A Framework for Generating Domain-Specific Heuristics for Evaluating Online Educational Websites

Roobaea AlRoobaea,
College of Computers & Info. Systems,
Taif University, Saudi Arabia and
School of Computing Sciences,
University of East Anglia, Norwich, UK
Email: r.alrobaea@uea.ac.uk

Ali H. Al-Badi,
Information Systems Dept., CEPS, SQU, Oman
Email: aalbadi@squ.edu.om

Pam J. Mayhew
School of Computing Sciences, University of East Anglia, Norwich, UK
Email: p.mayhew@uea.ac.uk

Abstract

Websites are growing in use worldwide but need to be continuously evaluated and monitored to measure their efficiency, effectiveness and user satisfaction, and ultimately to improve quality. For this purpose, heuristic evaluation methodologies, such as Nielsen's Heuristics, have become the accepted means for the usability evaluation of user interface designs; however, they are general, and unlikely to encompass all usability attributes for all website domains. The aim of this paper is to enhance one of the most-used usability evaluation methods by generating specific heuristics for the educational domain, and then to compare and contrast them against Nielsen's ten heuristics (as first validation stage for proposed framework) in terms of the number and severity of problems found, and of a number of usability measurements. The result show that the proposed framework succeeded in building a new set of heuristics for online educational websites, which managed to discover uniquely 55 (69%) of the usability problems in all chosen websites (80 problems in total), in comparison with Nielsen's heuristics, which discovered only 6 (8%). 19 problems (24%) were commonly discovered (overlapping or sharing) by both sets of heuristics. The time taken using Nielsen's heuristics was less than the time taken using the newly developed Educational Heuristics but this is because Nielsen's heuristics do not cover all the issues related to educational websites. It appears that the framework for generating context-specific heuristics did in fact produce an efficacious set of Educational Heuristics that covered the issues well in this domain.

Keywords- *Heuristic evaluation (HE), Framework, Nielsen's heuristics (NH), Educational Heuristics (EH), Usability problem and Severity rating*

I. INTRODUCTION

Dynamic websites such as online educational websites are now growing in use around the world but this new learning approach must be continuously evaluated and monitored to measure their success, levels of efficiency, effectiveness and satisfaction, and ultimately to improve their quality. Web design is a key factor in determining the success of e-learning websites, and users should be the priority in the designers' eyes because usability problems in e-learning websites can have serious ramifications, over and above the users failing to meet their objective. We need an effective methodology and appropriate heuristics to perform this task.

Various evaluation methods, tools and techniques can be employed to achieve this. The most well-known evaluation method used to assess user-interfaces and to discover usability problems is heuristic evaluation. This method depends on a set of usability principles or 'heuristics'. In 1994, Nielsen developed a set of guidelines (heuristics) for measuring usability, and they are still being used. It is a well-known method commonly used to find usability problems in a user interface [1]. However, these heuristics have been criticized for being too general and lacking in guidance for evaluators [2]. Educational websites must be effective if they are to meet the student's needs. Consequently, there is a need for new heuristic guidelines specifically designed for websites in particular domains to overcome the shortcomings of Nielsen's original heuristics. This paper aims to develop a set of heuristics specifically for assessing the usability of online educational websites, based on a recent framework developed by researchers for generating domain-specific heuristics [3]. Then, it proceeds to test the efficiency of the newly developed Educational Heuristics (EH) experimentally by applying them to evaluate three well-known online educational websites. Finally, we compare the performance of EH against Nielsen's ten heuristics (as first validation stage for proposed framework) in terms of the number and severity of problems discovered and their efficiency, thoroughness, validity, effectiveness and cost of use.

This paper is organized in the following way. Section 2 starts with a brief literature review, including a definition of usability problems and the concept of severity rating, and details some related work. Also, it describes the heuristics construction framework. Section 3 uses the proposed framework. Then, the new heuristics (EH) are applied in practice in three case studies. Section 4 and 5 provides an analysis and discussion of the results. Section 6 presents the conclusion and suggestions for future work.

II. LITERATURE REVIEW

Distance learning websites are becoming increasingly developed in the midst of the Internet revolution and ever-improving information technologies. In fact, the growth of the Internet has led to an explosion of educational website content, rising in accordance with demand. [4] asserts that e-learning occurs when students use the Web to proceed through the sequence of teaching, completing the learning activities and achieving learning results and objectives. He states, "users in any place and time can have access to distributed learning materials and processes over the Internet. This could be part of a winning strategy for particular needs, such as decongestion of overcrowded education facilities, and support for students or teachers and adult education" [5]. However, some of these websites are difficult to use due to the inexperience of many of the designers and the lack of effective, efficient and accurate appropriate guidelines for performing this task. Consequently, users spend more time learning how to use the website than learning the educational content. [6] states, "the quality is considered a crucial issue for education in general, and for e-learning in particular". The need for e-learning websites to be of sufficiently high quality cannot be overstated, and so it is extremely important, indeed necessary, to classify suitable criteria for addressing and assessing quality [7].

Emanating from the development of Web 2.0, there is now a need to study the usability of online educational websites. The 'usability' of an interface is one of the most significant aspects affecting the quality of the users' experience and ultimately the success of these websites. Poor websites may have a negative impact on various aspects of educational institutions. They do not allow users to achieve their goals efficiently, effectively, and with a high degree of satisfaction [8]. Also, users spend more time learning how to use these websites rather than on the content, causing frustration leading to abandonment of the site. ISO defines usability as "the extent to which the product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" [8]. Muir et al. defined pedagogic usability as a branch of usability that "affects educational website design and development, particularly in the context of supported open and distance learning" [9]. Usability is part of the Human-Computer Interaction (HCI) field, which has become a major area of research; it utilizes many and various techniques for evaluating the usability of interactive systems and the phenomena surrounding them [10]. These techniques are classified into inspection, inquiry and testing methods. The most well-known methods used to assess user interfaces and discover usability problems are heuristic evaluation and usability testing.

Heuristic evaluation was developed by [1]; it is an inspection method that guides the evaluator through a set of general usability principles or 'heuristics'. It can be defined as a process that requires a specific number of experts to use the heuristics in order to find usability problems in an interface in a short time and with little effort. [11] stated, "heuristic evaluation is a widely accepted method for diagnosing potential usability problems and is popular in both academia and industry". It is superior to all other methods in carrying out evaluation in less time and by a smaller number of expert evaluators; it does not require any special tasks or equipment. Also, it can be used early in the development process, and may be used throughout the development process [1]. The result of applying these heuristics is a list of usability problems [12]. These problems are classified into different groups to which a numeric scale is applied in order to measure the severity of each problem or issue; firstly, this issue is not a usability problem at all; secondly, this is a cosmetic problem that does not need to be fixed unless extra time is available on the project; next, this issue is a minor usability problem, and fixing this should be given low priority; then, this is a major usability problem, and it is important to fix this, so it should be given high priority; and finally, this issue is a usability catastrophe, and it is imperative to fix this before the product is released.

In the early years of computing, Nielsen's heuristics (NH) were widely applied in measuring the usability of Web interfaces and systems because it was the only such tool available. Many researchers then sought to compare and contrast their efficiency with other methods such as user testing. Some researchers used one or both Nielsen's heuristics and user testing in order to comprehend the success and popularity of the former and to assess the usability of a chosen system such as social networks, video games, online shopping websites and e-learning applications. Lately, their findings have been almost unanimous on certain aspects, namely that these heuristics have been criticized as being too generic, were not particularly useful and did not find a sufficiently high number of usability problems; in most cases, the extended or modified heuristics worked better [2, 13, 14]. For this purpose, many researchers have enhanced Nielsen's generic heuristics by extending or modifying them in order to make them applicable to the features of other interactive systems [15, 16, 17, 18,

19]. However, some researchers even found that the tested websites failed in certain respects, according to the extended or modified heuristics [14, 20]. Consequently, there is a need for new heuristic guidelines specifically designed for websites in particular domains to overcome the shortcomings of Nielsen's original heuristics, extended heuristics and modified heuristics. Such specific heuristics should be more efficient and cheaper, and should greatly facilitate the assessment of website usability. This was confirmed in a study conducted by [21], who claimed that the performance of the new set of heuristics was as good as, if not better than, Nielsen's set because the new set was specific to the domain tested.

Sifting the literature, these researchers have found that there is no comprehensive methodology or framework for generating context-specific heuristics. This finding and the criticality of website usability were the reasons that motivated these researchers to propose a framework for generating domain-specific heuristics for website evaluation. This framework must be applicable across numerous domains. It should be a method characterized by being optimal, cheap and quick, and by using only experts, although it should include user experience (thereby involving the advantages of user testing but avoiding its drawbacks). Consequently, heuristic evaluation was the starting point for achieving these targets. The development phase of the proposed framework consists of four main stages. It involves information gathering, comprehending the users' perspectives and requirements (through mini-usability testing and learning from errors), learning from the expertise and experience of and experts in the field (through a focus group/discussion, and finally analysing all the data generated. The following is an explanation of these stages:

Development Stage One (D1: Familiarisation): the aim of this stage is to review all published material in the area of usability evaluation methods with a specific focus on heuristic evaluation, and to collect any information related to the target domain.

Development Stage Two (D2: User Input): this stage consists of a mini-usability test (task scenarios and questionnaire). Users are asked to perform a set of tasks on a typical domain website, and then asked to fill out a questionnaire. The broad aim of this is to elicit feedback on a typical system in order to appreciate the users' perspectives, requirements and expectations.

Development Stage Three (D3: Expert Input (focus group/interview)): this stage entails a discussion amongst experts (in the domain and/or usability) in order to obtain a broader understanding of the specifics of the prospective domain. It is widely employed in various types of research such as product planning and usability studies. This assists in understanding the issues arising from the questionnaire results and also in garnering more information through conversations with expert evaluators (in the domain and/or usability)

Development Stage Four (D4: Draw Up Heuristics (data analysis)): the aim of this stage is to analyse all the data gathered from the previous three stages and highlight the characteristics of the target domain. The collected data consist of general areas of usability related to the selected domain. Then, the researchers establish the new heuristics for the selected domain, and a list of usability problem areas/attirbutes is constructed from the gathered information drawn from the three stages above. After constructing the framework, it is validated intensively through rigorous validation methods as follows:

1. **Heuristic Validation (Expert Evaluation) :** the aim of this stage is the validation of the newly proposed heuristics by conducting a heuristic evaluation. Expert evaluators need a familiarization session before the

actual evaluation. This session is important because if the application domain is unfamiliar to the evaluators, the session provides an opportunity to become acquainted with the domain and with the evaluation environment for testing the proposed heuristics on the targeted websites [12]. The collected data are then analysed empirically.

2. **Testing Validation :** (User Evaluation): the aim of this stage is to complement the results obtained from the expert evaluation, by carrying out usability lab testing on the same websites. Nielsen (1992) recommends conducting usability testing with heuristics evaluation because each one is complementary to the other [22]. Then, the efficiency and effectiveness of both sets of heuristics are assessed and compared against the lab testing in terms of discovering problems and their types. The details of the user evaluation will be outlined in another paper to follow[23]. In this paper, we focus on validation through heuristic evaluation and the second validation process will be future work. After proposing the framework, it was decided to evaluate its practicality by applying it to a real-life experiment.

From the literature review, not many studies have been conducted on usability evaluation methods for educational websites [24]. Thus, it is apparent that online educational websites represent a hot topic that has not yet been fully explored, nor have any specific heuristics been generated for this domain. The researchers chose three well-known websites in this domain: Skoool, AcademicEarth and BBC KS3bitesize.

III. RESEARCH METHODOLOGY

1. Selection of Usability Experts

The selection of usability experts is a very important phase for the success of this experiment. For this study, the researchers have decided to recruit eight expert evaluators, divided into two groups of four. Two out of four in each group are general usability experts (single) and the others are usability specialists (double) because usability specialists are better at finding usability problems through heuristic evaluation [22]. Also, they have a sufficient number of years of experience in this field.

2. Evaluation of the Practicality of the Framework

In the first stage, the researchers conducted a review of the literature related to usability, Usability Evaluation Methods (UEM), improving heuristic evaluation, as well as the requirements for online educational websites. In stage two, small user testing session with four task scenarios and a brief questionnaire was conducted by five users, to learn from their errors and gain an appreciation of which elements or features they expected to be present in online educational websites and their more general expectations of these sites. In stage three, a focus group discussion session was conducted with the experts in usability and the domain (education domain). In stage four, the researchers analysed the results of the three stages and incorporated the findings. After that, the usability problem areas/attributes were identified which are the most important areas in the educational websites, and the most usability problems were existed in them. Also, they can facilitate the process of evaluation and analysis, and to help designers and programmers identify the areas in their website that need improvement, as shows in Table1. Then, the proposed heuristics were established. It is common for online educational websites to take into account what is called ‘user experience’. The new heuristics are shown in Table 1. The researchers call the final draft Educational Heurisitcs, or EH.

Usability problem area/attributes	Educational heuristics
User usability	Supports modification and progress of evaluation
	Supports user tasks and avoids difficult concepts
	Feedback and support services
	Easy to remember
Motivational factors	Support leaner curiosity
	Learning content design and attractive screen design
	Motivation to learn
Content information and Process orientation	Relevant, correct, and adequate information
	Reliability and validity
	Privacy and security
Learning process	Assessment
	Interactivity
	Evoke mental images of the learner
	Resources
	Learning management
	Learnability
Design and media usability	Multimedia representations
	Accessibility and compatibility of hardware devices
	Functionality
	Navigation and Visual Clarity

Table 1: Final version of Educational Heuristics (EH)

In Validation Stage One, a pilot experiment was conducted to make sure that there was no spelling or grammatical errors and no ambiguous words or phrases, and that all of sentences in the instruments (heuristics and check-lists) were sufficiently clear to be used by the evaluators. A fewer minor improvements were made and testing proper began. Eight expert evaluators were divided into two groups that were carefully balanced in terms of experience. In each group, there were two double expert evaluators and one single expert evaluator. Each evaluator was asked to conduct his/her evaluation separately in order to ensure independent and unbiased evaluations [12]. Each group used two kinds of heuristics: Nielsen's set of heuristics (NH) and the newly developed educational heuristic set (EH) to evaluate the three different websites. The evaluation was carried out in a prescribed sequence, i.e. one group used EH on Skoool and BBC KS3bitesize, then NH on AcademicEarth, while the second group used NH on Skoool and BBC KS3bitesize, then EH on AcademicEarth. The researchers adopted this technique to avoid any bias in the results and also to avoid the risk of any expert reproducing his/her results in the second session through over-familiarity with one set of heuristics, i.e. each evaluation was conducted with a fresh frame of mind.

The Heuristics Validation Phase started with a training (familiarization) session for the eight expert evaluators. Then, the actual expert evaluation was conducted, followed by a debriefing session. The researchers analysed the results (usability problems and severity rating), and any disagreements between evaluators with regard to usability problems and/or their severity ratings were discussed in the debriefing session to remove any 'false positive' problems or any 'evaloutor subjective' problems.

IV. DATA ANALYSIS AND MEASUREMENTS

The results of the usability problems obtained from the NH and EH were analysed as follows:

1. Comparing the result of usability problems and their severity: for NH and EH the expert evaluators used checklists that had been developed by the researchers to facilitate the evaluation process, and Nielsen's

classification as mentioned above from 0 to 4 for rating the severity usability problems. The researchers extracted the problems from this checklist and removed all 'not real' and 'subjective' problems during the debriefing session. The problems agreed upon were merged into a unique master problem list and any problems upon which the evaluators disagreed were removed.

2. Comparing the performance of each method in terms of unique and overlapping problems: two independent evaluators were involved in removing the duplicated problems and comparing the remaining problems in order to identify all the unique and overlapping problems found.
3. Comparing the performance of each method in terms of discovering usability problems in the five areas\attributes.
4. Comparing the satisfaction scores of NH and EH by using System Usability Scale (SUS).
5. Making further comparisons between the performance of NH and EH in identifying usability problems; a set of metrics were used for examining their performance, as follows.

a) **Efficiency :** one measure of efficiency in UEMs is the "ratio between the number of usability problems detected to the total time spent on the inspection process" [25]. Efficiency = (No. of problems)/(Average time spent) [25].

b) **Thoroughness:** this is perhaps the most attractive measure; it is defined as a measure indicating the proportion of real problems found when using a UEM to the total number of known real problems (major and catastrophic) [25]. It can be calculated by the following formula: Thoroughness = (No. of real usability problems found)/(Total no. of real usability problems) [27].

c) **Validity:** this is the extent to which a UEM accurately identifies usability problems. It can be calculated by the following formula: Validity = (No. of real usability problems found)/(No. of issues identified as a usability problem) [27].

d) **Effectiveness:** this can be defined as the functional ability of a UEM to identify usability problems related to the user interface [28]. It is calculated by the following formula:

Effectiveness = Thoroughness × Validity [27]

e) **Cost of Employ:** the cost estimates can be done fairly simply by following Nielsen's equation: Cost estimate = (Number of evaluation hours) X (Estimate of the loaded hourly cost of each participant) [12]. He estimated the hourly loaded cost for professional staff at $100.

These measures will be calculated for NH and EH, and then compared against each other. This will allow us to determine whether our framework has generated a quality set of heuristics. Also, these result used for further comparison against user testing in second validation stage, and outlined in other paper [23].

V. RESULTS AND DISCUSSION

1. **Time spent:** the average time taken for doing the three experiments using NH was 24.33 minutes, whereas for EH the average was 42.66 minutes. The group who used NH managed to evaluate the website more quickly than the other group but discovered fewer usability problems, whereas, the group that used EH spent slightly more time evaluating the website, but discovered many more usability problems. Explanations for the differences in time spent and number of problems located were gleaned from the evaluators' feedback. They

said that NH was not particularly helpful, understandable or memorable for them. However, EH helped them to develop their skills in discovering usability problems in this application area; also, this set was more understandable and memorable during their evaluations and covered almost all broad areas. To further analyse these factors of time spent and number of problems discovered, efficiency metrics were applied.

2. **Numbers and types of usability problem areas:** the total number of problems discovered was 99, out of which 25 were identified using NH and 74 using EH. The EH discoverd 25, 28, 13 and 8 cosmetic, minor, major and catastrophic problems, respectively. However, NH discovered 7, 10, 5 and 3 cosmetic, minor, major and catastrophic problems, respectively. All the duplicated problems were removed and compared by two independent evaluators, to identify the unique and overlapping problems. When problems from the two evaluation groups were consolidated, there were 19 duplicates; we identified a total of 80 problems. The total number of uniquely identified problems was 61. The heuristic evaluation using EH identified 55 problems (69% of the 80 problems) that were not identified by NH, and there were 6 problems (8% out of 80) identified by NH that were not identified by EH. 19 problems out of 80 were discovered to be 'overlapping' between two methods as depicted in Figure 1). Furthermore, In terms of the usability problem areas/attributes detailed above, EH helped the expert evaluators to identify large numbers of usability problems in all of them, whereas NH failed to exposing any usability problems in two main usability problems areas/attributes, which are 'Motivational factors' and 'Learning process'. Also, NH failed to identify a sufficient number of usability problems in the 'Content information and Process orientation' area. These failings were because Nielsen's heuristics are too general and do not cover all usability features for interactive systems [29].

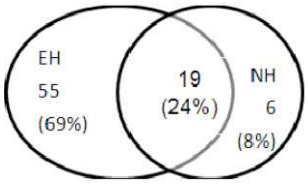

Figure 1: Overlap between Nielsen's heuristics and Educational Heuristics

3. **Efficiency, Thoroughness, Validity, Effectiveness, Satisfaction Score and Cost of Employ:** the above formulae were used and the results were that EH provided a more efficient result than NH through both groups in discovering usability problems (EH = 1.74 vs. NH = 1.03). This was because each group that used EH spent enough time (one hour, as estimated by Nielsen). This confirms what Nielsen stated in [12]: "the time spent on an evaluation can affect, to some extent, the number of usability problems found". Thoroughness for EH in identifying the number of real problems at Level 3 or 4 is higher than that of NH (0.1 vs. 0.04). There is a slight difference between the validity of both sets of heuristics in accurately identifying usability problems (EH = 0.12 vs. NH = 0.11). For effectiveness, EH (at 0.011) was slightly more effective in identifying usability problems than NH (0.004). In addition, these two methods differ significantly in terms of the number of identified usability problems and the average severity of these problems, and accordingly, where either thoroughness or validity is low, effectiveness will be low also. The researchers used the System Usability Scale (SUS) as previously mentioned. NH delivered a lower overall score (at 46), whereas EH delivered a much higher score (71). Finally, there was a slight difference in the cost of use in this research. This difference is due to the relative times consumed in the evaluations and analyses (NH=$3,228 vs. EH=$3,751).

VI. CONCLUSION AND FUTURE WORK

Heuristic evaluation is a widely accepted inspection method for diagnosing potential usability problems and is popular in both academia and industry. It can work more efficiently if appropriate heuristics are designed for the target domain. This research contributes to the advancement of knowledge in the field; its first contribution is the building of a framework for generating context-specific heuristics for the evaluation of websites. This framework was applied and validated in practice. The second contribution is the introduction of a new set of heuristics specific for educational websites. In order to consolidate and confirm the findings, future research could include: 1) Testing the proposed framework by developing new heuristics for different fields such as e-commerce and healthcare systems; 2) Complement the validation of the proposed framework by conducting user testing, and then compare and contrast the results obtained from user testing against Educational Heuristics and Nielsen's heuristics.

I. REFERENCES

1. Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces, Proc. *ACM HCI'90* (Seattle, WA, 1-5 April 1990), 249-256.
2. Chen, S. Y. and Macredie, R. D. (2005). The assessment of usability of electronic shopping: A heuristic evaluation, *International Journal of Information Management*, vol. 25 (6), pp. 516-532.
3. AlRoobaea, R. Al-Badi, A., Mayhew, P. (2012). Generating Domain-Specific Heuristics for Evaluating Social Network Websites, *MASAUM International Conference on Information Technology (MIC-IT'13).*
4. Abuzaid, R. (2010). Bridging the Gap between the E-Learning Environment and E-Resources: A case study in Saudi Arabia. Procedia-Social and Behavioral Sciences, 2(2): 1270-1275.
5. Ardito, C., Costabile, M., De Angeli, A. and Lanzilotti, R. (2006a). Systematic evaluation of e-learning systems: an experimental validation. In Proceedings of *The 4th Nordic Conference on Human-Computer Interaction: changing roles*, pp. 195-202. ACM.
6. Alkhattabi, M., Neagu, D. and Cullen, A. (2010). Information Quality Framework for E-Learning Systems. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, 2(4): 340-362.
7. Stracke, C. and Hildebrandt, B. (2007). Quality Development and Quality Standards in e-Learning: Adoption, Implementation and Adaptation. In Proceedings of *World Conference on Educational Multimedia, Hypermedia and Telecommunication 2007*, pp. 4158-4165.
8. ISO (1998), *ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability*.
9. Muir, A., Shield, L. and Kukulska-Hulme, A. (2003). The pyramid of usability: A framework for quality course websites. In Proceedings of *EDEN 12th Annual Conference of the European Distance Education Network, The Quality Dialogue: Integrating Quality Cultures in Flexible, Distance and e-Learning*, Rhodes, Greece, pp. 15-18.
10. Smith-Atakan, S. (2006), *Human-computer interaction*. Thomson Learning Emea
11. Magoulas, G. D., Chen, S. Y. and Papanikolaou, K. A. (2003). Integrating layered and heuristic evaluation for adaptive learning environments. In the proceeding of *UM2001*, 5-14.
12. Nielsen, J. (1994). Heuristic evaluation, *Usability Inspection Methods*, vol. 24, pp. 413.
13. Hart, J., Ridley, C., Taher, F., Sas, C. and Dix, A. (2008). Exploring the Facebook experience: a new approach to usability. In Proceedings of *The 5th Nordic Conference on Human-Computer Interaction: building bridges*, pages 471-474. ACM.

14. Thompson, A. and Kemp, E. (2009). Web 2.0: extending the framework for heuristic evaluation. In Proceedings of *The 10th International Conference NZ Chapter of the ACM's Special Interest Group on Human-Computer Interaction*, pp. 29-36. ACM.
15. Squires, D. and Preece, J. (1996). Usability and learning: evaluating the potential of educational software. *Computers & Education*, 27(1): 15-22.
16. Reeves, T. C., Benson, L., Elliott, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E. and Loh, S. (2002). Usability and instructional design heuristics for e-learning evaluation, In the proceeding of *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1615-162.
17. Alsumait, A. and Al-Osaimi, A. (2009). Usability heuristics evaluation for child e-learning applications. In Proceedings of *The 11th International Conference on Information Integration and Web-based Applications & Services*, pp. 425-430. ACM.
18. Pinelle, D., Wong, N. and Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design, In the proceeding of *The twenty-sixth annual SIGCHI conference on human factors in computing systems*, 1453-1462. New York, NY, USA,
19. Chattratichart, J. and Lindgaard, G., (2008), A comparative evaluation of heuristic-based usability inspection methods, In the proceeding of *CHI'08 extended abstracts on human factors in computing systems*, 2213-2220
20. Alrobai, A. AlRoobaea, R. Al-Badi, A., Mayhew, P. (2012). Investigating the usability of e-catalogue systems: modified heuristics vs. user testing, *Journal of Technology Research*.
21. Tan, W., Liu, D. and Bishu, R. (2009). Web evaluation: Heuristic evaluation vs. user testing, *International Journal of Industrial Ergonomics*, vol. 39 (4), pp. 621-627.
22. Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In Proceedings *ACM CHI'92 Conference* (Monterey, CA, May 3-7), pages 373-380. ACM.
23. AlRoobaea, R. Al-Badi, A., Mayhew, P. (2012). A Framework for Generating Domain-Specific Heuristics for Evaluating Online Educayional Websites- Further Validation, *2nd International conference on Human Computer Interaction Learning Technology (ICHCILT 2013).*
24. Bernerus, A. and Zhang, J. (2010). A Peek at the Position of Pedagogical Aspects in Usability Evaluation of E-learning System - A Literature Review of Usability Evaluation of E-learning System conducted since 2000. Report/*Department of Applied Information Technology* 2010: 085.
25. Fernandez, A., Insfran, E. and Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study, *Information and Software Technology*.
26. Liljegren, E. (2006). Usability in a medical technology context assessment of methods for usability evaluation of medical equipment, *International Journal of Industrial Ergonomics*, vol. 36 (4), pp. 345-352.
27. Sears, A. (1997). Heuristic walkthroughs: Finding the problems without the noise, *International Journal of Human-Computer Interaction*, vol. 9 (3), pp. 213-234.
28. Khajouei, R., Hasman, A. and Jaspers, M. (2011). Determination of the effectiveness of two methods for usability evaluation using a CPOE medication ordering system, *International Journal of Medical Informatics*, vol. 80 (5), pp. 341-350.
29. Masip, L., Granollers, T. and Oliva, M. (2011). A heuristic evaluation experiment to validate the new set of usability heuristics. In Proceedings of *The 2011 Eighth International Conference on Information Technology: New Generations*, pages 429-434. IEEE Computer Society.