Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make learn

Management,
Knowledge and Learning
International Conference 2014

HEURISTIC EVALUATION AND USER EXPERIENCE ASSESSMENT OF ONLINE SHOPPING PORTALS USING COGNITIVE WALKTHROUGH AND EXPERT METHOD

Malgorzata Plechawska-Wojcik
Lublin University of Technology, Poland
m.plechawska@pollub.pl

Abstract:
The paper presents a case study of GUI quality testing. The case study was performed on appropriately selected online shopping portals. The aim of the study was to assess the user experience and to collate it with expert opinion. The testing procedure was based on the expert method and cognitive walkthrough. It was adjusted to the nature of web portals. The paper outlines contain the method description and results including questionnaire summary and experts opinion. The most important users remarks concerning tested GUI is also presented.

Keywords: GUI usability, user experience, expert analysis, cognitive walkthrough

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make learn

Management,
Knowledge and Learning
International Conference 2014

1. INTRODUCTION

E-commerce is one of the fastest growing markets. Popularity growth of online retail and shopping is enormous. People can buy online almost everything. Every year the online retail market share is more and more extended (Zhou et. al, 2007). Online retail sales in 2012 reached $14,587.1bn (Cushman & Wakefield, 2013) and constant grow is estimated.

One can also find extended research on online consumer behaviours and attitude (Li & Zhang, 2002). Research are conducted on proliferation of online shopping regarding consumer- or a technology-oriented view (Jarvenpaa & Todd, 1997; Brown et al. 2003; Chau et al. 2002; Korgaonkar et al. 2004; Garbarino & Strabilevitz 2004; Huang et al. 2004).

Huge popularity of online shopping cause also growing competition among online retailers. Developers of shopping portals needs to put attention into functionality and ergonomics of the page. Consumers have access to increasingly wide range of online shopping and they became more and more demanding. Of course one need also to regard the huge impact of marketing. However, even the best advertisement is not enough to attract and to retain customers. That is why it is so important to focus on technology-oriented aspects of an online store. Such aspects are user interface features, web site content and design, and system usability (Zhou et. al, 2007).

The main objective of this paper is to present results of user experience test performed on several popular online retail and shopping portals. All tested portals have many functionalities, so well-designed graphical user interface is of great importance for users. Portals chosen for the use case are:
- http://www.amazon.com - retail and shopping portal offering many products
- http://www.ebay.com - popular online auction portal
- http://www.newegg.com - online shopping portal offering computers and electronics
- http://www.overstock.com - online shopping portal offering clothes and home goods
- http://www.walmart.com - online shopping portal offering many products such as electronics, home furnishings, video games
- http://www.bestbuy.com- online shopping portal offering electronics

The role of user experience in today online world is getting more and more important. Functionality and well-designed GUI is a key to success of each internet platform. Online retail and shopping portals usually offers to their clients such functionalities as extended searching and filtering of products, management of basket, managing promotions and special offers, managing wishlists, offering newsletter and loyalty programs, reporting etc.

Testing of GUI quality and user experience is a complex process, which should be planned and performed in iterations. Testing process should be performed on different stages of portal development: in the stages of design, implementation, evaluation (assessment), development of improvement guidelines. Testing may be performed using different methods performed on end-users (typical user tests, cognitive walkthrough) and IT experts (expert method) (Laskowski, 2012).

Testing process assumes the existence of feedback which can be used to support the work of interface designers. There are several different methods used for quality assessment of an application interface, which in general can be divided into two groups (Pan, 1999):
- automatic,
- experimental.

Methods involving end-users are considered to have several major drawbacks, which may result from the following issues (Laskowski 2012; Zabiński & Maczka, 2010; Abrahao et al, 2008):
- the selection of the best evaluation method (e.g. Krug's test, simplified cognitive walkthrough, user performance analysis, etc.) for the particular system;
- the cost and time-consumption of the survey (e.g. (Prenzel & Ringwelski, 2012) reported overall 105 hours of one system usability testing);
- the selection of users for the experimental group, including its size and profiles of the participants.

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make learn

Management,
Knowledge and Learning
International Conference 2014

Selection of research group members is especially important: end-users of online retail and shopping portals differ due to their needs. The research results on GUI quality depend strongly on the background of the examinees. It is preferable to have test subjects that reflect the expected profile of the end-users (Phung, 2007).

It is noteworthy, that although novice users interacting with a system for the first time prefer simple actions and ease of learning (Shneiderman, 1998), their behavior and work quality changes as their experience with the system increases. This is an important point of view when evaluating GUI quality over time, but setting up such a study requires observation over time and lots of resources (Prenzel, Ringwelski, 2012).

An expert analysis is one of the most reliable method GUI quality. As it is shown in the studies (Krug, 2000), in most cases group of 3 to 5 experts using the adequate methodology is able to detect and correct over 85% of errors in software – this applies also to errors in GUI quality. It should be noted that testing can never completely identify all the defects within software (Pan, 1999).

2. APPLIED TESTING METHODS

Graphical User Interface of retail and shopping portal should be checked and tested by both, end-users and experts. To check users' viewing experience GUI quality evaluating methods might be used. During such case study users experience also needs to be regarded.

An expert analysis in combination with cognitive walkthrough seems to be the most reliable method in case of assessment of responsive design web page GUI quality. As it is shown in the studies (Krug, 2000) in most cases group of several experts using the adequate methodology is able to detect and correct over 85% of errors in software – this applies also to errors in GUI quality. It should be noted that testing can never completely identify all the defects within software (Pan, 1999).

2.1. Cognitive walkthrough

Cognitive walkthrough emphasizes the ease of interface learning during the initial contact with the system (Laskowski, 2012). This method is based on defining a few tasks, which typical user should perform while working with an application (e.g. making a purchase) (Wharton et al., 1994). Those can be performed by application end-users. Each task is divided into individual components (steps), which are analyzed according to the following questions:
1.   *Does the user know what to do during the analyzed step?*
2.   *If the action performed by user is correct, is he aware of it?*
3.   *If the action performed by user is correct, does he feel like getting closer to reaching the goal?*

The difficulty of each step is evaluated using Likert scale of 1 to 5, with 1 meaning 'very easy', 5 – 'very difficult'.

In contrast to other GUI quality assessment methods, the main goal of cognitive walkthrough is to study the flow of the processes undertaken by user rather than assessing particular pages or screens of the application interface.

Cognitive walkthrough method has gained its popularity due to the ability of performing relatively quick analysis at low cost. Moreover, it may be used even at the beginning of the design phase. Although it should be noted that according to some research cognitive walkthrough more often allows experts to detect potential problems with the interface than existing ones (Koyani et al., (2004). Therefore, it is also used in conjunction with other general tests as an addendum identifying potential problem areas.

2.2. Expert analysis

Expert analysis is one of the most widely used method for application evaluation. While using the application, expert checks and marks the predefined areas in order to note for the compliance with interface design guidelines (such as Nielsen-Molich heuristics (Nielsen & Molich, 1990) and look for potential problems.

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make
learn

Management,
Knowledge and Learning
International Conference 2014

Each of those predefined areas can be divided into several more detailed sub-areas and be assigned with questions for the expert to answer while working with an application (Laskowski, 2012).

Experts run only the functional testing, as they focus on executing the system functions and examining their input and output data (Howden, 1987), treating whole tested system as a black box where only the inputs, outputs and specification are visible. Therefore, the system functionality is determined by observing the outputs to corresponding inputs and the implementation details are not considered (Pan, 1999).

Contrary to the widespread assumption, experts usually do not acquire better results in performing specific tasks in the tested system, as they usually do not know that system before the testing. But their expert status is based on experience with different kinds of software. This, as proven by studies, allows them to perform faster than the novices (Dillon & Song, 1997) and to spend less time handling the errors (despite making number of errors comparable to the novice users) (Jochen et al., 1991).

3. GUI QUALITY ASSESMENT PROCEDURE

The presented case study describes the analysis of GUI of several shopping portals. The main task of such portal is to enable easy and intuitive shopping regarding fast searching, managing a basket and use all additional features of the portal.

3.1. Cognitive walkthrough

The cognitive walkthrough involves several scenarios containing tasks to perform for users. Users performed all of them for each portal. Those scenarios are:
  – Using a wishlist – user needs to create a wishlist and add two products to it.
  – Subscribe and unsubscribe from a newsletter.
  – Working with the basket – user adds three items to the basket and then clean it using as few steps as possible.
  – Shipping internationally – user needs to find the answer for question about sending the products internationally.
  – Returning policy – user needs to find the answer for question about how many days user have to return a product.
  – Payment methods – user needs to find the answer for question about payment methods which are accepted.
  – Selling products - In case a portal enables to sell user's products, user needs to find the answer for question about the fee for selling them.
  – Recover password – user log out and use the option "recover password"
  – Stores – user needs to check how many stores do they have in Washington? (if any).

Users need to performed those task and answer several questions. They are:
  – Have you faced unnatural navigation?
  – Was the main menu easy to use?
  – Was the supportive menu easy to use?
  – Have you faced difficulties in use elements (maps, films, navigation elements...)?
  – Do you have any comments and notes about performing tasks?

3.2. Expert analysis

The most important issues concerning interface quality are organized by following areas: Application interface, Navigation, Data structure and Data input. The expert analysis criteria are presented in Table1 and Table2. Table 1 shows the detailed list (the list is called "LUT list") of areas and subareas with questions assigned to each point. Accordingly, Table 2 presents the grading scale used to assess each evaluated assessed area.

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make learn

Management,
Knowledge and Learning
International Conference 2014

Table 1: LUT list of predefined testing areas with questions assigned

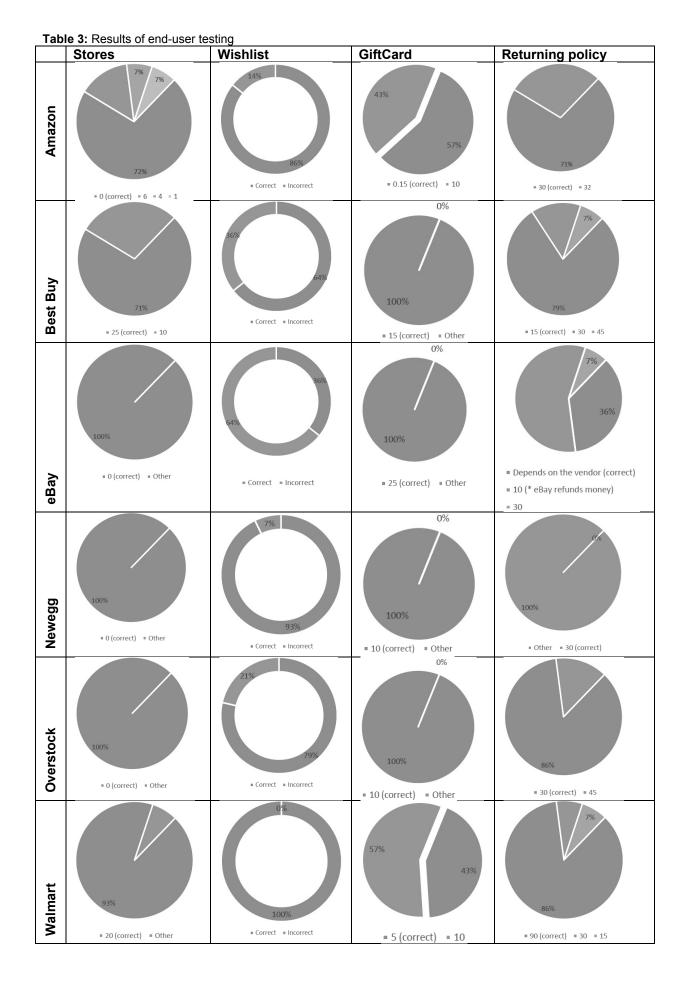| **Application interface**	
Layout	Is the layout readable?
	Is it adjusted to different resolutions?
	Is it adjusted to mobile devices?
	Is it consistent?
	Does it support task implementation?
Color scheme	Is there proper contrast between text and background ?
	Is the color scheme readable for people with color vision disorders?
	Is the color scheme readable on various kinds of displays?
Navigation and data structure	
Ease of use	Is the access to all sections of an application easy and intuitive?
	Is the access to all functions of an application easy and intuitive?
Information hierarky	Isn't the information hierarky too complicated?
Information structure	Is the information structure understandable for users?
	Is it consistent and well planned?
Screen elements	Do they support the navigation?
Feedback, system messages, user help	
System messages (general)	Do they provide enough information on the status of actions performed by user?
System messages (errors)	Do they contain hints on problem solution?
Feedback and user help	Does the information appear in places , where it may be needed?
	Is the provided information understandable for an average user?
	Is the provided information accessible for an average user?
	Is it possible for an average user to perform actions suggested by system help in order to solve the encountered problem?
Content	
Labels	Do the labels used in the interface provide enough information?
	Do all the interface elements have necessary labels?
Naming	Is the interface naming understandable for its users?
	Is the interface naming consistent?
Page text	Is it understandable for user?
Data input	
Data	Is the data validated by the form elements?
	Do the forms have elements acting as hints for the input data (e.g. on format or data range)?
	Can average user fill in the form easily?
Forms	Are they designed in a readable way?
	Are they adjusted to the mobile devices?
	Do they allow user to input all of the necessary data?

Table 2: Grading scale applied to LUT list

Grade	Description
1	Critical GUI errors were observed, preventing normal usage or discouraging user from using the application.
2	Serious GUI issues were encountered, which may prevent most users from task realization.
3	Minor usability GUI issues were observed, which if accumulated may have negative impact on user performance.
4	Single minor GUI issues were observed, which may have negative impact on user work quality (e.g. poor readability).
5	No GUI issues influencing either user performance or work quality were identified.

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make
learn

Management,
Knowledge and Learning
International Conference 2014

4. RESULTS OF THE CASE STUDY

The case study was performed in two groups. Members of the first group are typical end-users, who are not IT specialist. Their task is to go through all prepared tasks and share their user experience. The second group members are IT specialists who act as experts. The aim of the case study is to confront experts and novices and check the hypothesis: *The portal chosen as the best by end-users is also well assessed by experts*.

Results of the case study are divided into two parts. Results of the cognitive walkthrough shows assessment of GUI quality and shows user experience for each tested portal. Results of the expert analysis discuss the most important issues concerning interface quality from the evaluation report are organized by following areas: Application interface; Navigation and data structure; Messages, feedback, user help; Content; Data input. Each section was assessed separately during the tasks completion by experts. The results of the analysis contain grades (presented in Table 2) concerning each testing area (Table 1). Grades were assigned individually to each testing area of each section. The final section score is estimated on the basis of those remarks.

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

Management,
Knowledge and Learning
International Conference 2014

make
learn

Table 3: Results of end-user testing

	Stores	Wishlist	GiftCard	Returning policy
Amazon	7% / 7% / 72% ■ 0 (correct) ■ 6 ■ 4 ■ 1	14% / 86% ■ Correct ■ Incorrect	43% / 57% ■ 0.15 (correct) ■ 10	71% ■ 30 (correct) ■ 32
Best Buy	71% ■ 25 (correct) ■ 10	36% / 64% ■ Correct ■ Incorrect	0% / 100% ■ 15 (correct) ■ Other	7% / 79% ■ 15 (correct) ■ 30 ■ 45
eBay	100% ■ 0 (correct) ■ Other	36% / 64% ■ Correct ■ Incorrect	0% / 100% ■ 25 (correct) ■ Other	7% / 36% ■ Depends on the vendor (correct) ■ 10 (* eBay refunds money) ■ 30
Newegg	100% ■ 0 (correct) ■ Other	7% / 93% ■ Correct ■ Incorrect	0% / 100% ■ 10 (correct) ■ Other	0% / 100% ■ Other ■ 30 (correct)
Overstock	100% ■ 0 (correct) ■ Other	21% / 79% ■ Correct ■ Incorrect	0% / 100% ■ 10 (correct) ■ Other	86% ■ 30 (correct) ■ 45
Walmart	93% ■ 20 (correct) ■ Other	0% / 100% ■ Correct ■ Incorrect	57% / 43% ■ 5 (correct) ■ 10	7% / 86% ■ 90 (correct) ■ 30 ■ 15

473

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make learn

Management,
Knowledge and Learning
International Conference 2014

Detailed results of end-user testing are presented at Table 3. According these results Newegg portal is the best one. It is easy to manage the basket, create a wishlist, locate the stores, find information. Good results were achieved also for eBay and Overstock. According responders the worst results were achieved by Amazon.

Additional tested features were: Newsletter and shipping internationally. Newsletter testing occurred to be successful for Amazon (70%), Walmart (80%) and Overstock (only 50% of users succeeded). Shipping internationally is supported by Amazon, Best Buy, eBay, Newegg, Overstock and Walmart. Users reported *Main menu hard to use* problem for Amazon, Best Buy, eBay, Overstock and Walmart. *Supportive menu hard to use* were diagnosed for Walmart, Overstock and eBay.

Expert analysis was done according Table 1.
The best general notes (3.8) achieved Newegg portal. These results are consistent with end-user testing. According experts, Newegg achieved the best results in sections Application interface and Navigation and data structure, including Information hierarchy and Information structure. Experts found minor usability GUI issues in chaotic colour scheme and hard access to particular sections. Feedback, system messages, user help section was assessed at 3 points. Similar results (3.2) were achieved by eBay. Feedback, system messages, user help section was assessed as the best section. However, eBay's Information hierarchy and Information structure occurred to be complicated. Walmart, Overstock and Best Buy were ranked at 2.8. Data input occurred to be the worst rated section and Application interface - the best. The worst results were achieved for Amazon (2.4). Information hierarchy and Information structure were assessed as complicated and it is difficult to navigate the site. Screen elements do not support the navigation.

5. SUMMARY

Detailed results confirmed the paper hypothesis: The portal chosen as the best by end-users is also well assessed by experts. Both case studies were performed on various groups: novices and experts. Both group assessed the same portals but different scenarios and different set of questions. However, confrontation between experts and novices confirmed, that measured user experience of both group is confirmed.

Expert method allows to provide low cost GUI evaluation. Four experts have evaluated the SFA application interface in less than two and half hours. In addition, experts have developed improvement guidelines to the assessment of GUI quality. On the basis of those guidelines, different GUI elements were redesigned and adjusted.

Defined application sections and testing areas enabled checking the GUI in all important aspects. Moreover, this method helps not only to evaluate the application GUI, but also to point out its weaknesses and problems. This allows correcting the errors and achieving improvements in a relatively short period of time.

REFERENCE LIST

1. Abrahao S., Cachero C., & Matera M. (2008). Web usability and accessibility. Journal of Web Engineering, 7(4), 257-257.
2. Brown, M., Pope, N., & Voges, K. (2003). Buying or browsing? an exploration of shopping orientations and online purchase intention, European Journal of Marketing. 37(11/12), 1666-1685.
3. Chau, P.Y.K., Cole, M., Montoya-Weiss, M., & O'keefe, R. M. (2002). Cultural Differences in the Online Behavior of Consumers, Communications of the ACM, 45(10), 138-143.
4. Cushman & Wakefield (2013). Global perspective on retail: online retailing. A Cushman & Wakefield Research Publication. Retrieved from http://www.cushmanwakefield.com/en/research-and-insight/2013/global-perspectives-on-retail/
5. Dillon A., & Song M. (1997). An empirical comparison of the usability for novice and expert searchers of a textual and a graphic interface to an art-resource database. Digital Information.
6. Garbarino, E., & Strabilevitz, M. (2004). Gender Differences in the Perceived Risk of Buying Online and the Effects of Receiving a Site Recommendation, Journal of Business Research, 57, 768-775.

Human Capital without Borders:
Knowledge and Learning for Quality of Life
25–27 June 2014 · Portorož, Slovenia

make learn

Management,
Knowledge and Learning
International Conference 2014

7. Howden, W.E. (1987). Functional program Testing and Analysis. McGraw-Hill.
8. Huang, W.Y., Schrank, H., & Dubinsky, A.J. (2004). Effect of Brand Name on Consumers' Risk Perceptions of Online Shopping, Journal of Consumer Behavior, 4(1), 40-50.
9. Jarvenpaa, S.L., & Todd, P.A. (1997). Consumer reactions to electronic shopping on the world wide web, International Journal of Electronic Commerce, 2, 59-88.
10. Jochen P. et al. (1991). Errors in computerized office work: Difference between novice and expert users. ACM SIGCHI Bulletin.
11. Korgaonkar, P., Silverblatt, R., & Becerra, E. (2004). Hispanics and Patronage Preferences for Shopping From the Internet, Journal of Computer-Mediated Communication, 9(3), 2004.
12. Koyani S., Bailey R.W., & Nall J.R. (2004). Research-Based Web Design & Usability Guidelines, Computer Psychology.
13. Krug S. (2000) *Don't Make Me Think! A Common Sense Approach to Web Usability*. Que, New York.
14. Laskowski, M. (2012). Proposals for application interface testing methods. Informatyka Automatyka Pomiary w Gospodarce i Ochronie Środowiska. 4b, 21–24 (in Polish).
15. Li., N. & Zhang, P. (2002). Consumer online shopping attitudes and behavior: an assessment of research. Eighth Americas Conference on Information Systems, 508-517.
16. Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces, Proc. ACM CHI'90 Conf, 249-256.
17. Pan, J. (1999). Software Testing. Retrieved from http://www.ece.cmu.edu/~koopman/des_s99/sw_testing.
18. Prenzel A., & Ringwelski G. (2012). Design of Human-computer Interfaces in Scheduling Applications. Proceedings of ICEIS Conference, 219-228.
19. Phung Q.M. (2007). Identifying Usability Problems in relation to User Experience with System and Tests, http://projekter.aau.dk/projekter/files/61070563/1181058253.pdf.
20. Shneiderman, B. (1998). *Designing the user interface*. Addison-Wesley Longman.
21. Wharton C. et al. (1994). The cognitive walkthrough method: A practitioner's guide. Usability inspection methods. New York, NY: John Wiley & Sons Inc.
22. Zabiński, T. & Maczka T. (2010). Human system interface for manufacturing control - Industrial implementation. 3rd Conference on Human System Interactions (HSI), 350-355.
23. Zhou, L., Dai, L., & Zhang, D. (2007). Online shopping acceptance model - a critical survey of consumer factors in online shopping. Journal of Electronic Commerce Research, 8(1), 41-62.