Presented at the DLSU Research Congress 2017
De La Salle University, Manila, Philippines
June 20-22, 2017

Web Usability: A Literature Review

Giselle Joy Esmeria[1] and Rosemary R. Seva[2]
[1]*Industrial Engineering Department, De La Salle University Science and Technology Complex*
[2]*Industrial Engineering Department, De La Salle University Manila*
giselle.esmeria@dlsu.edu.ph

Abstract: The continuing growth in the use of internet has transformed the world into a virtual marketplace. The dependability of several web applications has become more important to the users. The fast growth of online applications available in the internet drives the user interface designers to become more competitively creative. Websites are expected to provide the users a satisfying experience at the least. There have been numerous research studies conducted focusing on the design of usability heuristics to evaluate the websites. However, few studies apply a methodology that would actually provide a deeper meaning to the results of usability evaluation. Usability evaluation can be meaningless without looking at the relationships between usability metrics and deriving its usability index. The results of the literature review show that there is no standard usability index that is derived to evaluate the usability of websites. This usability index is an important measure to determine the impact of the usability evaluation results to the management. This paper therefore aims to highlight the future research direction in the area of web usability measurements that would provide a meaningful standard score for all the usability metrics.

Key Words: Web usability, usability evaluation, usability index

1. INTRODUCTION

Human Computer Interaction (HCI) is continuously evolving with the fast change in technology. The use of internet becomes a way of life for most individuals. Internet is used for various purposes such as a means of communication thru emails and social networking, getting online education and doing research activities, online marketing and making financial transactions among other activities. It continues to expand and develop various websites where most users are expected to benefit from. The more websites created, the more options to choose. The users now have the freedom to influence the existence of the websites. For a website to be successful, the level of usability must be very high. Usability is a term that is derived from the term 'user friendly', but ISO 9241-11 provides a standard definition of the term. Usability is referred to as the extent to which product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction (Teoh, Ong, Lim, Liong, & Yap, 2009). Applying this definition to web usability, this simply means that the design of the websites must be easy to use and easy to understand. In designing the websites, there are three important criteria to be considered such as content visibility, ease of content access and the ease of content browsing (Mvungi & Tossy, 2015). The first criterion is necessary to catch the attention of the users.

The rationale in doing website usability is profitability (Rinder, 2012). The issues on website usability can be considered as an important management task on which success in business depends. The inadequate usability of a business website may adversely affect sales which results to lost business opportunities.

This paper presents a literature review on the different methodologies used for web usability analysis on various domains.

2. BACKGROUND

2.1 Usability Evaluation of Websites

Finding flights or booking for a hotel has never been easier. With the increasing popularity of travel websites, interface designers are compelled to provide a seamless experience for the users. Chiou et al. (Chiou, Lin, & Perng, 2011) introduced a five-stage strategic framework for website evaluation. This framework was further demonstrated in the two leading travel websites in Taiwan. Quality service in terms of convenient shopping is found to be relevant and providing more convenient payment methods is necessary to achieve this objective. Haidurova (2013) stressed that common problems in the design include bad presentation of search results, lack of clarity and difficulty in sorting and filtering and confusing display of price. These usability problems greatly influence the effective use of any travel websites.

Most electronic commerce (e-Commerce) applications are now Internet-based. People can communicate with web-based online stores like Amazon, e-Bay, Lazada, OLX and many other similar web-based businesses. The transition of doing business from the marketplace to marketspace provides bigger opportunities for interface designers. Interface design is becoming even more critical in the internet. Poor interface design may lead users to shift from one site to another and eventually lose business. Website quality has a tremendous effect on customer satisfaction and purchase intentions (Bai, Law, & Wen, 2008). One critical issue raised by Diaz et al. (2017) in the design of e-Commerce websites is the failure of interface designers to consider Hoftede's cultural dimensions. The cultural-oriented interface quality is an important element to attract global website users and local users.

The use of e-learning technology is highly recognized nowadays. The effectiveness of the system relies heavily on the design of the interface. Students may lose interest in using the technology when they have difficulty in understanding the interface. Retaining the interests of the students to use the e-learning websites provides too much of a challenge for interface designers as well. For instance, the study of Thowfeek and Salam (2014) revealed that there is a greater expectation from students with regards to the usability characteristics of e-learning websites. Students are attracted to use e-learning websites when it can provide better interactions. Penha et al. (Penha, Correia, Campos, & Barros, 2014) also conducted a study on similar website and the results show that violations on the basic principles of design and usability greatly affect the effective use of the e-learning system. The empirical-based study conducted by Harrati et al. (2016) show that positive user-experience and better usability of e-learning websites are important to the students who are considered the primary users.

2.2 Measures of Web Usability

Usability evaluation is significant in the area of user interface design. Results of usability evaluation are primarily used to aid user interface designers address the changing demands of users.

2.2.1 Heuristic Evaluation and User Testing

Tan et al. (Tan, Liu, & Bishu, 2009) recognized the two most popular usability evaluation techniques. These two usability evaluation techniques are known to be heuristic analysis and user testing. Both methods were compared for efficiency and effectiveness in evaluating four commercial websites. The results show that that both user testing and heuristic analysis complement each other. Neither of the two methods can be replaced by the other. The two methods address different usability problems and therefore, suggested to be used in different stages of user interface design process. Heuristic evaluation can be done in an earlier stage of the design process while user testing can be performed at a later stage of the design process. Heuristic evaluation when compared with other approaches is much easier to implement and less costly. Usability problems identified from the heuristic evaluation are associated with the usability heuristics. Quinones and Rusu (2017) reiterated the three scales to qualify as usability problems namely, severity, frequency and criticality. Severity is measured in terms of the ability of the interface to function, while frequency is measured based on the occurrence of the usability problem. Severity and frequency are then combined to get the criticality index.

Some of the issues that may arise in the use of heuristic evaluation is in the process of selecting the experts and the appropriate number of evaluators. Expertise of evaluators is of great importance in the evaluation of websites (Karoulis & Pombortsis, 2004). The expertise of evaluators has an impact in identifying the right number of evaluators required in a usability evaluation. From the experiment of Nielsen and Molich (1990), the performance of evaluators in determining usability problems may vary depending on the expertise and experience of the evaluators. Nielsen (1992) conducted another research to

determine the effect of expertise of usability evaluators. Evaluators are categorized into novice, regular specialists and double specialists. Novice evaluators are beginners in usability evaluation but are not necessarily new in the use of computers. Novice evaluators can only pinpoint about 50% of the total usability problems. The regular specialists are evaluators with experience in user interface design and evaluation but no special expertise in voice response systems. The regular specialists perform significantly better by identifying around 74% to 87% of usability problems. The double specialists are evaluators with expertise in both user interface and voice response systems. The third evaluators are expected to provide best level of heuristic evaluation and can identify from 81% to 90% of the total usability problems. With these results, the appropriate number of evaluators may be defined as follows:

1. **Simple or Novice evaluators.** About 15 evaluators are needed to find out 75% of the heuristically identifiable problems.
2. **Regular Specialist (HCI experts).** Three to five evaluators under this category can identify 75% of the heuristically identifiable problems. Mostly those belong to the major problems of the interface.
3. **Double Specialist (Double experts).** Only about two to three are necessary to point out 75% of the heuristically identifiable problems.

Traditional heuristics do not evaluate the specific features of particular applications which lead to the design of several usability heuristics to address specific domains.

A wide range of usability evaluation questionnaires have been developed and proposed by several researchers and it is quite challenging for practitioners to select the most appropriate questionnaires to use in a particular domain. Oztekin et al. (2010) developed a UseLearn checklist focusing on twelve high-level dimensions of usability which includes questions addressing error prevention, visibility, flexibility, course management, interactivity feedback and help, accessibility, consistency and functionality, assessment strategy, memorability, completeness, aesthetics, and reducing redundancy. Usability problems identified are analyzed using the criticality metric analysis by revealing the most significant usability problems to improve usability index. For the last few decades, there are a total of 24 standardized questionnaires used on HCI evaluation (Assila, Oliveira, & Ezzedine, 2016) . Out of this total, only five questionnaires are found to be dedicated to web applications namely WAMMI (Website Analysis Measurement Inventory), WEBUSE (Website Usability Evaluation Tool), WEQ (Website Evaluation Questionnaire), SUPR-Q (Standardized Universal Percentile rank questionnaire), and DEEP (Design-oriented Evaluation of Perceived Usability). Design of questionnaires for usability evaluation is a critical element to derive correctly the usability problems present in a user interface design. The five questionnaires proposed may be used as a baseline to determine website usability issues. Hermawati and Lawson (2016) further emphasized the need to develop usability heuristics for specific domain and establish standard measures to validate the results. The lack of validation of the heuristics does not clearly suggest the effectiveness of the heuristics.

2.2.2 Usability Index

Usability index is an important measure to determine the relationships of different usability metrics for the evaluation of websites. Results of heuristics evaluation can be meaningless without looking at the relationships of different usability metrics and deriving its usability index. Adapting a methodology to determine the usability index is necessary to aid user interface designers improve web quality. A Six Sigma quality approach was applied by Nielsen (2003) to compare the quality levels in terms of success rates of the public websites and the intranet. Around 139 public websites were tested for usability with 65% success rate. This corresponds to a 1.9 sigma quality level. For users tested on intranets, the success rate has reached 75% which is quite higher than the public websites. This corresponds to 2.2 sigma quality level. Higher success rate for intranet users is attributed to the familiarity of the functions and commands. Employees are more accustomed to the design of the intranets compared to the public websites on which the user interface designers have more freedom in the web design.

Sauro and Kindlund (2005) emphasized the limitations of traditional usability metrics and attempted to increase its meaning by adapting the Six Sigma methods. Common usability metrics were evaluated in terms of a standardized defect rate or quality level. Four usability metrics were converted into standardized forms. This conversion is necessary to determine the variation of each usability metric from the user-defined goal. Each standardized form is considered as the process sigma. Below are the common usability metrics converted into standardized forms and measured using process sigma:

1. **Task Completion.** Task completion is a special type of discrete data and measured in the form of binary data (i.e. success – complete task, failure – did not complete). Failure to complete the task is considered as a defect and the total of participants who attempted to do the task is the total opportunities for a defect to occur.
2. **Error Rates.** Error counts are considered also as discrete data. Each error is considered as a defect and the total opportunities for defect is calculated by multiplying the total number of participants with the number of task and sub-tasks.
3. **Satisfaction Scores.** Satisfaction scores are ordinal data which are normally derived from the Likert scale. The z-score is computed by getting the difference between the sample mean and the desired level of satisfaction divided by the standard deviation. The corresponding process sigma quality level is determined using the standardized normal table.
4. **Task Times.** Task times are continuous data which are normally derived from actual time study. The measurement of process sigma is based on the average actual time the task is completed by the user against the desired specification limit.

With the quality levels being measured from the standardized usability metrics, the analysts can easily identify the metrics which are falling behind the users' goals. Another study by Sauro and Kindlund (2005) attempts to derive a single measure for the construct of usability. The model considers the four usability metrics to measure efficiency, effectiveness and satisfaction. Efficiency is measured based on time and satisfaction as the average satisfaction scores derived from the survey. Effectiveness, on the other hand, is based on the count of errors and task completion time. To combine the metrics into a single usability score, correlation matrix of all four variables are set-up. The results show that task completion and satisfaction have moderate to significant correlation. In addition, there is a positive correlation between the subjective measures (i.e. satisfaction) and objective measures (i.e., task time, count of errors and task completion).

The Principal Components Analysis or PCA is used to create a single, standardized and summated usability metric. This is done by getting the average of the four standardized values based on the equal weights of the coefficients derived from the PCA.

3. METHODOLOGY

The literature review was carried out using the structured approach similar to the study of Chiou et. al (2010) which consists of four steps as follows:

1. Search a particular keyword in leading journal databases.

   The database used in the study included ScienceDirect and ACM Digital Library open access. The candidate articles were identified by searching titles with the phrases "website usability evaluation" and "measures of web usability".

2. Select the articles from important journals in ScienceDirect and ACM

   Articles published for the last 10 years were selected and reviewed (2007-2017). The studies that met the following conditions were included:

   - Research papers
   - Studies that contain web usability evaluation in a specific domain
   - Studies that present measures of web usability
   - Studies published between January 2007 to May 2017.
   - Papers written in English

   The following types of papers were excluded:

   - Studies that do not have clear approach or methodology for data analysis and interpretation
   - Theses that have not been published
   - Articles not focused on the application of usability evaluation and measures of web usability
   - Studies related to software usability and software design

3. Scan these articles by reading their titles and abstracts to select those relevant to website usability evaluation and measures of web usability

   Qualified articles were retained in the article list but irrelevant articles were eliminated. Using the defined inclusion criteria, a total of 42 studies (based on 166 articles) were

identified as relevant to the current view. About 124 studies were excluded using the exclusion criteria because the articles are beyond the scope of this review.

4. RESULTS AND DISCUSSION

Web usability evaluation has led to a wide number of studies in recent years. Figure 1 shows that the most number of studies related to web usability evaluation are published between 2012 to 2014.

Figure1. No. of Studies by Year

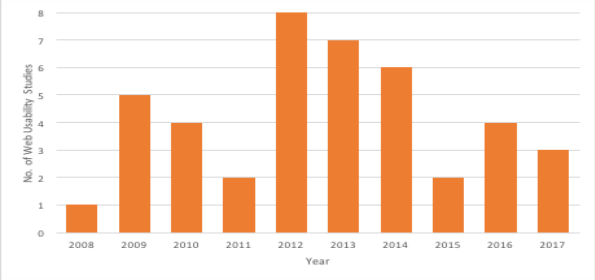
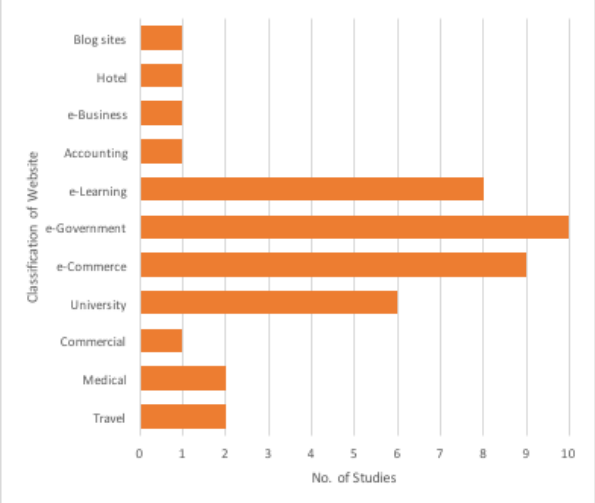


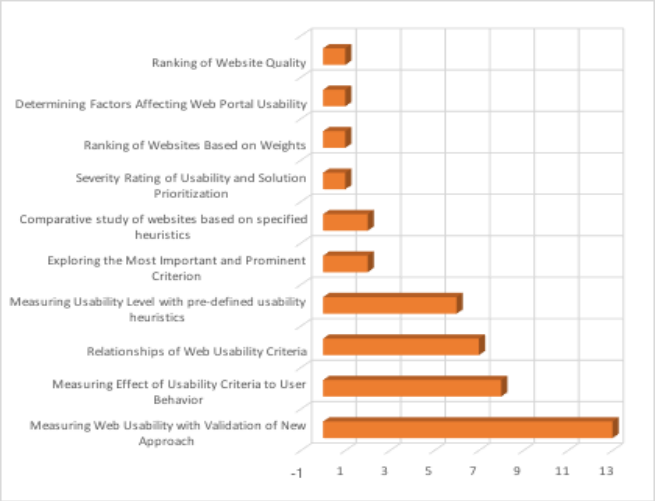
Figure 2, on the other hand, shows that several researchers show interest on evaluating e-Government websites followed by e-Commerce, e-Learning and University Websites.

Figure 2. No. of Studies by Website Category



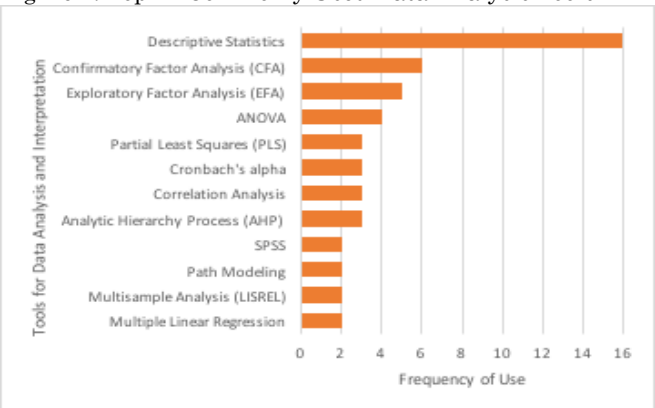
For the past ten years, most researchers focus their study on measuring web usability using newly-developed approach or framework of usability evaluation as shown in Figure 3.

Figure 3. Web Usability Study by Focus



There are a total of 46 data analysis tools identified from the literature review. These tools provide measures of web usability for different category of websites. The Top 12 commonly used data analysis is presented in Figure 4. These tools are good representation of the results in terms of the relevance of each web usability criteria.

Figure 4. Top 12 Commonly Used Data Analysis Tools



5. CONCLUSIONS

Websites are expected to provide the users a satisfying experience at the least. Different websites may require different usability characteristics. Therefore, interface designers are required to really be supplied with the correct information on the needs of the users. Defining the perfect usability heuristics to evaluate certain websites is far from being a crowded area for research. There is that constant need to design better heuristic that can actually translate the

way people see things.

Another area that researchers may look into is the identification and testing of the most appropriate usability measurements. Approaches such as AHP, multiple regression, and QFD are only applicable with the assumptions of independence and linearity in different usability dimensions. This cannot be generalized and must be proven true in all aspects. Some usability dimensions are notably dependent with one another. In addition, derivation of a single usability index that represents all areas of web usability is another field that needs to be explored. The use of six-sigma process and getting the weighted average may be used as baseline to derive a better measure of usability index.

5. REFERENCES

Akincilar, A., & Dagdeviren, M. (2014). A Hybrid Multi-Criteria Decision Making Model to Evaluate Hotel Websites. *International Journal of Hospitality Management, 36*, 263-271.

Allen, M., Currie, L. M., Bakkeh, S., Patel, V. L., & Cimino, J. J. (2006). Heuristic Evaluation of Paper-Based Web Pages: A Simplified Inspection Usability Methodology. *Journal of Biomedical Informatics, 39*, 412-423.

Assila, A., Oliveira, K. d., & Ezzedine, H. (2016). Standardized Usability Questionnaires: Features and Quality Focus. *Journal of Computer Science and Information Technology, 6*(1), 15-31.

Bai, B., Law, R., & Wen, I. (2008). The Impact of Website Quality on Customer Satisfaction and Purchase Intentions: Evidence from Chinese Online Visitors. *International Journal of Hospitality Management, 27*, 391-402.

Chiou, W.-C., Lin, C.-C., & Perng, C. (2011). A Startegic Website Evaluation of Online Travel Agencies. *Tourism Management, 32*, 1463-1473.

Delice, E. K., & Gungor, Z. (2009). The Usability Analysis with Heuristic Evaluation and Analytic Hierarchy Process. *International Journal of industrial Ergonomics, 39*, 934-939.

Diaz, J., Rusu, C., & Collazos, C. A. (2017). Experimental Validation of a Set of Cultural-Oriented Heuristics: e-Commerce Websites Evaluation. *Computer Standards & Interfaces, 50*, 160-178.

Haidurova, J. (2013). *Usability Evaluation of Ireland's Travel Websites*. Dublin Institute of Technology.

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2016). Exploring User Satisfaction for e-Learning Systems via Usage-Based Metrics and System Usability Scale Analysis. *Computers in Human Behavior, 61*, 463-471.

Hermawati, S., & Lawson, G. (2016). Establishing Usability Heuristics for Heuristic Evaluation in a Specific Domain: Is there a Consensus? *Applied Ergonomics, 56*, 34-51.

Hertzum, M. (2006). Problem prioritization in Usability Evaluation: From Severity Assessments Toward Impact on Design. *International Journal of Human-Computer Interaction, 21*(2), 125-146.

Leba, M., Ionica, A. C., & Edelhauser, E. (2013). QFD - Method for e-Learning Systems evaluation. *Procedia Social and Behavioral Sciences, 83*, 357-361.

Mvungi, J., & Tossy, T. (2015, July). Usability Evaluation Methods and Principles for the Web . *International Journal of Computer Science and Software Engineering, 4*(7), 165-171.

Nielsen, J. (2010, February 22). *Progress in Usability: Fast or Slow?* Retrieved April 7, 2017, from https://www.nngroup.com/articles/progress-in-usability-fast-or-slow/.

Oztekin, A., Delen, D., Turkyilmaz, A., & Zaim, S. (2013). A Machine Learning-Based Usability Evaluation Method for e-Learning Systems. *Decision Supports System, 56*, 63-73.

Oztekin, A., Kong, Z. J., & Uysal, O. (2010). UseLearn: A Novel Checklist and Usability Evaluation Method for E-Learning Systems by Criticality Analysis. *International Journal of Industrial Ergonomics, 40*, 455-469.

Penha, M., Correia, W. F., Campos, F., & Barros, M. (2014). Heuristic Evaluation of Usability - a Case Study with the Learning Management Systems (LMS) of IFPE. *International Journal of Humanities and Social Science, 4*(6), 295-303.

Quinones, D., & Rusu, C. (2017). How to Develop Usability Heuristics: A Systematic Review. *Computer Standards and Interfaces, 53*, 89-122.

Rinder, J. (2012). *The Importance of Website Usability Testing*. University of Oregon, Applied Information Management.

Sauro, J., & Kindlund, E. (2005). A Method to Standardize Usability Metrics into a Single Score. *Usability Professionals Association*. UPA.

Sauro, J., & Kindlund, E. (2005). Making Sense of Usability Metrics: Usability and Six Sigma. *Usability Professionals Association (UPA) Conference 2005* . UPA.

Syed, F. (2009, September 9). *Quality Function Deployment*. Retrieved April 5, 2017, from https://totalqualitymanagement.wordpress.com/2009/09/09/quality-function-deployment/.

Teoh, K., Ong, T., Lim, P., Liong, R. P., & Yap, C. (2009). Explorations on Web Usability. *American Journal of Applied Sciences, 6*(3), 424-429.

Thowfeek, M. H., & Salam, M. N. (2014). Students' Assessment on the Usability of e-Learning Websites. *Procedia - Social and Behavioral Sciences , 141*, 916-922.

Zhang, J., Johnson, T. r., Patel, V. L., Paige, D. L., & Kubose, T. (2003). Using Usability Heuristics to Evaluate Patient Safety of Medical Devices. *Journal of Biomedical Informatics*, 23-30.