The influence of game in e-learning: an empirical study

Dimitrios Rigas, Khaled Ayad

Abstract: A human-computer interface is an attempt to mimic human-human communication. In human-human communication, especially in learning, students interact emotionally either with each other or with their instructor in way that minimizes, to some extent, the formality of the learning arena/environment. In web based learning these emotions are usually not present within many of the types of e-learning environments. Researchers on the other hand have articulated that humour strengthens students' performance in a learning environment combined with amusement features. This mostly happens online were users in front of unadulterated educational screens. In this paper, we empirically investigated the role of edutainment applied avatar as a tool to represent the entertainment attributes in an e-learning framework. The empirical investigation aimed at measuring the usability of four experimental game-based interfaces; each of which is integrated with a combination of different multi-modal features which included; text, earcons, speech, and avatar. These four game-based learning interfaces were introduced in four phases; the first to be introduced consisted of text and speech only (TS), the second, text and earcons only (TE), the third, integrated with text, speech and earcons (TSE) and finally fourth game was with text, speech, earcons and avatar (TSEA). This combination of various multi-modal metaphors with e-learning systems were examined to determine the preferable set of multi-modal grouping that entertained and enhanced user's performance. Effectiveness and efficiency of these four environments were analyzed using an independent group of users. The outcomes showed a higher improvement rate in performance of students who learnt with the game interface integrated with the avatar than the other versions.

Keywords— Avatar, E-learning, Edutainment, Entertainment, Human Computer Interaction (HCI), Multi-modal.

I. INTRODUCTION

ADVANCES in technology affected many aspects of human existence, including our learning processes. In general learning processes of today involves the utilization of all available resources to obtain information as quickly as possible, and the learners involved, tend to work in groups to minimize the time required in the acquisition of information[1], [2], [3]. Researches in Human Computer Interaction (HCI) lay emphasis on enhancing systems functionality and measuring the user's capability to understand the process of how to be familiar with the system. Nowadays due to advancements in the field of information technology and integrated systems (Multimodal Interfaces), resulting in the development of diverse applications, developers of HCI now extend their concerns in design, to include issues regarding overall user experience, rather than only the common functional aspects/problems of IT systems[4]. This change has been prompted by issues raised by users familiar with IT systems through daily practice. These issues pinpoint desirable attributes of IT systems that influence a user's eagerness to practice and feel new experiences. As a result the new interfaces should be designed to provide a comfortable environment that allows the users to undergo the learning processes without being overwhelmed [5]. Moreover considering the skills and abilities of today's learners, new approaches must be invented to design and develop educational programmes and media [1]. The absence of face to face contact in e-learning, is one of the many problems that is encountered by this method of learning, since personal interactions and useful feedback messages, as well as real time information on their learning performance, are obviously absent [6]. Multi-modal features such as earcons, speech, and avatars, can be considered as substitutes for lack of face to face contact with a teacher in an online learning classroom. The presence of avatar, i.e talking virtual humans, in e-learning, with consideration to associated cognitive factors, might make the learning process more fun to users, increase their interest, motivation, and retention of concepts, and at the same time improve higher order thinking and reasoning skills [6, 7]. An avatar facial expression provides a more realistic interaction in human computer interfaces. An avatar can express emotions, feelings, and linguistic information that could be mimics human face expressions. Therefore work is in progress to enhance computer hardware with regards to high performance graphics and speed of cartoon-like and human-like synthesized faces for use in computer applications [8]. Avatars has been used in different areas, they are used in console games, multi-modal applications, or even teleconferencing environments. That is because of realistically offered that provides high level of realism in an interactive environment [9].

Play associated with educational games is another important aspect that can be used to enhance the achievement and motivation for today's generation of learners and interaction between users and e-learning systems [10]. The role of play in increasing motivation and stimulating curiosity has been proved by many studies and experiments. Play games enables the player to acquire motor skills, improves memory,

Manuscript received March 17, 2010;
Professor Dimitrios **Rigas** Head of Department of Media Technology. De Montfort University The Gateway Leicester, LE1 9BH UK +44 (0) 116 257 7485, email: d.rigas@DMU.ac.uk).
Khaled Ayad is with the Department of Media Technology, De Montfort University The Gateway Leicester, LE1 9BH UK, BD7 1DP UK (e-mail: kaaayad@DMU.ac.uk).

visualisation and problem solving [7] ,[8], [9], [10]. Playing games offers the learner a completely free medium of making mistakes and provides them with range of attempts which assists in embedding the information in the learner memory permanently [2].

This work combines various disciplines that use multi-modal attributes such as earcons, speech, and avatar to convey information or learning materials through gaming environments. The main aim is to enhance usability of user interface in e-learning systems, through experiencing learning by doing and learning through play environment. Four educational games were introduced here to assist in helping, orient, and enhance learner attention and enthusiasm. User's effectiveness and efficiency measurements showed clear evidence of the enjoyment experienced and the improvement in their performance that the users attained during their interactions with these learn by play medium.

II. Related studies

Numbers of studies with the aim of inspect the role of entertainment and positive emotions in enhancing learning in general and particularly on e-learning systems have been carried out by many researchers and resulted significant outcomes [11], [12], [13], [14], [15], [16], [17], 18].

Cartwright [19] accomplished work to find out effectiveness of employing games interface for accessing geographical information. More specifically the focus was how the geospatial sciences might make use of games technology and gaming. The experimental study involved three interfaces which have been tested to observe whether it enhances the access to information provided or not. "Nintendo" was the name of group of users that defined by [20] "as those users who have been exposed to computer games and thus potentially prefer access to computer delivered information via this type of interface". A comparison experiments was conducted through 4 stages for the same area and evaluations, first stage the Dorling Kindersley World Reference Atlas, produced both on CD-ROM and as a paper version was used for this test with about 50 first-year tertiary students. In stage 2 games-type interface was used in this part of the evaluation with smaller group (8 nominees), stage 3 was to compare two different methods of information access with same 8 candidates. Stage 4 which is final one was subjective methods typically used to determine the level of media quality required in applications with same 8 candidates as well. After the tasks had been performed for every stage, the candidates were asked to answer specific questions though questionnaire provided. The criteria addressed in the research include the general ability of the product to operate, which approach preferred by the users, which 2D or 3D interface is preferred by the users and the ability of the product to provide a better quality of picture.

As result the study demonstrated that most users preferred interactive multimedia in comparison to traditional paper maps. Moreover users who not familiar with computers preferred to use a game-like control, and with a 3D interface user's cognitive load significantly decreased compared to a 2D interface. In general the study came out with conclusion that using a games approach provided in mapping Geographical Information Systems (GIS) products provided a better understanding of a real place than a conventional map.

In another study [21] were the goal was to investigate multimedia instructions with combining varied modality of the text and the use of visual and audio cues. The study assumes that texts associated with diagrams as audio will decrease the extraneous load (modality effect). Moreover include visual cues to the diagrams will prevent visual search and also decrease extraneous cognitive load (split-attention effect). The study compared four different conditions: condition VN - visual text, no cues in the diagram; condition VC - visual text cues in the diagram; condition AN - audio text, no cues in the diagram; condition AC - audio text, cues in the diagram. The study does not only concentrate at the learning results but also at the mental effort the learners spent on the instructions. 151 second-year students from the Department of Education of the University of Gent in Belgium took place in the experiment and the topic examined in the experiment were part of a course on instructional design.

This experimental work produced negative results in general and do not answer the question stated to find out which approach is best for decreasing the extraneous load of multimedia instructions. Moreover merging cues to a diagram do not proved to be effective as expected. Additional negative results on learning emerged when replaced text with audio. Although previous researches provided positive outcomes of combining various modalities in educational multimedia, this study conflicted with them and proved the opposite. The author reasoned this failure to the lack of motivation students have to study the instructions, because the average mental effort during instructions is not high, even below average. Another explanation provided by author is long time spend on studying the instructional materials (approximately one hour), which is in turn decreases student's concentration and span of attention. Whatever the cause, several studies were proved the effective of combining different modalities to enhance usability in general and in particular in e-learning interfaces [22], [23], [24], [25], [26], [27], [28], [29], [30].

III. Experiment Methodology

In this experimental design four games based-learning interfaces were implemented, first game introduced with text and speech only (TS), second was with text and earcons only (TE), and third game integrated with text, speech and earcons (TSE) and finally fourth game was with text, speech, earcons and avatar (TSEA). Randomly dependent group of users were experienced these four games and to provide their feedback through questionnaire. Moreover four lessons rotated between the four platforms, each lesson kept the same amount of information. All four interfaces introduced to the users with selected topics from geology, lesson one was "what is volcano", lesson two was "earthquakes facts", lessons three was "types of rocks" and lesson 4 was "how oil is formed". Each user was provided with four tasks (questions) to perform in each platform and tasks level of difficulty was moderate.

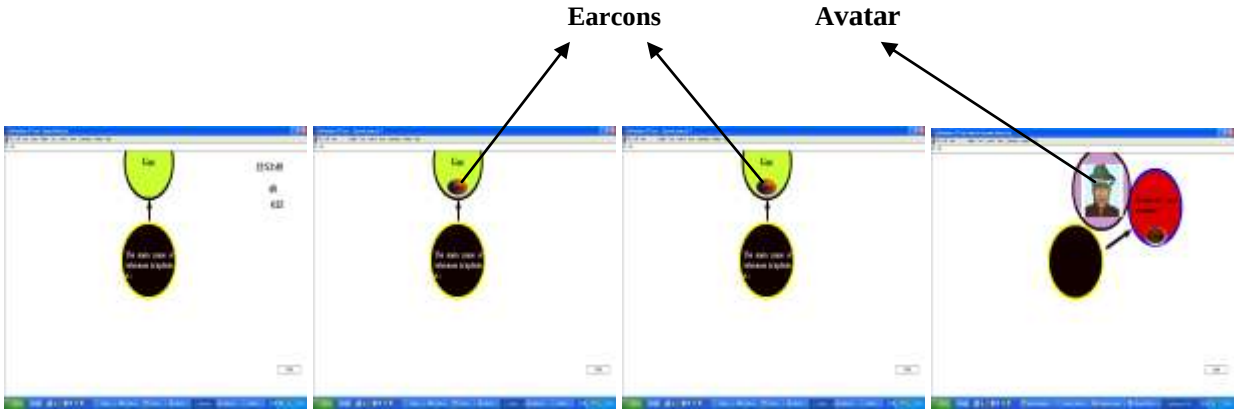


Fig. 1 example of the game interfaces (a) the game with speech only (b) the game with earcons only (c) the game with speech and earcons (d) the game with speech, earcons and avatars.

Nevertheless, the system provided users with three attempts for every task with a built-in clock. Each user was allowed to take up to 60 seconds to complete the task.

IV. Platforms Design

Each platform designed so that the user has to read and listen to specific lesson and directly must go to tasks (questions) which are proposed to be game. The mechanism of this game starts when the users click specific button to start the game, the first question appears as text with speech in the middle of the circle shape in the centre of the screen. As the user read and listen the question, straightaway the first answer comes out as text with speech in another circle on the top of the screen for about 3-4 seconds.

As soon as the second answer becomes visible with text and speech in another circle, the first answer (circle) disappears and so on with the remaining six answers. The correct answers are distributed randomly within random circles in specific task. By fast click on the chosen answer, the system generates immediate feedback notifying if this was the right or wrong answer. The system automatically jumps to the next question if the user answer was right and all circles disappears, and if the user answer was wrong , the game continues until the maximum time consumed, prompting user to go to next question and the game starts over with the second task. This game can be seen as wall watch and the answers rotate in clockwise direction. Snapshots of proposed game showed in Fig. 1 (a).

In the second interface, the same design is repeated, but here speech was replaced with earcons and also the way the game starts here are differs from the previous game. In this game the user must click on the middle of the circle to see the questions and click on the tone button to hear the tone and start the game in the same time, bearing in mind that the question tone will be heard only once by user.

The tone used in these earcons was generated by software called visual music, by half the tones to two portions, the first half allocated to question and the other half assigned to the answer.

In this design the user has two ways to answer, first by reading the text only or by using or click on the tone button grouped with each circle or answer.

The remaining first game procedures described and time restrictions were used with this game as well. Snapshots of proposed game showed in Fig.1 (b).

Third game interface is similar to interface two, but speech was included. In this game the user must click on the middle of the circle to see and listen to the question and click on the tone button to hear the tone and start the game in the same time, bearing in mind again that the question tone will be heard only once by user.

Fourth interface or game is a third interface enriched with an avatar. The avatar (sound and human like expressions) with the other modals (text, speech, earcons) used in this platform to introduce an extra edutainment aspects displayed simultaneously on the screen. Besides reading to the user the questions and answers, avatars are designed and targeted to entertain the learner by jokes, some funny expressions and gesture. These avatars were integrated with Crazy Talk V5.1 software with the author speech and personal and other friend's photographs.

V. Results and Discussion

A. *Users Profile*

The average age of users was 36 years with a standard deviation of 25.46% and 100 % were males. Users were generally high educated level, Doctoral degree was 45.65%, and Masters was 58.70%. In area study list, 39.18% of users were from computing and informatics department, 21.74% were from engineering in general, whereas communication and networking were 15.22 %. The remaining users were from different schools and department. 100% are using computer more than 10 hours per week. Whereas 95.65% using Internet more than 10 hours per week.

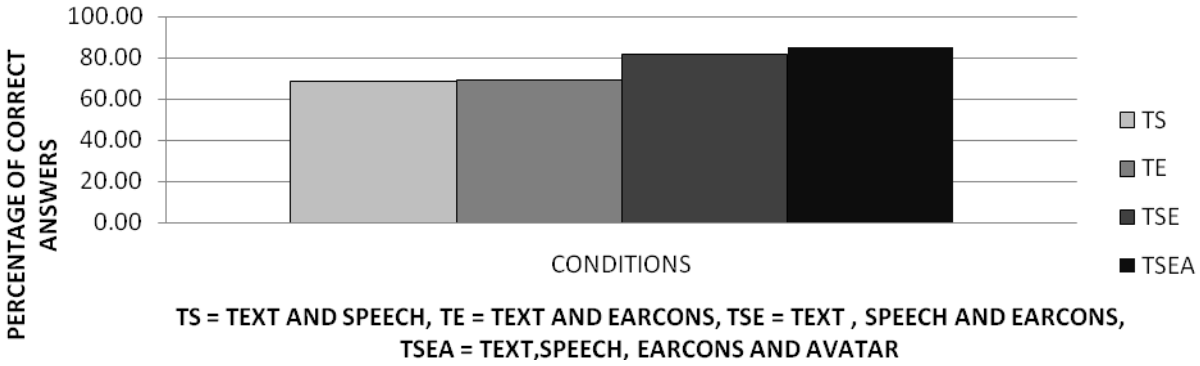


Fig. 2 Conditions TS, TE, TSE and TSEA average correct answers.

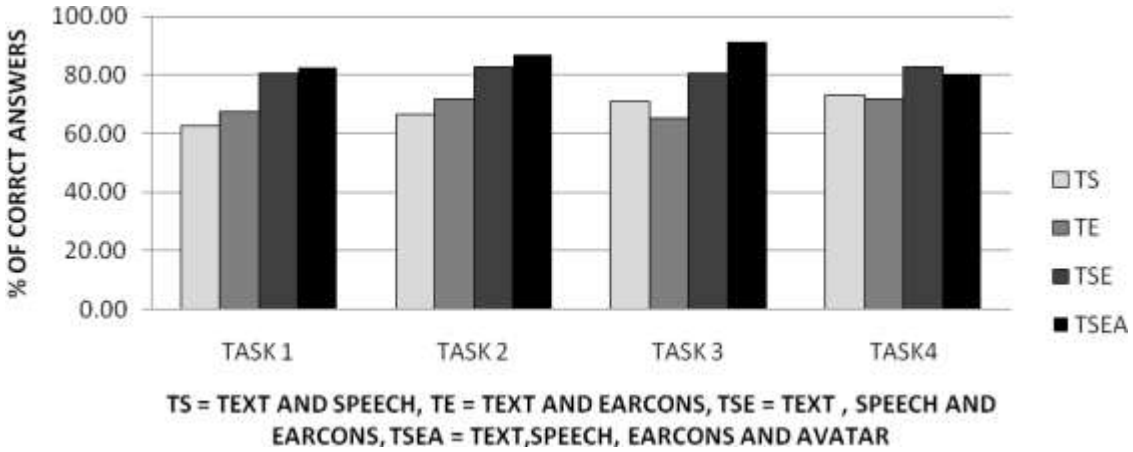


Fig. 3 Conditions TS, TE, TSE and TSEA correct answers for each task

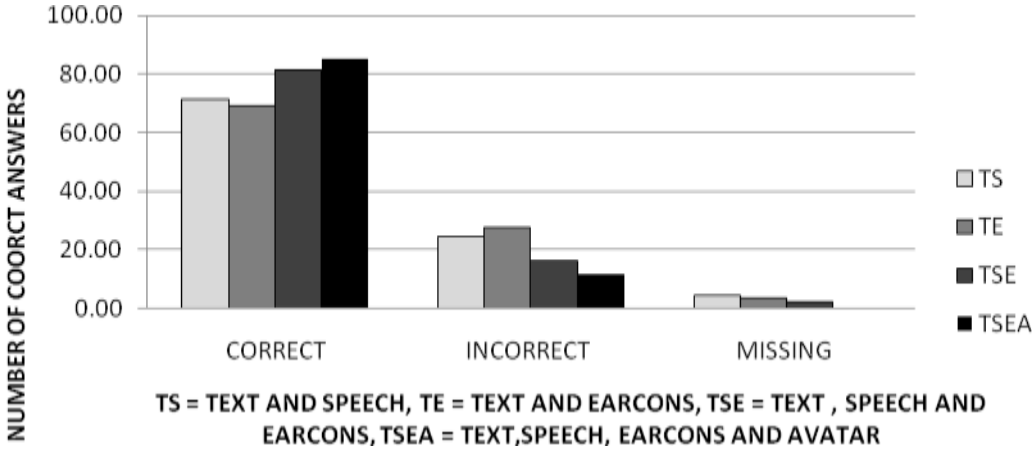


Fig. 4 Conditions TS, TE, TSE and TSEA correct, incorrect and missing answers.

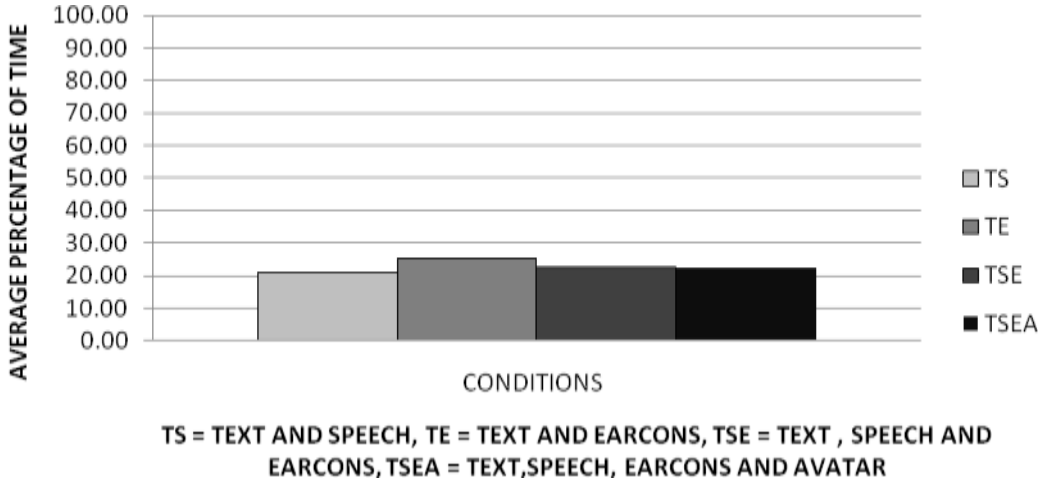


Fig. 5 Conditions TS, TE, TSE and TSEA average times

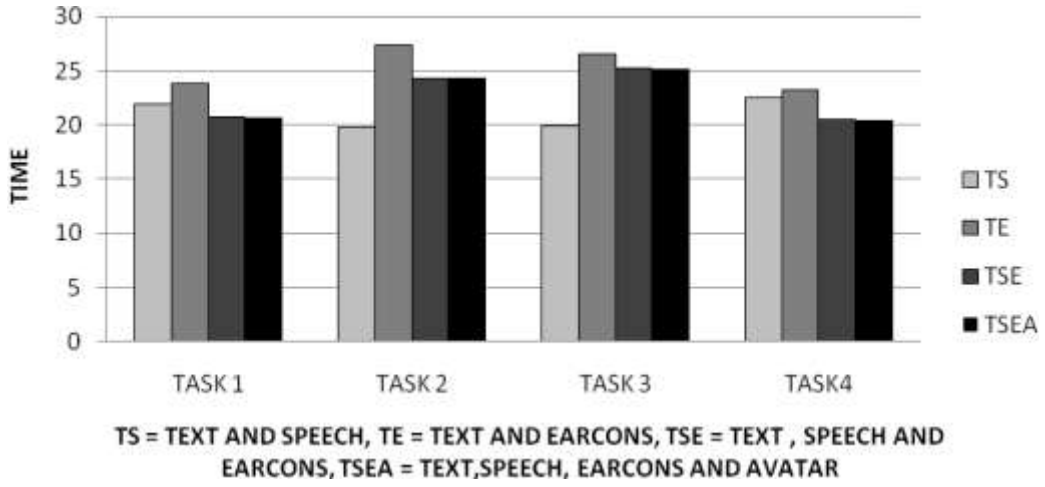


Fig. 6 Conditions TS, TE, TSE and TSEA average time for each task

4.35% of users only had excellent knowledge about Geology, 23.91% were good, 58.70% limited, and 13.04% had no knowledge at all. Users who had knowledge about e-learning were 28.26%. Concerning avatar 63.04% had not knowledge, 28.26% limited, 8.70% were good, 0% were excellent.

B. Effectiveness

A mean correct answer has been taken from each condition to show the general user performance, as illustrated in Fig. 2. It is obvious that TSEA condition had a better performance level than the other conditions. For TSEA, the figure was about 85.33%, whereas that for TSE was 81.52% and that for TE was 69.02%m and finally TS was 68.23%.

Furthermore TSEA condition is gained more percentage than others about correct answers, and the percentage of users whom answered incorrect was higher in condition TE followed by TS and finally TSE. On the task level, as shown in Fig. 3 below, task 1 mean correct answer percentage was 73.23%, task 2 was 76.99%, task 3 was 76.94% and task 4 was 76.92%. It is obvious that the higher correct answer percentage was in task 2 and the lowest was in task 1.

Moreover Fig. 4 states the number of correct answers compared to incorrect and missing answers, whereas the highest mean for the correct answer was in TSEA condition than the other conditions. In condition TSEA the correct answer mean was 85.33 %, TSE was 81.52 %, TE was 69.02% and TS was 71.20%.

These results are indication that users' performed better in condition TSE and TSEA than in conditions TS and TE because of the effect of multimodal metaphors as speech and earcons in condition TSE and the avatar in condition TSEA which interprets this successfulness.

C. Efficiency

Fig. 5 illustrates the average time (excluding reading, pre- and post-questionnaire) in all conditions. In condition TS the

average time was 21.04 seconds, condition TE was 25.26 seconds, condition TSE was 22.66 and condition TSEA was 22.47 seconds. The time spent in condition TE was a bit higher than in the other conditions. Moreover the time differences between tasks for the four conditions are depicted in Fig. 6.

Task 1 across all conditions mean time percentage was 22.34%, task 2 was 23.46%, task 3 was 23.67% and task 4 was 21.94%. It is clear that the time was approximately similar.

VI. Interfaces Preference

It is good idea to support and enrich the results documented by adding up single statement in the end of the questionnaire to let users choose the best interface was experienced. Fig. 7 displays the user's preference to each condition, were condition TSEA was the most platform chosen, followed by condition TSE, TE and condition TS was lowest in the graph.

VII. Interfaces preference order

Nevertheless, users required to order the interface according to their preference, in other words users have to place number 1,2,3 or 4 in the box provided in the questionnaire were is number 1 represents the best interface the user experienced and 2, 3 and 4 are the second, third and fourth. The results came keep going with interface 4 (TSEA condition) as shown in Fig. 8 which was the highest average 60.87% of user's choice. 19.57 % of users have chosen the condition TSE, 8.70% preferred condition TE and 10.87% chosen TS condition.

VIII. ANOVA Test

Based on the type of data and independent and dependent variables applied in this experiment, the prober test found and suggested by [31] is One-way repeated measures ANOVA which designed to compares how a within-subjects experimental group performs in three or more experimental conditions. The ANOVA compares whether the mean of any of the individual experimental conditions differ significantly from the total mean across the experimental conditions.

A. Friedman's ANOVA Test

Friedman's ANOVA test used for testing differences between conditions when there are more than two conditions and the same participant have been used in all conditions which is the similar with this experiment situation. Also the Friedman test is based on the fact that the samples should be randomly taken and independently of each other. Therefore this test has conducted to the time for the four conditions using the SPSS statistics software and the output of the result is shown in the Table 6.

Test Result:

The null hypotheses stated as follow:

H0 = there is no differences between the means of time of the four conditions.

$\mu1= \mu2=\mu3= \mu4$

H1 = there is differences between the means of time of the four conditions.

Since p-value = $0.18 \leq 0.05 = \alpha$, the null hypothesis accepted, and we can say that at the $\alpha = 0.05$ level of significance, there exists enough evidence to conclude that there is no difference in the true mean time recorded in rats for the four experimental conditions .

In addition, Friedman's ANOVA test has conducted to the number of correct - incorrect answers for the four conditions using the SPSS statistics software and the output of the result is shown in the Table 7.

Test Result:

The null hypotheses stated as follow:

H0 = there is no differences between the means of the correct - incorrect answers in the four conditions.

$\mu1= \mu2=\mu3= \mu4$

H1 = there is differences between the means of the correct - incorrect answers in four conditions.

Table I Friedman's ANOVA Test results for time

Test Statistics[a,b]

N	162.000
Chi-Square	4.793
df	3.000
Asymp. Sig.	.188

a. Some or all exact significances cannot be computed because there is insufficient memory.

b. Friedman Test

Table 2 Friedman's ANOVA Test results for correct -incorrect answers

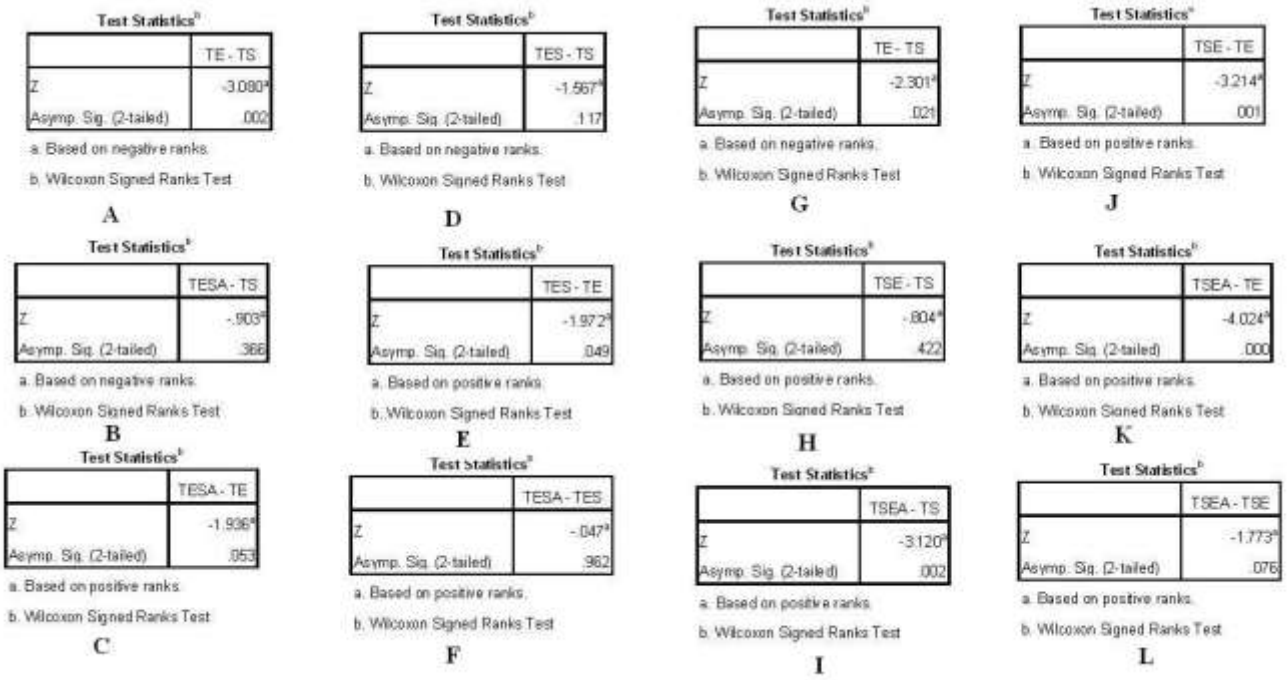
Test Statistics[a,b]

N	162.000
Chi-Square	29.803
df	3.000
Asymp. Sig.	.000

a. Some or all exact significances cannot be computed because there is insufficient memory.

b. Friedman Test

Fig. 9 Wilcoxon signed-rank test results



Test Statistics[b]

	TE - TS
Z	-3.080[a]
Asymp. Sig. (2-tailed)	.002

a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

A

Test Statistics[b]

	TES - TS
Z	-1.567[a]
Asymp. Sig. (2-tailed)	.117

a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

D

Test Statistics[b]

	TE - TS
Z	-2.301[a]
Asymp. Sig. (2-tailed)	.021

a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

G

Test Statistics[b]

	TSE - TE
Z	-3.214[a]
Asymp. Sig. (2-tailed)	.001

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

J

Test Statistics[b]

	TESA - TS
Z	-.903[a]
Asymp. Sig. (2-tailed)	.366

a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

B

Test Statistics[b]

	TES - TE
Z	-1.972[a]
Asymp. Sig. (2-tailed)	.049

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

E

Test Statistics[b]

	TSE - TS
Z	-.804[a]
Asymp. Sig. (2-tailed)	.422

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

H

Test Statistics[b]

	TSEA - TE
Z	-4.024[a]
Asymp. Sig. (2-tailed)	.000

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

K

Test Statistics[b]

	TESA - TE
Z	-1.936[a]
Asymp. Sig. (2-tailed)	.053

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

C

Test Statistics[b]

	TESA - TES
Z	-.047[a]
Asymp. Sig. (2-tailed)	.962

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

F

Test Statistics[b]

	TSEA - TS
Z	-3.120[a]
Asymp. Sig. (2-tailed)	.002

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

I

Test Statistics[b]

	TSEA - TSE
Z	-1.773[a]
Asymp. Sig. (2-tailed)	.076

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

L

Since p-value = 0.00 ≤ 0.01 = α, the null hypothesis rejected, and we can say that at the α = 0.01 level of significance, there exists enough evidence to conclude that there is a difference in the true mean correct - incorrect answers recorded in rats for the four experimental conditions and the treatment conditions used were effective.

B. Post hoc tests for Friedman's ANOVA

1) Wilcoxon signed-rank test

To see if there is significant difference in learning time using different training methods, Friedman's ANOVA employed compared whether the mean of any of the individual experimental conditions differ significantly from the aggregate mean across the experimental conditions, but this test do not give us indication of which condition is significantly the best, in other words the target of the study is to know the condition that the user perform significantly better in comparison with others. Wilcoxon signed-rank test is based on the differences between scores in the two conditions which going to compare. Once these differences have been calculated they are ranked, but the sign of the difference (positive or negative) is assigned to the rank. Therefore in this experiment each two condition have been tested separately, since four conditions are examined here, so the comparison distribution was as follow:

- Condition TS with condition TE
- Condition TS with condition TSE
- Condition TS with condition TSEA
- Condition TE with condition TSE
- Condition TE with condition TSEA
- Condition TSE with condition TSEA

Fig. 9 (A) below determines the test statistics, Z score are calculated from the T score that was the lowest value of sum of negative and positive scores, the value of Z was -3.080 and this value is significant at p= 0.002, therefore, because this value is based on the negative rank, it is safe to say that the time spent per user increased in the direction of TE than in the TS (Z= -3.080, p<.05) see Fig. 9 (A).

The value of Z in TS with TES comparison was -1.567 and this value is significant at p=0.117, therefore, because this value is based on the negative rank, therefore, the time used up was increased significantly in the TSE than in the TS (Z= -1.567, p<.05), Fig. 9 (D) summarises the test.

A comparison between TS and TSEA, were based on negative ranks, the time consumed per users is significantly increased in TSEA than in the TS (Z= -.903, p<.05) as shown in Fig. 9 (B).

The value of Z in TE with TSE comparison was -1.972 and this value is significant at p=0.049, therefore, because this value is based on the positive rank, therefore, the time used up was decreased significantly in the TE than in the TSE (Z= -1.972, p<.05), Fig. 9 (E) summarises the test.

The value of Z in TE with TSEA comparison was -1.936 and this value is significant at p=0.053, therefore, because this value is based on the positive rank, therefore, the time used up was decreased significantly in the TE than in the TSEA (Z= -1.936, p<.05), Fig. 9 (C) summarises the test.

The value of Z in TES with TSEA comparison was -0.047 and this value is significant at p=0.962, therefore, because this value is based on the positive rank, therefore, the time used up was decreased significantly in the TSE than in the TSEA (Z= -.049, p<.05), Fig. 9 (F) summarises the test.

Fig. 9 below also contains results of the test statistics of correct -incorrect answers for TS and TE condition, the value of Z was -2.301 and this value is significant at p= 0.021,

therefore, because this value is based on the positive rank, so the incorrectness decreased in the TS than in the TE (Z= -2.301, p<.05), see Fig. 9 (G).

In the case of TS with TSE comparison, Z was -.804 and this value is significant at p= .422, and because this score was based on negative ranks, the users made more incorrect answers in TS than in the TSE (Z= -.804, p<.05), Fig. 9 (H) summarises the test.

The value of Z in TS with TSEA comparison was -3.120 as shown in Fig. 9 (I) and this value is significant at p=0.002, therefore, because this value is based on the positive rank, therefore, the incorrectness was significantly lower in the TSEA than in the TS (Z= -3.120, p<.05).

In the case of TE with TSE comparison, Z was -3.214 and this value is significant at p= .001, and because this value is based on the positive rank, therefore, the incorrectness was significantly lower in the TSE than in the TE (Z= -3.214, p<.05), Fig. 9 (J) concludes the test.

Moreover the value of Z in TE with TSEA comparison was -4.024as shown in Fig. 9 (K) and this value is significant at p=0.000, therefore, because this value is based on the positive rank, therefore, the incorrectness was significantly lower in the TSEA than in the TE (Z= -4.024, p<.05).

Finally the value of Z in TSE with TSEA comparison was -1.773 as shown in Fig. 9 (L) and this value is significant at p=0.076, therefore, because this value is based on the positive rank, therefore, the incorrectness was significantly lower in the TSEA than in the TSE (Z= -1.773, p<.05).

IX. Discussion

This empirical study investigated the user's involvement whilst experiences the edutainment features in multimodal e-learning systems, this was throughout two quantitative aspects, in particular effectiveness and efficiency. Bearing in mind that the experimental circumstances is controlled to guarantee the validly, either by the platforms rotation method used or by lessons distribution mechanism provided. The experimental dependent variable measurements have been managed in order to acquire accurate results, that is has been done by incorporate fitted time design in all systems. The outcome reported was positive and the tests accomplished also resulted significant outcomes. Analyzed data collected award us many excellent points to be covered here supported with in depth explanation.

Users concerning effectiveness did very well in condition TSE and TSEA and further less in conditions TS and TE. These results are indication that users' performed better in condition TSE and TSEA than in conditions TS and TE because of the effect of multimodal metaphors as speech and earcons in condition TSE and the avatar in condition TSEA which interprets this successfulness.

The average time consumed by the users was less in conditions TS, TSE and TSEA and was higher in condition TE. Nearly equivalent time users spent in conditions TS, TSE and TSEA as an average was 22.05 seconds.

Moreover number of serial tests (ANOVA) made upon data in order to find out the differences between experimental platforms in terms of time and correctness. Fortunately these tests resulted significant variance and the null hypotheses were rejected. A Wilcoxon test resulted also differences between platforms and as results showed above, the incorrectness and time was significantly lower in condition TSEA than in the TSE, TE and TS conditions.

Concerning users' preference, condition TSEA was the mainly platform selected, followed by condition TSE, TSE was the third one and condition TS was the lowest. Nevertheless, a user Interfaces preference order also supports with interface TSEA (condition) which was the highest average 60.87% of user's option. 19.57 % of users have chosen the condition TSE and 8.70 % preferred condition TE and finally10.87% selected condition TS.

As result of the data analyses there is no doubt that the condition TSEA (The game with test, speech, earcons and avatars interface) was the superlative platform. The second platform was TSE (The game with test, speech, earcons interface), third position was for TS (The game with test, speech interface) and finally TE (The game with test, earcons interface) was the most undesired platform. Although all users enjoyed all conditions as power of the game in conveying the learning materials, however users preferred TSEA for the reason that of avatars which incorporated to add the fun elements that improves better the user's mode and reduces the stress during playing. This also strengths the power of game in enriching the entertainment environments with educational characteristics as many researchers has been proved.

X. Conclusion

This study evaluated four "edutainment" interfaces through learning by playing environment. These environment are TS (The game with text, speech interface only), the second platform was TE (The game with text, earcons interface only); third proposal was TES (The game with text, speech, earcons interface) and finally the forth interface was TSEA (The game with text, speech, earcons and avatars interface). Effectiveness and efficiency attributes were measured to investigate the role of integration of multi-modal metaphors within the game play environment on general users' enjoyment and consequently their performance.

These four interfaces and four lessons are randomly rotated between dependent groups of users. Users' effectiveness and efficiency results were collected and analysed and Likert five-point scale with 10-items table were also analysed and reported. Moreover ANOVAs test was applied to the data and user interface preference and order were also statistically measured.

The experimental results showed that the TSEA (The game with test, speech, earcons and avatars interface) interface outperformed all other interfaces on the subject of users' correctness and time consuming elements. In addition TSEA interface was the platform with the highest number of users' preference. Nevertheless, users' interface preference order also shows supports for the TSEA interface which had the highest average. The reason behind preferring TSEA by users can be referred to enjoyment and satisfactions users attained form the interaction with avatars accompanied with jokes and funny gesture. But also the results ensured that all conditions

were satisfied and this can be committing to the power of the game in conveying the learning materials strengthen with multi-modal metaphors which are already demonstrated its potential effect as mentioned earlier in enhancing usability of various interfaces.

REFERENCES

[1] Clark, K., et al., Making Games After-school: Participatory Game Design in Non-Formal Learning Environments. 2009.
[2] Prensky, M., Don't bother me, Mom, I'm learning!: How computer and video games are preparing your kids for 21st century success and how you can help. Paragon House, Minnesota, 2006.
[3] Prensky, M., Digital game-based learning. Computers in Entertainment (CIE), 2003. 1(1): p. 21.
[4] Wiberg, C., Usability and fun: An overview of relevant research in the HCI community
[5] Monk, A., Fun, communication and dependability: extending the concept of usability. People and computers XVI: memorable yet invisible, 2002: p. 3.
[6] Foroughi, R. Work in Progress: Giving Asynchronous Learning Platforms a Humanoid Head for Effective Communication with the Learners. 2006.
[7] Hogle, J., Considering games as cognitive tools: In search of effective "Edutainment.". Information Resources Report No. IR019225).(ERIC Document Reproduction Service No. ED425737), 1996.
[8] Beskow, J. Talking heads-communication, articulation and animation. 1996: Citeseer.
[9] Villa-Uriol, M., et al. Automatic creation of three-dimensional avatars. 2003: Citeseer.
[10] Squire, K. and H. Jenkins, Harnessing the power of games in education. Insight, 2003. 3(1): p. 5-33.
[11] Kara, Y. and S. Ye ilyurt, Comparing the Impacts of Tutorial and Edutainment Software Programs on Students' Achievements, Misconceptions, and Attitudes towards Biology. Journal of Science Education and Technology, 2008. 17(1): p. 32-41.
[12] Wiberg, C. and K. Jegers. Satisfaction and learnability in edutainment: a usability study of the knowledge game 'Laser Challenge'at the Nobel e-museum. 2003: Citeseer.
[13] Pinhanez, C., et al. Can Web entertainment be passive. 2001: Citeseer.
[14] Wiberg, K.J.C. FunTain: Design Implications for Edutainment Games. in In proceedings of ED-MEDIA 2003, Association for the Advancement of Computing in Education, Charlottesville, VA. 2003.
[15] Seiersen, K. and M. Nielsen, EuroPhysicsFun on the very frontier of science edutainment. Europhysics News, 2005. 36(4): p. 141-142.
[16] Sim, G., S. MacFarlane, and M. Horton. Evaluating Usability, Fun and Learning in Educational Software for Children. 2005.
[17] Gregor, S., Designing Websites for Learning and Enjoyment: A study of museum experiences. 2006.
[18] Wiberg, C. Fun in the home: Guidelines for evaluating interactive entertainment on the web. 2005.
[19] Cartwright, W., Exploring Games and Gameplay as a Means of Accessing and Using Geographical Information. Human IT, 2006. 8(3): p. 28-67.
[20] Ormeling, F. Ariadne's thread–structure in multimedia atlases. 1993.
[21] Tabbers, H., R. Martens, and J. Van Merrienboer, Multimedia instructions and cognitive load theory: Effects of modality and cueing. British Journal of Educational Psychology, 2004. 74(1): p. 71-81.
[22] Brewster, S., P. Wright, and A. Edwards. Experimentally derived guidelines for the creation of earcons. 1995.
[23] Rigas, D. and M. Alseid. Multi-modal aided Presentation of Learning Information: A Usability Comparative Study. 2008.
[24] Alsuraihi, M. and D. Rigas, Speech Displaces the Graphical Crowd. Journal of Computers, 2008. 3(6): p. 47.
[25] Alsuraihi, M. and D. Rigas. Efficiency of speech recognition for using interface design environments by novel designers. 2007: World Scientific and Engineering Academy and Society (WSEAS).
[26] Alseid, M. and D. Rigas. Efficiency of multimodal metaphors in the presentation of learning information. 2008.
[27] Alseid, M. and D. Rigas. Users' views of facial expressions and body gestures in e-learning interfaces: an empirical evaluation. 2009: World Scientific and Engineering Academy and Society (WSEAS) Stevens Point, Wisconsin, USA.
[28] Gazepidis, N. and D. Rigas, Evaluation of Facial Expressions and Body Gestures in Interactive Systems. NAUN International Journal of Computers. 1: p. 92-97.
[29] Brewster, S., Providing a structured method for integrating non-speech audio into human-computer interfaces. Department of Computer Science, 1994.
[30] Brewster, S., P. Wright, and A. Edwards. A detailed investigation into the effectiveness of earcons. 1994: Citeseer.
[31] Field, A., Discovering statistics using SPSS. 2009: Sage Publications Ltd.