Je-LKS
Journal of e-Learning and Knowledge Society
The Italian e-Learning Association Journal
PEER REVIEWED PAPERS
GAMIFICATION AND SERIOUS GAME FOR LEARNING
Vol. 11, n.3, 2015
ISSN: 1826-6223 | eISSN: 1971-8829

KNOWLEDGE-BASED ASSESSMENT IN SERIOUS GAMES: AN EXPERIENCE ON EMERGENCY TRAINING

Nicola Capuano[1]
Richard King[2]

[1]Dep. of Information Engineering, Electric Engineering and Applied Mathematics - University of Salerno, Italy - ncapuano@unisa.it
[2]Serious Games Institute -Coventry University Enterprises Ltd., UK richard.king@coventry.ac.uk

Keywords: Emergency training; Evidence centred design; Adaptive learning systems; Serious games; Knowledge representation.

Emergency preparedness is a promising application field for digital serious games enabling the simulation of real emergency scenarios and allowing a high learning transfer thanks to engagement and focus on specific tasks. Games can also play a role in the assessment that may happen without interrupting the learner, observing and evaluating what she is doing. Based on these premises we defined a serious game for evacuation training targeted to primary and secondary school students. The student is immersed in a virtual environment representing her school during an emergency with the aim of evacuating the building and adopting the correct behaviour. Any performed action is evaluated by the system, feedback is provided immediately and also when the game ends. Recovery micro-learning resources are then arranged and provided to the students to explain any errors they made and to help them reach better performances. The system is based on the application of a theoretical framework for evidence-based assessment where knowledge-

for citations:
Capuano N., King R. (2015), *Knowledge-based assessment in serious games: an experience on emergency training,* Journal of e-Learning and Knowledge Society, v.11, n.3, 117-132. ISSN: 1826-6223, e-ISSN:1971-8829

based structures have been used to represent emergency skills and to relate them to possible actions within the game. An experiment with students coming from four Italian schools has been also performed to validate the models and the prototype.

1 Introduction

Serious games are defined in (Michael & Chen, 2005) as "*games that do not have entertainment, enjoyment or fun as their primary purpose*". Even if the term was first introduced in 1970 (Abt, 1970), serious games were already used by the army to train military strategic skills. Thereafter serious games have been applied to a broad spectrum of application areas like government, education, corporate and healthcare.

According to several authors, like in (Corti, 2006) and (Susi *et al.*, 2007), serious games, when applied to education and training, have the potential of improving learning outcomes thanks to the level of engagement, motivation, role playing, and repeatability (i.e. failed strategies can be modified and tried again). Thanks to these benefits, the research about *Digital Game-Based Learning* is acquiring even more interest within e-learning, especially when directed at young learners.

As pointed out in (Susi *et al.*, *op. cit.*) in fact, the current generation of students are "*native speakers in the language of digital media*" and have experienced new forms of computer and video game play that "*has shaped their preferences and abilities and offers an enormous potential for their learning, both as children and as adults*".

In addition to the teaching function, serious games can play a role in assessment too. During the game, players take decisions that lead them to performing actions to overcome some challenge. Such actions can be recorded and used a-posteriori to appraise such decisions and, basing on that, to assess learning outcomes. This is the inspiring principle for the Information Trails framework defined in (Loh *et al.*, 2007).

In addition, the *Evidence Centred Design* (ECD) framework defined in (Mislevy & Haertel, 2006) helps linking actions performed by the learners in complex contexts (like serious games) and attributes that are object of assessment. By relying on these frameworks, an assessment can be tied to learners' actions within serious games, and can operate without interrupting what learners are doing or thinking (Shute, 2011).

In this paper we describe an e-learning application we developed for evacuation training, targeting primary and secondary school students, based on an immersive 3D serious game. According to ECD, the application relies on a formal definition of emergency preparedness skills to be acquired and on mapping them to possible actions the user may perform during the game. According to *Information Trails*, assessable actions are collected in the game through the implementation of specific *markers* placed in strategic places within each game scenario. The skill assessment is then used to automatically arrange micro-learning resources aimed at reinforcing missing or weak skills before a subsequent interaction with the game.

The paper is organized as follows. A brief discussion on related work is presented in section 2, and then the defined theoretical framework for evidence-based assessment is described in section 3. The prototype serious game is then discussed in section 4 while section 5 presents the results of the system experimentation with young learners coming from four Italian secondary schools. Final remarks close the paper.

2 Related work

The idea of using serious games in emergency training is not new. They are challenging, immersive, highly interactive, allow users play in realistic scenarios, as well as interest them and provide immediate feedback to performed actions. For these reasons serious games may become a great compendium to evacuation drills that are costly, rarely performed and focused on specific scenarios (Ribeiro *et al.*, 2012).

For example, (Chittaro & Ranon, 2009) proposed a serious game aimed at acquiring basic personal safety skills by reproducing a situation in which occupants of a University building (professors, students and administrative staff) have to evacuate in case of a fire emergency. The game is played from a first-person perspective and the character can move within the building and interact with available objects with the goal of evacuating the building following emergency signs and, at the same time, avoiding erroneous actions like taking an elevator, inhaling smoke, etc. When the game ends the system displays a report detailing the success or failure of the player, the time required to play the scenario, and the correct and incorrect actions, also explaining why they were deemed correct or incorrect.

A similar approach is presented in (Silva *et al.*, 2013), where a serious game named EVA, recreating a virtual fire evacuation drill is proposed. The player starts in a given room of a building on fire and has to traverse the building to go outside as quickly as possible, choosing from one of the possible exits. Several emergency signs are in place so as to help the player to identify the nearest exit to her current spot.

In such initiatives, games are seen as stand-alone training resources rather than parts of complete training plans on emergency and evacuation training. The assessment potential of such tools, the ability of discovering incorrect behaviours and the possibility to correct them with targeted training actions is then underused when not completely disregarded.

An example of integration of a game on emergency preparedness with a learning management system is reported in (Dunwell *et al.*, 2014). The main purpose of that study was to demonstrate the benefits of "*allowing learners to interact with immersive technologies as part of a broader structured learning experience*" supporting, in this way, an intuitive-guided approach to learning. In that work, even if the integration issue is challenged, it has only partially solved admitting the construction of training plans including games but disregarding the possibility to assess the learner within the game and using the evaluation to affect the overall training plan.

Our work aims at overcoming these limitations through the definition of a theoretical framework enabling learner assessment within a serious game and the use of assessment results outside game boundaries to customize a micro-learning path aimed at reinforcing weak skills. The framework has been implemented and experimented in conjunction with a serious game for evacuation training targeted to primary and secondary schools students.

3 The theoretical framework

The proposed system relies on the definition of an assessment framework inspired to ECD (Mislevy & Haertel, 2006). The framework is based on the following three models:

- The *skill model* defines and relates the skills we wish to measure and associates a measure of proficiency to each of them;
- The *evidence model* describes how actions performed within the game (observable variables) impact on the proficiency of each skill;
- The *task model* specifies how to detect useful actions within the game without interrupting the learner interaction.

The following sub-sections elaborate on each of these models. As we will see, they show a simpler formulation with respect to *ECD*, grounded on knowledge-based structures rather than on probabilistic approaches and obtained as extension of the model proposed in (Capuano & King, 2015). Then, section 4 describes the developed serious game that may be seen as an implementation of the *ECD presentation model*.

3.1 The skill model

The skill model describes and relates the skills needed to safely perform an evacuation during an emergency. It is defined as a graph *O* (*S*, *HP*, *IRB*) where *S* is a set of skills while *HP* (Has Part) and *IRB* (Is Required By) are set of edges corresponding to different types of relations (Capuano *et al.*, 2015), whose meaning is described below:

- *HP*(s_1, s_2) means that the skill s_2 is part of s_1 i.e. s_1 is acquired if and only if every s_2 so that s_2 is part of s_1 is acquired;
- *IRB* (s_1, s_2) means that the skill s_1 is required by the skill s_2 i.e. in order to acquire s_2 it is needed to have acquired s_1 before;

The defined model is shown in Figure 1. In the graph, the skill *emergency preparedness* is composed of three sub-skills: *evacuation conduct*, *security sigs awareness* and *behaviour in case of emergency*. The latter skill is in turn composed of three sub-skills: *earthquake behaviour*, *gas leak behaviour* and *fire behaviour*. Moreover, the *IRB* relations state that the *evacuation conduct* should be acquired before the *security sigs awareness* that, in turn, should be acquired before the *behaviour in case of emergency*.

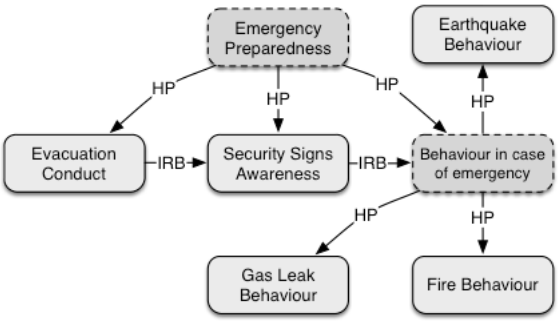

Fig. 1 - The defined skill model.

Edges belonging to *IRB* are not considered for assessment but play an important role in the subsequent arrangement of micro-learning resources for reinforcement (Capuano *et al.*, 2014) that is not object of this paper.

A *proficiency value* $e_s \in [-1, 1]$ is associated to each skill s for each learner. To consider a skill as acquired by a learner, its evaluation should be greater than 0. Skills without outgoing *HP* edges (named *leaf* skills) are evaluated according to the evidence model described in the next sub-section while the proficiency of higher level skills (i.e. skills with outgoing *HP* relations, shown with dashed borders in Figure 1) is obtained by mediating the values of the skills directly connected with an outgoing *HP* edge.

3.2 The evidence model

The evidence model connects correct and incorrect actions that may be performed within the game with skills that are object of evaluation. The Figure 2 shows actions connected with *leaf* skills taken from the skill model depicted in Figure 1. A connection between a skill *s* and an action *a* has an associated mark $m_{sa} \in [-1, 1]$, that specifies how much *a* is indicative of a high (in case of positive mark) or low (in case of negative mark) proficiency in the skill *s*. In Figure 2, solid lines indicate positive marks while dashed lines are negative.

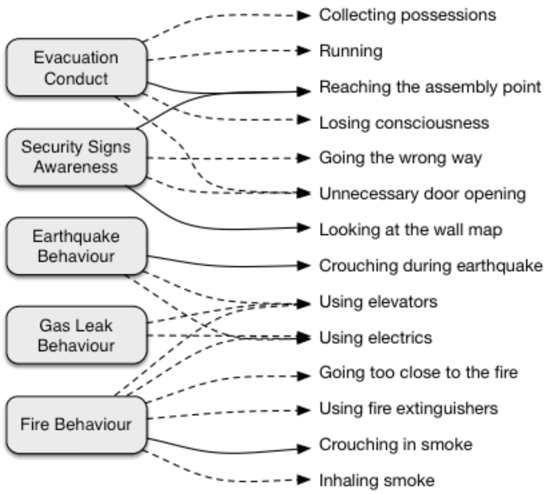

Fig. 2 - The defined evidence model.

The model foresees that both performing and avoiding performing an action is relevant. So, for each positive action *a*, a negative action $\bar{a}$ exists whose mark, for a given skill, is equal to the inverse of the corresponding positive action. For example, if the mark of the action *collecting possessions* for the skill *evacuation conduct* is equal to -1 then the mark of the implicit negative action *leaving possession behind* is equal to 1 for the same skill.

Starting from the evidence model, it is possible to estimate the value e_s for each skill *s* based on actions $A = \{a_1, \ldots, a_n\}$ performed by a learner in the serious game (including both positive and negative actions). Proficiency values are initially settled to 0 and, after each interaction with the game, a general evacuation mark *m* is calculated with the following equation:

$$m = \begin{cases} \max\left(0, \frac{\tau - t}{\tau}\right) & \textit{if the evacuation succeedes} \\ -1 & \textit{otherwise} \end{cases}$$

where t is the time spent by the player to successfully finish the game (i.e. to evacuate the building) while τ is the time limit expected for the played scenario.

Intuitively, the less time spent to complete evacuation, the higher is the score obtained. In case the player does not finish the game, m is settled to be -1. After that, the value of e_s is calculated, for each skill s, according to the following equation:

$$e_s = km + (1-k) \sum_{a \in A(s)} \frac{m_{sa}}{|A(s)|}$$

where $A(s) \in A$ is the subset of performed (positive or negative) actions that are connected to the skill s (as reported in the evidence model) and $k \in [0,1]$ is a constant that regulates the impact of the global performance with respect to evacuation time on the skill evaluation.

Skills showing an evaluation $e_s > 0$ are considered as acquired by the learner while skills so that $e_s \leq 0$ have to be supported by micro-learning resources before a subsequent play. Resources are arranged automatically basing on skill model relations according to the al algorithm discussed in (Capuano & King, 2015). A sample micro-learning resource connected with the skill *security signs awareness* is shown in Figure 3h.

3.3 The task model

The *task model* monitors and identifies actions performed by the learner within the game by relying on virtual *markers* placed in strategic points in each game scenario. Markers record learner time and state as she passes within a radius of a single point in the virtual space allowing, in this way, for a number of metrics.

For example total evacuation time is calculated in terms of time from starting to exit markers; the exit marker is also used to activate the action *reaching the assembly point*; placing multiple path markers, route selection is analysed and the action *going the wrong way* (or the related negative action) can be activated; markers linked to individual items, e.g. an elevator call (Figure 3g), indicate whether a player activates the item and registers the corresponding action for assessment.

The implementation of these markers is made directly in the game code and is invisible to the learner. Data gathered through markers is used both immediately to provide feedback to the player and at the end of the game to build the list of positive actions that is sent to the evidence model together with the total playing time. Negative actions are detected by observing the absence of the corresponding positive actions. The specific action reaching *the assembly point* states that the evacuation succeeded.

At the end of the game the general score *m* is calculated according to the evidence model and displayed to the learner. A personalized message is then built and displayed according to proficiency levels calculated for each skill. Micro-learning resources are then arranged according to performed actions and provided as a support to improve evacuation proficiency.

4 The Serious Game

The developed serious game adopts a three-dimensional environment where the player has freedom of movement to explore in a first-person perspective. Built with the *Unity 3D Game Engine*, the game has been designed from 2D architectural schematics of four different Italian schools and has been guided by photos taken inside school buildings in order to increase visual fidelity.

The architectural schematics have been imported into a 3D modelling application and then adapted to create a coherent structure suitable for first-person rendering. They were then imported into a game engine. The import path presented several challenges in adapting a model originally unintended for real-time rendering to be suitable for a real-time and exploratory environment: polygon counts, overlapping geometry and texture resolution had to be managed to allow adequate performance, and corrections made to the model to support freedom of movement through the building.

Collision meshes have been implemented as well as a simple physical model for the player camera, ensuring that the player is constrained by gravity and solid objects. Light sources have been positioned and walls have been textured using images coming from interior photos. Similarly the floor has been retextured, using specular and normal mapping to emphasise the divisions between individual boards for establishing a realistic environment.

The Figure 3(a-d) shows the process of creation of the school building from blueprints (a) to the 3D model (b). The interior (d) has been designed and textured to be as close as possible to available photos (c) also taking into account performance constraints.

In order to improve game performances, scenario paths have been drawn: the areas of the building through which the player is allowed to navigate have been mapped in advance for each scenario (see Figure 3a). In this way it was possible to increase the fidelity of these areas and remove extraneous content through occlusion culling.

The environment has been populated with simulative objects in the form of 3D models and textures. Objects like fire and smoke (see Figure 3e) are capable of inflicting damage to the player. Image post-processing is used in order to distort and obscure vision upon exposure to smoke, this is accompanied by a health bar that shows the level of smoke exposure. The implementation of these overlaid elements, which include a timer, are examples of game-based elements being introduced, reducing the fidelity of simulation towards the goal of creating an effective gameplay model.

Fig. 3 - Virtual model creation (a-d) and screenshots from the game (e-h).

Several other interactive elements have been included in the game like evacuation maps (Figure 3f), elevators (Figure 3g), extinguishers, fire alarms, doors, etc. To improve the fidelity, in the earthquake scenario ceiling tiles are added and given physical behaviours using the same model applied to the player; as the simulation starts these tiles fall and collide with the floor.

Communication between the *Unity 3D Engine* and the assessment and micro-learning components is performed through the *Unity Scripting API* invoked by *Javascript* code embedded in the player Web page. Assessment and micro-learning components have been developed in *C#* and integrated in an existing learning system (Capuano *et al.*, 2009) that is also in charge of the delivery of reinforcing resources (e.g. Figure 3h) arranged according to assessment outcomes.

5 Experimentation and Evaluation

In order to validate the defined approach and the related theoretical and technological components as well as to measure their effects in the learning process, an experiment has been organized with 45 students (and 4 tutors) coming from 4 different Italian secondary schools (8 students from schools 1, 3 and 4; 21 students from school 2, one tutor per school). Experiment sessions were organized in the PC lab of each school and spanned two hours each (in particular two sessions have been organized for school 2 with two sub-groups of 10 and 11 students).During the first 30 minutes the purpose of the experiment was explained by tutors and questionnaires to be filled were distributed. Then the students were left free to familiarize with the learning environment without accessing the game. During the subsequent 60 minutes the students were allowed to access the game as well as connected micro-learning resources arranged by the system automatically, according to student performances. During the last 30 minutes, students had to fill the questionnaires.

The questionnaires covered two different aspects of the game: perceived value and usability. The obtained results are described and discussed in the next two subsections. The effectiveness of the serious game as didactic resource was also evaluated through the collection of quantitative data about user performances as well as through interviews with the tutors. The result of this analysis is reported in the third sub-section.

5.1 Perceived value

To measure the level of engagement conveyed by the proposed learning approach in terms of perceived value and quality, the participating students have been asked to provide an answer to the following set of questions after a session of one hour of interaction with the system (including both the game and the supporting micro-learning resources).

1. How responsive was the game to actions that you performed?
2. How much did the graphical aspects of the game involve you?
3. How compelling was the experience of moving through your own school building?
4. How much did the hardware quality interfere or distract you from performing the assigned tasks?
5. To what extent have you been engaged in the solution of the game?
6. How easily did you control the interaction with the game?
7. Were the provided supporting learning resources consistent with correct and wrong actions performed within the game?

The possible answers for these questions, corresponding to the 5 points of the Likert scale, are: *not at all* (1), *a little* (2), *moderately* (3), *very* (4) and *completely* (5).While the majority of questions have a positive polarity, question 4 has a negative polarity (i.e. a high score corresponds to negative evaluation).

The Figure 4 shows the average students' evaluation with respect to each question. Many students have answered with a high score to question 5 (M 4,09 SD 1,05) showing that the proposed training method was compelling and has engaged the learners. The medium score obtained to question 2 (M 3,27 SD 0,86) shows that the visual aspect is only partially responsible for the high level of engagement. Other explanations could be found in the game playability and in the virtualization of the school building. This latter hypothesis is supported by the high score obtained to question 3 (M 3,84 SD 0,95).

A moderately low score has been obtained in correspondence of question 1 (M 2,84 SD 0,96) which could be justified by taking into account that the PCs that have been used in the experiment, belonging to school labs, didn't have any dedicated graphics hardware so in many cases the game has been played with a low frame-rate. This hypothesis is also confirmed by the relatively high score obtained on question 4 (M 3,42 SD 1,10). It is worth noting that the lowest values for question 1 have been obtained by the same schools (1 and 2) that obtained the highest score on question 4.

The score obtained on question 6 (M 3,44 SD 0,93) suggests that the user interaction with the game is good but may be improved. Even if this result may be affected by the hardware performance issue, it can be acceptable by considering the prototypical nature of the system. Finally, as shown by the moderately high score obtained on question 7 (M 3,93 SD 0,94), students felt the supporting micro-learning resources useful and consistent with the actions performed during the game. This sustains the effectiveness of the evidence-based assessment.

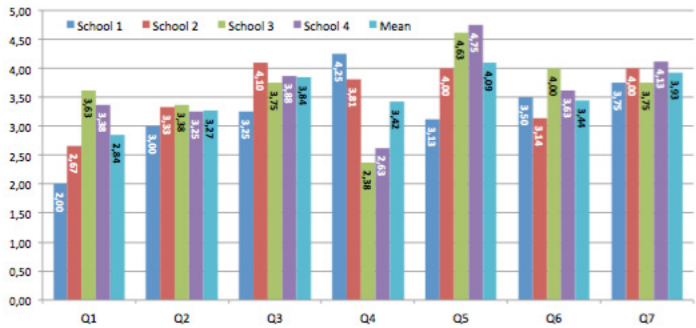

Fig. 4 - Mean score for questions on perceived quality (by school and overall)

5.2 Usability

To measure the usability of the developed prototype system we used the *System Usability Scale* (SUS) defined in (Brooke, 1996). SUS is based on a set of 10 questions whose answers are given on the 5-point *Likert* scale: *strongly disagree* (1), *disagree* (2), *neither/nor* (3), *agree* (4) and *strongly agree* (5). The 10 items that compose the SUS questions are:

1. I think that I would like to use this system frequently;
2. I found the system unnecessarily complex;
3. I thought the system was easy to use;
4. I think that I would need the support of a technical person to be able to use this system;
5. I found the various functions in this system were well integrated;
6. I thought there was too much inconsistency in this system;
7. I would imagine that most people would learn to use this system very quickly;
8. I found the system very cumbersome to use;
9. I felt very confident using the system;
10. I needed to learn a lot of things before I could get going with this system.

The SUS yields a single number representing a composite measure of the overall usability of the system. To calculate the SUS score it is necessary to sum the score contributions from each item. For items 1, 3, 5, 7 and 9 (that have positive polarity) the score contribution is the scale position minus 1. For items 2, 4, 6, 8 and 10 (that have negative polarity), the contribution is 5 minus the scale position. The final score, ranging from 0 to 100 is obtained by multiplying the sum of the scores by 2.5.

The mean SUS score obtained for the serious game, over all users, according to provided answers is 61.28 (SD 4,35) with a minimum of 52.5 (achieved twice) and a maximum of 70. This is an encouraging result given the prototypical nature of the system, also taking into account the low performances offered by the lab hardware.

5.3. Effectiveness as Didactic Resource

In order to analyse the effectiveness of the serious game as didactic resource, we investigated, for each student and classroom, the average number of accesses to the game and the results of the evidence-based assessment mediated over all defined skills and normalized between 0 and 10. The table 1 reports the obtained results.

As it can be seen the average evaluation for emergency skills is 6,44 (SD 1,65) that is 1,44 points above the threshold of 5 (normalized from 0) used by the micro-learning resources arrangement algorithm. The total number of accesses made by the 45 students is 173 with an average of 3,84 accesses per user (SD 1,81), a minimum of 1 and a maximum of 10.

To understand if the practice with the game helps in obtaining better performances on modelled skills the *Pearson* correlation index ρ_{XY} has been calculated between the number of accesses to the game by a student (X) and the overall evaluation obtained by the same student on emergency skills (Y). The obtained result of $\rho_{XY} = 0,69$ shows an *almost strong* direct correlation between the two variables (where a correlation is usually defined as *strong* when ρ_{XY} is greater than 0,7).

Interviews with the 4 tutors have also been performed with the aim of obtaining additional advice about the usefulness of the proposed approach. The tutors unanimously consider the serious game as a valuable approach to train emergency preparedness and, two of them, particularly appreciated the mixed approach relying on the engaging driver of the serious game but also exploiting traditional micro-learning resources to reinforce acquired skills. They have found the majority of students very participatory and appreciated an improvement on their understanding of the target topic.

Three of the interviewed tutors suggest the inclusion of the game in the standard school curricula as reinforcement to the mandatory training about security rules and one of them proposes to extend the experiment to teachers and staff members. As for negative aspects, two tutors over 4 suggested to improve the usability of the game that, on some older PCs, presented a low fluency making difficult a smooth interaction.

Table1
CORRELATION BETWEEN THE NUMBER OF ACCESSES AND THE ASSESMENT RESULTS

Subset	Accesses			Assesment		Accesses/Assessment Correlation
	Total	M	SD	M	SD	
School 1	33	4,13	1,25	6,63	1,19	0,71
School 2	86	4,10	2,28	6,19	1,86	0,76
School 3	27	3,38	1,19	6,75	1,28	0,63
School 4	27	3,38	1,41	6,63	1,92	0,75
Overall	173	3,84	1,81	6,44	1,65	0,69

Final Remarks

Serious games have proved their effectiveness when applied to education and training, especially with young learners. Moreover, in addition to the teaching function, they are also able to play a role in assessment that can be done when the learner is playing, without interrupting her, by observing and evaluating what she is doing. Emergency preparedness is a promising field of application for serious games. The immersive approach leads in fact to the player acting as if situated in a real scenario while, adding game elements, results in increased learning transfer and retention thanks to engagement and focus on specific tasks.

Based on these premises we developed an interactive serious game for evacuation training targeted to primary and secondary school students that relies on a theoretical framework able to model skills we wish to develop by analysing actions the learner may perform in the simulated environment and the impacts of performed actions on the proficiency level of each skill. In this way it is possible to discover incorrect behaviours based on the analysis of learner actions and to correct such behaviours through the on-the-fly arrangement of ad-hoc micro-learning resources.

The developed game has been customized on maps coming from four different Italian schools in order to provide the maximum level of fidelity also taking into account the position of the emergency signs and exits. Then, an experimentation involving 45 students and 4 tutors coming from such schools has been performed.

Evaluation results are promising both from students' and tutors' points of view. The developed game was considered as highly engaging by students that have been particularly attracted by the virtualization of their own school buildings. Tutors have appreciated the mixed mature of the approach relying also on traditional micro-learning resources. Results on assessment show a positive correlation between the number of accesses to the game and the proficiency on target skills; the study on usability shows a mildly positive score, which is a positive result when taking into account the prototypical nature of the system.

As for negative aspects we can report, in some cases, a low reactivity of the game, related to the performances of the used lab PCs, resulting in a low frame-rate compromising the playability of the game. Unfortunately, addressing this issue is non-trivial and can't be done by simply reducing the games visual fidelity. Creating games with lower fidelity to meet the specifications of lab hardware risks to make the games themselves little compelling for users familiar with high-fidelity entertainment gaming. So, a proper solution to this issue, would be a reduction of the gap in hardware availability between school and leisure times.

REFERENCES

Abt, C. C. (1970), *Serious Games*. Viking Press, 1970.

Brooke, J. (1996), *SUS: A "quick and dirty" usability scale*. In P. W. Jordan et al. (Eds.), Usability Evaluation in Industry, Taylor and Francis.

Capuano, N., Gaeta, M., Marengo, A., Miranda, S., Orciuoli, F., Ritrovato P. (2009), *LIA: an Intelligent Advisor for e-Learning. Interactive Learning Environments*, Taylor & Francis, 17(3), 221-239.

Capuano, N., Gaeta, M., Ritrovato, P., Salerno, S. (2014), *Elicitation of Latent Learning Needs Through Learning Goals Recommendation*. Computers in Human Behavior, Elsevier, 30, 663-673.

Capuano, N., King, R. (2015), *Adaptive Serious Games for Emergency Evacuation Training*. 7thInt. Conf. on Intelligent Networking and Collaborative Systems, IEEE.

Capuano, N., Longhi, A., Salerno, S., Toti, D. (2015) *Ontology-driven Generation of Training Paths in the Legal Domain*. Int. J. of Emerging Technologies in Learning, 10(7), 14-22.

Chittaro L., Ranon R. (2009), *Serious games for training occupants of a building in personal fire safety skills*. Int. Conf. on Games and Virtual Worlds for Serious Applications, IEEE, 76-83.

Corti, K. (2006), *Games-based Learning; a serious business application*. PIXELearning Limited.

Dunwell, I, Lameras, P., Petridis, P., Arnab, S., De Freitas, S. (2014), *Facilitating Intuitive-Guided Learning in a Serious Game through Integration with a Learning Content Management System*. 8thInt. Conf. on Complex, Intelligent, and Software Intensive Systems, IEEE, 265-272.

Loh, C.S., Anantachai, A., Byun, J., Lenox, J. (2007), *Assessing what players learned in serious games*. 10thInt. Conf. on Computer Games: AI, Animation, Mobile, Educational & Serious Games.

Michael, D., Chen, S. (2005), *Serious Games: Games That Educate, Train, and Inform*.

Course Technology PTR.

Mislevy, R.J., Haertel, G.D. (2006), *Implications of evidence-centered design for educational testing. Educational Measurement*: Issues and Practice, 25(4), 6-20.

Ribeiro, J. Almeida, J.E., Rossetti, R.J.F., Coelho, A. Coelho, L. (2012), *Using Serious Games to Train Evacuation Behaviour*. Conferencia Ibérica de Sistemas y Tecnologías de Información, 771-776.

Shute, V.J. (2011), *Stealth assessment in computer-based games to support learning*. In Computer games and instruction, Information age, 503-524.

Silva, J.F., Almeida, J.E., Rossetti, R.J.F., Leca Coelho, A. (2013), *Gamifying evacuation drills*. 8thIberian Conf. on Information Systems and Technologies, IEEE, 1-6.

Susi, T., Johannesson, M., Backlund, P. (2007), *Serious Games - an overview. Technical report*, HS-IKI-TR-07-001, University of Skövde.