Proposing a Methodology to Evaluate Usability of E-Commerce Websites: QUEM Model

Nasrin Dehbozorgi
School of e-Learning,
Shiraz University

Shahram Jafari
School of Electrical and Computer Engineering
Shiraz University

ABSTRACT

In this work we propose a methodology named QUEM (Quantitative Usability Evaluation Model) for quantitative usability evaluation of e-commerce websites. Many e-commerce websites lack user friendliness. The main factor that prevents these firms to conduct usability testing is the high costs and the need for usability experts. In this work, ISO/IEC 9126-1:2001 quality model was selected as a basis for defining usability characteristics of our model. Based on these standard usability characteristics a set of usability factors is proposed specifically for evaluation of e-commerce websites. The usability factors of QUEM model have been extracted from a wide range of usability guidelines and checklists. Since all of the usability factors do not have the same significance in the overall usability assessment of the web sites, the proposed factors have been weighted by adopting AHP (Analytical Hierarchical Processing) approach. In order to demonstrate the efficiency of QUEM model, usability of two e-commerce systems has been assessed by adopting this model. QUEM model enables the e-commerce firms to measure the usability of their sites in a minimum cost and span of time.

Keywords

Usability, Quality Model, e-commerce website, QUEM model, Analytical Hierarchy Process (AHP)

1.INTRODUCTION

With the vast development of information technology and global advances in this field, e-commerce plays an important role in operations of companies. In recent years much concern has been attracted towards increasing the usability of e-commerce websites and developing usability evaluation methods. Since e-commerce has become the main distributing channel between companies and the customers in B2C transactions, focusing on the quality of this channel and the service that it renders to customers, is very critical for the sustain of the business. Usability is one of the most important factors in system quality. There exists two approaches to evaluate the usability of any e-commerce system; the first approach is to evaluate the system based on user's point of view and the second one is based on the developer or experts' point of view. Therefore, the quality evaluation of an e-commerce system is mainly based on two approaches [1]: one is evaluating the quality of final developed system and the second approach is qualification of the process of system development. In this research, the first approach has been adopted to assess the usability of e-commerce websites. In order to evaluate the quality of developed systems, a set of quality characteristics and criteria are required as a basis to describe the system quality. This set of characteristics and the relationship between them is called the Quality Model [2].

In our research, after extensive research that has been conducted to analyze and compare different software quality models, the ISO/IEC 9126-1:2001 quality model has been adopted as a basis to focus on usability characteristics of e-commerce websites. For customizing these characteristics specifically for e-commerce websites, a wide range of usability guidelines and checklists were studied, which were mainly centered on works done by Nielsen, IBM and W3G guidelines. Hence, in our model -which is named QUEM model- we propose a generic set of sub-characteristics for evaluation of e-commerce websites.

In the following sections we will have an overview on existing quality evaluation methods and approaches. Then the research methodology will be discussed. Finally in the experimental result section, the presented QUEM model is applied on two e-commerce systems and their usability is assessed and quantified into numerical values and compared with each other.

2. REVIEW OF THE LITERATURE

Although from 1976 much research has been done by the experts to define a general quality model and an approach to evaluate the usability of websites, a general and inclusive accepted methodology for evaluating the quality of websites has not been presented yet. Some of the current approaches just define an overall measurement of the website's design, while they are not specifically developed for evaluation of e-commerce websites. There have been some researches that take into account the features of e-commerce websites, but they have a limited scope and do not include all usability features. Another shortcoming of presented evaluation methods is that they are just useful to assess the overall quality of websites without being able to evaluate special features of them. Some of the existing usability evaluation methods still require the usability experts to assess the usability of the site, while others consider the same importance for all usability features of the website. In the table 1, some of the most well-known methods, approaches and developed set of quality metrics are overviewed and compared against certain shortcomings which were discussed above.

Table 1. Brief Overview of Quality Evaluation Methodologies

Methodologies & Researches conducted by	Shortcomings					
	1*	2*	3*	4*	5*	6*
Chang, et al. [3]	✓			✓		
Garzotto, et al. [4]		✓				✓
Keevil [5]	✓		✓		✓	
Rutledge [6]	✓				✓	
Renaud, et al. [7]		✓				
Calongne [8]	✓					
Ivory, et al. [9]	✓					
Mich, et al. [10]	✓					
Opaluch [11]	✓					
Lightner [12]					✓	
Sauro , et al. [13]	✓				✓	

1*: Are generic and not for e-commerce websites
2*: Have limited scope
3*: Unable to assess special features of website
4*: Not measuring usability into numerical value
5*: Consider equal weight to all factors
6*: Require involvement of usability expert

The listed shortcomings in table 1 prevent these quality evaluation methods and guidelines to render effective results. Therefore, the aim by presenting QUEM model in our research is to fix the shortcomings of previous models. QUEM model with its hierarchal structure helps the e-commerce firms and developers to conduct usability testing without the need for experts or usability skills.

3. RESEARCH METHODOLOGY

As discussed in section 1, in order to identify a set of usability factors we have adopted the ISO/IEC 9126-1:2001 standard model as a basis for hierarchal structure of QUEM model. One of the main features of ISO/IEC 9126-1:2001 model is that it has been defined based on the international consensus and agreement of all country members of ISO organization. Another distinct feature of this model is that it's hierarchical structure and framework supports both top-down and bottom-up approaches in evaluating the quality of systems. As mentioned in [14], the most significant characteristics of ISO/IEC 9126-1:2001 quality model are:

- Having evaluation criteria
- Unambiguous terms and expressions
- Accurate definitions of characteristics
- One-to-many relationship between various layers of the model
- Hierarchical structure

In ISO/IEC 9126-1:2001 quality model, the usability attribute consists of three characteristics: Understandability, Learnability and Operability [14].

After validating ISO/IEC 9126-1:2001 quality model as the best choice for the basis of our methodology, we offer to customize it to be used specifically for evaluation of B2C e-commerce websites. In the next section our customized criteria are presented.

3.1 Identification of Usability Factors

In QUEM model a set of usability factors were indentified based on the extensive literature survey, and have been categorized in a hierarchical structure.
Some researchers claim that quality factors such as navigability, customizability are not included in ISO/IEC 9126-1:2001 quality model [14], while our model covers these criticisms by proposing these quality factors under their relative characteristic in the ISO/IEC 9126-1:2001 model. Based on discussions presented in [15, 16, 17, 18 and 19], the usability factors of QUEM model have been defined as presented in Table 2.

Table2. Usability Sub-factors of QUEM model

Alias	Understandability Factors
U1	Consistency and standardization of text and content through the website
U2	Effectiveness of text and information
U3	Consistency and standardization of graphical design, icons and site structure
U4	Aesthetics and minimalistic design and structure
U5	Considering user's disability and limitation in the design of the pages
U6	Error prevention
U7	Providing error feedback and handling the error message properly
U8	Visibility of system status for the user
Alias	**Learnability Factors**
L1	Effectiveness of help and guidelines
L2	Accessibility of help and guidelines through all pages
L3	Facilities to enhance learning for inexperienced users
Alias	**Operability Factors**
O1	Match between system structure and real world
O2	User control and freedom in the shopping process
O3	Flexibility of the system
O4	Considering cultural issues
O5	Proper navigation facilities
O6	Proper categorization and structure of information
O7	Enjoyment
O8	Keyboard and accessibility
O9	Content
O10	Enhancement of purchase process
O11	Support and interaction with users
O12	Customization y user and operability of design

3.2 Hypothesis Testing

In order to assess the validity of proposed usability factors in QUEM model, the Delphi technique was employed. In this method in order to gather required statistical information a structured survey should be prepared and given to participants, who are mainly experts. The advantage of this model is the avoidance of direct confrontation of the experts. [20]

In validation test phase of our research, a group of 16 experts were requested to join and participate in the survey. They were mainly experts in the fields of ecommerce, software engineering and Human Computer Interaction (HCI). Most of them were among the university lecturers with more than 10 years of experience. Afterwards, a questionnaire was developed which consisted of the usability factors of QUEM model and was sent to the experts. They were asked to give their idea whether they agree on these factors or not. In Delphi method, the survey should be performed in several rounds until all the experts agree on the finalized selected factors [21].
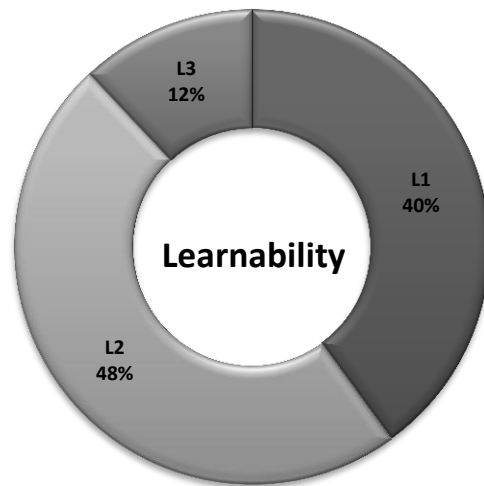
In our work this process was repeated in two rounds. The responses in the second round were very close to those of the first round. During these two rounds, one of the operability factors was omitted due to its lower ranking in both rounds. After proving the validity of the QUEM model, the reliability of the proposed usability factors was established by testing their internal inconsistency by computing the Cronbach's Alpha. The Cronbach's Alpha was computed separately for each usability characteristics. The computed values for Understandability, Learnability and Operability were 0.75, 0.63 and 0.81. The results demonstrate a marginal reliability for Learnability factor, while it has high reliability for the other two usability factors.

3.3 Weighting the Factors of QUEM Model

Usability characteristics are subjective concepts that can be evaluated based on different viewpoints, which makes for different results. Considering the fact that all of the usability factors do not have the same importance in overall assessment of the usability, the need arises to give a reasonable weight to these factors. In this work, the Analytical Hierarchy Process (AHP) has been adopted to assign weigh to the proposed usability factors. AHP is a powerful decision making tool for multiple evaluation criteria and uncertain situations, which makes complicated problems easier to be analyzed.

In AHP method, the relative importance of each factor is evaluated based on the comparison among all factors by the expert. Final outcome of this method is the average weight that is assigned to each usability factor. In order to weigh usability factors of QUEM model, the same group of the 16 experts was requested to participate. A checklist was prepared and presented to them to compare and rank the given usability factors. Based on the pair wise matrix, the eigenvector approach was adopted to calculate the weights of the factors. This was based on the eigenvector which corresponds to the largest Eigen value (λ_{max}).

Finally, by adopting the Saaty's consistency index (CI) [22], consistency ratio (CR) for each questionnaire was verified. Considering the fact that the best consistency ratio in AHP should be lower than 0.1, the questionnaires with higher consistency ratio were eliminated, so that all the results could be applied to AHP standard. The final weights of the usability factors were identified and the mean average of the given weights was considered to be a valid weight for each factor. The normalized resultant weights of the understandability, learnability and operability of the QUEM model are shown in Figure.1, 2 and 3.
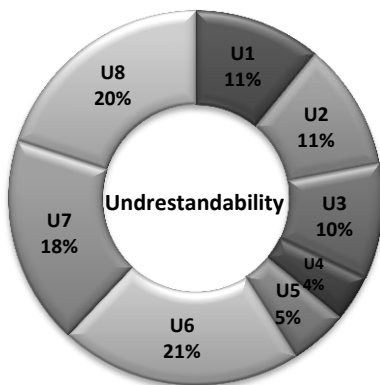

Factors	U1	U2	U3	U4	U5	U6	U7	U8
Weights	0.11	0.11	0.1	0.04	0.05	0.21	0.18	0.2

Figure1: Weights of Undrestandability Factors

Factors	L1	L2	L3
Weights	0.4	0.48	0.12

Figure2: Weights of Learnability Factors
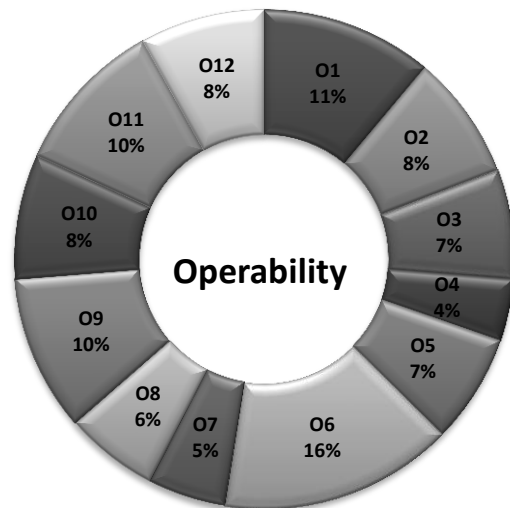
Factors	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	O12
Weights	0.11	0.08	0.07	0.04	0.07	0.15	0.05	0.06	0.1	0.08	0.1	0.08

Figure 3: Weights of Operability Factors

4. EXPERIMENTAL RESULT

In order to demonstrate the application of QUEM model and its validation, this model was applied on two of the most well-known e-commerce websites; Amazon[23] and Ebay[24]. A survey was conducted among a group of university students to rank the features of these two websites based on the checklists that was presented to them.

The participants of the survey were consisted of 47 university students among different fields of study. The

questionnaires were composed of the usability factors of QUEM model, and the students were asked to rank these factors out of 100 score.

After collecting the questionnaires, the results were normalized and the average score for each of the factors for both websites was computed as shown in Table 3.

Table3. Usability values of Amazon and Ebay Portals

Usability Sub-factors	Amazon	Ebay
Consistency and standardization of text and content through the website	80.45	67.64
Effectiveness of text and information	92.27	100
Consistency and standardization of graphical design, icons and site structure	72.90	72.50
Aesthetics and minimalistic design and structure	71.50	86.75
Considering user's disability and limitation in the design of the pages	51.20	45.2
Error prevention	52.90	60.76
Providing error feedback and handling the error message properly	63.28	61.00
Visibility of system status for the user	75.35	69.65
Effectiveness of help and guidelines	69.20	78.85
Accessibility of help and guidelines through all pages	61.83	71.33
Facilities to enhance learning for inexperienced users	44.83	64.00
Match between system structure and real world	65.00	68.00
User control and freedom in the shopping process	57.13	61.38
Flexibility of the system	54.86	50.57
Considering cultural issues	78.25	69.00
Proper navigation facilities	79.86	62.29
Proper categorization and structure of information	74.53	90.13
Enjoyment	72.00	78.00
Keyboard and accessibility	54.17	38.83
Content	85.70	80.60
Enhancement of purchase process	78.00	86.63
Support and interaction with users	76.70	79.10
Customization y user and operability of design	65.25	52.38

The calculated values of Table 3 are the scores given to usability factors of these websites, regardless of the weights of the factors. This does not render a precise usability evaluation result.

Therefore, to assess the usability of these websites by considering the acquired weights of the factors, equation (1) was employed on the collected data, and the finalized score of each usability characteristic for both portals was calculated.

$$\text{Value of each Characteristics} = \frac{\sum_{k=i}^{j}(\text{weight}_k * S_k)}{\sum_{k=i}^{j}(\text{weight}_k)} \quad (1)$$

Weight$_K$ = Weight of Factor k
i, j ∈ {U_1,..,U_8, L_1,..,L_3, O_1,..,O_{12}}
S = Score of usability factors

Therefore according to the aliases in Table 2:
For Understandability, (i, j) = (U_1, U_8)

For Learnability, (i, j) = (L_1, L_3)
For Operability, (i, j) = (O_1, O_{12})

By adopting the QUEM model, the final values of usability characteristics for both websites was computed. In Amazon website the values for Understandability, Learnability and operability were: 69.28, 62.74 and 70.01, while the score of the same factors in Ebay website were: 70.02, 73.46 and 69.89.

Figure 4 shows the comparative graph between the finalized scores of both websites.
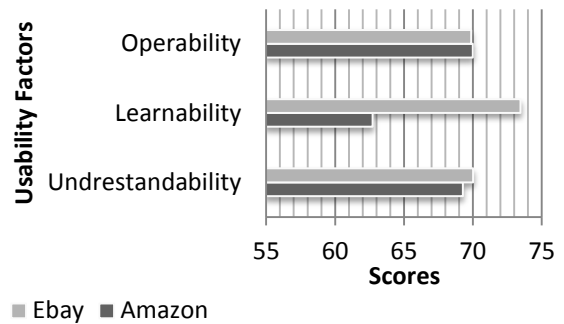
Figure 4: Usability scores of Amazon and Ebay websites

As expected, one can see that generally Amazon and Ebay have the same rankings. The only exception is for the learnability factor, in which Ebay has a higher score.

5. CONCLUSION

In this research we propose a methodology named QUEM, for quantitative usability evaluation of e-commerce websites. By analyzing most well-known usability guidelines and quality models, the ISO/IEC 9126-1:2001 quality model was selected as a basis for our model. In order to propose an inclusive and effective usability evaluation method, a set of usability factors were identified and extracted from a wide range of usability guidelines and checklists. After conducting hypothesis testing to check the validly and reliability of the QUEM model, the weights of the factors have been calculated by adopting the AHP approach. To apply the model to real environment and prove both its efficiency and conclusiveness, two e-commerce portals were selected to be evaluated as case studies.

The QUEM model by covering the shortcomings of previous methodologies and minimizing the required cost and effort enables the e-commerce firms and website developers to conduct a precise usability testing for their website without the need for experts.

6. ACKNOWLEDGEMENT

7. REFERENCES

[1] Pressman R .2000. Software Engineering, a Practitioner's Approach, 5th edn. Mcgraw-Hill

[2] Botella P, Burgues X, Carvallo J.P, Franch X, Quer C .2004. Using Quality Models for Assessing COTS Selection. In: Proceeding of MPEC'04 and ICSE'04.

[3] Chang, E.J., Dillon, T.S., & Cook, D. 1997. An Intelligent System Based Usability Evaluation Metric. Proceedings of the 1997 IASTED International Conference onintelligent Information Systems (IIS '97) Held In December 8-10, 1997, Grand Bahama Island, Bahamas, 218-226.

[4] Garzotto, F., Matera, M., & Paolini, P. 1998. Model-Based Heuristic Evaluation of Hypermedia Usability. Proceedings of the Working Conference on Advanced Visual Interfaces Held in May 24-27, 1998, L'Aquila, Italy, 134-145.

[5] Keevil, B.1998. Measuring the Usability Index of Your Web Site. Proceedings of the ACM 16th Annual International Conference On Computer Documentation Held in September 23-26, 1998, Quebec, Canada, 271-277.

[6] Rutledge, S.T. 1998. Good Sites: A Rubric for Evaluating Student Web Sites. Retrieved April 18, 2005, from Http://lrs.ed.uiuc.edu/students/srutledg/goodsites8.html.

[7] Renaud, K., & Van Dyk, T..2001. Tailoring E-commerce Sites to Ease Recovery After Disruptions. Proceedings of the 25th Annual International Computer Software and Applications Conference (COMPSAC'01) Held in October, 8-12, 2001, Chicago,IL, 603-612

[8] Calongne, C.M. .2001. Designing for Web Site Usability. The Journal of Computing Sciences in Colleges, 16(3), 39-45.

[9] Ivory, M. Y., & Hearst, M.A. 2002. Statistical Profiles of Highly-Rated Web Sites. Proceedings of the ACM CHI 2002 Conference Held in April 20-25, 2002, Minneapolis, MN, 367-374.

[10] Mich, L., Franch, M., & Gaio, L. 2003. Evaluating and Designing Web Site Quality. IEEE Multimedia, 10(1), 34-43

[11] Opaluch, R. 2003. Usability Metrics. In Ratner, J. (Ed.), Human Factors and Web Development (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Publisher

[12] Lightner, N.J. 2004. Evaluating E-Commerce Functionality with a Focus on Customer Service. Communications of the ACM, 47(10), 88-92.

[13] Sauro, J. & Kindlund, E. 2005. A Method to Standardize Usability Metrics Into a Single Score. Proceedings of the CHI 2005 Conference on Human Factors and Computing Systems Held in April 2-7, 2005, Portland, OR, 401-409.

[14] Behkamal B, Kahani M, Kazem Akbari M .2009. Customizing ISO 9126 quality model for evaluation Of B2B applications. Journal of Information software Technology 51- 599-609

[15] Rogers Y, Sharp H, Preece J .2011. Interaction Design: beyond human-computer interaction, 3rd Edition. Wiley

[16] Mit, Ercim, Keio .2008. Web Content Accessibility Guidelines (WCAG) 2.0, W3C Recommendation, www.W3C.org, Accessed 13th June 2011

[17] Nielsen, J., Molich, R., Snyder, C. And Farrell,S. 2000b. E-Commerce user experience: Methodology. Nielsen Norman Group Http://www.nngroup.com/reports/ecommerce/

[18] Rohn, J.A. .1998. Creating Usable E-Commerce Sites. ACM's Standardview. 6(3), 110- 115.

[19] Barnard L. And Wesson J. 2004. A Trust Model for E-commerce in South Africa, Proceedings Of SAICSIT 2004, Pages 23 – 32

[20] Okoli C., Pawlowski S.D. 2004, The Delphi Method as a research tool: an example, design Considerations and applications , Information & Management 42 , 15–29

[21] Grisham T. 2009. The Delphi technique: a Method For testing complex and Multifaceted topics, International Journal of Managing Projects in Business, Vol. 2 No. 1, Pp. 112-130

[22] Saaty T.L, Vargas L.G (1984) The Ligimacy of Rank Reversal. OMEGA, 12(5):513-516

[23] Amazon. Www.Azazon.com. Accessed 24th May 2012

[24] Ebay. Www.Ebay.com. Accessed 24th May 2012