

Seamless Evaluation Integration into IDS Educational Games

Andreea Molnar
Brunel University
Kingston Lane, Uxbridge, UB8 3PH
United Kingdom
Andreea.Molnar@brunel.ac.uk

Patty Kostkova
University College London
Gower Street, London WC1E 6BT
United Kingdom
P.Kostkova@cs.ucl.ac.uk

ABSTRACT

Serious games have recently enjoyed unprecedented interest among researchers. Despite the fact that the main focus of this domain remains on player entertainment and engagement, serious games have established themselves as an alternative educational paradigm. However, little interest has been given to investigating methods for the assessment of their effectiveness. Game evaluation focus remains on investigating usability and technical aspects. Those attempting to evaluate the educational impact typically rely on traditional questionnaires that typically negatively affecting user immersion and the overall game enjoyment. Among game genres, Interactive Digital Storytelling (IDS) is a fast growing genre that merges computer games, multimedia and cinematic storytelling with entertaining education. Like any other teaching intervention, educational IDS's have to be evaluated against a defined set of Learning Objectives (LOs). In this paper we present a *User-Centred Seamless Evaluation Framework* for IDS games and describe the algorithm for the integration of knowledge assessment to measure knowledge improvement against given LOs without affecting player enjoyment. We implement this algorithm in an educational game and show that the framework has been perceived by the majority of the players as a positive enhancement to the game, 94% of the participants reported preferring this form of assessment as opposed to the more "traditional" methods. Statistically significant knowledge improvements were obtained after the game play.

Categories and Subject Descriptors

K.3.1 [Computing Milieux]: Computer Usages in – *computer assisted instruction (CAI)*, *computer-managed instruction (CMI)*, *distance learning*.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

FDG'12, May 14-17 2012, Crete, Greece.
Copyright 2012 ACM 1-58113-000-0/00/0010...\$10.00.

General Terms

Measurement, Design, Experimentation, Human Factors.

Keywords

Assessment, educational games for health, hygiene education, usability, IDS games, usability

1. INTRODUCTION

Games and gamification of online content are emerging as a multimedia domain of increasing interest from both research and industry arenas [3]. However, most of the research has focused either on various technical aspects of the game [13] or on the entertainment potential [2, 5] only few studies are focused on the educational potential [1].

Moreover, the evaluation of game potential in education is rare and subject to drawbacks and flaws [1]. Evaluations are either implemented as a separate part of the game, and when integrated, they are performed in an intrusive manner disturbing the game flow, leading to a reduction of enjoyment for game players [22]. Previous research has shown that when participation in separate evaluation activities is mandated, it can be viewed as burdensome for the participant and it decreases the players' enjoyment [18]. Therefore, there is a need for new methods of assessment that do not negatively affect the game enjoyment element. In this paper we present *User-Centred Seamless Evaluation Framework* for IDS games by seamlessly integrating the assessment into the IDS game flow. The advantages of having an embedded assessment are twofold: it removes or seriously reduces the test anxiety without affecting the validity or reliability of the results [19], and it engages players with the subject taught by providing immediate feedback [10].

Among existing genres, IDS games are growing as a hybrid multimedia discipline that merges computer games and cinematic storytelling. IDS games are recognised to provide a motivating and engaging experience for players [11]. Moreover, in the educational arena, it can improve their problem solving abilities and their ability to organize knowledge [11]. Due to their dynamicity, IDS games also provide the player with a better feeling of control over the game and directly influences the unfolding of the story.

The rest of this paper is organized as follows. First, an overview of previous research is presented in Section 2. Section 3 presents the *User-Centred Seamless Evaluation Framework* and Section 4 the

seamless evaluation integration into the framework Conversation Layer. Section 5 presents the GHD Game that is our case study for assessment integration. Section 6 presents the evaluation study design and results. It starts with the assessment of the seamless evaluation. Afterwards, it presents the effectiveness of the game in conveying the LOs with the assessment integrated. Section 7 highlights our future research directions. The last section ends the paper with a summary of our research and some reflections on the findings.

2. RELATED WORK

IDS games are a niche multimedia domain covering digital games and cinematic storytelling. IDS games could cover different areas, and can be used for different purposes such as helping people to cope with traumatic situations [21], while providing a motivating and engaging experience for players [11]. In addition, positive results have been obtained when using IDS for educational purposes [11]. However, a “large proportion” of evaluations performed in IDS studies have a low number of participants [6]. Additionally, very few studies have explored different means to evaluate the games, and none, to the best of our knowledge, have done so by having the evaluation seamlessly embedded in the game flow. Although studies such as Mobile Urban Drama collect some data during the game play, which at a later stage they will use in a classroom evaluation independent of the game play [9], in this research we are referring to studies in which the evaluation is fully performed though the game. Therefore, to set our evaluation method in the wider educational evaluation context, we will broaden the related work discussion to cover game projects outside of the IDS genre.

Among the few games that we found that attempt to evaluate the effectiveness within the game flow, examples include Bugs Kingdom [17], EducaMovil [14], Winterfest [23], and Global Conflicts [7].

First, Bugs Kingdom [17] is a platform game, part of the DG SANCO funded E-Bug project, aimed at reinforcing knowledge related to microbe transmission and food hygiene. The assessment of the LO is performed through a quiz, similar in style to “How to be Millionaire”, inserted between levels of the game. The quizzes are inserted at the end of each level of the game. A facilitator asks questions and the player has to choose among three options in the game. Feedback is given after the results are submitted.

Second, EducaMovil [14] is a suite of open source mobile learning games, through which snippets of educational content and quizzes are integrated. A certain action of the game triggers a certain lesson to open, and a question is given to the learner after the educational content snippet is shown. Feedback, either positive or negative is offered immediately after the lesson is finished.

Third, Winterfest [23] aims at improving player arithmetic skills. The assessment is integrated in the game story, and the feedback is given instantly. However, due to the rigidity with which the evaluation is performed, it disturbs users from the game play [23].

Fourth, Global Conflicts [7] is an adventure/role-playing game. The player has to collect facts. There is no feedback provided until the end of the game, when the player is given feedback based on the collected facts.

Although all these games have the evaluation embedded in the game, our approach is different on several accounts: first by

addressing a different game genre than the games discussed in this section; and second by seamlessly embedding the LOs assessment into the game flow. Moreover, our approach is among the few that ensure the player is given feedback when the first set of questions is asked, during the game play, in order to realise what s/he needs to improve, further enhancing the educational element. Finally, this research not only assesses the effectiveness of the game in conveying the LOs but also the players’ opinion/perceptions of the integrated assessment without affecting player/user experience. This has not been previously performed to the best of our knowledge.

3. PROPOSED USER-CENTERED SEAMLESS EVALUATION FRAMEWORK FOR IDS GAMES

In this section we will set the scene by providing the overview of the proposed *User-Centred Seamless Evaluation Framework* for IDS games. IDS games are aimed at delivering both entertainment and educational content. In this context, the user centred evaluation assessing player knowledge gain against a set of LO’s is an important issue in educational games, providing information regarding the efficiency of the game, as well as feedback for the player about his/her knowledge. This further enhances the educational value of the game and contributes to player knowledge. It has been shown that feedback is vital for learning, as it is necessary to encourage ‘deep’ learning and engage students with the subject [10].

This section will first introduce a novel educational IDS framework, *User-Centred Seamless Evaluation Framework* that caters for seamless evaluation integrated into the game. Although previous IDS frameworks exists, such as [20] based on which this research was based, most of them do not cater for IDS games for educational interventions, and to the best of our knowledge there is no framework that formalises the seamless evaluation. The IDS game framework (see Figure 1) has five layers: *Presentation Layer*, *Conversation Layer*, *Quest Layer*, *Mission Layer*, and *Educational Layer*.

All the layers are briefly introduced below, as in a standalone IDS educational framework that can be independent of the seamless evaluation. Then the *Conversation Layer* is further extended to show how the seamless evaluation fits in the framework.

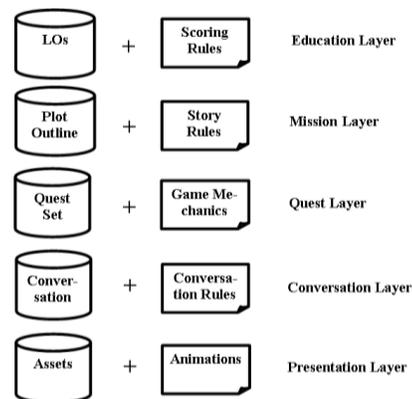


Figure 1. User-Centred Seamless Evaluation Framework for IDS Games.

3.1 Framework Layers Description

The five layers of the framework are presented below.

3.1.1 Presentation Layer

The *Presentation Layer* contains the assets/animations needed to deliver the IDS. It consists of animations for characters, rooms, items etc, and the motion models that are used to describe how the non-playing characters move or behave. Together they form the game animation function.

3.1.2 Conversation Layer

The layer on top of the *Presentation Layer* is the *Conversation Layer*. In IDS, conversation is the main means of interaction and content presentation. This layer consists of *Conversation Nodes* and *Conversation Rules*. A *Conversation Node* (CN) is a line of text/or a sentence recited by a player character. The *Conversation Rules* show which player is saying what, and in which context they are saying it. For example, a rule could be: a game (or so called 'non-playing') character greets the player at the beginning of the game. Another rule states that the non-playing character greets the player only if the player does it first. This part has the power to generate adaptive conversations based on the previous players' actions.

The *Conversation Nodes* component contains all the conversations taking place in the IDS, and the transition between a CN to another one in a dialogue can be defined as:

Def 1: Let CNs be a non empty set of conversation nodes, C a set of Characters, $c \in C$, then a conversation node $c_n \in CNs$ is defined $c_n = (c, Text, \{c_{n+1}, \dots, c_{n+x}\})$ where $\{c_{n+1}, \dots, c_{n+x}\}$ are elements of CNs , determining the text dialog line to be displayed after the conversation node c_m , by the character c . If c and $\{c_{n+1}, \dots, c_{n+x}\}$ are an empty set then the current conversation is over.

In the *Conversation Rules* section, we also define the transition function δ , that interprets the conversation rules, and invokes the right CN(s).

Def 2: Let P be a set of game specific conditions, $p \in P$, R a set of rooms, $r \in R$, C a set of Characters, $c \in C$, and CN a set of nodes, $c_m, c_{n+1} \in CN$, the transition function, δ , is defined as:

$$\delta: (p, r, c, c_n) \rightarrow (p, r, c, c_{n+1}).$$

3.1.3 Quest Layer

The *Quest Layer* contains the *Quest Set* and the *Game Mechanics*. Quest in the context of this research refers to any 'story element' of the game that requires activation when certain conditions are met, a series of states visited according to a transition function and finished in the quest end state, based on a set of conditions that are met. It contains the *Game Mechanics* that determines the operation of the game world and deals with player interactions with the game.

3.1.4 Mission Layer

It is the top level of the story. The mission, an ultimate quest starting with the game initiation state and finishing when the IDS story reaches the finish state. It has the highest level of abstraction in the IDS.

3.1.5 Education Layer

The *Education Layer* consists of *LOs* and *Scoring Rules*. The *LOs* contain a high level description of the LO delivered through the

game. For example a LO could be: "One should only use antibiotics with a doctor's permission". The *Scoring Rules* consist of rules describing how the LO evaluation, contributes to the player's score. For example, how many points the player gets for answering correctly one of the questions in the game.

4. SEAMLESS EVALUATION INTEGRATION

In order to insert the LOs into the game we have to define the algorithm enhancing the IDS framework. The *Conversation Nodes* and *Conversation Rules* have to be enhanced to evaluate the educational impact of the game so that the educational purposes of the game are satisfied. The *User-Centred Seamless Evaluation Framework* consists of additionally 3 types of CNs and 3 conversation rules as follows.

To allow the seamless evaluation to be integrated in the educational IDS the following has to be added to the game:

1. new CNs through which we deliver the seamless evaluation,
2. rules to describe by whom and in which context the CNs concerning the evaluation are delivered.

The *Conversation Nodes* and *Conversation Rules* have to be enhanced to evaluate the educational impact of the game so that the educational purposes of the game are satisfied. For example, the *User-Centred Seamless Evaluation Framework* consists of an additional 3 types of CNs and 3 conversation rules as follows.

Conversation Nodes enhancements include three CN entities: questions, options and feedback:

Def 3: A set of questions, $Q \in CN$, textual version of the LOs: let $Q_i \in Q$ be the textual version of the LO_i, and $CN_x \in CN$ is the conversation node after which the question is inserted, before the actual teaching of the LO_i, and $CN_{x+m} \in CN$ is the conversation node after which the post-evaluation is inserted after LO_i is taught

$$CN_x = (c, Text, Q_i)$$

$$CN_{x+m} = (c, Text, Q_i)$$

- **Def 4:** A set of options, $O \in CN$, the players has to choose from, containing correct and incorrect option(s). O_1, \dots, O_k are options for LO_i presented after Q_i . Only one option is correct. The options could be precise, or more general (e.g. they could contain right and wrong answers, or something as general as, if the user believes that this fact is right, or if it is wrong), 'You' $\in C$ is the player character actively choosing the next step in the game while his/her knowledge is tested:

$$Q_i = (You, QuestionText, \{O_1, \dots, O_k\})$$

- **Def 5:** A set of feedback replies, $F \in CN$ that are given to the player after s/he replies, $F_i \in F$ for option O_i . Immediate feedback provides the player with a sense of control over the task and improves the player concentration [22].

$$F_i = (c, FeedbackText, CN_{m+1})$$

The *Conversation Rules* contain rules for the LOs, how to deliver them and in which context and how many options are available to the user. Except for these several other rules have to be added:

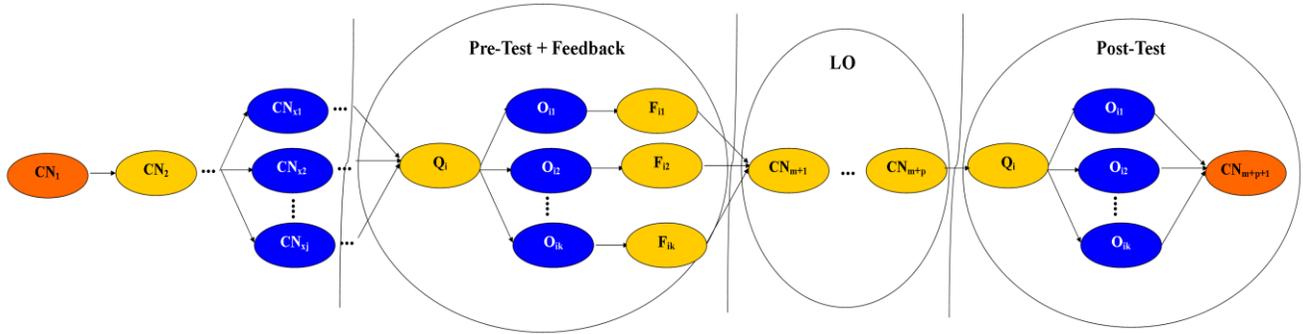


Figure 2. A section of game which highlights how Pre-Test, feedback, LO, and Post-Test is integrated in the game. CN – conversation node, Q - question, F – feedback, O - option. The colors represent a different character, in this case with blue is the player, who has to select among the different options, while with other colors are different characters.

- For $\forall Q_i, i \in \{1, \dots, n\}$, where n is the number of questions asked before the LO is delivered, another Q_i is asked at a later stage after the LO is delivered. The necessity of having the question asked twice is to determine the knowledge gained after player has explored the LOs through the game against the baseline knowledge before the exposure (see Def. 3).
- For $\forall Q_i, i \in \{1, \dots, n\}$, where n is the number of questions, there should be different O_1 to O_k , options, where k is the number of possible options for the players reply.
- For \forall option $O_i, i \in \{1, \dots, k\}$, where k is the number of possible options, a feedback F_i exists. The feedback is mandatory only for the first time when the question is asked (see Figure 2).
- Immediately after the player selects the option $O_i, i \in \{1, \dots, k\}$, where k is the number of possible options, the feedback F_i is immediately displayed. $pb, pa \in P$ represent a set of conditions “before” and “after” determining whether the player has not been taught the LO, yet the baseline knowledge is tested (the feedback is displayed) or whether the player has been taught the LO and the post-intervention knowledge is tested (then no feedback is required).

$$\delta: (pb, r, You, Oi) \rightarrow (p, r, c, F_i)$$

$$\delta: (pa, r, You, Oi) \rightarrow (p, r, c, CN_{m+p+1})$$

You in this case is the player (Figure 3 and Figure 4).

Figure 2 presents the pieces of the game, focusing mostly on the *Conversation Nodes* layer and when multiple options are available on the *Scene Layers*. The lines that separate in the figure show that there is a discontinuity in our description of the story. The evaluation is presented just for a single LO, but in the game multiple LOs can be presented. We see that the pre-test, that involves the question that is asked, the options that the user has to choose and the feedback will be first, before the LO is taught. LOs can be taught through multiple CNs. At a later stage in the game the player is asked the same questions, and the same options are given. The question can be asked by the same or a different character.

5. CASE STUDY

The seamless evaluation questions, as presented above, were integrated into a role-playing game: the GHD (Global Handwashing Day) Game [15]. The game is an educational IDS game that relies heavily on the narrative. It aims to reinforce the importance of hygiene, focusing on hand washing.

The plot of the game is as follows. First the player is placed in the e-Bug agency and s/he is introduced to her/his boss, Big C. Also here, the player meets Alyx who will be the player’s partner and would help him during the investigation. After the introductions are made, Big C introduces the problem Hugh Gaego, a famous actor, is supposedly poisoned, and the player has to decipher the mystery: whether it was a case of an alleged poisoning or not, and who the guilty party is, if any, for poisoning Hugh. The state space of the game is quite vast, allowing the players to explore different parts of the game, by making it non linear and allowing different options during the investigation. Not all the paths lead to an answer and they are not all mandatory for solving the mystery.

During the investigation, Alyx is always ready to help the player, asking questions related to the investigation and assisting with evidence that was collected. Although totally integrated into the game flow, the questions assess the educational content presented (Figure 3). The questions are spread throughout the game, and asked so that they will fit in the context of the game. However, these questions are asked before the player is exposed to game mechanics through which s/he can learn about the objectives being asked. The questions are asked in an abstract manner, in order to see whether the player understands the scientific concept, and if



Figure 3. Example of an evaluation question integrated in the game



Figure 4. Example of a question at the end of the game, during the debriefing

what s/he learns is generalised. We do so because a previous study has shown that some of the skills learned through games are not necessarily broad and general, and the player is able to use the skills in the same environment but has problems translating them in a real-world environment [1]. However, the LOs are delivered both in an abstract manner as well as through the game mechanics. If the player gets an answer wrong, the correct answer is given to the player immediately after the player gives her/his answers, in order to correct misconceptions and allow the player to improve its knowledge during the game play.

Because we have a pre and post set of questions to assess the knowledge before and after the LO is delivered, and due to the CSI (Crime Scene Investigation) nature of the game, we decided that for the given story, it is suitable to ask the second round of questions towards the end of the game, when all the LOs are taught. Therefore, when the investigation is over and the player and his partner return to the headquarters for debriefing Big C, she asks the player the same set of questions and the player has to select among the same set of options as when the questions were asked for the first time (Figure 4).

6. EVALUATION

The aim of the GHD Game evaluation is twofold:

- To assess the player perception of the new assessment method – the seamless and its effect on user experience
- To assess the effectiveness of the game with the seamless evaluation integrated in conveying the educational content

The key to understanding the impact of the seamless evaluation is the first assessment; however the seamless evaluation can only demonstrate useful results if the positive education impact is not affected.

6.1 Method

The participants played the game either in a controlled environment (in a school with a teacher present with 50 minutes to finish the game), or online at their convenience. The schools at which the evaluation took place were located in London and Glasgow, UK. All the participants were given incentives to

participate in the study. The game session was followed by a survey which was not mandatory and assessed the seamless evaluation, game usability, and different aspects of the game design.

The *Seamless Evaluation* was assessed through a mixed method, combining survey, performed at the end of the game playing session, and observations done during the playing sessions.

The *effectiveness of the game in conveying the LOs* was assessed through the experimental studies in which the participants have to play the game beginning through to end. For measuring the statistical significance of the effectiveness of the game in conveying the educational content, we used a paired t-test [12].

6.2 Participants

All the participants that took part in the study were asked to fill a questionnaire with demographic data at the beginning of the game playing session, then to play the game, and at the end to complete a survey. The pre-questionnaire and the survey were not mandatory. Moreover the end survey was not given to all the participants, therefore not all the participants who played the game completed the questionnaire at the end of the game playing session. However since the effectiveness of the game in delivering the LOs was independent of the participants answering the survey, was considered relevant to include also these participants in order to provide a better overview of the teaching potential of the game.. For this reason, in order to provide a better overview of the participants' demographic data, this section discusses first the demographic data for the participants that participate in the game playing session and afterwards the demographic data of those participants who did both (participate in the game playing session and also completed the survey afterwards).

The participants were either students who played the game during school visits, researchers and students who volunteered to participate, or people who found and played the game online on edugames4all website [4]. The website was re-launched in October 2011, and GHD Game was added during that period. The website was since then promoted during the Global Handwashing Day 2011 in UK [8], and through other means (mailing lists, social networking websites). The traffic comes from 73 countries, however most of it comes from English speaking countries: UK (~60%), US (~10%), Ireland (~9%).

145 participants were considered for the evaluation. The participants were selected based on whether they finish the game or not. The main reason for this decision is the fact that the evaluation is integrated in the game, and the post evaluation is towards the end of the game, therefore for a player who did not finish the game, the results of the evaluation were not available.

The end survey was completely filled by 21 participants (the ones who left them incomplete were not considered).

6.3 Seamless Evaluation

The seamless evaluation was assessed through a survey filled in by the players. The survey assessed the player opinion with regards to the integration of the evaluation in the game flow. The players were not previously informed that their knowledge would be assessed during the game play.

As a part of the survey, the players were first asked whether or not they realised that they have options to choose from. 95% of the

players realised that they have to choose one of the options presented.

The players who realised that they have to select one of the presented options, were asked to rate how these affect their game experience on a 5 point Likert scale. The options and the percentage of players choosing a certain option are presented in Table 1. As it can be seen half of the players consider the questions to be a good addition to the game: 12% considering that they enrich their game experience, while 44% that they make the game more interesting. Among the rest, 24% were not affected in any way by the questions integrated, and the rest were affected in a negative manner. This can lead to the conclusion that for most of the players, the integrated evaluation does not only facilitate the assessment but can also improve the game experience.

Table 1. The results for how people perceived that the seamless evaluation affected their experience

#	Option	%
1	They obstructed my game experience	6
2	It wasn't too bad, they didn't discourage me but I would prefer not to have them	12
3	They did not affect me in any way	24
4	It was good having them, they had made the game more interesting	44
5	They enriched my game experience, they engage me more into the game	12

The last question explains the scope of having the assessment integrated in the game flow. Afterwards, the players were asked whether or not they would like to have the educational content assessment integrated in the game or preferred the more “traditional” version of taking a test. Among the players who answered this question, 94% preferred to have the evaluation integrated in the game flow.

The subjects also had the option to leave comments to these questions. One of the comments left was that, once an option was selected as an answer, the player could not change his option. This constraint was imposed, as immediately after the player answer, s/he is provided with feedback from one of the non-playing characters in the game (except during the post-test when no feedback is provided). Letting the player change the answer, would allow them to modify their answer as a result of the feedback received resulting in a flawed evaluation of their knowledge. A simple solution to this problem could have been to have the player confirm the answer, but it was considered that repetitive requests for confirmation would interrupt the game flow (conversation flow), our aim was to avoid this as much as possible.

Another comment was from a subject that justified his choice, as having preferred the evaluation integrated in the game, as otherwise it would destroy the “immersiveness” element of the game play and it would prove disruptive.

Overall, the results of the seamless evaluation assessment indicate the players’ strong preference of this method. Moreover, most players considered the questions as an enhancement to the game. Additionally, even the players who were less than positive about the addition of the questions into the game preferred this method to a test. There is still place for improvement, as some of the

players did not notice that they could choose among the different options. However, overall the results are positive, and they indicate that this method could not only facilitate the assessment but also improve the game.

6.4 Game Effectiveness of Conveying LOs with Seamless Evaluation Integrated

One of the most important aspects of the educational game is improving knowledge as a result of the game play. Only few studies have been performed so far that measure the educational value added beyond mere entertainment [5].

Therefore this section evaluates the change in players’ LOs knowledge as a result of playing the game. We wanted to see whether the player could learn by playing this game, or if the knowledge assimilation is affected as a result of the novel assessment mechanism. Eleven LOs were assessed:

- LO-1: Microbes found in food can transfer to humans
- LO-2: Separate utensils should be used for raw meat, and vegetables
- LO-3: Bacteria from raw meat can make a person sick
- LO-4: Food cooked properly should be free of bacteria
- LO-5: Vomiting viruses are unpleasant but usually not dangerous
- LO-6: Vomiting viruses can spread through sneezing, coughing or just particles of vomit that are in the air after someone is sick
- LO-7: Vomiting viruses and E. coli can spread through bad hygiene
- LO-8: It is not always necessary to take medicine when dealing with E. coli and vomiting viruses infections
- LO-9: E.coli is commonly found in the lower intestine
- LO-10: E.coli can spread through the ‘faecal-oral’ route or poor food preparation hygiene
- LO-11: If eaten, bacteria from raw meat can make a person sick

The effectiveness of the game at conveying the educational content was performed using a paired t-test [12] on the number of correct answers the players had on the pre and post questionnaire. A 95% confidence interval was considered statistically significant. The results show that the difference between the players pre and post questionnaire questions is statistically significant ($p=0.01$, $\sigma=2.20$).

7. FUTURE WORK AND DISCUSSIONS

This research has led to several new research directions. We want to build an intelligent mechanism in the game that could detect if the players are stuck in the game and provide them with help. If the players drop out is due to the game difficulty this could help reduce the dropout rate. Moreover, adding different game mechanics into the game, so that the game adapts better to player’s knowledge, can be a way forward, and it would be interesting to see whether this has any effect on players’ retention to playing the game. At the moment the game provides the players with a tutorial that can be consulted during the game playing session, and a non-

mandatory training mission that could be used to teach the game mechanics [16].

A further direction might be to ask players questions that are not generalisable and explore whether this obtains better results as it has been shown that the players might have problem translating the skill in another environment [1].

We have also noticed that when players know that there is a score, and do not hold only the internal motivation of solving the mystery of the game, they get more involved in the game play and behave competitively. Having the score integrated into the game, and making the game more competitive, could be a next step in making the game more engaging for the players.

Due to the design of the study no conclusion can be made regarding the behavioural change as a result of the game playing session, or the long term effects of the game playing on the players acquired knowledge. However, on the short term, positive effects can be noticed and short term changes are a precondition for long terms changes to occur [5]. Nevertheless, the aim of this evaluation was to assess the seamless evaluation and positive results were obtained in this area.

As a next step we will explore assessing what effect it would it have building a community on the players' engagement and their willingness to re-play the game, or similar games. This would mean creating a community in which the people could connect to each other and share their ideas about the game. Moreover, we want to extend the game to support multiple players. This could lead to the creation of different assessment strategies, such as players asking/helping each other when they get stuck somewhere in the game.

Furthermore, we will explore how seamless evaluation could be integrated in other types of games in which conversation is not necessarily present, or present to such extend as in IDS games.

8. CONCLUSIONS

IDS is an emerging multimedia domain that embeds cinematic storytelling with video games. Although positive results have been shown while using them for education [11], the evaluation was typically performed as a separate part of the game. This paper investigates a method for enhancing the Seamless Evaluation IDS framework with knowledge assessment. The algorithm implements pre and post assessment of players knowledge against given LOs before the LO is delivered providing a baseline data and afterwards to measure a knowledge gain. Three sets of CNs were introduced: Q, O and F were introduced before and after the LO is taught, and the same Q and O were given afterwards. The rules needed to plug these CN in the game flow were also discussed.

The framework implemented and tested in GHD Game, an IDS that aims to reinforce issues pertinent to hand and food hygiene. This game was used as a case study for the game evaluation.

The evaluation follows several aspects:

- To assess the seamless evaluation
- To assess the effectiveness of the game with the seamless evaluation integrated in conveying the LOs

The seamless evaluation was assessed using a survey filled by the participants after the game playing session. Most of the players considered that the results improved their game experience (56%

of the participants), and 24% consider that they were not affected by the evaluation in any way. Moreover, 94% of the participants prefer the seamless evaluation to taking a test.

The result of the evaluation that assessed the efficiency of the game in conveying the LOs showed that the players knowledge has improved during the game, and the changed was statistically significant ($p=0.01$).

Overall the results of the evaluation were positive: the seamless evaluation doesn't only have the educational potential, but also improve the gaming experience. Moreover players knowledge of the LOs covered in the game improved as a result of playing the game, and the results are statistical significant.

9. ACKNOWLEDGMENTS

The authors would like to thank e-Bug project members and DG SANCO for providing funding for this EU project, David Farrell for designing and programming the initial games, and Dr. Gabriel-Miro Muntean for helping with the evaluation.

10. REFERENCES

- [1] Bavelier, D., Green, C. S., Han, D.H., Renshaw, P. F., Merzenich, M. M., and Gentile, A. A. 2011. Brains on video games. *Nat. Rev. Neurosci.* 12, (December 2011), 763-768. DOI=10.1038/nrn3135.
- [2] Cox, A., Cairns, P., Shah, P., and Carroll, M. 2012. Not doing but thinking: the role of challenge in the gaming experience. In *Proceedings of the ACM International Conference on Multimedia* (Nara, Japan, October 29 – November 02, 2012), MM'12. ACM, New York, NY, 79-88. DOI=<http://doi.acm.org/10.1145/2207676.2207689>.
- [3] DuHadway, C., and Shetty, S. 2012. *Gamification for Improved Search Ranking for YouTube Topics*. <http://googleresearch.blogspot.com.es/2012/03/gamification-for-improved-search.html>
- [4] Edugames4all. 2013. www.edugames4all.org
- [5] Fuchslocher, A., Niesenhaus, J., and Krämer, N. 2011. Serious games for health: An empirical study of the game "Balance" for teenagers with diabetes mellitus, *Entertainment Computing*, 2, 2 (2011), 97-101. DOI = <http://dx.doi.org/10.1016/j.entcom.2010.12.001>
- [6] Garzotto, F., Paolini, P., and Sabiescu, A. 2010. Interactive storytelling for children. In *Proceedings of the 9th International Conference on Interaction Design and Children* (Barcelona, Spain, June 9-12, 2010), IDC '10. ACM, New York, NY, USA, 356-359. DOI=<http://doi.acm.org/10.1145/1810543.1810613>
- [7] Global Conflicts. www.globalconflicts.eu
- [8] Global Handwashing Day. 2011. <http://www.globalhandwashingday.org.uk/>
- [9] Hansen, F. A., Kortbek, K. J., and Grønbaek, K. 2010. Mobile urban drama for multimedia-based out-of-school learning. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia* (Limassol, Cyprus, December 1-3, 2010.), MUM'10. ACM, New York, NY, USA, Article No. 17. DOI=<http://doi.acm.org/10.1145/1899475.1899492>

- [10] Higgins, R., Hartley, P., and Skelton, A. 2002. The conscientious consumer: Reconsidering the role of assessment feedback in student learning, *Studies in Higher Education*, 27, 1, 53-64. DOI = 10.1080/03075070120099368
- [11] Hodhod, R., Cairns, P., and Kudenko, D. 2011. Innovative integrated architecture for educational games: challenges and merits. In *Transactions on edutainment V*, Zhigeng Pan, Adrian David Cheok, Wolfgang Müller, and Xubo Yang (Eds.). Springer-Verlag, Berlin, Heidelberg 1-34.
- [12] Hsu, H., and Lachenbruch, P. A. 2008. *Paired t Test*. Wiley Encyclopedia of Clinical Trials.
- [13] Liu, S., Chen, Q., Dong, J., Yan, S., Xu, C., and Lu, H. 2011. Snap & play: auto-generate personalized find-the-difference mobile game. In *Proceedings of the 19th ACM international conference on Multimedia* (Scottsdale, Arizona, USA, November 28 - December 1), MM'11. ACM, New York, NY, USA, 993-996. DOI=
<http://doi.acm.org/10.1145/2072298.2071921>
- [14] Molnar, A., and Frías-Martínez, V. 2011. EducaMovil: Mobile Educational Games Made Easy. In *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Lisbon, Portugal, June 27 - July 1, 2010), Ed-Media'11. Chesapeake, VA: AACE, 3684-3689.
- [15] Molnar, A., Farrell, D., and Kostkova, P. 2012. Who poisoned Hugh? - The STAR framework: integrating learning objectives with storytelling. In *Proceedings of the 5th International Conference on Interactive Storytelling* (San Sebastian, Spain, November 12-15, 2012), ICIDS'12. David Oyarzun, Federico Peinado, R. Michael Young, Ane Elizalde, and Gonzalo Méndez (Eds.). Springer-Verlag, Berlin, Heidelberg, 60-71.
DOI=http://dx.doi.org/10.1007/978-3-642-34851-8_6
- [16] Molnar, A., and Kostkova, P. 2013. If you build it would they play? Challenges and solutions in adopting health games for children. In *Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems, Let's talk about Failures: Why was the Game for Children not a Success?* (Paris, France, April 27).
- [17] Molnar, A., and Kostkova, P. 2013. On effective integration of educational content in serious games. In *Proceedings of 13th IEEE International Conference on Advanced Learning Technologies* (Beijing, China, July 15-18, 2013), ICALT'13.
- [18] Sapouna, M., Wolke, D., Vannini, N., Watson, S., Woods, S., Schneider, W., Enz, S., Hall, L., Paiva, A., Andre, E., Dautenhahn, K., and Aylett, R. 2010. Virtual learning intervention to reduce bullying victimization in primary school: a controlled trial. *Journal of Child Psychology and Psychiatry*, 51, 1, (2010), 104-112.
- [19] Shute, V. J., Hansen, E. G., and Almond, R. G. 2008. You can't fatten a hog by weighing it—Or can you? Evaluating an assessment for learning system called ACED. *International Journal of Artificial Intelligence and Education*, 18, 4, (December 2008), 289-316
- [20] Spierling, U. Interactive digital storytelling: Towards a hybrid conceptual approach. In *Proceedings of Digital Games Research Conference* (Vancouver, British Columbia, Canada, June 16-20, 2005), DIGRA'05.
- [21] Szilas, N., Richle, U., Boggini, T., and Dumas, J. 2010. Using highly interactive drama to help young people cope with traumatic situations. In *Proceedings of International Conference on Interactive Storytelling* (Edinburgh, UK, November 01 – 03, 2010), ICIDIS'10, Ruth Aylett, Mei Yui Lim, Sandy Louchart, Paolo Petta, and Mark Riedl (Eds.). Springer-Verlag, Berlin, Heidelberg, 279-282.
- [22] Wendel, V., Göbel, S., and Steinmetz, R. 2011. Seamless learning in serious games - How to improve seamless learning-content integration in serious games. In *Proceedings of 3rd International Conference on Computer Supported Education* (Noordwijkerhout, Netherlands, May 06 -08, 2011), CSEDU'11, 219-224.
- [23] Winterfest, www.learnspiel-winterfest.de