Pertanika J. Sci. & Technol. 21 (1): 247 - 260 (2013)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Usability of Educational Computer Game (UsaECG): A Quantitative Approach

Hasiah Mohamed@Omar[1]*, Azizah Jaafar[2] and Rohana Yusoff[1]

[1]*Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, 23000 Dungun, Terengganu, Malaysia*

[2]*Institute of Visual Informatics, Universiti Kebangsaan Malaysia , 43600 Bangi, Selangor, Malaysia*

ABSTRACT

Heuristic Evaluation (HE) is used as a basis in developing a new technique to evaluate usability or educational computer games known as Playability Heuristic Evaluation for Educational Computer Game (PHEG). PHEG was developed to identify usability problems that accommodate five heuristics, namely, interface, educational elements, content, playability and multimedia. In HE process, usability problems are rated based on severity score and this is followed by presentation of a mean value. The mean value is used to determine the level of usability problems; however, in some cases, this value may not accurate because it will ignore the most critical problems found in a specific part. In developing PHEG, a new quantitative approach was proposed in analyzing usability problems data. Numbers of sub-heuristics for each heuristic involved were taken into account in calculating percentage for each heuristic. Functions to calculate critical problems were also introduced. Evaluation for one educational game that was still in development process was conducted and the results showed that most of the critical problems were found in educational elements and content heuristics (57.14%), while the least usability problems were found in playability heuristic. In particular, the mean value in this analysis can be used as an indicator in identifying critical problems for educational computer games.

Keywords: Usability of educational computer game, tool, interface, playability, multimedia

INTRODUCTION

In application development process, evaluation plays an important and integral part. This evaluation can be carried out either during the development (formative) or once the development is completed (summative). Expert evaluators may involve in any stage either formative or summative and real (potential) users are normally involved during summative evaluation. Meanwhile, selection

Article history:
Received: 31 March 2012
Accepted: 31 August 2012

E-mail addresses:
hasia980@tganu.uitm.edu.my (Hasiah Mohamed@Omar),
aj@ftsm.ukm.my (Azizah Jaafar),
rohanayu@tganu.uitm.edu.my (Rohana Yusoff)
*Corresponding Author

ISSN: 0128-7680

of the evaluators depends on the types of application, evaluation techniques and tools. One of the many popular applications nowadays is computer games. The popularity of computer games leads to the increment of computer game development that integrates educational elements in it. There are a lot of researches conducted on the integrations of computer games and educational elements in the perspective of impact, implication and effects (Yee Leng *et al*., 2010). The terms used for these applications are known as game based learning, computer games, educational computer games and digital game based learning (Kato, 2010; Papastergiou, 2009; Robertson & Howells, 2008).

The integration of fun to be played by the users and also the ability to contribute to the teaching and learning processes have become vital elements in any educational games development process. In order to merge these elements, comprehensive evaluation technique is needed during the development process. One of the evaluation techniques that is normally used by expert evaluators is Heuristic Evaluation (HE). In particular, HE is used by expert evaluators to examine the interface of any applications during interactive design process. The experts' involvement in the evaluation process is able to help developers to detect usability problem before the game can be released (Hasiah & Azizah, 2011b). The ability and the characteristics of HE have been used as a basis in developing specific heuristic technique to evaluate educational computer games (ECG). This technique is known as Playability Heuristic Evaluation for Educational Computer Game (PHEG), which consists of five heuristics, namely, interface (IN), educational element (ED), content (CN), playability (PL) and multimedia (MM) (Hasiah & Azizah, 2010). The experts who are involved in the evaluation process are from various backgrounds based on the heuristics provided, such as interface expert (for IN), educational technologies (for ED), subject matter experts (for CN), multimedia experts (for MM) and game developers (for PL).

The experts' involvement in the evaluation process shows a significant impact in identifying usability problems based on their knowledge and experiences. On the other hand, gathering experts in one place to conduct an evaluation is not an easy task since experts' work commitments need to be taken into consideration. In order to overcome this problem, an evaluation system known as AHP_HeGES was developed to assist the evaluation process. This online system can be used by the experts to conduct the evaluation and it is capable of handling the experts from various backgrounds at one time (Hasiah & Azizah, 2011b). A pilot study was conducted to test the system with the involvement of an expert for each heuristic. All the experts were able to accomplish the evaluation process accordingly, as well as to identify and list down usability problems based on the sub-heuristics involved and rate severity scale.

In analyzing the HE data or known as usability problems, severity rating plays important roles in helping developers to predict the level of usability problems in any application that is being evaluated. The experts normally identified all usability problems after inspecting the interface of the applications and rating the severity rating score based on the severity that was introduced by Nielsen (1995). Most of the papers have reported that the presentations of severity rating were calculated based on the mean of the severity rating (Nielsen, 1995; Pinelle *et al*., 2009; Ssemugabi & Villiers, 2007; Tan, Liu & Bishu, 2009).

The aim of this study was to propose potential quantitative analysis approach for usability of ECG based on the usability problems (results) presented in Hasiah and Azizah (2011b).

This paper reports the potential quantitative analysis for PHEG data in order to estimate the usability level of ECG. This research is important to facilitate game developers to get evaluators' feedback and usability problems of the ECG that is still in the development process.

HEURISTIC EVALUATION

Heuristics evaluation (HE) is a design guideline which serves as a useful evaluation tool for both product designers and usability professional (Nielsen & Molich, 1990). HE is an inspection evaluation technique that is normally used by an expert to find any usability problem in any product or system (Mureen *et al*.,2007; Nielsen & Molich, 1990). In particular, HE is commonly used for formative evaluation where the product or system is still in a development process. HE involves a small number of evaluators (expert in specific field) who have been assigned to inspect a system according to heuristics or guidelines that are relevant and focused on the interface of the system. HE is a light-weight process that can be cheap, fast, and easy to apply in an evaluation process (Nielsen, 1994). It can be used both in the design and evaluation phases of development and can even be applied to paper-based designs before the first working prototype is created (Nielsen & Molich, 1990). Studies on HCI have shown that using five evaluators may be enough to find most usability problems, while adding more will reduce the benefit to the cost ratio, and hence, suggesting that three may suffice (Nielsen & Molich, 1990). It can be used both in design and evaluation phases of development and can even can be applied to paper-based designs before the first working prototype is created. The HE technique has emerged from an evaluation of software (system and products) to one of the most popular applications nowadays, that is, games (Hasiah & Azizah, 2010).

In HE, there is a list of heuristics attributes that cover common criteria for any system that focuses on user's interface and interaction elements. These elements cover all the perspective of the system in general but in terms of educational computer game (ECG), to the best of our knowledge, there are no specific heuristics which accommodate all the elements in ECG, such as educational design and contents. Therefore, another set of heuristics that focuses on ECG is required. The argument for the requirement is the usability in ECG should deal with several elements of education if they are to be applied in teaching and learning officially. Hence, the elements of education, such as content and educational design, should be taken into consideration in the evaluation. A specific evaluation technique that is dealing with all the important criteria of educational computer games is known as Heuristic Evaluation for Educational Computer Game (PHEG) (Hasiah & Azizah, 2011a). In particular, PHEG accommodates five heuristics in evaluating the usability of educational computer games (UsaECG), interface, educational element, content, playability and multimedia.

EVALUATION PROCESS IN HE

Heuristic evaluation is a discount usability engineering method for quick, cheap, and easy evaluation of a user interface design (Kirmani, 2008). The goal of the evaluation process in HE is to find the usability problems in the design so that they can be attended to as parts of an iterative design process (Kirmani, 2008). The HE, developed by Nielsen and Molich in 1990 (Nielsen & Molich, 1990), is a technique used to evaluate the usability, with the inspection

being carried out mainly by evaluators, normally referred to as expert evaluators. The studies by Nielsen (1995) have shown that a number between 3 and 5 evaluators is enough. "It detects approximately 42% of serious design problems and 32% of minor problems, depending on the number of evaluators who reviewed the site" (Nielsen, 1995). In order to make a HE efficient and to provide quality results, the phases below should be taken into consideration:

1. Prior training: The evaluator must become familiar with the interface for a few minutes to learn the website and to be able to carry out the HE agilely.
2. Evaluations: The evaluator follows the set of heuristics to find deficiencies or to catalogue the website as usable. He can write comments.
3. Rate severity: The evaluator should determine the severity of each of the problems encountered. It is therefore appropriate that the priority of the problems is rated. He suggests three parameters: Frequency of problems occurs (are the users affected by the occurrence of the problem, and persistence of the problem). Is it a one-time problem and can the users overcome once they know about it or will the users repeatedly be bothered by the problem? (Nielsen & Molich, 1990). The problems in each parameter can score on a scale of 0 (not a usability problem) to 4 (catastrophe: it is obligatory to fix it) (Nielsen & Molich, 1990).
4. Review: In order to analyze each of the evaluations made to present a report with all the problems and possible resolutions by taking into account the qualitative analysis obtained (Nielsen & Molich, 1990).

Presentation of the usability problems have been made more meaningful based on rating scale (Nielsen, 1995). Rating scale is used to rate the severity of usability problems that have been identified during the evaluation process. The rating scale are based on the number with specific meaning, namely; 4 (Usability catastrophe: imperative to fix this before product can be released), 3 (Major usability problem: important to fix, so should be given high priority), 2 (Minor usability problem: fixing this should be given low priority), 1 (Cosmetic problem only: need not be fixed unless extra time is available on project) and 0 (I don't agree that this is a usability problem at all) (Nielsen, 1995).

CALCULATION OF THE USABILITY PROBLEM FOUND IN HE

A comparative evaluation study that investigated the extent to which HE identifies usability problems in a web-based learning application and compares the results with those of the survey evaluations among end-users (learners) was conducted (Ssemugabi & Villiers, 2007). Four evaluators (experts) in different expertise (user interface design, instructional/educational design and teaching) were invited and they agreed to participate in the evaluation process. There were a total of 58 problems identified by the experts and the severity rating was used to categorize the problems found. All the usability problem found were then divided into 2 categories, namely, major problem for severity scores 4 and 5 and minor for severity scores 1 and 2. Data were presented based on the numbers of problem founds and the mean of severity rating.

Network Game Heuristic (NGH) was developed by Pinelle *et al*. (2009), who later conducted an evaluation process with ten participations who had previous experiences with

usability evaluation and experience with multiplayer networked games (double experts). The participants evaluated and rated the usability problem using the Nielsen's severity rating (Nielsen, 1995). The results are shown in a form of table that summarizes the heuristic, total problem and mean severity. The evaluation results showed that the new heuristics were effective at specializing the evaluation process for multiplayer network games. The argument here was the way the results are being presented in term of severity rating. By having the mean value for the severity rating, it is debatable that it will not show the real level of the usability problems in the network game that is being evaluated.

Web evaluation using Heuristic Evaluation and user testing was also conducted (Tan *et al*., 2009). For HE, nine expert evaluators were recruited and defined as who had graduated level coursework in human computer interaction and in human factors of web design. The evaluators independently examined the interfaces and judged their compliance with a set of heuristics. After the evaluation process had been completed, all the findings (usability problems) were compiled and reported. The problems were classified on the basis of severity. A set of severity criteria was established to rate the severity of the problems. The three different severity ratings included were severe, medium and mild problems. Severe problems included catastrophic usability problems, where users were unable to do their work and major problems where users had difficulty, but were still able to find workarounds. Hence, fixing the problems is a mandatory. Medium problems included medium usability problems where users stumbled over the problem, but could quickly adapt to it. Fixing the problems should be given medium priority. Mild problems included minor usability problems, where users could easily work around the problem. Fixing the problems should be given low priority. Meanwhile, usability problem found was presented based on the numbers and the type of severity (namely, severe, medium and mild).

Quantitative analysis for a heuristic evaluation was adapted from González *et al*. (2009). To the best of our knowledge, this is one of the few studies that has attempted to quantitatively analyze heuristic evaluation data (usability problems). Research conducted by González *et al*. (2009) was based on UsabAIPO project that had initiated new experiment to obtain quantitative result after a heuristic evaluation process. Function of UsabAIPO was introduced and this gave the estimation of the degree or the level of usability of the website. The number of heuristics and sub-heuristics involved play an important role in developing the functions and overall calculation. The results are presented in the form of percentage of the overall usability. Usability level can be considered as good when its value is higher than 80% and 100%. This also means that all the sub-heuristics are satisfied or fulfilled (González *et al*., 2009).

PLAYABILITY HEURISTIC EVALUATION FOR EDUCATIONAL GAMES (PHEG)

Heuristic Evaluation (HE) is used as a basis in developing specific heuristic technique to evaluate educational computer games known as Playability Heuristic Evaluation for Educational Computer Game (PHEG) consisting of five heuristics, namely, interface (IN), educational element (ED), content (CN), playability (PL) and multimedia (MM) (Hasiah & Azizah, 2011a). The experts who are involved in the evaluation process came from various backgrounds,

and in this case, interface expert (for IN), educational technologies (for ED), subject matter experts (for CN), multimedia experts (for MM) and game developers (for PL). The criteria for usability evaluation of educational computer games (UsaECG) consist of 5 heuristics and 37 sub-heuristics. Table 1 below shows the PHEG.

TABLE 1: Playability Heuristic for Educational Computer Games (PHEG)
(Source: Hasiah & Azizah, 2011a)

Heuristic and Sub heuristics	
Interface (IN)	
IN1	Visibility of system status.
IN2	Match between system and the real world.
IN3	User control and freedom.
IN4	Consistency and standards.
IN5	Error prevention.
IN6	Recognition rather than recall.
IN7	Flexibility and efficiency of use.
IN8	Aesthetic and minimalist design.
IN9	Help users recognize, diagnose, and recover from errors.
IN10	Help and documentation.
Educational Element (ED)	
ED1	Clear learning objectives.
ED2	Suitable for learning process.
ED3	Functions as self directed learning tools.
ED4	Considers the individual learning level differences.
ED5	Provide feedback about the knowledge being constructed.
ED6	Offers the ability to select the level of difficulty in games.
Content (CN)	
CN1	Reliable and proven content with correct syllabus flow.
CN2	Clear structure of content.
CN3	Screen navigation is precise.
CN4	Supporting learning materials is relevant.
CN5	Content materials are engaging.
CN6	The content is chunk based on topic and subtopic.
Playability (PL)	
PL1	Provide enough information to get started to play.
PL2	Control keys follow standard conventions.
PL3	Users should always be able to identify their score in the game.
PL4	Users able to save games in different states.
PL5	Successful users in completing all the activities in a module are rewarded.
PL6	Challenges provided are positive game experiences.
PL7	The game is enjoyable to replay.

TABLE 1: (Continued)

Heuristic and Sub heuristics	
Multimedia (MM)	
MM1	Each multimedia element used serves a clear purpose.
MM2	Usage of multimedia elements is suitable with the content.
MM3	Combinations of multimedia elements are adequate.
MM4	The presentation of multimedia elements is well managed.
MM5	Numbers of multimedia elements for each screen is not more than 2 elements.
MM6	The use of multimedia elements support meaningfully the information provided.
MM7	The quality of multimedia elements used is good.
MM8	The use of multimedia elements enhances the content presentation.

RESEARCH METHODOLOGY

This study was carried out based on the previous set of heuristic and sub-heuristics that were developed to evaluate UsaECG (Hasiah & Azizah, 2011a) known as Playability Heuristic for Educational Computer Games (PHEG). PHEG consists of 5 heuristics such as (Interface (IN), Educational Element (ED), Content (CN), Playability (PL) and Multimedia (MM)). Heuristic for Interface consists of 10 sub-heuristics, and this is followed by heuristic for Educational Element (6 sub-heuristics), heuristic for Content (6 sub-heuristic), heuristic for Playability (7 sub-heuristics) and heuristic for Multimedia (8 sub-heuristics). Each heuristic was weighted according to its sub-heuristics, represented in the form of percentage of the sub-heuristic weight corresponding to the heuristic, as shown in Table 2.

Table 2 shows the heuristics, the number of sub-heuristics for each heuristic and its weighting used to calculate the total percentage of the educational computer games usability. The formula, called UsaECG(x), is as follows:

$$\text{UsaECG}(x) = ((\text{IN}/0.2073) + (\text{ED}/0.1622) + (\text{CN}/0.1622) + (\text{PL}/0.1892) + (\text{MM}/0.2162))/5 \quad (1)$$

where, IN represents the score of Interface heuristic, ED is Educational Element heuristic, CN is Content heuristic, PL is Playability heuristic and MM is Multimedia heuristic. UsaECG refers to as the weighted mean and it can refer to an indicator for the overall usability of ECG. Each of these variables (IN, ED, CN, PL and MM) obtains the corresponding value when applying the next formula:

$$F(x) = (\sum H / \sum Ht) \times P \quad (2)$$

where, $\sum H$ is the summation of the severity scores for each sub-heuristic group, P is the percentage for the current group, and $\sum Ht$ represents the summation of the sub-heuristic group in the worst case (in the event that all severity ratings are 4).

TABLE 2: Percentage assigned for each sub-heuristic in UsaECG

Heuristic	Total Sub heuristic	Weight each sub heuristic	Weight each sub heuristic (%)
Interface (IN)	10	0.2703	27.03
Educational Element (ED)	6	0.1622	16.2
Content (CN)	6	0.1622	16.2
Playability (PL)	7	0.1892	18.92
Multimedia (MM)	8	0.2162	21.62
Total	37	1	100

The usability of educational computer games [usaECG(x)] gives an estimation of the degree or the level of critical usability problem found, namely, the value of the overall critical usability problem found of the educational computer games. In Table 2, function (1) and function (2) are derived from González *et al*. (2009), which we found as one of the promising attempt to quantitatively analyze the results of a usability evaluation based on the HE method. Function (2) was then modified in order to simplify the calculation.

EVALUATION PROCESS

There are six steps involved in the evaluation process. First, the experts were identified and contacted through email. The experts who agreed to join the evaluation process replied and they were provided with the URL of the evaluation system (AHP_HeGES). Then, the experts performed the evaluation based on the provided PHEG with the explanation on how to conduct the process. The experts identified usability problem and rated severity scale. Once the evaluation was completed, the admin was able to view the usability problems found (data) and then performed the analysis. Fig.1 below shows the flow of the said evaluation process.

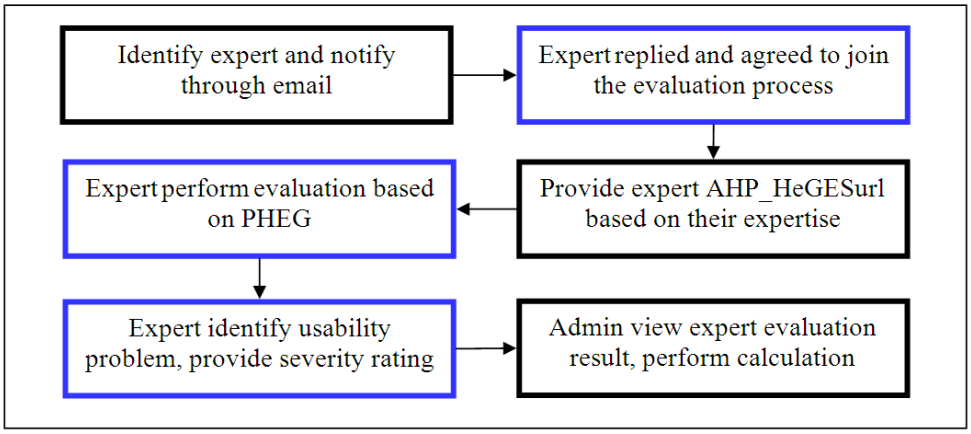

Fig.1: Evaluation flow (Source: Hasiah & Azizah, 2011b)

There were five expert evaluators involved in the process which consisted of a HCI expert, an educational element expert, a content expert, a playability expert (game developer) and a multimedia expert. The educational game used for the evaluation process was DatabaseFun game. Table 3 shows the profile of the experts involved and Table 4 presents the number of usability problems found by all the experts.

TABLE 3: Profile of Expert Evaluators

Expert Evaluators	Highest Qualification	Professional Role	Duties/course taught (relevant to this study)
IN	MSc	Senior lecturer in Science System	Experience in teaching HCI for the past 5 years
ED	MSc	Senior lecturer in Education faculty	Teaches Instructional Design and Technology
CN	MSc	Senior lecturer in System Science	Teaches Database for 3 years
PL	Bsc	Game developer	Has been involved in developing game for the past 3 years
MM	MSc	Senior lecturer in IT	Teaches multimedia for 4 years

TABLE 4: Usability Problems and Rating Found by Experts (Source: Hasiah & Azizah, 2011b)

Expert	Usability Problems	Severity Rating				
		4	3	2	1	0
IN	10	2	4	3	1	0
ED	8	3	2	1	1	1
CN	9	1	3	3	2	0
PL	8	3	2	2	0	1
MM	9	3	2	2	2	0

Extracting Result Analysis

Based on the results of usability problem found in Table 3, the calculation for ∑H, P and ∑Ht was conducted. Table 4 shows the calculation of ∑H, P and ∑Ht based on the usability problem found for each heuristic. ∑H is the summation of the severity scores for each sub-heuristic group, P is the percentage for the current group, and ∑Ht represents the summation of the whole group in the worst case (in the event that all ratings were 4). The usability problem found for the calculation purposes focused on the criticality of usability problem found, i.e. where severity score was 4. An example of the calculation on how the value of ∑H for expert IN is as follows:

$$\sum H = (4*2) + (3*4) + (2*3) + (1*1)$$
$$= 8 + 12 + 6 + 1$$
$$= 27$$

The example of the calculation on how the value of ∑Ht for expert IN is as follows:

$$\sum Ht = 4*2$$
$$= 8$$

The results in Table 4 represent the value for each heuristics; for example, F(IN) = (8/27) * 27.03 is F(IN) = 8.0089. The value of 29.63% represents the critical usability problem found in the interface for the ECG evaluated. In term of usability level, it was about 70%, which meant that it still could not be considered good, as mentioned by González *et al*. (2009), whereby the usability level could be considered as good when its value higher is than 80% and 100% or when all the sub-heuristics are fulfilled.

TABLE 5: Calculation for Each Heuristic

Expert	Usability Problems Found	Example of Calculation				
		∑H	∑Ht	P	F(x)	F(x)%
IN	10	27	8	27.03	8.0089	29.63
ED	8	21	12	16.2	9.2571	57.14
CN	9	21	12	16.2	9.2571	57.14
PL	8	22	4	18.92	3.4400	18.18
MM	9	24	12	21.62	10.8100	50.00
Mean (indicator)						42.42

Evaluation Results

The function for UsaECG (x) was used to calculate the overall critical usability problem found in the ECG evaluated.

$$UsaECG(x) = ((IN/0.2073) + (ED/0.1622) + (CN/0.1622) + (PL/0.1892) + (MM/0.2162)) / 5$$

Based on the function, UsaECG(x), the critical problem for each heuristic was calculated and presented in the form of a bar graph, as shown in Fig.2. Here, UsaECG(x) refers to weighted mean and it is an indicator for the overall usability of ECG.

$$UsaECG(x) = ((8.0089/0.2073) + (9.2571/0.1622) + (9.2571/0.1622) + 3.4400/0.1892) + (10.8100/0.2162)) / 5$$
$$UsaECG(x) = (29.63 + 54.14 + 57.14 + 18.18 + 50.00) / 5$$
$$UsaECG(x) = 42.42$$

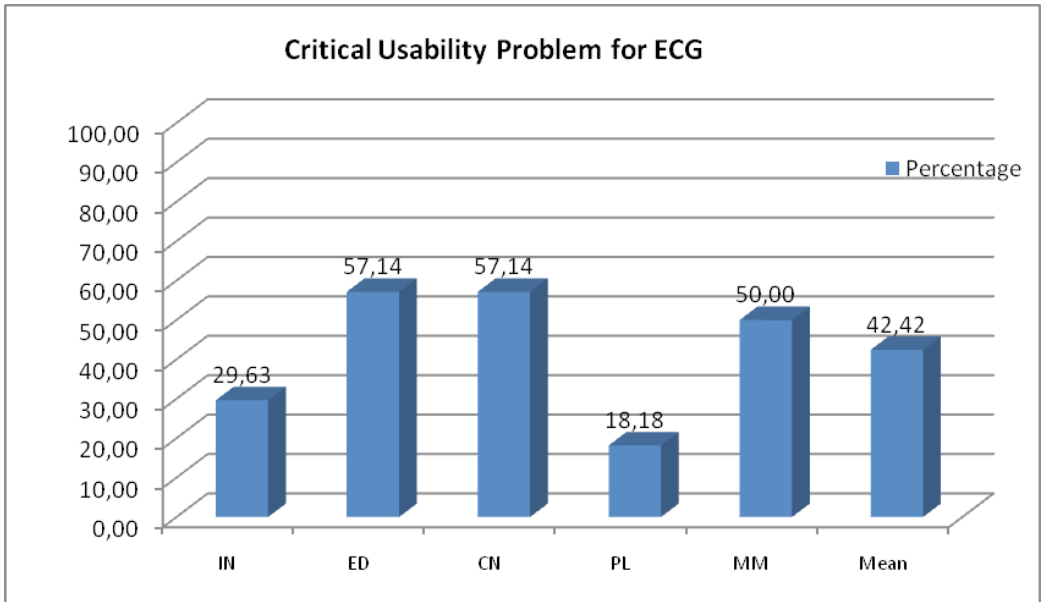

Fig.2: The Percentage of Critical Usability Problem for ECG

Fig.2 shows the percentage of the critical problem found for each heuristic. The most critical problem was found in Educational Element and Content heuristics, whereby there was 57.14% critical problems found for each heuristic. The least critical problem found was in Playability heuristic, whereby the value was 18.18%, indicating that in term of usability level for Playability, it is considered as good (González *et al*., 2009) because the value is 81.82%. Even though the usability level for Playability can be considered as good, there is a need to look in details for each usability problem found in order to help developers to improve ECG.

The mean value (42.42%) for the overall critical problem found in ECG represents an indicator for the overall usability level of ECG. Meanwhile, the overall result shown in Fig.3 should be able to help developers to grasp the idea on the part that needs to be improved accordingly. The Educational Element and Content heuristics need to be given priority in term of improvement, followed by Multimedia, Interface and Playability.

DISCUSSION

Usability problems found in most of the heuristic evaluation processes presented are based on severity rating. Several studies have presented severity rating based on the mean of severity scale, i.e. the value represents the score level of usability problems. In some cases, the mean value may not be accurate to represent the usability problems found. This is because the mean value will ignore the most critical problems found in specific part. One of the possible solutions to overcome this problem is by analyzing the numbers of severity rating for the critical problems found. The quantitative analysis for HE (González *et al*., 2009) was adapted as a basis of this analysis. Some modifications were done in order to simplify the analysis process. The function for UsaECG(x) and F(x) was used to help in the analysis of the data.

The number of the sub-heuristics for each heuristic was used in order to calculate the percentage for each heuristic. Based on the percentage presented, further calculation and function were developed. The values (%) for each critical problem found were extracted from the developed functions. This value should be able to help developers to simply get the results of the evaluation for the ECG that is still in the development process. The presentation of the results in the form of percentage is one of the distinctive results from this process. This will help game developers to shorten the process of analyzing the usability problem found.

CONCLUSION

Evaluating educational computer games using Playability Heuristic Evaluation for Educational Computer Games (PHEG) technique shows the ability of the new technique in identifying usability problems. Meanwhile, the involvement of expert evaluators demonstrates the numbers of usability problems found and the score of the usability problems. The new approach in analyzing critical usability problems found was introduced in order to fabricate more presentable results. The functions to calculate critical problems, based on the number of sub-heuristics, were created. The presentation of the critical usability problems should be able to help ECG developers to easily grab the most critical part that needs to be improved. This particular approach contributes to the body of knowledge in quantitatively analysing usability problems. The outcome of the analysis can be used to indicate the overall critical usability problems of educational computer games. Thus, a future study is suggested to be carried out on the evaluation processes involving expert evaluators (3 to 5) for each heuristic to detect more usability problems. The analysis process needs to take into consideration the number of problems found in term of its uniqueness. Based on the uniqueness of the problems, the overall critical problems can be calculated accurately.

REFERENCES

Allen, Mureen, Currie, Leanne M., Bakken, Suzanne, Patel, Vimla L., & Cimino, J. J. (2006). Heuristic evaluation of paper-based Web pages: A simplified inspection usability methodology. *Journal of Biomedical Informatics, 39*(4), 412-423.

González, M., Masip, L., Granollers, A., & Marta, O. (2009). Quantitative analysis in a heuristic evaluation experiment. *Advances in Engineering Software, 40*(12), 1271-1278.

Hasiah M.O. and Azizah J. (2010). *Challenges in the evaluation of educational computer games.* Paper presented at the International Symposium on Information Technology (ITSim), Kuala Lumpur, Malaysia.

Hasiah M. O., & Azizah J. (2011a). *Methodology to evaluate interface of educational computer game* Paper presented at the Pattern Analysis and Intelligent Robotics (ICPAIR), Kuala Lumpur, Malaysia.

Hasiah M. O., & Azizah J. (2011b). *Tools to Evaluate Usability of Educational Computer Game (UsaECG).* Paper presented at the 2nd International Conference on User Science and Engineering (i-USEr 2011), Kuala Lumpur, Malaysia.

Hvannberg, E. T., Law, E. L. -C., & Lárusdóttir, M. K. (2007). Heuristic evaluation: Comparing ways of finding and reporting usability problems. *Interacting with Computers, 19*(2), 225-240.

Kato, P. M. (2010). Video games in health care: Closing the gap. *Review of General Psychology, 14*(2), 113-121.

Kirmani, S. (2008). Heuristic Evaluation Quality Score (HEQS): Defining Heuristic Expertise. *Journal of Usability Studies, 4*(1), 49-59.

Nielsen, J. (1995). Severity ratings for usability problems Retrieved 15 February, 2012 from http://www.useit.com/papers/heuristic/severityrating.html.

Nielsen, J., & Molich, R. (1990). *Heuristic evaluation of user interfaces*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people, Seattle, Washington, United States.

Papastergiou, M. (2009). Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education, 52*(1), 1-12.

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009). *Usability heuristics for networked multiplayer games*. Paper presented at the Proceedings of the ACM 2009 international conference on Supporting group work, Sanibel Island, Florida, USA.

Robertson, J., & Howells, C. (2008). Computer game design: Opportunities for successful learning. *Computers & Education, 50*(2), 559-578.

Ssemugabi, S., & Villiers, R. D. (2007). *A comparative study of two usability evaluation methods using a web-based e-learning application*. Paper presented at the Annual Research Conference of the South African institute of Computer Scientists and Information Technologists on IT Research in Developing Countries (SAICSIT), Port Elizabeth, South Africa.

Tan, W. -S, Liu, D., & Bishu, R. (2009). Web evaluation: Heuristic evaluation vs. user testing. *International Journal of Industrial Ergonomics, 39*(4), 621-627.

Yee L. E., Zah bte Wan Ali, Wan bt. Mahmud, Rosnaini, & Baki Roselan. (2010). Computer games development experience and appreciative learning approach for creative process enhancement. *Computers & Education, 55*(3), 1131-1144.