

User-Interface Usability Evaluation

Ali Mohamed

*Faculty of Engineering, Computer Science and Engineering
American University of Sharjah
Sharjah, UAE*

b00066590@aus.edu

Tarik Ozkul

*Faculty of Engineering, Computer Science and Engineering
American University of Sharjah
Sharjah, UAE*

tozkul@aus.edu

Abstract

Nowadays, computers and internet are playing the major role in the development of business and different aspects of human lives; hence, the quality of user-computer interface became an important issue. User interface (UI) can become an Achilles heel in a well-functioning system; due to the fact that most users judge the quality of a product by its usability. The UI layout design improves the usability of a product and accordingly may determine its success; so, due to this and more, the need of an objective way of evaluation of UI has arisen. This paper discusses various UI usability evaluation techniques and shows the recent developments in this field.

Keywords: User Interface, User Interface Evaluation, Software Interface, Machine Intelligence, Usability, Usability Evaluation.

1. INTRODUCTION

Computers are becoming the cornerstones in most our life activates nowadays, and unlike the past almost everyone is a computer user; so adopting a complex user interface of a product will eventually result in catastrophic failure. The user interface has to be: simple, consistent, conventional, and familiar. Also, user interactions can be used to judge the usability of the system [1].

The usability of a UI has been a trending topic and many researches have been conducted on this topic. Also usability testing is the feedback part of the prototype interface design process, until the final product is out and even after the production [2].

Evaluating the usability of a user interface can be done in a subjective way by depending on the opinions of users and experts to determine the quality of a system, or in an objective way by using certain rules and metrics to decide on the quality of a given system [3]. Both the opinions of users and data used by rules to evaluate the system can be collected either manually or automatically. In this paper section 2 covers the subjective evaluation of UI, and section 3 covers the objective evaluation of the system whereas section 4 covers the automation of both of the evaluation approaches. In section 5 a methodology that lies in the intersection of the two approaches is covered, in which the evaluation is done by either taking the opinion of group of targeted users "subjective" or using data collected from them "objective" to determine the quality of a system that is most appropriate to be used by them.

2. SUBJECTIVE EVALUATION

Subjective evaluation of a user interface is done by analyzing the opinions of users and experts to give a sound judgment accordingly. One type of the subjective evaluation of the user interface is the heuristic evaluation which is an inspectional method, where a certain number of experienced

evaluators determine to which extends the design of the UI follow an established guidelines (Heuristics) [4]. Those heuristics are set of rules defined by J. Nielsen who is the author of the technique [5]. Another type of subjective evaluation like usability scale method [6] uses well defined and standardized questionnaires filled by the users post using the targeted user interface, and then results are analyzed to evaluate the user interface according to those users' experience.

In work done by F. Paz et.al [4]; they conducted a heuristic evaluation on a website interface with a team of experts, and then compared the results with those obtained from a usability test done by normal users. Their results showed that 90% of the UI problems identified by the normal users were already identified by the experts' team in the heuristic evaluation stage. But, they also, showed that only 20% of the problems marked as critical by the experts in the heuristic evaluation process were relevant to the normal users. Also, they did not cover the time and cost of the heuristic process, and they did not cover the coverage of the technique in an elaborative way.

Mansor Yushiana et.al [7] showed that heuristic evaluation can be used to assess the interface design of websites based on online catalogues like libraries' websites. In their research they focused on three heuristics out of the ten heuristics of J. Nielsen, which are aesthetic design, visibility, and match with reality, and then they trained ten users to use a proposed library interface in order to assess it according to the three heuristics adopted. After the users provided their opinions individually, the results showed that the proposed interface had 70% conformance with the heuristics and they were able to identify several interface design violations.

Dorian Davis et.al [8] have conducted comparative study of the usability of type2 diabetes education websites using heuristic evaluation based on the J. Nielsen heuristics. In their study, they used Alexa [9] which is a web traffic analyzer of data collected from sample of all internet users in the world to find the most used search engine at the time the research was being conducted. Later, they have used that search engine to search for the top three type2 diabetes education websites using their defined search phrases. After this, they consulted five heuristic evaluation experts to provide judgment on the three websites' interfaces. After combining the findings they were able to provide comparative judgment on the three websites. Also, they identified the top five mutual violations from the three websites, and provided recommendations according to them. They suggested that their evaluation needs to be supported with users' interaction testing as a complement to the heuristic evaluation.

Freddy Paz et.al [10] showed that the classical J. Nielsen heuristics fail when it comes to evaluating transactional websites, so, they came up with a set of fifteen heuristics to cover such types of websites. They have conducted a comparative study by asking a group of students to perform the heuristic evaluation of a transactional website using the two different sets of heuristics after being trained to do so. After that, they have analyzed the surveys obtained from the students after the evaluation, and the results showed that their proposed heuristics surpassed the traditional ones in terms of "perceived ease of use" and "perceived usefulness" and "intention of use" which are the aspects need to be compared based on the suggestion of the technology acceptance model (TAM) [11]. However, they have declared that the difference is not significant enough to revise those low scoring heuristics.

Ewa Callahan et.al [12] have adopted the computer usability satisfaction questionnaires for interface evaluation. In their study, they have used the computer usability satisfaction questionnaires (CUSQ) [13], which is a questionnaire developed according to psychometric factors to evaluate an online products catalog website interface. In their research, they hired a group of users and gave them certain search tasks to conduct on an online product catalog website. After that, the users were asked to take the questionnaire to give their feedback about the interface of the website. From their results, they recommended one type of catalog where the selection of product is done through the user navigation in the website instead of entering specific attributes and then getting the recommendations accordingly.

3. OBJECTIVE EVALUATION

Observing the behaviors of the user while interacting with a UI can reflect the actual feedback of that user, and in turn give objective evaluation of the interface. If enough users are involved in such evaluation, an index can be generated from their experiences to compare different user interfaces.

Ahmet et.al [14]; showed that such an evaluation can be used as an indicator of the level of the machine intelligence since it tests the ease of use in terms of user interactions.

In the work done by Ahmet et.al [14]; they have compared the results obtained from user interactions evaluation with the results of survey based evaluation. They have collected data form users with similar background and experience while the tested systems doing the same tasks. Then, they have used a fuzzy logic system to evaluate the systems. The fuzzy logic inference system designed was a function of:

- Complexity of each subtask in the main task;
- UIQ (User Interface Quotient) data;
- Total number of subtasks;
- Difficulty of data transfer between the machine and human.

Based on this, they got an index for each UI objectively and they compare it with the results from a survey that was given to the same users involved in the first test. Their results showed that their methodology matched the survey in 70% of the cases. The researchers declared that the coverage of the methodology can be improved if the user interactions logging is done automatically, and more users are involved in the process.

Cheolil Lim et. al [15] showed that user interface evaluation can be used as a middle stage for design enhancement. They have used heuristic evaluation done by expert professors to revise their design of an interface for digital textbook platform for schools, and after that, they have used the recommendations obtained from the stage of heuristic evaluation to modify their prototype. They have used both surveys and log files of students interacting with their modified version of the interface, and introduced changes to their prototype interactively in this stage. After the two levels of revisions, their results showed that the students expressed that the modified interface is user-friendly, which proves that both subjective and objective evaluation methods can be used to enhance the interface design.

4. AUTOMATION USABILITY EVALUATION

The results of usability testing done by different evaluators can vary even if they use the same technique; so, the either the number of the evaluators participating in the process must be increased, or the process must be automated. Capturing the usability data can be automated to bring up many advantages like [2]:

- Reducing the cost of evaluation: automation decreases the time of the process and the people hired, which in terms reduce the cost.
- Increasing consistency: automation can help in uncovering all the errors in the system.
- Increasing quality of evaluation: automation increases amount of testing data which leads to improvement in quality and coverage of the evaluation.
- Reducing the need of expert evaluators: the need of expert evaluators can be mitigated as a result of increasing the amount of testing data by automation.
- Enabling comparison between different alternatives: automation can enable comparison by analytical modeling.

4.1 Automation of Heuristic Evaluation

The heuristic evaluation of a UI can also employ automatic capturing of the level of compliance of the UI with the heuristics, this can reduce the time and cost of the evaluation process, but the

tradeoff is the complexity of developing the automated system and the difficulty of covering all types of UI specific heuristics.

In work done by Alexiei et.al [16], the researchers proposed a system that automates the heuristic evaluation of a website. They developed the set of heuristics to be followed in the evaluation process, moreover they allow modification of the heuristics used by their system which made it scalable. They ran their system on 10 different websites and compared it with the results of testing to normal J. Nielsen heuristic evaluation on the same 10 websites. Their results showed that their system was able to surpass the J. Nielsen evaluation in terms of usability violations detection; their results demonstrated that their system detected 28.15% more usability violations than the manual J. Nielsen heuristic evaluation. But their automated system failed to cover site-specific recommendations as part of the heuristics, which were included by J. Nielsen as Nielsen's site-specific recommendations. Also, their heuristics that they have developed is targeted towards business websites. They have declared that if their system is to be used for evaluating another type of websites, the set of heuristics must be revised.

4.2 Automation of User-Machine Interaction Evaluation

User interaction with the UI can be monitored automatically to increase the coverage of the method, also, the quality of the evaluation increases and the time of the process decreases due to automatization. Automatic capturing of user behaviors can be done using intelligent agents then upon that this data can be analyzed to give judgement about the UI.

Intelligent agents: An agent is a software abstraction used to describe complex entities in the system. The agent is defined by its behavior. Its importance is due to the fact that programming an agent-based system is done by specifying its behaviors instead of defining classes, methods and attributes [17].

Properties of a typical software agent:

- **Autonomous:** software agent is capable of operating as a standalone process and without involving the user.
- **Communicative:** software agent communicates with the user, other agents, and other processes in the system.
- **Perceptive:** software agent can adapt in the system by perceiving and responding to changes occurring in the system.

Software agents can be used in usability testing where the computation is distributed in many testing machines those capture and communicate their statistics to be analyzed.

In work done by Eduardo et al. [3], they have developed an agent-based evaluation system of web UI consisting of two types of agents. One type of agents to analyze the code of the web page itself and detect any problem that can be found from the code, such as browser specific tags and the size of the page itself which makes it take more time to load and so. The other type of agents monitors user interaction with the system while navigating from a single page to another combined with phrasal analysis of the content of the web pages involved in the overall task. They tested their system on a website and their results showed they all the issues detected by their system matched the experts' evaluation. But their study focused only on HTML based websites and their system does not analyze the part of the code that is residing in the server, so, their system can be enhanced to cover this issue.

5. USER-CENTRIC EVALUATION

The user interface must be designed in a way that mimics the users' way of thinking out the targeted task [18]; so, from this concept the need of user-centered interface design and evaluation has arisen, to come up with systems those cover the needs of users in an effective and optimized way.

In work done by Sookyung et.al [19], they have developed a nursing documentation system's interface to handle nursing documentation like nursing admission assessment and blood administration, etc..., and in order to evaluate their user interface design. Five nurses were hired and trained for a month, then they were asked to use the system and fill a questionnaire accordingly. Their results showed that the nurses' feedback matched with a heuristic evaluation that they have implemented by five experts in a previous stage.

In work done by L. O. Yusuf et.al [20], the researchers have designed a prototype for reinforced concrete design software (RCD) after consulting RCD experts to ease the design of the software. They presented their proposed interface to usability experts to get their feedback in order to enhance their design. At this stage they didn't mention the reasoning of opinions of those experts giving recommendations. After that, they gave their prototype to twenty-five RCD experts to perform certain tasks on it, while the participants were asked to perform the same set of tasks on an existing software in order to compare the two. Their results showed that twenty-two out of the total agreed on the preference of their prototype over the existing solution.

6. CONCLUSION

The design of a user interface is the major factor that determines the user experience and the user's decision on whether to keep on using a certain product or abandon it. So, there has been a quite heavy research effort done in the user interface design and evaluation area. This paper covers the different techniques used in user interface evaluation. This is also considered as the feedback part of the design process which helps to improve and strengthen the interface to effectively cover user's requirements.

The paper covered both the subjective and objective evaluation approaches including both old and recent works in that area. The paper also covers the user-centered interface evaluation as an approach that can be implemented by both subjective and objective means of evaluation.

We conclude that, both subjective evaluation in terms of opinions of experts, and objective evaluation in terms of collecting data about user interaction with the system are important tools to judge the quality of a certain user interface design. Obviously some issues may only surface when the real users start interacting with the system and may be missed by the experts. There is both cost and time needed to implement users' interaction evaluation with an acceptable coverage, and this makes it convenient to employ experts in the evaluation process as they may do it faster and with a better coverage of than tests conducted by limited number of users. Moreover, automating both approaches reduces the cost and time and also can provide better coverage in users' interaction test; since more users will be involved with a reduced cost. However the complexity of automating the evaluation process and making the automated versions as accurate as the normal ones is still an issue and there exist a trade off between the benefits of automation versus development complexity and effectiveness.

7. REFERENCES

- [1] What My 2.5 Year-Old's First Encounter With An iPad Can Teach the Tech Industry, Todd Laplin <http://www.cbsnews.com/news/what-my-25-year-olds-first-encounter-with-an-ipad-can-teach-the-tech-industry/>.
- [2] M. Y. Ivory and M. A. Hearst, "The state of the art in automating usability evaluation of user interfaces," *ACM Computing Surveys (CSUR)*, vol. 33, pp. 470-516, 2001.
- [3] E. Mosqueira-Rey, D. Alonso-Ríos, A. Vázquez-García, B. B. del Río, and V. Moret-Bonillo, "A multi-agent system based on evolutionary learning for the usability analysis of websites," in *Intelligent Agents in the Evolution of Web and Applications*, ed: Springer, 2009, pp. 11-34.

- [4] F. Paz, F. A. Paz, D. Villanueva, and J. A. Pow-Sang, "Heuristic Evaluation as a Complement to Usability Testing: A Case Study in Web Domain," in *Information Technology-New Generations (ITNG)*, 2015 12th International Conference on, 2015, pp. 546-551.
- [5] J. Nielsen, "10 usability heuristics for user interface design," Fremont: Nielsen Norman Group.[Consult. 20 maio 2014]. Disponível na Internet, 1995.
- [6] J. Brooke, "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, pp. 4-7, 1996.
- [7] A. Schmetzke, E. Greifeneder, M. Yushiana, and W. Abdul Rani, "Heuristic evaluation of interface usability for a web-based OPAC," *Library Hi Tech*, vol. 25, pp. 538-549, 2007.
- [8] D. Davis and S. Jiang, "Usability evaluation of web-based interfaces for Type2 Diabetes Mellitus," in *Industrial Engineering and Operations Management (IEOM)*, 2015 International Conference on, 2015, pp. 1-6.
- [9] "How Are Alexa's Traffic Rankings Determined?" Alexa. Alexa Support, n.d. Web. 29 Apr. 2016.
- [10] F. Paz, F. A. Paz, and J. A. Pow-Sang, "Evaluation of Usability Heuristics for Transactional Web Sites: A Comparative Study," in *Information Technolog: New Generations*, ed: Springer, 2016, pp. 1063-1073.
- [11] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS quarterly*, pp. 319-340, 1989.
- [12] E. Callahan and J. Koenemann, "A comparative usability evaluation of user interfaces for online product catalog," in *Proceedings of the 2nd ACM conference on Electronic commerce*, 2000, pp. 197-206.
- [13] J. R. Lewis, "IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use," *International Journal of Human-Computer Interaction*, vol. 7, pp. 57-78, 1995.
- [14] Al Zarqa, Ahmet, Tarik Ozkul, and A. R. Al-Ali. "A Study toward Development of an Assessment Method for Measuring Computational Intelligence of Smart Device Interfaces." *INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTERS IN SIMULATION* (2014): n. pag. Web.
- [15] C. Lim, H.-D. Song, and Y. Lee, "Improving the usability of the user interface for a digital textbook platform for elementary-school students," *Educational Technology Research and Development*, vol. 60, pp. 159-173, 2012.
- [16] A. Dingli and J. Mifsud, "Useful: A framework to mainstream web site usability through automated evaluation," *International Journal of Human Computer Interaction (IJHCI)*, vol. 2, p. 10, 2011.
- [17] "Why, When, and Where to Use Software Agents." Agent Builder. N.p., n.d. Web. 15 Apr. 2016. <http://www.agentbuilder.com/Documentation/whyAgents.html> /.
- [18] L. E. Wood, *User interface design: Bridging the gap from user requirements to design*: CRC Press, 1997.

- [19] S. Hyun, S. B. Johnson, P. D. Stetson, and S. Bakken, "Development and evaluation of nursing user interface screens using multiple methods," *Journal of biomedical informatics*, vol. 42, pp. 1004-1012, 2009.
- [20] L. Yusuf, O. Folorunso, A. Akinwale, and I. Adejumobi, "User Interface Design and Usability Testing of a Reinforced Concrete Design (RCD) Beam Interface," *British Journal of Mathematics & Computer Science*, vol. 1, p. 16, 2011.