Tan, S. C. et al. (Eds.) (2013). Workshop Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Development and Evaluation of a Problem Solving Oriented Game-Based Learning System

Hsin-Yi LIANG[a], Song-Yu MEI[a], Yu-Syuan WANG[b], Jhih-Liang JIANG[b], Gwo-Haur HWANG[c*] & Chen-Yu LEE[c]

[a] *Graduate Institute of Network Learning Technology, National Central University, Taiwan*
[b] *Department of Information Technology, Ling Tung University, Taiwan*
[c] *Department of Information Networking and System Administration, Ling Tung University, Taiwan*
*ghhwang@mail.ltu.edu.tw

Abstract: Problem solving is an intellectual skill to achieve effective learning, and it can be widely applied to many domains. In order to improve the problem solving abilities, previous studies had shown the significant effects of game-based learning to promote students' learning. Besides, previous researches also suggested the cognitive style plays an essential role to affect the usability of game-based learning, which significantly influences the learning effectiveness. Therefore, the differences of cognitive styles on usability evaluation are considered in this study. Two quests are designed in our game-based learning system to improve students' problem solving abilities. The first quest is helpful to promote the mathematical logic and reasoning abilities, while the second quest is helpful to promote the verbal logical reasoning ability. 49 students from two universities in Taiwan participate in this experiment. According to the analysis of cognitive style questionnaires, there are 9 serialist participants and 28 holist participants for the valid samples. In order to improve the system design, the Nielsen's heuristic evaluation questionnaires are applied. The major result indicates that the Nielsen's eighth heuristic (aesthetic and minimalist design) is most satisfied by the participants, while the Nielsen's sixth heuristic (recognition rather than recall) is most dissatisfied. Even some differences are observed, there are no significant differences of the usability evaluation between serialist participants and holist participants.

Keywords: System development, system evaluation, problem solving, game-based learning

1. Introduction

Problem solving is an intellectual skill to achieve effective learning (Liu, Cheng, & Huang, 2011). Previous research modeled there are four stages of problem solving, that is, to identify and understand the problem, to organize the proper strategies, to carry out the plan, and to look back the problem solving process (Polya, 2008). For the reason, problem solving requires multiple abilities, such as analysis of related information, organization, creative thinking, and critical evaluations (Publishing, 2009). Besides, problem solving abilities can be widely applied to many domains, for example, computer sciences, mathematics, social sciences, and design (Hwang, Wu, & Chen, 2012). Previous research also indicated problem solving abilities are positively related to the learning performance and high-level thinking abilities (Hwang et al., 2012). Hence, there address the need to improve students' problem solving abilities.

To improve the problem solving abilities, there is a need to provide students enough opportunities and practices through related problem solving activities (Polya, 2008). However, many studies argued the effectiveness of traditional learning (Lee & Chen, 2009). More specifically, in a traditional learning context, students learn to passively copy the standard solution methods step by step. Therefore, students easily forget the procedures due to lack of self-awareness and feedback (Polya, 2008). Moreover, the insufficient thinking and reasoning process also limited the improvement of problem solving abilities (Lee & Chen, 2009). As a result, students have difficulties to apply proper strategies in solving problems in novel situations.

For the reason, there is a need to encourage students actively reflect on learning and provide sufficient feedback to support learning. On the other hand, previous studies have shown the significant effects of game-based learning to promote students' learning (Gee, 2003; Pahl & Rowsell, 2012). The interactive environment not only improves the playfulness of learning, but also gives students immediate feedback to reflect on their learning (Tao, Yeh, & Hung, 2012). Moreover, game-based learning provides students a flexible learning environment, which promotes students developing various strategies to solve problems (Elia, van den Heuvel-Panhuizen, & Kolovou, 2009). In other words, game-based learning can significantly improve students' performance of problem solving through non-routine problem solving activities (Lee & Chen, 2009). Therefore, this study developed a game-based learning system to promote students' problem solving abilities.

Despite of such advantages of game-based learning to improve the problem solving abilities, there are some design issues which affect the learning effectiveness, such as disoriented problems (Webster & Ahuja, 2006), difficulties of manipulation (Kiili, 2005), and influences of multimedia (Hastings & Tracey, 2004). Besides, previous research also suggested cognitive style significantly affects student's preferences and behaviors, which plays an essential role to affect learning effectiveness in game-based learning (Ford, 1985; Frias-Martinez, Chen, Macredie, & Liu, 2007). Therefore, there is a need to take cognitive style into consideration to improve the learning effectiveness.

Regarding the system design, previous studies have shown usability is a strong predictor of such design issues (Schell, 2008). Indeed, the usability highly affects students' performance and perceptions (Virvou & Katsionis, 2008). In this vein, usability is evaluated in this study to improve the design of system. Nielsen's heuristic approach is selected because it is the most commonly used and can be effectively applied by both novices and experts (Nielsen, 1994; Nielsen & Mack, 1994). Moreover, Nielsen's heuristic approach evaluates the system from various aspects, for instance, the interface design, help and instruction, and the feedback of interactions. Therefore, Nielsen's heuristic approach is more effective to identify the design problems comparing to the other methods of usability evaluations, such as user testing and cognitive walkthrough (Fu, Salvendy, & Turley, 2002).

In brief, a game-based learning system is developed in this study to improve students' problem solving abilities. Nielsen's heuristic evaluation of usability is used to improve the design of this system. Besides, the differences of cognitive styles on usability evaluation are also discussed to satisfy individual needs. The methodology and result will be discussed in the following sections. Besides, the design guideline will be posed for the future studies and the improvement of this system.

2. Development of System

2.1 System Architecture

Adobe Flash is selected to develop our proposed game-based learning system. This is due to the fact that Adobe Flash has been widely used in the development of game-based learning programs, e.g., mathematics education (Shafie & Ahmad, 2010), college education (Kuk, Milentijević, Rančić, & Spalević, 2012), and energy efficient education (Cowley, Moutinho, Bateman, & Oliveira, 2011). Additionally, it includes many attractive features, such as the ease of learning. Furthermore, it provides with strong graphic capabilities, which are not available in other standard programming languages (Lee & Lee. 2007). In order to improve students' problem solving abilities, there are two quests in our game-based learning system. The first one is a mathematical logic and reasoning quest, and the second one is a verbal logical reasoning quest. The details of two quests are described below.

2.2 The First Quest

In the first quest, as shown in Figure 1(a), the road towards to a town is blocked by a landslide because of the heavy rain. Therefore, players are required to clear the blocks to enter the town. A dog guard is able to help players finding out the instructions from the landslide to operate the machine to clear the blocks. According to the instructions, players can infer the answer and improve the mathematical logic and reasoning abilities. After the start of the quest, there are a timer and a give up button on the top screen, and a goal button and a hint button are on the bottom left screen. The setting is illustrated as shown in Figure 1(b). If players click the instructions directly without the help of the dog guard, it will result in a collapse and game over as shown in Figure 1(c). When players click the give up button, a dialog box including continue button and quit button will be appeared as shown in Figure 1(d). If players click the goal button, the goal of the quest will be popped up as shown in Figure 1(e). If players click the hint button, the instruction to pass the quest will be popped up as shown in Figure 1(f). By the help of the dog guard, the instructions including the logical reasoning rules will be obtained and appeared on the top of the panel of machine as shown in Figure 1(g). When passing the quest, it will pop up players' information such as gained experiences, gained game coins, gained logical reasoning abilities, time to spend, and times to click hints in a congratulation dialog box as shown in Figure 1(h).

(a) (b) (c) (d) (e) (f) (g) (h)

Figure 1. The interface in the first quest.

2.3 The Second Quest

There are two related topics in the second quest. At first topic, players have to find out the characters of the highest grade. The goal of second topic is to find out the character that destroyed the piggy's toy car. For these two topics, some questions are needed to be solved by the players. If players give wrong answers for three times, then they will fail to conquer this quest, and the game is over. Because of similarity of these two topics, only related figures of the first topic are illustrated as follows.

The piggy is crying because his toy car is broken, players have to find out the destroyer as shown in Figure 2(a). A lot of buttons are provided with caption boxes as shown in Figure 2(b). The hearts represent a player’s life, and it means the number of times left to answer the question. If players want to give up the quest, the give up button is available. If the item buttons and NPC buttons are pressed, some verbal clues will be appeared. The functions of the goal button and the hint button are similar to those in the first quest, and the illustrations are shown in Figure 2(c) and Figure 2(d), respectively. After reading the clues, players can click the start answer button to start answering questions as shown in Figure 2(e). While all questions being answered correctly, the explanation of the answers will be clarified to promote the verbal logical reasoning ability as shown in Figure 2(f).

Figure 2. The interface in the second quest.

3. Methodology

The process of the experiment is listed as shown in Figure 3. At first, all participants have to fill out cognitive style questionnaires developed by Ford (1985), and then the whole process is explained. Afterwards, the participants login the assigned website to play the game. After passing all quests, the participants have to fill out Nielsen's Heuristic Evaluation questionnaires.

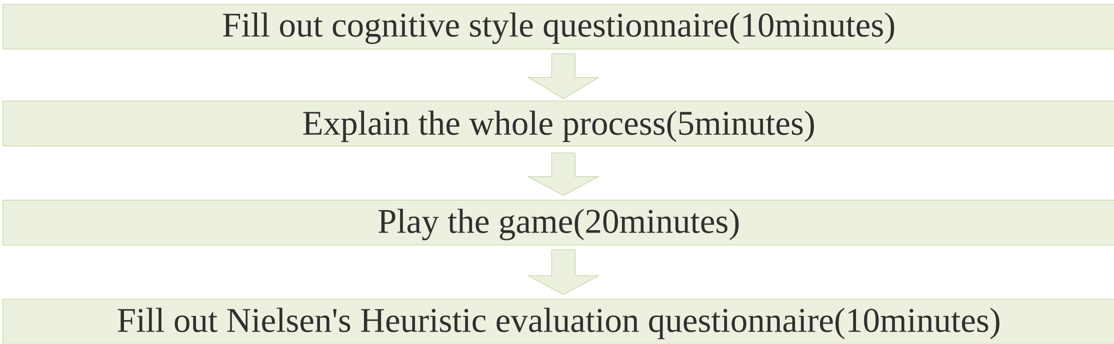

Figure 3. The experiment process.

The participants of the experiment are 49 students from two universities in Taiwan, and their major is information technology. There are 37 valid samples and 12 invalid ones. In this study, the statistical software of SPSS 19 is used to analyze the valid samples, and the overall reliability is .914. According to the analysis of cognitive style questionnaires, there are 9 serialist participants and 28 holist participants for the valid samples as shown in Table 1.

Table 1: The cognitive style of participants.

	Serialist	Holist	Total
Valid samples	9	28	37
Invalid samples	3	9	12
Total	12	37	49

4. Result and Discussion

The descriptive statistics of valid samples as well as their rank of Nielsen’s ten heuristics is listed in Table 2. According to the analysis, the top rank of the heuristic is H8 (aesthetic and minimalist design) with the average score of 3.74. It means the visual design of this game is most satisfied by the players. On the other hand, the least satisfied heuristic is H6 (Recognition rather than recall) with the average score of 2.91. In other words, most players consider insufficient available information to be provided in this system.

Table 2: Descriptive statistics of valid samples and their rank of Nielsen’s ten heuristics.

Nielsen’s ten heuristics	N	Mean	S.D.	Rank
H1:Visibility of system status	37	3.21	.71	9
H2:Match between system and the real world	37	3.26	.78	6
H3:User control and freedom	37	3.32	.69	5
H4:Consistency and standards	37	3.23	.74	7
H5:Error prevention	37	3.35	.61	3
H6:Recognition rather than recall	37	2.91	.73	10
H7:Flexibility and efficiency of use	37	3.41	.76	2

H8:Aesthetic and minimalist design	37	3.74	.60	1
H9:Help users recognise, diagnose and recover from errors	37	3.23	.79	8
H10:Help and documentation	37	3.35	.83	4

Furthermore, the influence of the cognitive style is also considered as shown in Table 3. The top three ranks of heuristics satisfaction for the serialist group are H8, H10 (help and documentation), and H1 (visibility of system status), and the bottom three ones are H6, H9 (help users recognise, diagnose and recover from errors), and H5 (error prevention). On the other hand, the top three ranks of heuristics satisfaction for the holist group are H8, H7 (flexibility and efficiency of use), and H5, and the bottom three ones are H6, H1, and H4 (consistency and standards). For both groups of serialist and holist, the top rank of heuristics satisfaction is H8 (aesthetic and minimalist design), and the bottom rank of heuristics satisfaction is H6 (recognition rather than recall).

However, it observed some noticeable differences of satisfaction for H1 and H5 between serialist group and holist group. The heuristic H1 is at the third rank for the serialist group, but it is at the ninth rank for the holist group. In other words, the visibility of system status (H1) is much satisfied by the serialist group, but the contrary result is provided by the holist group. The heuristic H5 is at the eighth rank for the serialist group, but it is at the third rank for the holist group. That means error prevention (H5) is much satisfied by the holist group, but the contrary result is provided by the serialist group.

Table 3: Descriptive statistics and the ranks of Nielsen's ten heuristics for the serialist group and the holist group.

Nielsen's ten heuristics	Serialist				Holist			
	N	Mean	S.D.	Rank	N	Mean	S.D.	Rank
H1	9	3.33	.67	3	28	3.17	.73	9
H2	9	3.22	.55	7	28	3.27	.85	6
H3	9	3.26	.46	5	28	3.33	.75	4
H4	9	3.26	.55	6	28	3.23	.80	8
H5	9	3.22	.52	8	28	3.39	.64	3
H6	9	2.92	.52	10	28	2.90	.80	10
H7	9	3.33	.60	4	28	3.44	.81	2
H8	9	3.63	.59	1	28	3.77	.61	1
H9	9	3.15	.71	9	28	3.26	.83	7
H10	9	3.41	.86	2	28	3.33	.84	5

Thus, in order to verify the differences of the usability evaluation between two groups, *t* test is used to analyze. The result indicates there are no significant differences between the serialist group and the holist group for Nielsen's ten heuristics as shown is Table 4.

Table 4: *t* test for the differences of heuristics satisfaction between the serialist group and the holist group.

Nielsen's ten heuristics	Cognitive style	N	Mean	S.D.	*t*
H1:Visibility of system status	serialist	9	3.33	.67	.547
	holist	28	3.17	.73	
H2:Match between system and the real world	serialist	9	3.22	.55	.865
	holist	28	3.27	.85	
H3:User control and freedom	serialist	9	3.26	.46	.784
	holist	28	3.33	.75	
H4:Consistency and standards	serialist	9	3.26	.55	.909

	holist	28	3.23	.80	
H5:Error prevention	serialist	9	3.22	.53	.472
	holist	28	3.39	.64	
H6:Recognition rather than recall	serialist	9	2.93	.52	.941
	holist	28	2.90	.80	
H7:Flexibility and efficiency of use	serialist	9	3.33	.60	.718
	holist	28	3.44	.81	
H8:Aesthetic and minimalist design	serialist	9	3.63	.59	.537
	holist	28	3.77	.60	
H9:Help users recognise, diagnose and recover from errors	serialist	9	3.15	.71	.714
	holist	28	3.26	.83	
H10:Help and documentation	serialist	9	3.41	.86	.820
	holist	28	3.33	.84	

Besides, in order to obtain the advanced realization of participants' perceptions and suggestions toward this system, an interview with individual participants is held after the experiment. After organization, it generates 8 codes by the synthesis of interviewers' perceptions and suggestions toward this system as shown in Table 5. These codes are used to complement the inadequate part of the quantitative analysis.

Table 5: The coded interview results.

Code	Code description
A1	Invisibility of the button
A2	Difficulty to realize the hint
A3	Less information to play the game
A4	Consistency of quest button design
A5	Slow Internet connection
A6	Expectation of more contents in the game
A7	Addition of sound effects and background music
A8	Provision of a logout button

The coded interview results mapping to Nielsen's heuristics for the serialist group and the holist group in the first quest is listed and denoted by the format of code(a list of participants' username) as shown in Table 6. Most of participants in the seialist group mention that they cannot understand how to do in the first quest (A3 mapping to H10). They need more information to play the game. One participant in the seialist group mentions that the hint is helpless due to the difficulty to realize (A2 mapping to H9). In the other hand, 8 participants in the holist group mention to the difficulty to realize the hint (A2 mapping to H9). Four participants in the holist group mention to less information to be provided (A3 mapping to H10). One participant mentions to the invisibility of the button (A1 mapping to H1), and another one participant mentions to the consistency of quest button design (A4 mapping to H4).

Table 6: The code for the serialist group and the holist group in the first quest.

Nielsen's heuristics	Serialist	Holist
H1	-	A1(test08)
H4	-	A4(test11)
H9	A2(test09)	A2(test06, test11, test20, test22, test31, test41, test42, test57)
H10	A3(test01, test09, test12, test26, test44)	A3(test05, test10, test41, test47)

The coded interview results mapping to Nielsen's heuristics for the serialist group and the holist group in the second quest is listed as shown in Table 7. Most of participants in the seialist group mention to less information to be provided (A3 mapping to H10). In the other hand, 5 participants in the holist group mention to the difficulty to realize the hint (A2 mapping to H9). Four participants in the holist group mention to the invisibility of the button (A1 mapping to H1), and another two participants think they mention to less information to be provided (A3 mapping to H10).

Table 7: The code for the serialist group and the holist group in the second quest.

Nielsen's heuristics	Serialist	Holist
H1	-	A1(test08, test20, test41, test42)
H9	-	A2(test11, test14, test22, test31, test41)
H10	A3(test01, test09, test12, Test26, test44)	A3(test42, test47)

The coded interview results without mapping to Nielsen's heuristics for the serialist group and the holist group is listed and taken as the suggestions to improve the system design as shown in Table 8. Three participants in the serialist group expect more contents are provided in the game (A6), and one participant in the serialist group expects to add sound effects and background music (A7). In the other hand, 4 participants in the holist group mention to slow Internet connection (A5). Moreover, one participant in the holist group expects to add sound effects and background music (A7), and another one participant in the holist group expects that a logout button is provided (A8).

Table 8: The suggestions proposed by serialist and holist participants to improve the system design.

Serialist	Holist
A6(test04, test10, test14) A7(test38)	A5(test05, test06, test07, test08), A7(test16), A8(test14)

5. Conclusion

A game-based learning system is developed in this study to promote students' problem solving abilities. The usability of Nielsen's heuristic evaluation is used to examine the design of this system. In addition, students' cognitive styles are also considered to measure their influences on the usability evaluation. The major result indicates that the heuristic of aesthetic and minimalist design (H8) is most satisfied while the heuristic of recognition rather than recall (H6) is most dissatisfied. Furthermore, when the influence of the cognitive style is also considered, even some differences are observed, the result shows that there are no significant differences of the usability evaluation between the serialist group and the holist group. Regarding to the result, It is necessary to provide available information as more as possible to improve the system design.

According to the observations of the coded interview results mapping to Nielsen's heuristics, the heuristics of H9 (help users recognize, diagnose and recover from errors) and H10 (help and documentation) are referred to by most participants. Moreover, system performance and attractive playfulness, such as the respond time of Internet connection and audio effects, are also mentioned by some participants.

The results are helpful to improve the proposed game-based learning system in this study. However, some limitations and future studies are listed below. Small sample size is a possible reason to lead to no significant differences of the usability evaluation between the serialists and the holists, thus, a larger sample size is expected in the future studies. Moreover, possible influences of prior students' experiences to play games and the analysis of the students' usage profile are not considered in this study, it is expected to be included in the future studies.

Acknowledgements

This study is sponsored by the National Science Council in Taiwan under the contracts no. NSC 101-2511-S-008-010-MY3 and NSC 102-2511-S-275-003.

References

Cowley, B., Moutinho, J. L., Bateman, C., & Oliveira, A. (2011). Learning principles and interaction design for 'green my place': a massively multiplayer serious game. *Entertainment Computing, 2*(2), 103-113.

Elia, I., van den Heuvel-Panhuizen, M., & Kolovou, A. (2009). Exploring strategy use and strategy flexibility in non-routine problem solving by primary school high achievers in mathematics. *ZDM, 41*(5), 605-618.

Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology, 16*(1), 65-77.

Frias-Martinez, E., Chen, S. Y., Macredie, R. D., & Liu, X. (2007). The role of human factors in stereotyping behavior and perception of digital library users: a robust clustering approach. *User Modeling and User-Adapted Interaction, 17*(3), 305-337.

Fu, L., Salvendy, G., & Turley, L. (2002). Effectiveness of user testing and heuristic evaluation as a function of performance classification. *Behaviour & Information Technology, 21*(2), 137-143.

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE), 1*(1), 20-20.

Hastings, N. B., & Tracey, M. W. (2004). Does media affect learning: Where are we now? *TechTrends, 49*(2), 28-30.

Hwang, G., Wu, P., & Chen, C. (2012). An online game approach for improving students' learning performance in web-based problem-solving activities. *Computers & Education, 59*(4), 1246-1256.

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *Internet and Higher Education, 8*, 13-24.

Kuk, K., Milentijević, I., Rančić, D., & Spalević, P. (2012). Pedagogical agent in Multimedia Interactive Modules for Learning–MIMLE. *Expert Systems with Applications, 39*(9), 8051-8058.

Lee, C., & Chen, M. (2009). A computer game as a context for non-routine mathematical problem solving: The effects of type of question prompt and level of prior knowledge. *Computers & Education, 52*(3), 530-542.

Lee, K. C., & Lee, J. (2007). Programming physics softwares in Flash. *Computer Physics Communications, 177*(1), 195-198.

Liu, C., Cheng, Y., & Huang, C. (2011). The effect of simulation games on the learning of computational problem solving. *Computers & Education, 57*(3), 1907-1918.

Nielsen, J. (1994). Usability inspection methods. *Conference Companion on Human Factors in Computing Systems*, 413-414.

Nielsen, J., & Mack, R. L. (Eds.). (1994). *Usability inspection methods*. New York, NY: John Wiley & Sons.

Pahl, K., & Rowsell, J. (2012). *Literacy and education.* London, England: Sage Publications.

Polya, G. (2008). *How to solve it: A new aspect of mathematical method.* Princeton, NJ: Princeton University Press.

Publishing, F. (2009). In Mackall J. (Eds.), *Problem solving (3rd edition)*. New York, NY: Infobase Publishing.

Schell, J. (2008). *The art of game design: A book of lenses.* Burlington, MA: Morgan Kaufmann Publishers.

Shafie, A., & Ahmad, W. F. W. (2010). Design and heuristic evaluation of MathQuest: A role-playing game for numbers. *Procedia-Social and Behavioral Sciences, 8*, 620-625.

Tao, Y., Yeh, C. R., & Hung, K. C. (2012). Effects of the heterogeneity of game complexity and user population in learning performance of business simulation games. *Computers & Education, 59*(4), 1350-1360.

Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education, 50*(1), 154-178.

Webster, J., & Ahuja, J. S. (2006). Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. *MIS Quarterly, 30*(3), 661-678.