A Comparison of Learning Effectiveness Among Serious Games with Varying Degrees of Playability

Ching-Sheng Yang [a*], Shu-Hua Lin [b], Kuo-Hua Wang [a]
[a] *Graduate Institute of Science Education, National Changhua University of Education, Taiwan*
[b] *Department of Biology, National Changhua University of Education, Taiwan*
*pureman720730@gmail.com

Abstract: Serious games are composed of three components, namely, learning contents, learning strategies, and game elements. To emphasize the significance of game elements on learning effectiveness in serious games, this study proposed a design pattern for courses of serious games, based on serious game courses which addressed the endocrine unit for students in the seventh grade. Since the courses were first developed and put to use three years ago, the courses have been implemented and modified many times. The courses adopted in this study were the fifth version. Employing a quasi-experimental research design, an experiment of two-week courses was organized in three groups, namely the traditional instruction group, the serious game group, and the low playability game group, which used courses with less playabilities modified on the basis of serious games. A total of 234 students participated in this research. The research instruments included achievement test. The results showed that serious game courses with sufficient playabilities outperformed the low playability group and traditional group in both the achievement test, indicating that playability was an important factor influencing the effectiveness of serious game courses. Additionally, with respect to the categorical variables, gender differences had no significant influence over serious game courses. Learning background had no significant influence on the outcome of achievement test of the serious game group.

Keywords: Serious game, Endocrine, Playability, digital game

Introduction

1. Research motivation and background

Digital gaming accounts for a substantial portion of the time spent on computers by students in Taiwan, and a similar phenomenon occurs in other countries (McFarlane, Sparrowhawk, & Heald, 2002). Naturally, computer games play a significant role in the everyday life of students. When applying digital gaming to learning practices, we must pay attention to the integration of game playability, in addition to learning contents. The games that maintain a balance between learning contents and playabilities are referred to as serious games. A serious game may contain a few key elements, such as meaningful learning content incorporating game elements, learning strategies guiding students' exploration of learning content, and game elements enhancing learning. Serious games are digital games characterized by the incorporation of both entertaining and instructional functions, while avoiding the boredom of lite games (Michael & Chen, 2006; Zyda, 2005).

Serious games are composed of three components, where the playability is defined by many scholars as the qualities enabling a game to be fun and entertaining to play (Alessi & Trollip, 1985; Malone, 1981; Prensky, 2001). Researcher summary these qualities and divid

them into two reciprocally influencing categories, namely, “game attributes” and “design elements,” and then used them to describe the elements necessary in the development of digital game-based instruction. The two groups of course designs with different playabilities are distinctive from each other in their design elements.

Table 1
Essential design and game attributes of digital game

Design / Game factor	Competition & Challenge	Result & Feedback	Rule & Goal	Story
Problem solving	X	X	X	
Adaptation	X	X	X	
Successes	X	X	X	
Playfulness	X	X		
Interactive	X	X		
Funny	X	X		X
Fantasy				X

Learning contents need to be reconstructed by employing a combination of appropriate game elements and learning strategies, because simply mixing concepts in the textbooks into games may lead to page turner styled learning. Without appropriate learning strategies, serious games probably turn into nothing but entertainment. Because of time limits in course delivery, appropriate learning strategies may help shorten the time between flow and break down (Roth, 2006), preventing learners from wandering around aimlessly. This also prevents learning activities from being turned into purely entertaining games. One of the key characteristics of games is its problem solving. It is of practical significance to integrate problem solving strategies into digital gaming based courses. According to previous studies, a problem solving process may consist of six steps: emergence of problems, identification of problems, sifting through all kinds of information to seek key information for problem solving, seeking reasonable interpretation to develop an answer, judging or verifying the precision of the answer, and solving the problem by finishing a workable conclusion.

Previous studies have discussed the difference of learning performance using commercial entertainment games and lite games with variable playabilities. However, little research has been done to explore the effects of playabilities on learning performance through serious game courses. The reason for this is that previous studies have been unable to modify the particulars or programs of the serious game courses because they were not entirely designed by themselves. In this study, the serious game courses were jointly designed by the researchers and a graphic designer, thus we could modify and adjust any part of the course content as needed.

2. Research objectives and questions to be answered

In summary, this study develops serious game courses as well as low playability digital gaming courses modified on the basis of serious game courses, to understand the effects of playability on learning effectiveness. In addition, this study also explores the influences of background variables, which are commonly examined in digital gaming related research, such as gender differences and learning backgrounds, to understand whether these attributes influence the learning styles of different strategies. Previous studies have reported that different genders have different preferences toward game genres(Bonanno & Kommers,

2008 ; Wang & Wang, 2008) . This study is also concerned about whether such differences may influence the performance of different learning strategies. Hence, this study attempts to investigate the following questions:

- What is the difference in student achievement test using narrative, low playability games and serious game courses?
- How do gender differences and learning background influence achievement test?

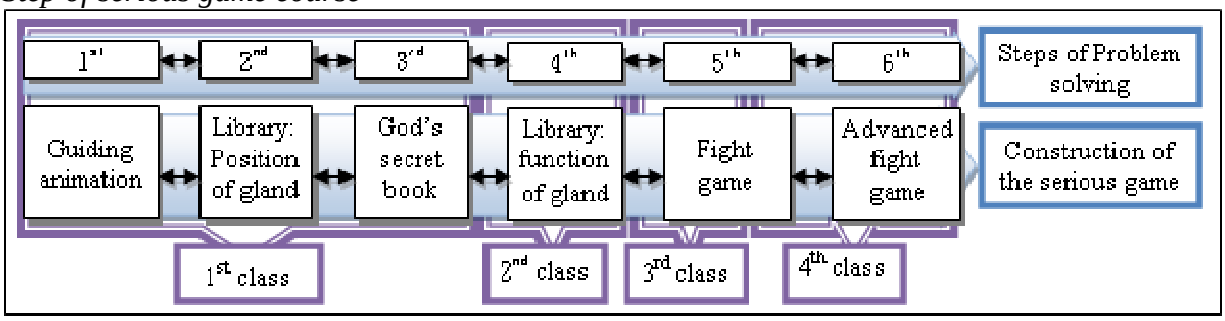
3. Research Methods

3.1 Serious game course

Adopting the design pattern of serious game courses presented in this study, the researchers developed a serious game course entitled, “Endocrine Warrior.” The development tools were: Adobe FlashCS4, PHP5, & MySQL. Since its first release, the module has undergone four revisions within three years. Three professors specializing in education science and 32 junior high school science teachers participated. The courses have been implemented seven times, in 16 classes of the seventh grade and 14 classes of the ninth grade. Since its implementation, the courses have undergone three major revisions to reflect the comments and suggestions made by the teachers and education experts. The serious game adopted in this study was the fifth revision.

The serious game course were composed of five components, including the “God’s secret book & Library,” “Lab: Position of gland,” “Lab: function of gland,” and “Fight games and recorder of gaming experience.” The instructional knowledge contained in each component was reconstructed according to their respective and adapted game elements. The six components were linked through an animated startup story, and the course design followed the serious game model designed by the researcher. The implementation of course lasted for two weeks and was delivered in four sessions.

Fig 3
Step of serious game course



3.2 Research Instruments

The main research instruments used by this study included an achievement test on endocrine concepts. The achievement test on endocrine concepts was compiled by the researchers themselves, and evaluated by ten science teachers from senior junior high schools and three professors specializing in education science. The test papers were pre-tested with trials to verify reliability, degree of difficulty, and discrimination. Altogether, 570 ninth grade students participated in the pre-test. The valid sample was 562 students, the KR20 value was 0.84, the mean degree of difficulty was 0.66, and the mean discrimination was 0.50.

3.3 Research Design

This study employed a quasi-experimental control-group design with repeated measures.. The research subjects were 234 students of the seventh grade from seven regular classes, who were then divided into three groups, namely the narrative course (*n*=66), serious game course (*n*=101), and low playability game-based course (*n*=67). The narrative group (NG) was the control group, in which traditional narratives were used for course delivery. The serious game group (SGG) delivered a complete set of serious game course. The low playability group (LPG) substituted the third cycle ofserious game with seminars to implement the last two steps of the problem solving (see Fig 2). With respect to the game design (see Table 2), the games in this group were less competitive and challenging than those in the serious game group.

All three groups took a pre-test of achievement on endocrine concepts prior to course implementation. Then, two weeks of course instruction was implemented. A post-test of achievement was taken two days after the courses were finished. The 1st delay-test of achievement was administered two weeks after the courses finished. Finally, eight weeks after the delay-test, a 2nd delay-test of achievement on endocrine concepts were administered to test the long-term memory retaining effects.

Table 2
The research design

Group	2 days ago	Course of implement (2 weeks)	2 days later	2 weeks later	10 weeks later
NG	pre-test	traditional narratives	post-test	1st delay-test	2nd delay-test
SGG		serious game			
LPG		Low playability of serious game			

3.4 Data Processing and Analysis

The classification of students' prior learning backgrounds was based on the average score earned in the last two monthly examinations on Biology preceding the endocrine unit. The students were divided into three groups according to their learning background. The students whose average grades ranked in the first 27% were regarded as high achievement (HA) students; the ones whose average grades ranked in the last 27% were regarded as low achievement (LA) students ; the remaining students were regarded as medium achievement (MA) students .

As to the data analysis aspect, the achievement test was administered four times. Each test randomly sorted the results. Given the same test paper was used, the data analysis should have been carried out using Repeated measures analysis of variance (ANOVA), rather than analyzing individual assessment, to prevent statistical errors.

4. Research Findings and Discussion

This study focuses on the essentiality of playability to the effectiveness of serious game. The subjects of the study were divided into three groups, namely the narrative group (NG), the serious game group (SGG), and low playability group (LPG).. The effectiveness of learning was measured through the scores obtained from a series of assessment, including a post-test, 1st, and 2nd delay-test. Moreover, the influences of gender difference and prior learning background on learning effectiveness were also explored.

Repetitive measurement was corrected using the Huynh-Feldt correction, with the primary effects of the research control reaching a significant level: F (2, 234) = 4.32， p = 0.01(Table 4). The post-hoc comparison was made using the Bonferroni correction. By and large, the SGG significantly outperformed the LPG and NG. However, the effect of LPG was not significantly higher than NG; instead, the achievements of the two were very close.

Table 4
Three-way repeated measures ANOVA results

Source	*SS*	*df*	*MS*	*F*	
Between group					
Achievement test (Repeated measure, A)	86237.82	2.93	29481.15	210.77	**
Control (B)	4196.59	2.00	2098.30	4.32	*
Gender (C)	61.83	1.00	61.83	0.13	
Background (D)	101664.70	2.00	50832.35	104.67	**
*A * B*	3194.18	5.85	545.98	3.90	**
*A * C*	27.59	2.93	9.43	0.07	
*A * D*	17286.03	5.85	2954.69	21.12	**
*A * B * C*	337.35	5.85	57.66	0.41	
*A * B * D*	2050.50	11.70	175.25	1.25	
*A * C * D*	624.62	5.85	106.77	0.76	
*A * B * C * D*	1122.70	11.70	95.95	0.69	
Within group					
Subject	104901.57	216.00	485.66		
residual	51209.71	430.00	119.09		
Total	372915.21	703.65			

** $p<0.01$; * $p<0.05$;

As interaction was present between the achievement and research control, analysis of simple main effect was performed (Table 5). The researchers were concerned about which didactics had the best performance. In the pre-test, no significant difference was found among the three groups, indicating that the students' prior knowledge before engaging in learning was not statistically different. Significant differences were identified in the post-test, the 1st delay-test, and the 2nd delay-test, which were put to a post-hoc comparison subsequently. The post-hoc comparison was conducted using the Sheffe technique, due to the unequal number of samples. By and large, the serious game group significantly outperformed the LPG in both the post-test and the 1st delay-test, and significantly outperformed the NG in the 2nd delay-test, suggesting the serious group had excellent long-term memory retention. Moreover, although the LPG without combat games and the NG did not reach a significant level, this suggests that without combat games, providing the same learning effects as traditional instruction is possible. If combat games can be provided, the learning performance can be further improved through game-based teaching.

Table 5
Analysis of simple main effect: different course and achievement test

Simple main effect	*SS*	*df*	*MS*	*F*		*Post-hoc*
Different achievement test						
NG	22550.28	2.20	10263.77	46.88	**	—
SGG	60010.58	2.56	23419.63	144.36	**	—
LPG	25153.26	2.58	9759.26	42.95	**	—
Different course						—
Pre-test	418.83	2	209.41	1.47		—
Post-test	3738.66	2	1869.33	5.26	*	2>3*; 2>1 ; 1>3
1st delay-test	6197.14	2	3098.57	7.32	**	2>3*; 2>1 ; 1>3
2nd delay-test	6119.89	2	3059.94	5.58	**	2>1*; 2>3*; 3>1

** $p<0.01$; * $p<0.05$; 1:NG; 2:SGG; 3:LPG

In the achievement test of repeated measurement, no interaction was present among the three factors: control, gender, and achievement grouping. Additionally, no interaction was present between the control and gender or achievement grouping, suggesting both achievement grouping and gender exercise little influence on learning effectiveness. In other words, whether male or female, and regardless of learning background, learning through serious games, or low playability games, or traditional instructions rendered no significant difference in the achievement tests.

However, an in-depth analysis revealed that in the 1st delay-test, the male students in the serious game group significantly outperformed the male students in the low playability game group ($F(2, 120)=6.14$, $p < 0.00$), suggesting that unlike female students, male students may achieve better learning effectiveness through more immersive game experiences.

Furthermore, the adoption of various didactics may create significant differences in the 1st delay-test for low achievement students ($F(2, 63)=3.99$, $p = 0.02$). Through post-hoc comparison, we saw that the students in the serious game group with lower achievement significantly outperformed the students in the LPG of the same achievement group. In addition, high achievement students in different groups had significant differences in the 2nd delay-test ($F(2, 66)=4.14$, $p = 0.02$). The post-hoc comparison revealed that through combat game-based learning, the high achievements students had better long-term memory retention than those in the NG. Therefore, higher playability was beneficial, to a certain extent, for both low achievement and high achievement students. However, because of the small sample size of background groupings, the analytical results by achievement groupings are pending further verification.

5. Conclusion and Suggestions for Future Research

The purpose of this study is to explore the role played by serious games in achieving learning effectiveness. Test results indicate that serious games with an adequate level of playability outperformed digital game-based courses with lower playabilities or traditional courses with no playabilities at all. Previous studies on digital gaming tended to emphasize hands-on skills (ex: Shen & O'Neil, 2006; Alkan & Cagiltay, 2007; Ko, 2002; Vogel, Greenwood-Ericksen, Cannon-Bowers, & Bowers, 2006; Robertson & Howells, 2008); however, serious game not only met the needs of teachers who were under the pressure of

raising enrollment quotas (Kebritchi, Hirumi, Kappers, & Henry, 2009), but also transcended the effectiveness of traditional teaching and learning, eliminating the concerns of high critical tests surrounding onsite teaching.

No significant difference existed in the outcome of serious game between male and female students; however, male students slightly showed more demands for playability in the game-based courses. This is probably because of the game genres. This study primarily adopts role-playing games, simulation games, and strategy games. These genres are typically male oriented (Bonanno & Kommers, 2008 ; Wang & Wang, 2008). Therefore, it would stimulate male students to spend more time studying if more playabilities could be added to the games. This is why the serious game group outperformed the others in the delay-test. The gender difference, nevertheless, had little impact on the overall performance of the learning efforts.

Previous studies have discovered that learning background is inversely relational to digital game experience. Learners with mid-lower achievements are familiar with digital gaming, in line with the affordance of computers (Wijekumar, Meyer, Wagoner, & Ferguson, 2006). Therefore, for learners with mid-lower achievements, serious games with sufficient playabilities are more engaging, and are thus more facilitative for learning processes. By contrast, they have little impact on students with high achievement.

By and large, the courses developed on the basis of the design architecture for serious games in this study can exercise significant influences on learning achievements. Gender difference presents little influence on playability. As far as prior learning background is concerned, courses with more playabilities are more suitable for students with mid-lower achievements, while no significant difference was observed for students with high achievement.

Regarding the direction for future research, we make three suggestions based on the findings of this study as follows. First, serious game courses should be carried out progressively, with playability being added gradually, and the content being deepened continually. Because playability plays a critical role in the learning performance of serious game, playability must be incorporated into the content to prevent a page turner type of design. Given that the design architecture presented in this study is not difficult to accomplish, it can probably be used as a reference to further development of serious game courses. Second, more factors of influence can be taken into consideration in future studies to further explore the effects of playability, such as learning styles and proficiency of digital gaming experiences. Third, future researchers may experiment with more diverse learning modules to develop more serious games, such that the learning effects can be better observed.

Acknowledgements

The authors are grateful to the National Science Council for its grant for this study. The authors are also indebted to all of the experts, professors, teachers, and students who have taken part in the implementation of this serious game. We appreciate their kind comments and suggestions about which game can be further modified.

References

[1] Alessi, S. M. & Trollip, S. R. (1985).*Computer-based instruction: Methods and Development*. NJ: Prentice-Hall.

[2] Alkan, S., & Cagiltay, K. (2007). Studying computer game leaning experience through eye tracking. *British Journal of Educational Technology, 38*, 538-542.

[3] Bonanno, P., & Kommers, P. A. M. (2008). Exploring the influence of gender and gaming competence on attitudes towards using instructional games. *British Journal of Educational Technology, 39*, 97-109.
[4] Kebritchi, M., Hirumi, A., Kappers, W., & Henry, R. (2009). Analysis of the supporting websites for the use of instructional games in K-12 setting. *British Journal of Educational Technology, 40*,733-754.
[5] Ko, S. (2002). An empirical analysis of children's thinking and learning in a computer game context. *Educational Psychology,22*, 219-233.
[6] Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive science, 5*, 333-369.
[7] McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games. report of TEEM (Teachers Evaluating Educational Multimedia).* Retrieved November 8, 2007, from *http://www.teem.org.uk/publications/teem_gamesined_full.pdf*
[8] Michael, D., & Chen, S. (2006). *Serious games: Games that educate, train, and inform.* Boston, MA.: Thomson Course Technology.
[9] Prensky, M. (2001). *Digital Game-Based Learning*. New York: McGraw Hill.
[10] Robertson, J., & Howells, C. (2008). Computer game design: Opportunities for successful learning. *Computer & Education, 50*, 559-578.
[11] Roth, W. M. (2006). *Learning science: a singular plural perspective.* Rotterdam: Sense Publishers.
[12] Shen, C. –Y., & O' Neil, H. (2006). *The effectiveness of worked examples in a game-base learning environment.* American Educational Research Association, San Francisco, CA.
[13] Vogel, J. J., Greenwood-Ericksen, A., Cannon-Bowers, J., & Bowers, C. A.(2006). Using virtual reality with and without game attributes for academic achievement, *Journal of research on technology in education, 39*, 105-118.
[14] Wang, H. –Y., & Wang, Y. –S. (2008). Gender differences in the perception and acceptance of online games. *British Journal of Educational Technology, 39*, 787-806.
[15] Wijekumar, K. J., Meyer, B. J. F., Wagoner, D., & Fergusin, L. (2006). Technology affordances: the "real story" in research with K-12 and undergraduate learners. *British Journal of Educational Technology, 37*, 191-209.
[16] Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer, 38 (9),* 25-32.