Journal of Applied Sciences 15 (6): 917-922, 2015
ISSN 1812-5654

ansinet
Asian Network for Scientific Information

RESEARCH ARTICLE OPEN ACCESS

DOI: 10.3923/jas.2015.917.922

Improving the Effectiveness of Developed Heuristic Evaluation Method

Majed A. Alshamari
Department of Information System, College of Computer Sciences and Information Technology (IT), King Faisal University, P.O. Box 400, Al-Ahsa 31982, Kingdom of Saudi Arabia

ARTICLE INFO
Article History:
Received: April 13, 2015
Accepted: April 14, 2015

Corresponding Author:
Majed A. Alshamari,
Department of Information System,
College of Computer Sciences and Information Technology (IT),
King Faisal University,
P.O. Box 400, Al-Ahsa 31982,
Kingdom of Saudi Arabia

ABSTRACT
Use of web-based systems are considered seriously since the revolution of the Internet. This can be justified through the large number of users with different types of skills, knowledge, culture and also due to the increase number of business movement to the web rather than the traditional way. Heuristic evaluation is one of methods previously proposed and used to assess a system usability. It is considered as a cheap, fast and effective usability evaluation method. It is conducted by a set of experts for examining a system interface through "heuristics". This method has been challenged due to its broad and general heuristics. The aim of this study is to examine the effectiveness of this method after developing the heuristics. Participatory Heuristics Evaluation (PHE) with more guidelines and descriptions have been adopted in this study. The usability testing was used in this research as a benchmark. This study results showed that there is an improvement in the method effectiveness in terms of the identified usability problems number, in particular disaster and major usability problems. The results further stresses the need for serious development for heuristics in terms of its procedure or components.

Key words: Usability testing, heuristic evaluation, developed heuristics evaluation, usability effectiveness

INTRODUCTION

Usability is identified as a very broad term in system design. It is mainly concerned as how much usable a system is. Ensuring system usability can be described as a positive quality assurance practice. Ignoring usability can cost organizations significant resources in terms of financial, time and user satisfaction. Therefore, recently, evaluating systems' usability is seriously considered. Different methods are used to assess and evaluate systems' usability as it's important to assess usability criteria in different stages of the system. Different approaches are usually adopted in order to evaluate the usability such as usability testing, cognitive walkthrough and heuristic evaluation in addition to the inquiry methods; surveys and interviews. Therefore, this study aims to examine the effectiveness of the heuristics evaluation after adopting participatory heuristics evaluation and breaking them into guidelines with more details and explanations for each set of heuristics.

Usability testing is one of the common usability evaluation methods. Although it is claimed to be costly and time consuming, yet it usually produces satisfactory results (Lindgaard and Chattratichart, 2007; Isa *et al*., 2014). Also it needs a set of tasks to be performed by a number of users. Nielsen (2000) and Bevan *et al*. (2003) suggested that 5 users are enough to reveal 75% of the usability problems. However, this argument was challenged by a number of researches Mueller *et al*. (2009) and Lindgaard and Chattratichart (2007) who found that 5 users are only able to discover up to 35% of the usability problems. But at the same time, 5 users are enough to evaluate usability problems if the cost and time are seriously considered. Cognitive walkthrough system is another usability evaluation method which needs a number of experts based on a predefined set of tasks. It is also described as quick and efficient method. Besides, it is not popular as heuristic evaluation and again needs experts with different background. Heuristic evaluation is a popular method among usability practitioners and was proposed by Nielsen (1995). Some

researchers described it as a fast, cost-effective and effective method for revealing usability issues (Shazeeye and Shanmugam, 2007; Geng and Tian, 2015). Allen *et al*. (2006) and Hvannberg *et al*. (2007) stated that 3-5 evaluators are required to perform heuristic evaluation session in order to check the targeted interface or system based on a set of heuristics. Also, they are supposed to have HCI background in order to eliminate any issues of misconducting the evaluation process. Recently, this method was criticized due to its general level of application and being an old heuristic so its applicability is questioned (Geng and Tian, 2015). Experts with different background is expected to offer more validity to the results as it is claimed to be a subjective assessment method (Allen *et al*., 2006; Hvannberg *et al*., 2007).

Integrating one or more than one usability evaluation methods will certainly offer better quality results. But it is considered as a costly decision if multiple usability evaluation methods are adopted. Therefore, it is recommended to improve one of the limitations of these methods in order to increase the effectiveness of any method involved. Therefore, this study aimed to adopt a Participatory Heuristics Evaluation (PHE) with more guidelines and descriptions to examine the effectiveness of this method.

Sivaji *et al*. (2011) explored usability evaluation methods and the ways heuristic evaluations can be improved. They proposed an automated tool to save evaluator's time and procedure automation was the clear objective. They also claimed that the effectiveness of heuristic evaluation method is feasible after automation of the evaluation procedures. However, they needed a group discussion to finalize the evaluator's comments and evaluations. They further claimed that the saved time could be used by the evaluators to detect more usability problems. However, this contribution did not improve the method itself, but enhanced the procedure and the method application. Peute *et al*. (2011) conducted a study on how participatory approach can help heuristics evaluation in health information system. They obtained low quality results when it is compared to the achieved results of usability testing method. On the other hand, the achieved results may be justified due to the nature of the system as it's a health information system. They also clearly stated that research is needed to assess the value of the domain specific classifications applied in this study to extend the HE knowledge to guideline based HIS. Besides, they validated their results using usability testing and think aloud techniques as benchmark.

Edwards *et al*. (2008) presented a case study to assess a commercial electronic health record usability. They attempted to improve the heuristic evaluation effectiveness through adopting heuristic walkthrough method. The evaluation process was divided into two stages. The first one is a task-based evaluation. The second stage is "free-form" where the evaluators check how each heuristic is applied in the targeted system. They also used the achieved results as an input for the second round of the system design to avoid future usability problems. They recommended paired evaluators', knowledge about the targeted system and different experts' backgrounds to improve the effectiveness of the used method. While, Masip *et al*. (2011) conducted an interesting research to apply the current heuristics on two real experiences to examine the applicability of the current heuristics. They were not able to utilise all the heuristics as some of them were not applicable. Moreover, they produced 16 categories after extensive literature review. Later on, a set of experiments were conducted to examine the proposed heuristic validity. They found that the new set of heuristics can perform better but it needs more detailed explanations in terms of its features. Mazlan *et al*. (2012) criticized the usability practitioners for re-using the same heuristics without taking into consideration its suitability for different product domains or developing the traditional heuristics.

Therefore, it can be seen now clearly the need for improved heuristic that can be used for usability evaluation. The above studies showed different types of development of HE but it still needs improvement. However, the improvement process can occur in different ways either in its procedure (Sivaji *et al*., 2011) or in its components such as the heuristics themselves or the evaluators.

Evaluation of a system's usability can be achieved through applying different usability evaluation techniques such as usability testing, think aloud, usability guidelines and heuristic evaluation. Adopting more than one technique can prove a costly approach in terms of time needed, financial burden and resources required. However, the effectiveness of an individual technique application can also questioned due to a number of limitations for each of the usability evaluation techniques. Thus, improving the effectiveness of one of these techniques can be described as a very useful way. This can be achieved by developing the traditional heuristics by adding more heuristics and breaking each heuristics with more explanation. This approach has been recommended by a number of researches in different ways and criticized the traditional heuristics due to its broad level and inadequate details (Chen and Macredie, 2005; Muller *et al*., 1998; Alroobaea *et al*., 2013).

The developed heuristic evaluation is based on the traditional heuristic evaluation and participatory heuristic evaluation which includes more descriptions for each guideline. This was done through an intensive literature review to the related work (Chen and Macredie, 2005; Muller *et al*., 1998; Alroobaea *et al*., 2013; Masip *et al*., 2011; Sivaji *et al*., 2011). The developed guidelines adopt both process oriented and product oriented paradigm in order to cover the various aspects of the system and to support the process of human work as described in more detail in Table 1. The explanation for each heuristic can also maximize the opportunity of usability problems detection (Chen and Macredie, 2005). Therefore, this study aims to improve the effectiveness of usability evaluation method through improving the drawbacks of traditional heuristic evaluation.

Table 1: Developed heuristic evaluation

Heuristics and guideline descriptions: The system
System status
Informs users of their current location
Informs users of what can be done at the current location
Informs users to go if an action is taken
Match between system and the real world match
Speaks users' language
Follows real world conventions
Presents the information in a natural and logical order
User control and freedom
Allows users to perform , undo and redo tasks
Avoids unexpected actions
Consistency and standards
Has a consist set of words and actions
Standardizes its terminologies and actions (color, size, font)
Error prevention
Prevents from happening
Provides clear error messages
Guides users in case an error has occurred
Recognition rather than recall
Provides clear options and actions
Helps users not to memorize or remember
Externalizes the information through visualization
Flexibility and efficiency of use
Tailors actions based on users preferences
Can be used by different type of users (disabled, low experiences)
Has shortcut for frequent tasks and actions
Aesthetic and minimalist design
Shows only the needed information
Does not distract users through irrelevant actions
Help users recognize and recover from errors
Uses precise error messages
Indicates the errors occurred
Uses plain, polite and informative language
Help and documentation
Provides help whenever its needed
Provides easy to search features
Has embedded help in its contents
Emergency exit
Allows users to leave the unwanted state
Allows users to make their own decisions regarding the cost of exiting the current location
Users skills
Supports and enhances users skills
Extends and supplements users knowledge and background
Pleasure and respectful interaction with the user
Enhances users experiences
Treats users with respect
Reflects the users professional role and personal intention
Quality work
Supports the user in delivering quality work
Has accurate, timelines, aesthetic appeal and completeness attributes
Privacy
Protects personal information, Treats users information privacy

MATERIALS AND METHODS

To achieve the study objectives, 5 usability experts were recruited who have at least computer sciences degree and have previously conducted a usability evaluation. They were introduced to the proposed work and to the targeted system which includes the tactics to approach to the concerned person and to explain the study objectives for their welfare. They were provided the developed heuristic evaluation sheets in order to carry out the study in their preferred time. They were also asked to return the evaluation sheets back within a week. The following severity rating scales were used by each evaluator. The heuristic evaluation procedure followed was that recommended by Nielsen (1995) who stated that usually 3-5 evaluators are required to conduct heuristic evaluation. The questionnaire developed for personal interview and data recording contained the items shown in Table 2.

Table 2: Items in questionnaire developed for personal interviews

Rating	Description
0	Not a usability problem at all
1	Superficial problem
2	Minor problem
3	Major problem
4	Usability disaster

Each problem was rated by each evaluator based on the above scale. In case there is no agreement on the rating as expected, then the average of the evaluator's ratings was calculated as a common practice in usability evaluation as suggested by Shazeeye and Shanmugam (2007) and Allen *et al.* (2006).

RESULTS AND DISCUSSION

This section presented the study results with special reference to the performance of each evaluator and the type of problems discovered. It also illustrates the encountered problems based on heuristics. In addition, the usability testing was used as a benchmark for this study.

Evaluator's performance: Overall, 38 usability problems were identified by all the 5 evaluators. Out of these, 7 usability problems were described as usability disaster. While, another 7 usability problems were rated as major problems. The evaluators also reported that 10 and 14 usability problems were defined as minor and superficial problems. Table 3 shows the evaluator's performance.

It can be seen from Table 3 that evaluator No. 5 scored the lowest performance as the detection rate was 50%. This type of explained guidelines enabled the evaluators to identify usability problems. It was claimed by 2 evaluators that the guidelines and their subsequent explanation guided them to more usability issues. It also helped them to concentrate more on each aspect of the targeted interface. Although the range of the detection rate was not high (50-68%), yet it is acceptable for evaluator's performance in the traditional heuristic evaluation method where 4 evaluators were able only to reveal 20% of the found usability problems (Nielsen *et al*., 1998). The improved heuristic can be claimed if it maximizes the opportunity for the evaluators to explore more about the targeted system. The study results may suggest a solution to overcome one of the main criticisms of the traditional heuristic evaluation as its broad and general guidelines. Moreover, this can mislead the evaluators where subjective assessment occurs.

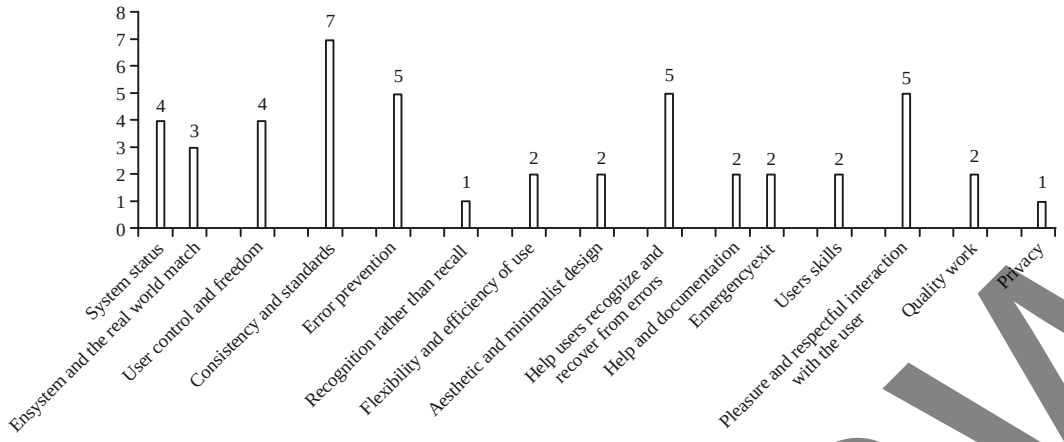

Fig. 1: Number of usability problems found based on heuristics

Table 3: "Evaluators'' performance based on the developed heuristics evaluation

	Evaluator				
Problem rating	1	2	3	4	5
Disaster	7	6	6	2	6
Major	7	3	3	1	1
Minor	9	9	8	7	1
Superficial	3	7	6	14	11
Total	26	25	23	24	19
Detection rate (%)	68	65	60	63	50

Problems found based on heuristics: Each usability problem identified belongs to one or more heuristics. In this study, 7 usability problems were linked to consistency and standards. These problems were noticed by majority of the evaluators. This could be justified as the targeted system was developed in-house. This might be the reason for large number of usability problems related to consistency and standards. For example: The users are required to agree on the terms and conditions each time they log in to the system. Although it does not need time as it is obviously a check box but it violates the design principles and consistency standards. Error prevention, help users and pleasure guidelines are linked with 5 usability problems for each of them. Although each usability heuristic is supposed to be unique yet there are a number of usability problems belong to more than heuristic such as error messages are not informative. For example, in this case, the system failed to prevent the error and also failed to provide help to users in terms of what to do next. Evaluators have clearly mentioned that users did not feel respectful if two or more such usability issues are encountered in a single task. The data entry fields in the targeted system did not have any constraints. This can be described as the main reason for these issues related to error prevention, help users and pleasure interaction heuristics. Figure 1 shows that how many usability problems were found by each heuristics. Table 4 shows interesting results as there were usability problems linked to more than heuristic. The evaluators declared that the description for each heuristic helped them to uncover more usability problems. They explained that the detailed description seemed to draw their attention to inspect the interface deeper and also further.

User testing: User testing is usually adopted to ensure that all the aspects of a system are assessed either through heuristic evaluation or usability testing. In this study, user testing was utilized as a bench mark to examine the effectiveness of the developed heuristic evaluation. As reported earlier in this study, 38 usability problems were detected but clearly it is really difficulty to know how many of these 38 usability problems represent the actual number of usability problems in the targeted system. Therefore, usability testing technique was utilized to offer more valid results. Out of the total users, five were recruited as a representative sample of the targeted users for the system. A total of 6 tasks representing the main system functions were assigned to 5 users to be performed. They were supervised by a usability expert to take all the notes and their comments. Quantitative usability data of the user testing was ignored as the main purpose of adopting this method was to use it as a bench mark for the developed heuristic evaluation. The usability problem number and type were the main objective of applying such technique. Also, 43 usability problems were detected and only 7 unique usability problems were reported. But they were neither usability disasters or major usability problems. All the usability disaster and major problems detected by the developed heuristic evaluation were also identified by the usability testing. This implies the effectiveness of both the usability testing and the developed heuristic evaluation methods. Although the study results can not be generalized but it indicated satisfactory results can be achieved from heuristic evaluation method.

Overall, the developed heuristic evaluation in this study found that 84% of the total usability problems were found by both methods. If this percentage is compared to a number of available studies that adopted only the traditional heuristic evaluation method according to Thyvalikakath *et al*. (2009) and Tana *et al*. (2008), then it can clearly be claimed that adopting participatory heuristic evaluation with guidelines explanations can seriously influence the effectiveness of

Table 4: Found usability problems based on participatory heuristics evaluation

Heuristics	Problems' descriptions
1	It does not tell the users his location and how to use the system as it's the first system to be used for its purpose
4-13	Users are expected to agree for terms and conditions each time they visit the main system homepage
4	Inappropriate use of color for log in and log out buttons (red and bold colour used)
4-7-13	Users who are not able to use the Arabic version, they have been given a different link to login to the English version
5-9	Error messages are not informative. They inform the users "you are not allowed to book an appointment with a certain doctor in a certain time" Users are expected to try again and work out the reasons
3	Leaving appointment page will cause disappearance of all the data have been entered by the user
5-9	Conflicting messages. If a user is booking an appointment, "there are no available appointments" message is always appeared
15	The system does not inform users all the data and information are private and secured as the system automatically shows family members details to the user who login
3-4	All the system news and announcements are designed poorly (using a flash technology with a new page). Once its clicked , user needs to close the new browser and go back to his current status
4	Sign out button is positioned in two places (at the top right and next to the user log in details)

heuristic evaluation. In this study, it's clear that the addition of more heuristics with their descriptions can benefit the traditional heuristic evaluation. The interesting finding in this study was that both the methods applied were able to reveal all the usability disaster and major usability problems. This can be also seen as an indication of the effectiveness of the developed heuristic evaluation.

CONCLUSION

This study has attempted to examine the effectiveness of heuristics evaluations method through developing the guidelines. The results of this study showed that the effectiveness of this method can be improved if the heuristics were developed and explained extensively to the evaluators. Also, in this study, the developed heuristics were able to offer interesting results in identifying disaster and major usability problems. This study results highlighted the potential for further investigations of a customized guidelines. The customization of the guidelines should be based on the specific-domain as it may offer better results as the traditional heuristics evaluation effectiveness is still questionable.

ACKNOWLEDGMENTS

The author is grateful to Deanship of Research, King Faisal University, Hofuf, Al-Ahsa, Kingdom of Saudi Arabia for financial assistance to carry this study.

REFERENCES

Allen, M., L.M. Currie, S. Bakken, V.L. Patel and J.J. Cimino, 2006. Heuristic evaluation of paper-based Web pages: A simplified inspection usability methodology. J. Biomed. Inform., 39: 412-423.

Alroobaea, R.S., A.H. Al-Badi and P.J. Mayhew, 2013. A framework for generating a domain specific inspection evaluation method: A comparative study on social networking websites. Proceedings of the Science and Information Conference, October 7-9, 2013, London, pp: 757-767.

Bevan, N., C. Barnum, G. Cockton, J. Nielsen, J. Spool and D. Wixon, 2003. The magic number 5: Is it enough for web testing? Proceedings of the CHI '03 Extended Abstracts on Human Factors in Computing Systems, April 5-10, 2003, Ft. Lauderdale, FL, USA., pp: 698-699.

Chen, S.Y. and R.D. Macredie, 2005. The assessment of usability of electronic shopping: A heuristic evaluation. Int. J. Inform. Manage., 25: 516-532.

Edwards, P.J., K.P. Moloney, J.A. Jacko and F. Sainfort, 2008. Evaluating usability of a commercial electronic health record: A case study. Int. J. Hum. Comput. Stud., 66: 718-728.

Geng, R. and J. Tian, 2015. Improving web navigation usability by comparing actual and anticipated usage. IEEE Trans. Hum. Mach. Syst., 45: 84-94.

Hvannberg, E.T., E.L.C. Law and M.K. Lerusdottir, 2007. Heuristic evaluation: Comparing ways of finding and reporting usability problems. Interacting Comput., 19: 225-240.

Isa, W.A.R.W.M., A.M. Lokman, E.S.A. Wahid and R. Sulaiman, 2014. Usability testing research framework: Case of handicraft web-based system. Proceedings of the 2nd International Conference on Information and Communication Technology, May 28-30, 2014, Bandung, pp: 199-204.

Lindgaard, G. and J. Chattratichart, 2007. Usability testing: What have we overlooked? Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 30-May 3, 2007, San Jose, CA., pp: 1415-1424.

Masip, L., T. Granollers and M. Oliva, 2011. A heuristic evaluation experiment to validate the new set of usability heuristics. Proceedings of the 8th International Conference Information Technology: New Generations, April 11-13, 2011, Las Vegas, NV., pp: 429-434.

Mazlan, M.F.B., A. Sivaji, S.S. Tzuaan and A.M. Lokman, 2012. Enhancing the Heuristic Evaluation (HE) by development and validation of a Collaborative Design Measurement System (CDMS): Collaborative design and measurement system for designers and testers. Proceedings of the IEEE Colloquium on Humanities, Science and Engineering (CHUSER), December 3-4, 2012, Kota Kinabalu, pp: 473-478.

Mueller, C.J., D. Tamir, O.V. Komogortsev and L. Feldman, 2009. An economical approach to usability testing. Proceedings of the 33rd Annual IEEE International Computer Software and Applications Conference, July 20-24, 2009, Seattle, WA., pp: 124-129.

Muller, M.J., L. Matheson, C. Page and R. Gallup, 1998. Methods and tools: Participatory heuristic evaluation. Interaction, 5: 13-18.

Nielsen, J., 1995. Severity ratings for usability problems. http://katsvision.com/canm606/session_2/M2_reading4.pdf

Nielsen, J., M. Hertzum and B. John, 1998. The evaluator effect in usability studies: Problem detection and severity judgements. Proceedings of the 42nd Annual Meeting Human Factors and Ergonomics Society, October 5-9, 1998, Chicago, Illinois, pp: 1336-1340.

Nielsen, J., 2000. Why you only need to test with 5 users. Nielsen Norman Group, March 19, 2000. http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Peute, L.W.P., M.M. van Engen-Verheul, E. Kilsdonk, N. Peek and M.W.M. Jaspers, 2011. Effectiveness of participatory heuristic evaluation: Case study with a guideline based health information system. Proceedings of the 5th International Symposium on Human Factors Engineering in Health Informatics, August 26-27, 2011, Trondheim, pp: 207-212.

Shazeeye, K. and R. Shanmugam, 2007. Heuristic Evaluation Quality Scores (HEQS): A measure of heuristic evaluation skills. J. Usability Stud., 2: 61-75.

Sivaji, A., A. Abdullah and A.G. Downe, 2011. Usability testing methodology: Effectiveness of heuristic evaluation in e-government website development. Proceedings of the 2011 5th Asia Modelling Symposium, May 24-26, 2011, Kuala Lumpur, pp: 68-72.

Tana, W.S., D. Liu and R. Bishu, 2008. Web evaluation: Heuristic evaluation vs. user testing. Int. J. Ind. Ergon., 39: 621-627.

Thyvalikakath, T.P., V. Monaco, H. Thambuganipalle and T. Schleyer, 2009. Comparative study of heuristic evaluation and usability testing methods. Adv. Inform. Technol. Commun. Health, 143: 323-327.