Abstract—E-learning is a sophisticated computerised learning issue that requires exploration from many points of views. It involves various domains, such as social, emotional, psychological, and other related areas. In this work a multi edutainment platforms were evaluated in order to explore usability aspects of edutainment in e-learning. These three platforms were Virtual classroom, Game-based and Storytelling. Entertainment features represented by speech, avatars with graphics and game activity was combined in these three different platforms. Moreover earcons were incorporated to help users resolve problems given efficiently and quickly. A dependent group was used to measure users’ performance in terms of user achievement, correct answers and satisfaction. The experimental results collected indicated that the Game-based approach was the best in terms of users’ achievement effectiveness as well as satisfaction with the interface and overall learning experience.

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

I. INTRODUCTION

Educatonal institutions today are obliged to comply with globalisation. Learners, on the other hand, are required to pursue the development of digital technology to become a part of modern society. This new means of information flow has changed learners’ attitude to online educational resources [1]. Countless studies from various disciplines, such as human-computer interaction and social and cognitive psychology, demonstrate that students’ attitude and emotions have a very important effect on student behaviour in online learning [2], [3], [4], [5], [6], [7], [8], [10], [13], [16], [18], [19], [20], [21], [22]. Therefore, e-learning as a field of recent concern brings as much change to instructors as it does to students, requiring a new set of skills that replaces traditional practices [4].

Furthermore, e-learning is not only an educational issue but also a complex experience that involves various domains such as social, emotional, psychological, and other related areas [23]. Many e-learning systems still present insufficient information, forcing students to navigate too quickly to less text-based and more interactive sites [24]. Online learners today require more than mere information; they need to fill the gap missing in e-learning systems, an environment similar to that experienced in traditional learning, to interact with someone who could represent their teacher, to feel that they are in a classroom learning from other students, experiencing enjoyment during their learning, emotionally involved with the lesson and the instructor and learning from the teacher’s facial expressions, gestures, hand movements and other activities that can convey the message[25].

There are, in fact, two types of learners today, those who have experienced digital learning and those who are starting to do so [1]. Digital learners spend the majority of their time on games and this could be referred to as the enjoyment they are obtaining. Where mature learners are still not comfortable with this new medium, however, this might be due to their missing the element of interaction or to the developer paying less attention to this class of society [26].

The intention of this study is to design edutainment platforms that can be used for the empirical evaluation of entertainment; with this as the focus; tools were required to demonstrate the amount of entertainment experienced by participating users. The motivation was to measure effectiveness, efficiency and user satisfaction and to evaluate how memorable and educationally valuable were certain aspects of edutainment. The results obtained from the experimental studies were compared and discussed in order to produce empirically derived conclusions as to which platform is preferable as a standard for e-learning applications.

II. PREVIOUS WORK

A number of studies have been conducted to investigate the role of entertainment and positive emotions in enhancing learning in general, such as [29], [30], [31], [32], [33], [34], [35], [36], [37]. In this study a selected number of important related work is presented in the next section.

One work [29] has investigated museum websites that present educational materials. These websites are designed to provide the general public with educational materials through enjoyment. In this study five people participated, from the field of museum-website design experts in Taiwan. The approach used in this study was descriptive-qualitative and
was based on semi-structured in-depth interviews and expert interviews as the primary method.

The interviewees gave some important suggestions for motivating learning through a museum’s e-learning websites. This include: 1) the importance of an attractive appearance; 2) the value of increased interaction with learners; 3) ease of use; 4) asynchronous accessibility; 5) the value of relaxation and performing short tasks; and 6) the usefulness of hyperlinks. Furthermore, the study derived five development guidelines for the design of learning through enjoyment in museum websites and noted a number of aspects, including: 1) the adoption of multimedia and interactive technologies; 2) consideration of the characteristics of self-directed learning; 3) the importance of qualified staff and adequate financial support; 4) identifying a target audience; and 5) the importance of making information more sharable. Although these findings are targeted at museum websites, they can be applied generally to e-learning systems.

In another study [30], the target was to explore the influence of tutorial and edutainment design of educational software programs that present the topic of “cell division” on student achievements, misconceptions and attitudes. The cell-division achievement test (CAT), cell-division concept test (CCT) and biology attitude scale (BAS) were applied at the beginning and at the end of the research. A total of 72 students took part in the study (age range 14-15 years). The users were distributed randomly in three groups within three ninth-grade classes of a public secondary school. Two were experimental groups and one was the control group. In each group there were 24 students. Users in all groups had knowledge about computers, but no experience of learning with CAI. The control group practised a traditional teaching method, while the experimental groups were educated through computer-based learning.

The study showed increases in general achievement in CAT for the experimental groups. Students understood the general functions of mitosis and meiosis easily since instructional software programs were obvious and effective. In addition, the study discovered that misconception in the experimental groups is not entirely eliminated even after the treatment. Nevertheless the study confirmed that using edutainment software program noticeably alters students’ attitudes towards biology materials.

III. EXPERIMENTS AND METHODS

Two interfaces have been designed and developed. The first multi-modal E-learning platform encompassed edutainment elements. The second interface was non-edutainment interface. Everything in both interfaces is similar in terms of content and number of tasks but differs in the means of presenting the information or material. Please refer to Table 1 shows the features incorporated in both interfaces.

Each platform contains two chapters (Case 1, Case 2) in which difficulty gradually increases from Case 1 to Case 2. To avoid any familiarity with the topic and the interface sampling in the experiment, random rotation technique was applied between platforms and also in terms of chapters (Cases). As seen in Table 2, the user may starts from Case 1 in edutainment interface and finishes in Case 2 in non-edutainment interface, and the opposite is true. Tasks were divided into 2 groups, the first recognition questions and the second recall. Again tasks increased in term of difficulty (easy - moderate - hard). Table 3 summarised the tasks executed by users’.

The user must use both interfaces and they have to decide which interface is better and enjoyable through a feedback. The study targeted high level educational students (Master & PhD Students), and the subject matter examined as E-learning content was Human Computer Interaction (HCI).

Considering that, HCI as a subject matter is mainly theoretical. Learners need something to watch and listen to improve their imagination and engagement. This can be done by entertaining users with the learning materials. To achieve this goal, an idea employed was to adapt Avatar (sound and human like expressions) as an assistance tool to convey the message to the student by incorporating amusing elements such as jokes, facial expression and body gesture.

The word avatar comes from the Sanskrit language [20] and can be translated as God’s Incarnation on Earth. In the virtual reality community, avatars are 3D humanoid characters inhabiting virtual space, with varying degrees of animation and behavioral abilities. Avatars typically represent humans who visit the space virtually. Each visitor controls their avatar and is aware of other visitors’ avatars and their actions. The avatar moves and gestures in the mirror as the experimental subject moves and gestures in the physical room [20], [21]. Furthermore interesting possibilities can be offered by avatar world for online learning, collaboration, discovering new environments and attracting learners to keep progressing. Both educator and students can also build their own virtual worlds. This creates a sense of realism that is often absent from distance learning, which has been considered a benefit of educational three dimensional virtual [22].

IV. HYPOTHESIS AND METHODOLOGY

The Null hypothesis H0 stated in this study was “Multimodal E-learning and Edutainment Systems will be the same (does not enhance students memory)”. To verify the proposed hypothesis, three interfaces were designed and built, the first one was Virtual class, presenting the information in way similar to being in a classroom with a teacher and a students’ discussion blackboard with amusing questions and answers that make the environment more interactive. The second interface presented the subject in a Game-based learning interface. Storytelling was the third interface used in this experiment to present information.

The users must use these three interfaces (dependent group) and they have to say which interface is the best and enjoyable through their feedback.

V. PILOT STUDY

Three different geology lessons were rotated between the three platforms Table 1. Each lesson had similar volume of information. Six people took part in the pilot study. All the users The experiment was recorded using software called Camtasia in order to review the session and discover any mistakes later. Users completed the final questionnaire. They
then expressed their comments on aspects of the experiment. All users commented that the lessons were relatively long, although they felt that the contents were relevant. Additionally, the exercise was not easy to answer even for those with experience in the subject matter. However, most users agreed that the interface was easy to use and enjoyable. On the questionnaire itself, most commented that it was easy and well organised. Changes were made in response to the users’ feedback and lessons were shortened to avoid overload. Exercises were reviewed and modified to match the lessons’ contents, considering the differing users.

VI. EXPERIMENTAL TASKS

Each user was provided with six tasks (questions) to perform in each platform. Tasks were moderate in terms of difficulty. 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.

VII. PRE-EXPERIMENTAL AND TRAINING TIME

To avoid users becoming confused or anxious, and also because they were from various disciplines, such as biology or medicine, it was important to get them to relax, by informing them of the goal of the experiment and encouraging them not to worry about whether they had any knowledge about geology or not, and to calm them by telling them that this was not a test of their abilities, but an investigation of the differences between the platforms developed. As training, users were given five to 10 minutes to discover the three environments before beginning the experiment. Moreover, it was felt helpful to explain to the users orally what he/she had to do after they had read the instructions for the questionnaire. Although everything was written clearly and simply in the questionnaire, users were reminded orally about time limitation and the order of the tasks and answers were provided to specific questions that users wanted to ask.

VIII. EXPERIMENTAL DESIGN

A. Virtual classroom

Avatar (sound and human-like expressions) was the main modal used in this platform to introduce the entertainment aspects with text as an additional modal displayed simultaneously on the screen. Here the idea was to create a learning environment similar to a real classroom where teacher and students interact with each other during the class to help in understanding the lesson, and to enrich the system with a third dimension, that of entertainment. This was done by interchanging the questions and answers between a teacher and the students, building valuable dialogue and sometimes amusing conversation that might increase student learning quality. The Virtual class designed to mimic the real classroom included a teacher with board, chairs, tables, floor carpet and wall colour arranged and organised in a clear manner.

Avatars representing the teacher and students were put forward to read to the user the subject matter. At the same time learners were thus entertained by jokes, facial expressions, gestures and short messages as feedback from the teacher and sometimes as questions or comments from students. See Fig 1 (a), (b) and (c).

Speech was included to enrich the system with necessary learning elements. Nevertheless, the system switched between teacher and student in a rotational manner according to the priority. The avatar dressed in a dark colour, such as blue or dark grey against a light brown colour for the board. Under the avatar was a clock counting down to show the user the time remaining for avatar to complete the task and a special button to stop the avatar when the user chose not to continue. There was also an extra button, of a different colour and shape, for entering the quiz section.

To test the knowledge retained by the learners, users were asked six questions per lesson. These were built into the system, each question considered as one task, and to answer the question the user had a maximum of 60 seconds. Whatever the user believed was correct, he/she had to click on the proper answer immediately to check the response in the form of amusing sounds and text. If he/she answered correctly, the user had to click on the next button to go to the second question (task). In the same manner, the user had to complete all the remaining tasks.

B. Game-based learning

The method used in this platform differed completely from the previous interface. Put simply, the user had to read and listen to a specific lesson and go directly to tasks (questions) designed as a square game, since it was composed of squares. This game was implemented by the same software mentioned above.

In the early phase, the game appears as six squares (boxes), with the main box (rectangular), in the top middle of the screen, allocated for the question. The question appeared only when the user shifted the mouse over the box. As the user read the question, he/she immediately had to move the pointer over the remaining boxes (six squares) to find the correct answer which was distributed randomly in the boxes. Besides the text-based answer, the game provided earcons to help in answering questions should the user be unsure of the correct answer. The tone used in these earcons was developed by software called Visual Music. By halving the tones to two portions, the first half was allocated to the question and the second half to the answer. Therefore, to find the correct tone, the user must click on any of the tone buttons allocated to each square or answer. The system automatically jumped to the next question if the user’s answer was correct and the box disappeared. Conversely, the number of boxes rose to seven if the user’s answer was wrong. The number of boxes would continue to increase if the answer remained incorrect, with nine boxes as a maximum, prompting the user to go to the next question and the game started over.

This made the probability of finding the correct answer low, forcing the user to find the correct answers during the first few attempts. Fig 1 (d), (e) and (f) illustrate snapshots of the proposed game.
C. Storytelling

Storytelling, as a means of conveying information to people, is considered one of the most effective ways, in the edutainment aspect of the learning medium, of teaching students different subjects. The idea adopted from some researches, such as Egan, #15; Gruen, #21, and tested here as an edutainment interface, presented the science information (geology) in an interactive way. The system articulates the information vocally with the text as an extra channel, associated with pictures, graphs, dynamic diagrams and movies that mimic and disclose the information offered. Certain circumstances determine the attractiveness of the means of conveying the information. For example, the presenter should stop for a while after every part of information, or raise or change his voice at times when the information is more important. So this interface is designed to hold one lesson, this lesson being divided to pages, each page containing some information as speech, text and graphs. The user must press the next button to navigate to the next portion of information. Additionally the learner is provided with the facility of playing, stopping and going back to a specific point at any time. Fig 1(g), (h) and (i), demonstrate a snapshot of this interface. As in the procedure employed in the Virtual class interface, the same type of timed tasks was allocated to this platform.

IX. RESULTS AND ANALYSIS

Overall, 48 volunteers took part in the study, held over four weeks, mostly in the Bradford University research laboratory. All the users utilised the three lessons and three edutainment conditions. Conditions and lessons were distributed randomly but were the same within each user.

X. USERS PROFILE AND PERSONALITY

Users were aged 18-54; the average age was 36 with a standard deviation of 25.26. A total of 95.83% of users were male. Users were generally highly educated. A total of 43.75% had a doctoral degree while 56.25% had a Masters, whereas 2.08% were undergraduates. In terms of the area of study, 43.75% were from computing and informatics and 12.9% were from engineering. The remaining users had other specialisations. A total of 99.75% used computers for more than 10 hours per week, whereas 91.67% used the Internet more than 10 hours per week.

Only 2.08% of users had an advanced knowledge of geology, 25% had some knowledge, 54.17% a very limited knowledge and 18.75% no knowledge at all. A total of 33.33% of users had knowledge of e-learning.

XI. EXPERIMENTAL SESSIONS

Experimental sessions took 30-45 minutes with an average time of 37.5 minutes (standard deviation 10.60 minutes), including pre- and post-questionnaire.

The time was distributed as follows: participants started with the pre-session questionnaire for an average of three minutes and then read the tasks for an average of four minutes, performed the experiment for an average of eight minutes, followed by the feedback questionnaire for an average of two minutes. This time distribution was similar for the other platforms.

XII. TASKS ACHIEVEMENT

Fig. 2 shows that the mean user achievement for Game-based was higher in comparison to Virtual class and Storytelling platforms. Respectively, the figures are 79.86%, 97.9% and 85.76%. The proportion of users who completed their tasks without any mistakes was as follows: Virtual class 77.08%; Game-based 89.6%; and Storytelling 45.80%. In general, as depicted in Fig. 2, better user performance observed in Game-based condition.

In terms of each task on its own, as shown in Fig. 2, the mean percentage of students who completed tasks one, two, three and four were 94.81 in Virtual class condition and slightly less in tasks five and six.

Overall, participants performed slightly better in the first four questions in all conditions compared with tasks four and five. But figures were slightly lower in Storytelling where the average was 79.17% for tasks four, five and six, as seen in Fig. 2.

XIII. EFFECTIVENESS

A mean correct answer has been taken from each condition to show the general user performance, as illustrated in Fig 3. It is obvious that Game-based here had a better performance level than the other conditions.

For Game-based, the figure was about 78.82%, whereas that for Virtual class was 53.82% and that for Storytelling was 44.10%. On the task level, the number of correct answers in Game-based was higher than in the other two conditions. As shown in Fig 3 below, Virtual class mean correct answer percentage was 53.82. In contrast, the figure for Game-based was 78.82% and that for Storytelling was 44.10%. It is obvious that the correct answer percentage decreases gradually from Game-based to Virtual class and finally to Storytelling.

XIV. SATISFACTION

Uses satisfactions was measured using a questionnaire. The questionnaire used Likert five-point scale with 10-items [40]. Users were asked to express their agreement with specific statements. The average score for conditions respectively was 75.26%, 73.4%, 64.48, see Fig 4. In addition to the standard condition, the Likert five-point scale was enriched by an extra five statements that also expressed user opinion scored as a normal average. In Virtual class, the average user score was 3.37, in Game-based it was 3.75 and in Storytelling it was 3.17, please look at Fig 5.
Fig. 1 (a) Virtual class condition (teacher); (b) Virtual class condition (students); (c) Virtual class condition (female student); (d) The storytelling interface (example 1); (e) The storytelling interface (example 2); (f) The storytelling interface (example 3); (g) The game in early phase; (h) The game when the user answers correctly; (i) The game when the user answers incorrectly.

Fig. 2 Achievement for all conditions in general, achievement for users who completed all tasks in all conditions, achievement for all conditions for each task separately.
XV. INTERFACE PREFERENCE

It is a good idea to support and enrich the documented results by including responses to a single statement at the end of the questionnaire allowing users to choose the best interface experienced. Fig 6 displays the users’ preferences for each condition, where Game-based had the most platforms chosen, followed by conditions one and three.

XVI. INTERFACE PREFERENCE ORDER

Users were required to order interfaces according to their preference, by placing numbers one, two or three in the box provided in the questionnaire, where number one represents the best interface the user experienced and numbers two and three their second and third preferences.

The results show that game-based interface had the highest average of users’ choices, at 58.33%, as shown in Fig 6. A total of 25% of users chose Virtual class and 16.67% preferred Storytelling.

XVII. ANOVA TEST

Based on the type of data and independent and dependent variables applied in this experiment, the proper test, found and suggested by [41], is a one-way repeated-measures ANOVA which is designed to compare 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 differs significantly from the total mean across the experimental conditions.
A. Friedman’s ANOVA test

Friedman’s ANOVA test has been conducted to the number of incorrect answers for the three conditions.

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

B. Post-hoc tests for Friedman’s ANOVA

1) Wilcoxon signed-rank test

To see if there is a significant difference in incorrect answers of the different training methods used, Wilcoxon signed-rank test has been done for each two conditions separately. Three conditions were examined and the comparison distribution was as follows:

1. Virtual class against Game-based;
2. Game-based against Storytelling;
3. Virtual class against Storytelling.

The results of the test for incorrect answers of $Z$ value was $-5.97$ and this value is significant at $p= 0.0$. Therefore, because this value is based on the positive rank, so the incorrectness decreased in the game compared with the Virtual class ($Z=-5.97$, $p<.05$). In the case of the Storytelling/Virtual class comparison, $Z = -7.83$ and this value is significant at $p= .433$. Because this score was based on negative ranks, the users made more incorrect answers in Storytelling than in the Virtual class ($Z=-.783$, $p<.05$).

The value of $Z$ in comparing Storytelling with Game-based was $-6.290$ and this value is significant at $p=0.0$. Therefore, because this value is based on the positive rank, the incorrectness was significantly lower in the Game-based than in the Storytelling ($Z=-6.290$, $p<.05$).

XVIII. DISCUSSION

In general, better user performance is noticeable in Game-based, but in Virtual class the result was less but still acceptable. It can be reasoned that this positive achievement of users in the Game-based quiz was due to earcons enhanced that helped users to retain information over longer period of time. Respectively, these were 79.86%, 97.9% and 85.76%. Whereas 77.08% of users completed their tasks without any mistakes in Virtual class, in Game-based this figure was 89.6% and in Storytelling was 45.80%.

Moreover, users did very well in Game-based, being approximately 78.82% correct, whereas in Virtual class this figure was considerably less, at approximately 53.82%, and even less in Storytelling, at 44.10%. These results are an indication that users performed better in Game-based than in Virtual class and Storytelling. In addition, it is obvious that the correct answers percentage decreases gradually from Game-based to Virtual class and Storytelling. Although most users had a limited knowledge of geology, these results are, to some extent, reasonable, taking into account the effect of earcons in Game-based which interprets the level of success.

An SUS-score [40] comparison shows that this was higher in Virtual class than in the other conditions. The average score for Virtual class was 75.26%, for Game-based was 73.39% and for Storytelling was 64.47%. On the other hand the additional statements provided showed that user satisfaction was higher 3.75 in Game-based, then 3.37 in Virtual class and 3.17 in Storytelling. On overall, the users’ satisfaction was almost as high in both Virtual-class and Game interfaces, but considerably less in Storytelling interface.

Moreover, a number of serial tests (ANOVA) were made on the data in order to find out the differences between experimental platforms in terms of correctness. Fortunately, these tests resulted in significant variance and the null hypotheses were rejected for correctness. Nevertheless, a Wilcoxon incorrectness test result was significantly lower in the game than in the Storytelling and Virtual class conditions.

As far as users’ preferences are concerned, Game-based was the platform with the highest number of users’ preference, followed by Virtual class and Storytelling. Nevertheless, users’ interface preference order also shows supports for the Game-based which had the highest average (58.33%) of users’ options. In comparison, 25% of users chose Virtual class and 16.67% preferred Storytelling.

As a result of the data analysis, there is no doubt that condition two (Game interface) was the best platform compared with the Virtual class and Storytelling. Although the Virtual class came second, users preferred Game-based approach rather than listening passively to the lessons. This also emphasises the power of the game in conveying learning materials, as many other researchers have suggested.

XIX. CONCLUSION

The focus of this experiment was to investigate users’ involvement while experiencing the edutainment features in multimodal e-learning systems, through three quantitative aspects, in particular users’ achievement, correctness, and satisfaction.

The experiment’s circumstances were controlled to guarantee its validity, either by the rotation method used or by the lessons distribution mechanism provided. The experimental dependent-variable measurements were managed in order to acquire accurate results, which were achieved through incorporating a fitted time design in all systems and a webcam was employed to capture users’ responses and record their expressions. The outcome reported was positive and the tests also resulted in other significant outcomes.

The experiment’s results established that the Game-based learning interface outperformed all other interfaces. The use of game as an educational channel proved to be valuable and had a reduced rate of error towards the subject sample. Besides teaching people raw facts, the Game developed desirable pleasure and enjoyment that assisted in achieving the users’ aspirations. This was the target that our research set out to achieve.
Fig 6: Users’ interface preference order; users interface preference.

REFERENCES

[28] M. Perrett, “Integrating Learning and Fun: Expanding the Educational Game Genre, 2007”
[38] K. Egan, “Supplement to TEACHING AS STORY TELLING.”
