The Evaluation of Interface Elements and Game Elements in Serious Games: A Peer Assessment Approach

Chia-Chi YEH, Hsun-Chu CHEN, Jen-Hang WANG* & Sherry Y. CHEN
Graduate Institute of Network Learning Technology National Central University, Jhongli Taiwan
* harry@cl.ncu.edu.tw

Abstract: Previous studies indicated that game elements may make players have joyful experience but some research indicated that players' learning performance may not be as good as what was expected. This might be due to the fact that poor user interface presented in serious games hindered students' learning performance. Therefore, designers should not only consider game elements, but also take into account interface elements. However, there is a lack of such studies that pay attention to both game elements and interface elements when developing serious games. To this end, this study aims to investigate how interface elements and game elements are considered in the design of serious games. A peer assessment approach is applied to achieve this aim. The results from the peer assessment indicated that designers generally tended to focus on game elements and paid less attention to interface elements. Furthermore, all game elements were similarly considered but interface elements were not equally treated. In particular, designers seemed to ignore the importance of error prevention and recovery from errors.

Keywords: Serious Game, User Interface, Game Elements, Nielsen Heuristics

1. Introduction

Over the past decades, digital games have become very popular. Among various types of games, serious games are widely accepted as effective learning tools, which can be applied for inquiry, multitasking, collaboration, creativity problem-solving, and decision-making (David and Watson, 2010, Gee, 2003). This may be due to the fact that serious games provide a joyful environment, which contains multiple game elements, including storyline, game play, artistic/graphics, and sound/special effects, AI. The significance of these game elements is described in Table 1.

Table 1: Game Elements

Game Elements	Significance
Storyline	To facilitate players to develop a quick understanding of unexplored virtual worlds (Park, 2010; Busselle and Bilandzic, 2008)
Gameplay	To use rules to ensure that all players take same paths and make them feel excited about playing the game (Prensky, 2001).
Artistic/Graphic	To improve display characteristics to attract players (Prensky, 2001)
Sound/special effects,	To affect players' task performance and stress response (Hébert et al., 2005; North and Hargreaves, 1999)
AI	To provide players with feedback so that they can know their current status (Dobrev, 2012)

The game elements presented in Table 1 may make players have joyful experience but players' learning performance may not be as good as what we expected. For instance, Schrader and Bastiaens (2012) indicated that some serious games decreased students' learning performance. This might be due to the

fact that user interface is not properly presented in the serious games so that some learners may need to spend additional time learning how to play these serious games. In other words, only taking into account game elements is not sufficient and there is also a need to consider usability guidelines.

Usability is widely used in system design and the aim of usability refers to making systems easier to use, pleasant to use, and matching them more closely to users' requirements (Nielsen, 1994). A number of usability guidelines can be applied to evaluate usability (e.g., Shneiderman's Eight Golden Rules of Interface Design, Norman's Design Principles and Nielsen's Heuristics). Among these usability guidelines, Nielsen's heuristics (Table 2) were first formally described in presentations in the Human–Computer Interaction conference through papers published by Nielson and Molich (1990). Since then, they have refined the heuristics based on a factor analysis of 249 usability problems to derive a revised set of heuristics with maximum explanatory power. Such heuristics are most commonly used because they can be used effectively by novices and experts alike and can be performed at any stages of the development lifecycle (Nielsen, 1994). In particular, recent studies attempted to incorporate Nielsen's heuristics into the interface design of serious games. For example, Hsieh, Su, Chen and Chen (2015) took a user-centered design approach to develop a Robot-based Learning Companion based on Nielsen's heuristics. Additionally, Mei, Ku and Chen (2015) used Nielsen's heuristics to develop problem-solving games that can accommodate learners' gaming experience.

The aforementioned studies suggested that incorporating Nielsen's heuristics into the interface design of serious games is feasible. Therefore, designers should not only consider game elements, but also take into account interface elements. However, there is a lack of such studies that pay attention to both game elements and interface elements when developing digital games, especially serious games. To this end, the study presented in this paper aims to investigate how interface elements and game elements are considered in the design of serious games. A peer assessment approach is applied to investigate game elements and interface elements provided by serious games and relationships between game elements and interface elements are also examined in this study.

Table 2: Nielsen's ten heuristics (1994)

Heuristics	Explanations
H1: Visibility of system status	The system should always keep user informed about what is going on by providing appropriate feedback within reasonable time
H2: Match between system and the real world	The system should speak the user's language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order
H3: User control and freedom	Users should be free to develop their own strategies, select and sequence tasks, and undo and redo activities that they have done, rather than having the system do these for them
H4: Consistency and standards	Users should not have to wonder whether different words, situations, or actions mean the same thing and the system should follow platform conventions
H5: Error prevention	Even better than good error messages is a careful design, which prevents a problem from occurring in the first place
H6: Recognition rather than recall	Make objects, actions, and options visible. The users should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate
H7: Flexibility and efficiency of use	Allow users to tailor frequent actions. Provide alternative means of access and operation for users who differ from the ''average'' user (e.g., physical or cognitive ability, culture, language, etc.)
H8: Aesthetic and minimalist design	Dialogues should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility

H9: Help users recognize, diagnose and recover from errors	Error messages should precisely indicate the problem and constructively suggest a solution. They should be expressed in plain language
H10: Help and documentation	Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large

2. Methodology

2.1.1 Participants

A total of 41 individuals participated in this study. Participants were students from the some universities in Taiwan, aged between 22 and 25 years old, and they volunteered to take part in the study. These participants were randomly assigned into seven groups, each of which included five or six individuals. A request was issued to learners in lectures, and further by email, making clear the nature of our study and their participation. All participants had sufficient computing skills to act as designers to design serious games in this study.

2.1.2 Experimental Procedure

During the experiment, each group was requested to design a serious game with the Game Maker, which was chosen because it has been used as a rapid development tool in several studies (e.g., Moreno-Ger et al., 2007). Moreover, the Game Maker takes an object-oriented and event-driven approach (Overmars, 2004) so it may be easy to use for most designers. The experimental procedure was divided into two steps: (1) preliminary training and (2) peer assessment (Figure 1). The details are described below.

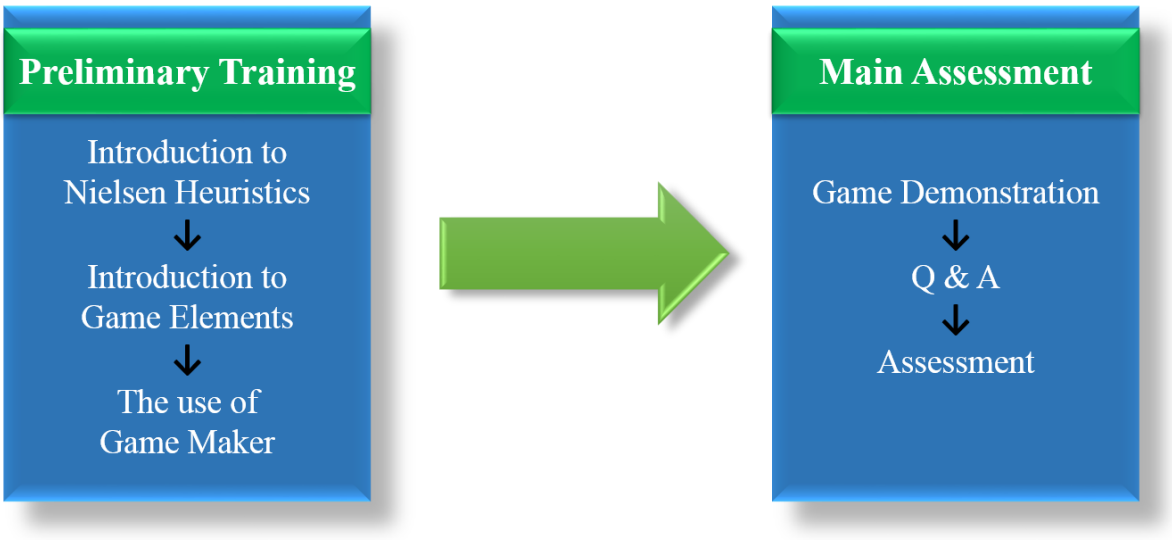

Figure 1. The experimental procedure

2.1.3 Preliminary Training

In order to ensure that every participant has proper domain knowledge, the participants were provided with preliminary training sessions, which covered three topics. Table 3 describes the details of each topic. Regardless of the topics, the participants needed to do some exercises in the end of each session so that they could review what they had learnt in each session.

Table 3: The preliminary training procedure

Topics	Learning Objectives	Length
Introduction to Nielsen's heuristics	The participants should have a basic understanding of how to use Nielsen's heuristics to assess usability	Three 60-minute sessions
The Design of Game Elements	The participants should develop skills of how to incorporate various game elements	Two 60-minute sessions
The Use of the Game Maker	The participants needed to know how to use tools and logics provided by the Game Maker	Nine 60-minute sessions

2.1.4 Peer Assessment

The assessment covered two parts. One was related to game elements while the other was concerned with user interface. During the peer assessment, each group was required to demonstrate serious games developed by them in front of remaining six groups. Additionally, they needed to present how they incorporated the game elements and interface elements into their serious games. Such demonstration and presentation took 20 minutes or so. Finally, there was a five-minute Q&A session, where the participants needed to answer questions raised by the other groups.

After listening to the demonstration and presentation and the answers to the questions, the other groups started to make assessment. More specifically, the other groups needed to fill out a mark sheet for each group. The first part of the mark sheet included five items, which were committed to assess whether the the game elements were properly implemented, i.e., G1: Storyline, G2: Game Play, G3: Artistic/Graphics, G4: Sound/Special effects, G5: Artificial Intelligence. The second part included 10 items, which were committed to evaluate whether user interface was properly designed based on Nielsen's ten heuristics. To facilitate assessment, detailed criteria were produced according to each item. Finally, the first part contained 18 criteria while the second part included 31 criteria. Irrespective of the number of criteria in each part, 20 points were assigned to each item in the first part while 10 points were given to each item in the second part. By doing so, the perfect score of each part is 100. Additionally, the higher the score was given, the more the design matches with the criteria, and vice versa.

2.1.5 Data Analysis

In this study, the independent variables are interface elements and game elements and the dependent variables are the scores obtained from each interface element and each game element. An independent t test was applied to identify differences between interface elements and game elements. Furthermore, an ANalysis Of VAriance (ANOVA), suitable to test the significant differences of three or more categories (Stephen & Hornby, 1997), was applied to analyze differences among interface elements and among game elements. By doing so, the interface elements and game elements that designers paid the most attention and those that designers paid the least attention could be identified. Additionally, Pearson's correlations were also applied to analyze relationships between the scores from game elements and those from interface elements so that the associations between these two types of elements could also be discovered.

3. Results and Discussion

Table 4 describes the mean values and standard deviations of the scores from game elements and those from interface elements. We further used an independent t test to analyze their significant difference. The results indicated that a significant difference ($t(40)=2.421$, $p<.05$) existed between the scores from game elements (Mean = 71.81, SD=17.73) and those from interface elements (Mean = 60.14, SD=13.17). In other words, designers tended to pay more attention to game elements than interface

elements when they developed serious games. We further analyze how each game element (Section 3.1) and each interface element (Section 3.2) were implemented in serious games and how game elements are associated with interface elements (Section 3.3).

Table 4: The mean and standard deviation of game and interface

	Game Elements	Interface Elements
Mean	71.81	60.14
SD	17.73	13.17
Significance	t(40) = -2.421*	

Keys: * p < .05, ** p < .01

3.1 The evaluation of game elements

The results from the ANOVA indicated that no significant differences exist among scores obtained from each game element (F(4)=.38, p>.05). This finding suggested that all game elements were almost equally taken into account when designers developed their serious games (Mean = 14.36, SD=0.89). This might be due to the fact that serious games became very popular in recent years (Johnson, Vilhjálmsson, and Marsella, 2005). Accordingly, designers could have a clear understanding of all game elements so that most of them could effectively integrate these elements into the design of serious games.

Even though there was no significant difference, we still found some interesting results (Table 5). More specifically, the highest score and the lowest standard deviation (Mean=15.59, SD=2.10) were found in G2 while the lowest score and the highest standard deviation (Mean=13.46, SD=6.48) were demonstrated in G4. In other words, most of designers paid attention to G2, which refers to Gameplay, including game rules and the compatibility of the game. This finding revealed that most of the designers considered Gameplay as an important factor in the design of serious games. On the other hand, not all of designers were concerned with G4, which refers to Sound/Special effects. This finding implied that not all of designers thought that Sound/Special effects were essential for the design of serious games. In particular, one group totally ignored the Sound and Special effects and obtained zero score for this game element. In other words, designers had diverse views towards the Sound and Special effects presented in serious games.

Table 5: The score from each game element

	G1	G2	G3	G4	G5
Mean	14.76	15.59	14.44	13.46	13.56
SD	3.25	2.10	2.98	6.48	2.70

3.2 The Assessment of Interface elements

One difference between the game elements and interface elements is that a similar level of attention was paid to the former while different levels of attention were given to the latter. More specifically, the results from the ANOVA indicated that there was a significant difference among scores obtained from each interface element (F(9)=10.63, p=.000). The results from the post-hoc analyses indicated that the scores from H5 and H9 were significantly lower than those from H2, H3, H4, H6, H7, and H8. In other words, designers paid more attention to H2 (match between system and the real world), H3 (user control and freedom), H4 (consistency and standards), H6 (recognition rather than recall), H7 (flexibility and efficiency of use), and H8 (aesthetic and minimalist design), among which H4 received the most attention. Conversely, they ignored H5 (Error prevention) and H9 (Help users recognize, diagnose, and recover from errors) and H 9 was received the least attention.

The other difference lied within the fact that the score of each interface element was low (Mean = 6.01, SD=1.65). In particular, the scores given for H5 and H9 were lower than the mean value (Table 6). Additionally, high SDs existed in these two interface elements. These findings suggested that designers

not only tended to ignore H5 and H9, but had diverse opinions for the assessment of these two interface elements. This might be because H5 and H9 are concerned with the reduction of errors that users may make. However, such errors need to be identified by field works. The lack of evidence from the field works may make designers have difficulties in knowing errors that users may make and helping them recover from such errors. This might be the reason why these two interface elements were ignored by the designers.

Table 6: The score from each interface element

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
Mean	6.17	7.11	7.16	7.96	3.69	7.12	6.88	7.35	3.13	4.22
SD	1.09	0.81	1.58	0.98	1.95	1.34	1.14	1.25	1.03	1.91

3.3 Game VS. Interface

Pearson correlations were applied to analyze relationships between each game element and interface element (Table 7). As shown in Table 7, positive relationships existed between H4 and G1 ($r=0.80$, $p<0.05$), between H7 and G2 ($r=0.78$, $p<0.05$), and between H8 and G3 ($r=0.76$, $p<0.05$). The positive relationship between H4 (Consistency and standards) and G1 (Storyline) suggest that user interface designed in a consistent way can make players immerse in the storyline of the serious game. This might be owing to the fact that the coherence of the storyline of a serious game could be clearly and easily demonstrated in consistent user interface. Accordingly, such coherence could increase the immersions of the players.

Table 7: Pearson correlation between game and interface

r	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
G1	.42	.60	.11	**.80***	.34	.20	.69	.35	.14	.04
G2	.21	.29	.65	.63	.73	.58	**.78***	.63	.09	-.02
G3	-.04	-.23	.60	.63	.41	.29	.69	**.76***	-.07	-.42
G4	-.01	.01	-.01	.71	.27	-.10	.71	.34	.07	-.16
G5	-.13	-.08	.34	.73	.42	.10	.74	.57	.10	-.30

Keys: * $p < .05$, ** $p < .01$

Regarding the relationship between H7 (Flexibility and efficiency of use) and G2 (Game Play), the former refers to flexible manipulation and efficiency of use while the latter refers to the experience of players reaching the game's objectives (Papaloukas, Patriarcheas and Xenos, 2011). Therefore, this finding suggested that players could smoothly pursue the game's objectives when they were allowed to play the game in more efficient and flexible way. Regarding the relationships between H8 (Aesthetic and minimalist design) and G3 (Artistic/Graphics), both elements are concerned with the prettiness of user interface. This may be the reason why a positive relationship existed between these two elements. In other words, they were related to each other.

In summary, the interface design of serious games is closely associated with the entertainment of serious games. Therefore, designers may need to take into account both game elements and interface elements to improve players' gaming experience.

4. Conclusions

This study used a peer assessment approach to investigate how designers perceived the importance of interface elements and game elements in the design of serious games. Figure 2 summarizes the findings of this study. As shown in Figure 2, the results of this study indicated that designers generally tended to focus on game elements and paid less attention to interface elements. Furthermore, each game element was similarly considered though G2 (Game Play) was paid the most attention. On the other hand, each interface element was not equally treated. More specifically, designers were more concerned with H2 (Match between System and the Real World), H3 (User Control and Freedom), H4 (Consistency and

Standards), H6 (Recognition rather than Recall), H7 (Flexibility and Efficiency of use), and H8 (Aesthetic and Minimalist design). However, they seemed to take less notice of H5 (Error Prevention) and H9 (Help users Recognize, Diagnose and Recover from errors). In other words, they might ignore the importance of error prevention and recovery from error. These findings implied that future works should examine how to help designers develop a deep understanding of the importance of interface elements so that they could know how to design usable user interface for serious games. The ultimate goal is that players are able to know how to initiate serious games effectively so that their learning performance can be enhanced. In spite of the fact that the findings from this study provide the aforementioned guidance, this is an exploratory study, which included a small sample. Thus, there is a need to consider a larger sample to provide additional evidence in the future. The other limitation is that this study only takes into account Nielsen's heuristics. Thus, it is necessary for further studies to consider other usability guideline, e.g. Shneiderman's "Eight Golden Rules of Interface Design" (Schneiderman, 1986).

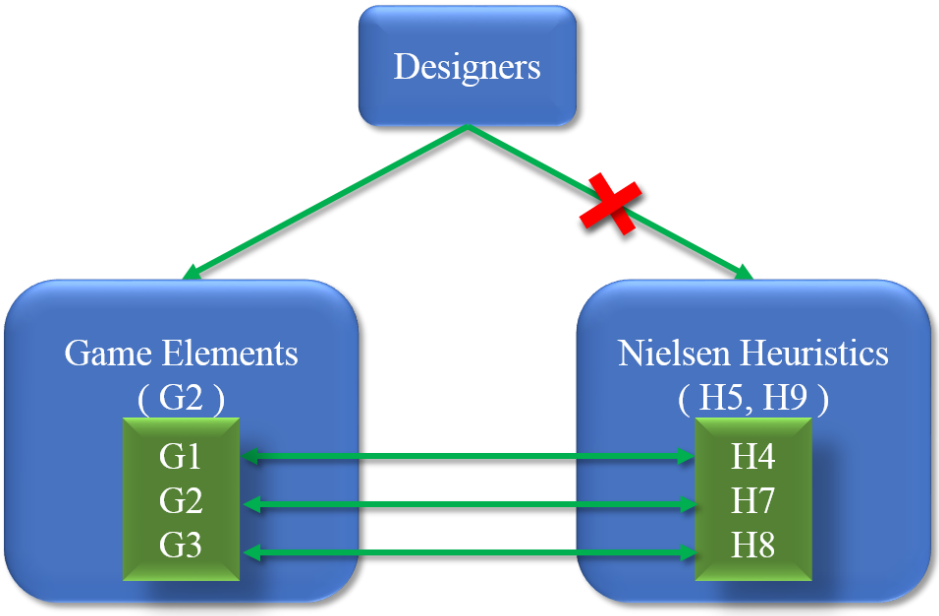

Figure 2. The Summary of the Findings

Acknowledgements

This study is supported in part by the Ministry of Science and Technology of the Republic of China under Contract No. MOST 103-2511-S-008 -010 -MY3, MOST 104-2511-S-008 -008 -MY3, and MOST 104-2811-S-008-007.

References

David, Maria Manuela, & Watson, Anne. (2010). Participating in what? Using situated cognition theory to illuminate differences in classroom practices. *New directions for situated cognition in mathematics education* (pp. 31-57): Springer.

Dobrev, D. (2012). A Definition of Artificial Intelligence. *arXiv preprint arXiv:1210.1568*.

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

Hébert, S., Béland, R., Dionne-Fournelle, O., Crête, M., & Lupien, S. J. (2005). Physiological stress response to video-game playing: the contribution of built-in music. *Life sciences, 76*(20), 2371-2380.

Hsieh, Y.Z., Su, M.C., Chen, S. Y. & Chen, G. D. (2015) The Development of a Robot-based Learning Companion: a User-Centered Design approach. *Interactive Learning Environments*. 23(3), 356-372

Hwang, G. J., Yang, L. H., Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education, 69*, 121-130

Johnson, W. L., Vilhjálmsson, H. H., & Marsella, S. (2005). Serious games for language learning: How much

game, how much AI?. In *AIED* (Vol. 125, pp. 306-313).
Ke, F. (2008). Computer games application within alternative classroom goal structures: cognitive, metacognitive, and affective evaluation. *Educational Technology Research and Development*, *56*(5-6), 539-556.
Kuo, M. J. (2007, March). How does an online game based learning environment promote students' intrinsic motivation for learning natural science and how does it affect their learning outcomes?. In *Digital Game and Intelligent Toy Enhanced Learning, 2007. DIGITEL'07. The First IEEE International Workshop on* (pp. 135-142). IEEE.
Moreno-Ger, P., Sierra, J. L., Martínez-Ortiz, I., & Fernández-Manjón, B. (2007). A documental approach to adventure game development. *Science of Computer Programming*, *67*(1), 3-31.
Nielsen, J. & Molich, R. (1990). Heuristic evaluation of user interfaces. In Proceedings of CHI 90,249-256.

Nielsen, J. (1995). 10 Usability Heuristics for User Interface Design. *Fremont: Nielsen Norman Group.[Consult. 20 maio 2014]. Disponível na Internet*.
Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.
North, A. C., & Hargreaves, D. J. (1999). Music and driving game performance. *Scandinavian Journal of Psychology*, *40*(4), 285-292.
Mei, S. Y., Ku, O., Chen, S. Y. (2015) The Evaluation of the Interface of a Game-Based Learning System: Experts vs. novices. *4th International Congress on Advanced Applied Informatics.*
Overmars, M. (2004). Teaching computer science through game design. *IEEE Computer*, 37(4), 81-83
Park, N., Lee, K. M., Jin, S. A. A., & Kang, S. (2010). Effects of pre-game stories on feelings of presence and evaluation of computer games .*International journal of human-computer studies*, *68*(11), 822-833.
Prensky, M. (2001). Fun, play and games: What makes games engaging .*Digital game-based learning*, *5*, 1-05.
S. Papaloukas, K. Patriarcheas, and M. Xenos, "Games' Usability and Learning - The Educational Videogame "Be The Manager!"", *Proceedings of the 5th European Conference on Games Based Learning* (ECGBL 2011), Academic Publishing International (API) Press, pp.449-456, 2011.
Schneiderman, B. (1986). Eight golden rules of interface design. *Disponible en*.
Schrader, Claudia, & Bastiaens, Theo J. (2012). The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. *Computers in Human Behavior*, *28*(2), 648-658.
Stephen, P., & Hornby, S. (1997). *Simple statistics for library and information professionals.* London: Library Association.