

WEB SITE USABILITY STUDY FOR ACTUAL LINK SITE BY HCI

Adem Karahoca, Dilek Karahoca, Senay Yalçın

Department of Computer Engineering, Bahcesehir University, Bahcesehir, Istanbul, Turkey
akarahoca@bahcesehir.edu.tr, dkarahoca@bahcesehir.edu.tr, syalcin@bahcesehir.edu.tr

Abstract: *This study investigates the rapid application development process of an actual link site that visited for free mp3, sms, basic software programs etc. by generally students. We try to determine how effective the usability of www.linkindir.com which is the represent of actual link sites and usage convenience by university students. Linkindir.com has been developed by students of Computer Engineering Department in CSE3401 Software Engineering I, at Bahcesehir University. Software Usability Measurement Inventory (SUMI) inquiry method is used to measure overall usability of the software according to the ISO 9241 standard. SUMI is used to determine affectedness, efficiency, helpfulness, control, and learnability of the software. Cognitive walkthrough used with a detailed procedure to identify each task step to help and to direct the end users to take answers for the SUMI questionnaire. By this way, software usability factors are determined in the early stages of the software development process to minimize software development lifetime and costs. We also, investigate different software indicators both IQ, personal factor tests and software usability measurement inventory.*

Keywords: *HCI, Software Usability, Cognitive Walkthrough, IQ Tests, Personality Tests*

1. Introduction

Rapidly growing Internet has lots of data in the web sites. But most of the web sites have problematic usability issues. Nielsen(2002, 2003) defines usability by giving five quality components; learnability, efficiency, memorability, errors and satisfaction. In addition to these, Jones(1997) defines usability like the total effort required to learn, operate and use software or hardware. Also, end user's cognitive abilities, cognitive styles, personal properties based on the behavioural aspects, and interaction between personal properties, are effects software usage performance as important like software quality level.

Nowadays software quality concept is specialized and move to web site usability issue. Software quality factors are functionality, reliability, usability, efficiency, maintainability and portability considered in ISO/IEC 9126 (1998). However, software development life cycle is critical at the requirements analysis with the customer needs. At this stage, software designing inputs come from the requirement analysis. Actually user interface designs give a shape for the usability of the software. For this purpose, early stages in the rapid application development projects, which are called web portal projects, include some usability handicaps inside (Folmer & Bosch, 2004).

Software quality metrics are overlapping software usage effectiveness and efficiency with behavioural relationships, but, psychometric test usage are not considered enough. As proposed from this viewpoint, we targeted to find starting point of the behavioural metrics and psychometrics in the ISO standards to support each other. By this way, software usability measuring is realized with extension metric concepts. Also, we analyzed that not enough only to implement software designing steps successfully, but also, we need to take care of the end users psychometric test results into the software designing steps especially for the usability purposes.

In this study firstly, we designed and developed an intellectual content including web site (www.linkindir.com) in our CSE3401 courses at 2004 fall semester. Firstly we analyzed web site usability by Kirakowski's SUMI survey and as mentioned in (Dix et al., 2004,pp.239) (Whiteside, Bennett, & Holtzblatt,1988) are proposed 21 usability metrics which are related with ISO 9241 metrics. These surveys cover both learning performance and cognitive abilities of the students. Also, usability factors are interested with directly cognitive abilities. Thus, we evaluated our test group with IQ (Serebriakoff, 1994) and personality tests (Serebriakoff, 1996).

The content of the paper is organized as follows. In section 2, research aims, methods, participants and questionnaires are considered. Section 3 summarizes the results of the survey. Discussions, conclusions and future works are given in section 4.

2. The Research

Actually computer engineering students are selected by ÖSS(Student Selection Exam) in Turkey. But, we have to observe the cognitive abilities of the students and then observe them in the web usability steps to gain thinking about the correlation between cognitive abilities with software usability indicators. Web site usability conditions can be handled in the analysis and design steps in the development phase of the software development life cycle. According to the (Calcaterra et al., 2005) hypermedia navigation behaviour is linked to computer skills rather than to the cognitive styles. But our assumption is cognitive abilities directly related with software usage success ratio. Cognitive abilities and cognitive styles are different things and therefore, students are selected from the same faculty and department, especially from computer engineering department. So, we assumed that Computer Engineering students have averagely same cognitive abilities as mentioned in the conclusions.

In our study, our team started to develop web based actual link site project. In the early prototyping development stages, we used inquiry methods to measure overall usability according to the ISO 9241 standard. Software Usability Measurement Inventory (Kirakowski, & Corbett, 1993) is used to determine affectedness, efficiency, helpfulness, control and learn-ability of the software. Also at this point, cognitive walkthrough (Dix, et al, 2004) used with a detailed procedure to identify each task step to help and to direct the end users to take answers for the SUMI questionnaire.

2.1. The Aim

The aim of the survey was determining correlation between behavioural and IQ test values with software usability test results to obtain new metrics or indicators to define web usability. Our research assumptions can be listed as follows:

1. How are the three different types of intelligence characteristics (numerical, geometrical and verbal) and the behavioural characteristics are effective or ineffective for the end users on the usability of the software?
2. Is the verbal understand-ability of the end users, effective in definitive time periods?
3. Is the software usability related with the end users personal factors?
4. Are the expert computer user's software usability success more than the other novice university students who have higher IQ values?
5. Should usability evaluation be applied on the ISO 9241 and must be adding new usability indicators which are related with IQ and personality tests?
6. What are the effects of the personality test results for the determining the software usability?

2.2. Method

HCI (human computer interaction) is an interdisciplinary science branch for the interaction of people and systems and the ways they influence each other. We can use HCI methods to determine how to design a system according to the needs of users, including their abilities, limitations, and work environment settings. We used Cognitive Walkthrough methodology to observe the end user reactions to the software in a detailed task list scenario within the HCI approach. Cognitive Walkthrough methodology is performed at any stage of design step using a prototype, a conceptual design document, or the final product. This is a more specific version of a design walkthrough, focusing on cognitive principles (Dix, et al, 2004).

Based on a user's goals, a group of evaluators steps through tasks, evaluating at each step how difficult it is for the user to identify and operate the interface element most relevant to their current sub goal and how clearly the system provides feedback to that action. Cognitive walkthroughs take into consideration the user's thought processes that contribute to decision making, such as memory load and ability to reason. The cognitive walkthrough is a technique for evaluating the design of a user interface, with special attention to how well the interface supports "exploratory learning," i.e., first-time use without formal training. The evaluation can be performed by the system's designers in the early stages of design, before empirical user testing is possible. Early versions of the walkthrough method relied on a detailed series of questions, to be

answered on paper or electronic forms. This tutorial presents a simpler method, founded in an understanding of the cognitive theory that describes a user's interactions with a system. The tutorial refines the method on the basis of recent empirical and theoretical studies of exploratory learning with display-based interfaces. The strengths and limitations of the walkthrough method are considered, and it is placed into the context of a more complete design approach (Chan, 2002).

This approach is intended especially to help understand the usability of a system for first-time or infrequent users, that is, for users in an exploratory learning mode. We applied personality factors tests, IQ and software usability tests to users with this approach.

2.3. Participants

This work is performed with 24 students, 18 boy and 6 girls, aged 20-22 years ($M=21.01$, $SD=0.89$) who are registered to the "CSE3807 Cognitive Science" course. Software usability tests, IQ and personal factor tests are applied to the whole students. There are near to 1000 items in the IQ, personality and software usability surveys. Nielsen's discount technique shows us that a test with at least 15 users enough to discover all the usability problems in the design stage (Nielsen, 1994, 2000).

2.4. Questionnaires

There are sort of software usability questionnaires to determine user satisfactions about the software. SUMI (software usability measuring inventory) and QUIS (questionnaire for user interface satisfaction) are developed by (Kirakowski, 1993). Also, web based software solutions can be evaluated with MUMMS (measuring the usability of multi-media) and WAMMI (website analysis and measurement inventory) which are prepared by (Levi & Conrad, 2001).

We had chosen SUMI, because it has three answer type Likert scale and good grouping for the ISO 9241 usability criteria inside. Software usability questionnaire includes 50 rules which are identified by 3-point scale (agree, undecided, disagree) and easy to get response in a five minutes. Personality survey (Serebriakoff, 1996), includes 4 different stages.

- a. Personality Factor 1: 20 item includes with two choices (a or b) should be filled in a 5 minutes,
- b. Personality Factor 2: 30 item includes with two choices (a or b) should be filled in a 10 minutes,
- c. Personality Factor 3: 15 geometrical shape completion test should be filled in a 5 minutes,
- d. Personality Factor 4: 210 item with yes or no questions should be filled in a 15 minutes.

Personality Factor 1 is used for determining extraverted as opposed to introverted person type. Personality Factor 2 is measure the neuroticism as opposed to stability personality type. Personality Factor 3 helps us to determine creativity of the students. This section of the survey includes geometrical shapes for the completion purposes. Personality Factor 4 tries to reveal stability of the persons. (Serebriakoff, 1994)'s IQ inventory is divided into three sub section: a) Numerical, b) Verbal, c) Geometrical. Each section has 50 questions inside.

2.5. Usability Measuring Scenario

Following scenario is used as a procedure for the students to testing www.linkindir.com web site. Each student was applied to the following work procedures.

2.5.1. E-Card Sending Procedure

1. Click the e-card sending link,
2. Select the comics section from the open page,
3. Select a image which you like,
4. Write your name and your e-mail address,
5. Write your friends e-mail address(es) into "to section", if you have two or more sender address put coma inside the mail addresses,
6. Put a message into message section with suitable format,
7. If you wish put smiles to your message,
8. Push the preview link to see your message form. If it is like you wish send it, otherwise put required modifications and then send it.

3. Results

We evaluated survey results into two groups: a) IQ and Personality factors b) Software usability factors according to the ISO 9241 and SUMI. IQ survey includes verbal, numerical and geometrical tests to evaluate the students (Serebriakoff, 1994).

3.1. Usability Metrics from Whiteside, Bennett, & Holtzblatt in (Dix, et.al., 2004, pp.239) for ISO 9241

21 criteria can be used to determine the measuring method for usability attribute and the possible ways to set the worst/best case and planned/now level targets. These measurements are named as usability metrics. Following part includes the answers of the each criterion.

1) *Time to complete a task*

E-card sending procedure implemented with 24 students. Procedure completion average time is 3 minutes, minimum completion time is 1 minutes and 20 seconds and maximum completion time is 6 minute and 33 seconds.

2) *Per cent of task completed*

All students are completed the test procedure except 3 students. Three students are not connected to the web site, reasons of the network problems. Procedure applicability percentage 87.5% is measured.

3) *Per cent of task completed per unit time*

A job completion time is average 3 minutes and if we accept unit time as a minute, 33.3% percentage is a job completion in a unit time.

4) *Ratio of successes to failures*

21 students are completed procedure successfully and 3 students are unsuccessfully. The rate of success over failure is 7.

5) *Time spent in errors*

The average time is 1 minutes and 30 seconds for spending to errors.

6) *Per cent or number of errors*

E-card sending procedure is completed without encountering with errors.

7) *Per cent or number of competitors better than it*

In 24 student, only 5 student completed procedure without done any error. According to this value, 20.8% is procedure completion percentage without making any error.

8) *The number of commands used*

5 commands used to complete the e-card sending procedure.

1. *Go to the web page,*
2. *Click the link,*
3. *Click for e-card image sending,*
4. *Filling the information for the e-card,*
5. *Clicking the send e-card button.*

9) *Frequency of help and documentation use*

Total 24 students attend to the test and help and documentation usage frequency is observed as 11. Students have said that there are enough help and documentation in the web site.

10) *Per cent of favorable/unfavorable user comments*

Half of the test users are comments the favorable and other half comments are unfavorable. Only 3 users (12.5%) found the web site useless.

11) *Number of repetitions of failed commands*

Average number of repetitions of failed commands is only 3.

12) *Number of runs of successes and of failures*

There aren't any unsuccessful commands in this work by e-card sending scenario. All commands performed the user requests depends on the hardware performance.

13) *Number of times interface misleads the user*

There aren't any interface misleads the user in the e-card sending procedure activities.

14) *Number of good and bad features recalled by users*

Number of good features is 9 and bad features are 6 named by users in the scenario.

15) *Number of available commands not invoked*

There are 64 unused commands (links) in the home page.

16) *Number of regressive behaviors*

Regressive behaviors number is observed as 9.

17) *Number of users preferring your system*

14 students prefer to use this web site after than the tests. After this work completed, we published this web site and following table show us the number of the visitor and their country.

Table III. Visitors' hit results for the linkindir.com

Country	Total Hit	Percent
Turkey	129.189	89.7%
Germany	5.099	3.5%
Netherlands, The	1.572	1.1%
United States	928	0.6%
United Kingdom	731	0.5%
France	717	0.5%
Switzerland	660	0.5%
Austria	488	0.3%
Belgium	383	0.3%
Sweden	329	0.2%
Unknown	1.969	1.4%
The rest	1.895	1.3%
Total	143.960	100.0%

18) *Number of times or average number of users need to work around a problem*

When this test performed by the students a few problems are encountered and average problem solving times is between 45 seconds and 1 minute.

19) *Number of times the user is disrupted from a work task*

Average number of times the user is disrupted from a work task is 5. These problems generally depend on hardware performance and solved in a minute.

20) *Number of times user loses control of the system*

In the website usability testing stage, none lose the software control.

21) *Number of times user expresses frustration of satisfaction*

75% per cent users are satisfied from the web site. But, there are critical unsatisfied properties in the web site. For example, complexity, insufficient categorization, low speed of the web server and insufficient sources are main categories of the frustrations.

3.2. Software Usability Measurement Inventory Results

SUMI (Software Usability Measurement Inventory) was developed by the Human Factors Research Group (HFRG) at University College Cork, Ireland. SUMI is a 50-item questionnaire for assessing software system usability. It has five subscales (the descriptions of the subscales are based on Kirakowski, 1993):

1. Efficiency: the degree to which users feel the software assists them in their work,
2. Affect: users' general emotional response to the software,
3. Helpfulness: the degree to which users feel the software assists them in using it,
4. Control: the degree to which users feel they, and not the software, are in control,
5. Learnability: the ease with which users feel they have been able to get started using the software and learn new features,

Each subscale has 10 items. Each item is rated on a 3-point Likert response scale with the points "agree," "don't know" and "disagree". Comparing the items of each scale with the descriptions of the seven dialog principles in ISO 9241 Part 10, Kirakowski (1993) found that four of the five subscales seem to correspond directly to dialog principles in ISO 9241 Part 10. The fifth subscale seems to be related to another dialog principle.

ISO 9241 Part 10	SUMI
Suitability for the task	Efficiency
Self-descriptiveness	Helpfulness
Controllability	Control
Conformity with user expectations	affect, efficiency
Error tolerance	
Suitability for individualization	
Suitability for learning	Learnability

	Successful (%)	Don't know(%)	Unsuccessful(%)
Affect	0,48	0,28	0,24
Control	0,51	0,32	0,17
Efficiency	0,50	0,25	0,25
Helpfulness	0,55	0,30	0,15
Learnability	0,63	0,24	0,13

4. Conclusions

Web site usability is measured both SUMI survey and usability metrics from (Whiteside, Bennett, & Holtzblatt). As detailed in Table III, learnability of web site is 63%, helpfulness is 55%, controllability is 51%, efficiency is 50% and affectivity is 48% is measured according to the SUMI survey. Students' IQ survey results are answered correctly in 70% for geometrical section, 74% for verbal section and 72.88% for numerical section. Average IQ score is 121 for both of the genders. Successful students who are correctly answered the geometrical questions more than the others have moderate and creativity personality types. Stable personal type of students has higher numerical section score than the others.

Intelligence characteristics and the behavioural characteristics are effective for the end users on the usability of the software. Higher IQ scores support the introverted, stable and generally creativity personality types. All students can be accepted novice for the web site. Thus, learnability of the commands and links directly based on students intelligence and learning abilities. By this way, we tried to evaluate a web site which is analyzed and developed in our course. Future works can be related about integration of the cognitive abilities, individual differences, cognitive styles and software usability measuring only one inventory. For this purpose, validation and confidence of the inventory has to proof by sample test.

References

1. Calcaterra, A., Antonietti, A., Underwood, J.(2005). Cognitive style, hypermedia navigation and learning, *Computers and Education*, 44, 441-457
2. Chan, W. (2002). Increasing the success of physician order entry through human factors engineering, *J. Healthcare Inf. Mang.*, vol. 16, no. 1, pp. 71-79
3. Dix, A., Finley, J., Abowd, G., Beale, A., (2004). *Human-Computer Interaction*, Prentice Hall.
4. Folmer, E., Bosch, J.(2004). Architecting for usability: a survey, *Journal of Systems and Software*, 70, 61-78.
5. ISO 9241-11 (1997). Draft International Standard on Ergonomics Requirements for office work with visual display terminals (VDT), *Part 11: Guidance on Usability*, ISO, 1997.
6. ISO/IEC 9126 (1998). Information Technology – Software product quality, *Part 1: Quality model*, ISO
7. Jones, C. (1997). Software quality analysis and guidelines for success, International Thomson Computer Press, 492.
8. Kirakowski, J., & Corbett, M., (1993). SUMI: The Software Usability Measurement Inventory, *British Journal of Educational Technology*, 24(3), 210-214
9. Levi, M., D., Conrad, F.,G. (2001). Usability Testing of World Wide Web Sites. Retrieved on March 9 2005 from http://www.bls.gov/ore/htm_papers/st9601050.htm
10. Nielsen J. (2002). *Designing Web Usability: The Practice of Simplicity*, 7th print, New Riders, Indiana.
11. Nielsen J.,R.,L., Mack (1994). *Usability Inspection Methods*, Wiley, New York, 1994.
12. Nielsen,J., (2000). Why You Only Need to Test With 5 Users, Retrieved on March 9 2005 from <http://www.useit.com/alertbox/20000319.html>
13. Nielsen,J., (2003). Usability 101: Introduction to Usability, Retrieved on March 9 2005 from <http://www.useit.com/alertbox/20030825.html>
14. Serebriakoff, V.(1996). “*Personality Characteristics Surveys*”, Turkish version.
15. Serebriakoff, V.(1994). “*Test your I.Q.*”, Chancellor Press, London, United Kingdom.