table 2 shows the suggested matrix and the combination of web metrics that could be used in each area.

An example of the use of combined metrics to identify a specific usability problem is as follows: If a site has low values for *average number of page views per visits* and *percentage of high or medium click depth visits metrics* together with high values for *bounce rate*, *average searches per visits* and *percentage of visits using search metrics*, then this indicates a navigational problem in the site.

The results, however, indicated the limitations of employing the metrics in the evaluation of the usability of e-commerce websites. These related to the fact that the web metrics could not provide in-depth detail about specific problems that might be present on a page.

## 6.2  User Testing and Heuristic Evaluation Methods

The results showed that the user testing and heuristic evaluation methods, unlike the GA method, identified specific usability problems on specific areas and pages on the websites. The usability problems identified by the user testing and heuristic evaluation methods were classified by their severity: major and minor. Major problems included those where a user made an error and was unable to recover and complete the task within the time limit which was assigned for each task. Minor problems included those where a user made a mistake but was able to recover and complete the task in the allotted time. Heuristic evaluators were asked to give their opinion as to whether an issue was major or minor.

Table 2: Web metrics indicating the overall usability of a site.

| Usability Problem Area | Web Metrics |
|---|---|
| Navigation | Bounce rate, average number of page views per visit, average searches per visit, percentage of visits using search, percentage of click depth visits. |
| Internal Search | Average searches per visit, percentage of visits using search, number of page views per visit, percentage of click depth visits, search results to site exits ratio. |
| Architecture | Percentage of time spent on visits, average searches per visit, percentage of visits using search, percentage of click depth visits, average number of page views per visit. |
| Content/Design | Percentage of click depth visits, percentage of time spent visits, bounce rate, top landing pages metrics: bounce rate, entrance searches and entrance keywords. |
| Purchasing Process | Order conversion rate, percentage of time spent visits, cart completion rate, checkout completion rate, cart start rate, checkout start rate and the funnel report. |
| Customer Service | Information find conversion rate. |

The Appendix summarises, with regard to the ten problem themes and 44 sub-themes that were generated by the analysis of the methods, the effectiveness of the user testing and heuristic evaluation methods in identifying each problem sub-theme based on the number of problems identified by these methods and their severity level. The Appendix shows the method(s) that could identify each problem sub-area, that might fail to identify some problems in the area, or that could not identify these problems.

The results showed that most of the problems that were uniquely identified by user testing were major ones which prevented real users from interacting with and purchasing products from e-commerce sites. Conversely, most of the problems that were uniquely identified by the heuristic

evaluators were minor; these could be used to improve different aspects of an e-commerce site.

# 7 AN EVALUATION FRAMEWORK

The results suggested a framework that could be used to evaluate the usability of e-commerce sites, see Figure 1.

The importance of this framework relates mainly to two issues: the reduction of the cost of employing user testing and heuristic evaluation methods, and the identification of the specific types of problem that could be identified by these two methods.
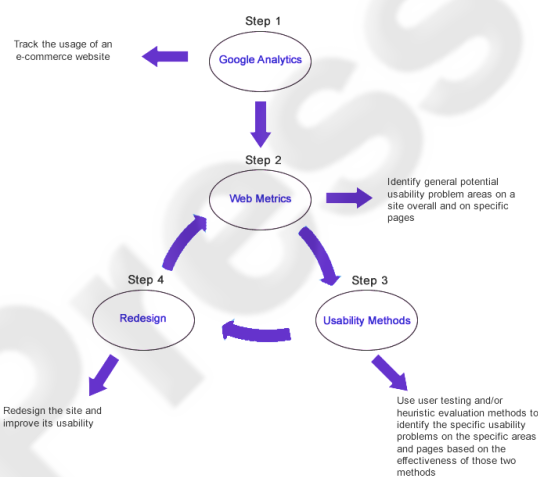


Figure 1: A framework to evaluate the usability of an e-commerce website.

## 7.1 Reduction of Cost

The cost of employing the three methods was estimated in terms of the time spent designing and analysing each of these methods. The approximate time taken to design and analyse the heuristic evaluation, user testing and GA methods was 247 hours, 326 hours and 360 hours, respectively. The approximate time taken to set up and design the GA method included 232 hours that were spent identifying the key metrics that indicated areas with usability problems, and 120 hours calculating the web metrics, and interpreting the metrics' values.

Despite the fact that the GA method required the highest total time in comparison to the user testing and heuristic evaluation methods, this method cost less in comparison to the other methods. This is because it did not require the involvement of users or experts, or the design of specific users' tasks,

questionnaires or guidelines as was the case with the user testing and heuristic evaluation methods. Furthermore, the long time that was spent on the analysis of this method was related to the fact that a specific matrix of web metrics that might indicate areas of usability problems had to be first created. However, if the time for this is ignored (because the matrix would not need to created again), then the time taken by the GA method was considerably less (120 hours).

## 7.2    Specific Types of Problem

The suggested framework describes the specific types of usability problem that could be identified by the user testing and heuristic evaluation methods.

The suggested framework is shown in Figure 1 and involves the following steps:

**Step 1:** This is a preparatory step in order to use GA software to track the traffic flows of a website. It includes inserting GA code in the pages to be tracked and configuring GA software. After this, GA can be used to start tracking users' interactions with the site for a specific time.

**Step 2:** This step involves the use of the suggested matrix of web metrics (summarised in Table 2) to measure the site's usage in order to obtain a clear picture of the general usability problems on the site overall and on specific important pages.

When using the matrix of metrics, the idea is that the evaluator identifies metrics with values that may indicate problems (i.e. a high value for bounce rate). Then, by noting which metrics are problematic, Table 2 can be used to identify the likely problem area, for example, navigational, search-related, etc.

**Step 3:** This step involves employing user testing and/or the heuristic evaluation method in order to identify specific usability problems in particular areas and pages (resulting from Step 2). The decision regarding which method(s) to employ (i.e. user testing, heuristic evaluation or these two methods together) is based on understanding the effectiveness of these methods in identifying specific minor and major usability problem areas, as illustrated in the Appendix. The Appendix helps companies choose appropriate methods and tasks for the evaluators. For instance, if Step 2 suggests a navigational problem, then the evaluator should make a judgment on whether this may be related to misleading or broken links; if it is related to misleading links then the Appendix indicates that this should be investigated by user testing but if it relates to broken links then the Appendix indicates

that this should be investigated by heuristic evaluation.

**Step 4:** This step involves redesigning the site and improving the usability problems identified by Step 3. Then, the usage of the site is tracked, moving to Step 2 in order to investigate improvements in the financial performance of the site and/or to identify new usability problems.

## 8    CONCLUSIONS

This research developed a framework to evaluate the usability of e-commerce websites which involved user testing and heuristic evaluation methods together with GA software.

The framework utilised the advantage of GA software using the specific web metrics that were suggested in this research. This is related to reducing the cost of employing the user testing and/or heuristic evaluation methods by highlighting the areas on an e-commerce site that appear to have usability problems. Then, and because of the limitations of these web metrics, the framework complements the limitations by suggesting the use of user testing and/or heuristic evaluation to provide details regarding the specific usability problem areas on a site. The decision regarding whether to use user testing and/or heuristic evaluation to identify specific problems on the site depends on understanding the advantages and disadvantages of these methods in terms of their ability to identify specific minor and major problems related to the 44 specific usability problems areas identified in this research. Therefore, the suggested framework enables specific usability problems to be identified quickly and cheaply by fully understanding the advantages and disadvantages of the three usability evaluation methods.

The framework offers a base for future research. The next step will be to evaluate the applicability and usefulness of the framework on further e-commerce companies.

## REFERENCES

Barnard, L., Wesson, J. (2004). A Trust Model for E-commerce in South Africa. *SAICSIT 2004*, pp. 23-32.

Chen, S.Y., Macredie, R.D. (2005). An Assessment of Usability of Electronic Shopping: a Heuristic Evaluation. J. *International Journal of Information Management*. 25, 516-532.

Fang W. (2007). Using Google Analytics for Improving Library Website Content and Design: A Case Study. *J. Library Philosophy and Practice*. 1-17.

Freeman, M.B., Hyland, P. (2003). Australian Online Supermarket Usability. *Technical Report*, Decision Systems Lab, University of Wollongong.

Inan, H. (2006). *Measuring the success of your website: a customer-centric approach to website managemen*t (Software Book). Hurol Inan.

McFadden, C. (2005). *Optimizing the online business channel with web analytics*. <http://www.webanalyticsassociation.org/en/art/?9 >.

Nielsen, J. (2003). *Usability 101: Introduction to usability*. Useit.com. <useit.com/alertbox/20030825.html>

Nielsen, J., Norman D. (2000). *Web-Site Usability: Usability on The Web Isn't A Luxury. Information Week*, http://www.informationweek.com/773/web.htm

Phippen, A., Sheppard, L. & Furnell, S. (2004). A practical evaluation of web analytics. *Internet Research*, 14(4), 284-293.

Prom C. (2007). Understanding On-line Archival Use through Web Analytics. *ICA-SUV Seminar*, Dundee, Scotland. <http://www.library.uiuc.edu/archives/ workpap/PromSUV2007.pdf>

Tilson, R., Dong, J., Martin, S., Kieke E. (1998). Factors and Principles Affecting the Usability of Four E-commerce Sites. 4th Conference on Human Factors and the Web (CHFW), AT&TLabs, USA.

# APPENDIX

| Usability Problem Area | Usability Problem Sub-Area | User Testing | | Heuristic Evaluation | |
|---|---|---|---|---|---|
| | | Mn | Mj | Mn | Mj |
| **Navigation Problems** | Misleading links | √ | √√ | √√ | |
| | Links were not obvious | | √√ | √√ | √ |
| | Broken links | √ | | √√ | |
| | Weak navigation support | | √ | | √√ |
| | Orphan pages | √ | | √√ | |
| **Internal Search Problems** | Inaccurate results | | √√ | √√ | √√ |
| | Limited options | √√ | | √√ | |
| | Poor visibility of search position | | | √√ | |
| **Architecture Problems** | Poor structure | | √√ | | √√ |
| | Illogical order of menu items | | | √√ | |
| | Illogical categorisation of menu items | | | √√ | |
| **Content Problems** | Irrelevant content | √ | √ | √√ | √√ |
| | Inaccurate information | √ | | √ | √√ |
| | Grammatical accuracy problems | | | √√ | |
| | Missing information about the company | | | √√ | |
| | Missing information about the products | √ | | √√ | |
| **Design Problems** | Misleading images | √ | | | |
| | Inappropriate page design | √ | √√ | √√ | √ |
| | Unaesthetic design | | | √√ | |
| | Inappropriate quality of images | | | √√ | |
| | Missing alternative texts | | | √√ | |
| | Broken images | | | √√ | |
| | Inappropriate choice of fonts and colours | √ | | √√ | |
| | Inappropriate page titles | | | √√ | |
| **Purchasing Process Problems** | Difficulty in knowing what was required for some fields | √√ | | √ | |
| | Difficulty in distinguishing between required and non-required fields | | √√ | | |
| | Difficulty in knowing what links needed to be clicked | | √√ | | |
| | Long ordering process | √√ | | √√ | |
| | Session problem | | √√ | | √√ |
| | Not easy to log on to the site | | | | √√ |
| | Lack of confirmation if users deleted an item from their shopping cart | | | | √√ |
| | Long registration page | | | | √√ |
| | Compulsory registration | | | | √√ |
| | Illogical required fields | √√ | | √√ | |
| | Expected information not displayed after adding products to cart | √√ | √√ | | |
| **Security and Privacy Problems** | Lack of confidence in security and privacy | | | | √√ |
| **Accessibility and Customer Service Problems** | Not easy to find help/customer support information | | √√ | √√ | |
| | Not supporting more than one language | √√ | | √√ | |
| | Not supporting more than one currency | | | √√ | |
| | Inappropriate information provided within a help section/customer service | √ | | √√ | |
| | Not easy to find and access the site from search engines | | | √√ | |
| **Inconsistency Problems** | Inconsistent page layout or style/colours/terminology/content | √ | | √√ | |
| **Missing capabilities** | Missing functions/information | √ | | √√ | |

**Mn: Minor problems**
**Mj: Major problems**
**√√: Good identification of the specific problem area**
**√:  Missed identification of some of the specific problem areas**
**Blank**: **Could not identify the specific problem area**