Evaluating Design Elements for Digital Educational Games on Programming: A Pilot Study

Stephanie Heintz
Department of Computer Science
University of Leicester
Leicester, LE1 7RH, UK
sah91@leicester.ac.uk

Effie Lai-Chong Law
Department of Computer Science
University of Leicester
Leicester, LE1 7RH, UK
elaw@mcs.le.ac.uk

Digital educational games (DEGs) are increasingly recognized as a promising tool for learning. To deepen the understanding of how two key components - content and player - of DEG contribute to learning effects, we developed a model on game elements. It informed the creation of a mini-game on programming, which was evaluated with 50 computer science undergraduates as a pilot study. Data were collected with questionnaires on background, domain-specific knowledge as well as user perception and with a screen cast recorder. Results show that most of the design requirements derived from the model were met. Pre-knowledge was found to be a significant factor influencing user perception. Implications to future work on implementing design elements such as feedback are drawn.

Digital educational games, Programming, Design elements, Design requirements, Learning effect

1. INTRODUCTION

Digital educational games (DEGs), given their potential of motivating and engaging players, are regarded as a promising tool for learning. This supposition is substantiated by an increasing number of research studies in the field of HCI and technology-enhanced learning (TEL) (e.g. Svingby, 2011; Dondlinger, 2007)

Despite the progress, there is still a lack of a clear understanding how the components of a DEG contribute to learning efficacy. In this paper we focus on two key components, namely the learning content to be delivered through a DEG and learner whom it is directed to. To maximise the impact of a DEG, it is necessary to understand how the learning and gaming part can seamlessly be integrated and how the game is perceived by learners.

Some research studies on the two game components have been conducted. For instance, Amory and associates (1999) examined learners' preferences for different game types for educational purposes by evaluating their perception of a game in terms of its elements such as sound, graphics, and story. However, their study was on entertainment games and therefore it is unclear if the finding about

preferred game types for teaching is applicable to DEGs. Another related research study on learners' perception of game types was undertaken by Chong and associates (2005). They focussed on different learning styles with reference to entertainment games. Furthermore, Prensky (2005) studied the relationship between learning content and learning activities in the context of DEGs. However, his arguments are more on a theoretical rather than an empirical basis. Rapeepisarn and associates (2008) combined Chong et al's and Prensky's work to derive a set of possible connections between contents, games and learners. Frazer and associates (2008) analysed the educational affordances of different gaming genres with a set of requirements adopted as well as adapted from the evaluation of learning environments. Their finding might be used as guidance on what game could be chosen for a certain type of learning content, but the work was based on the analysis of entertainment games and did not consider learners' perception.

In summary, the number of research studies analysing the learners' perception of the key elements of DEGs as well as the design decisions based on them is limited. To address this gap, it entails the development of a design model comprising critical game elements of DEGs. Such a model

may enable different games to be compared at the element level. Moreover, if the design concepts built upon the game elements are perceived in the way they are intended to and can impart significant learning effects, the model can serve as the reference framework for developing DEGs that meet their basic requirements for learning and gaming.

Based on the review of the related literature, albeit limited, we have developed a preliminary version of a design model for DEGs (Figure 2). To validate the model, we have implemented a pilot study which serves as a preparation for a larger study comparing different DEGs with different game elements. The pilot study has several objectives:

- to understand the relationship between certain characteristics of the target group and their perceptions of a DEG;
- to evaluate how the design decisions made about the intended perception of the DEG, based on the game elements, are actually perceived;
- to determine the learning outcome based on the pre-knowledge of the participants.

In the ensuing text, we describe this model and the pilot study in detail.

2. DESIGN OF AN EDUCATIONAL GAME

When creating an educational game from scratch that is supposed to teach knowledge useful for its player in real life, how to integrate such knowledge into the game is the main design challenge. Specifically, a conceptual framework (Figure 1) and a design model (Figure 2) are introduced which can be used to identify and define requirements for the implementation of such a game considering its entertaining as well as its educational nature.

2.1. Context of an Educational Game

Understanding an educational game as a combination of gaming and learning enables us to identify key factors influencing the success of the game. Gaming needs at least a player and a game to take place and learning involves a learner and learning content. It is necessary to combine both components to build an educational game: content and game need to be integrated while learner and player become one (see Figure 1).

For an educational game to be successful the content needs to be part of the game and delivered through the game to the learner in such a way that the game is enjoyable and the content gets extracted at the same time. The first step is to identify how

Figure 1: A conceptual framework of basic components of digital educational games.

game and content can be combined in a way that it meets learners' needs. The second step is to analyse learners' perception of the game and verify if the design elements are perceived as intended. The final step is to examine the actual learning outcome. These three steps have been conducted in a pilot study and some main results will be presented in the subsequent sections. When designing a DEG, content and target group are usually known and the game should fit these components, which define the context of the game.

For our pilot study, the domain from which the content is derived is computer programming. The selection is driven by the observation that a number of university students find learning programming difficult as well as boring and tend to give it up. We assume that the motivational effect of DEGs can provide a solution to this issue.

University students with some basic knowledge in Java have been chosen as the target group of the study. Specifically, first-year undergraduates from the computer science department of a university in the UK were recruited as participants and a topic from their Java course was selected to serve as learning content. This topic was a framework helping the students understand how lists are processed and trees are accessed through recursion. The functionality of the framework is to handle a string as a list of characters, by trapping it in an infinite loop, which repetitively processes the first character of the string and takes it off until the string is empty. To understand the role of pre-knowledge in the user perception as well as in the learning effect of the game, a second group of students from another university were also recruited; they had a similar level of Java knowledge, but not yet learned the use of this framework during their course (see Section 3).

2.2. Game Elements

According to Rollings and Adams (2003), a game comprises a number of elements such as Challenges, Gameplay, Setting, Interaction Model, Perspective, Modes and Structure. By integrating most of these elements and others derived from other resources (e.g. Dondlinger, 2007; Prensky, 2001; Fullerton, 2008), we have compiled a model representing the architecture of a game (Figure 2).

Figure 2: Model of the elements a game consists of.

The gameplay as the central part of the game consists of two elements: challenges and the actions available to overcome them. The gameplay is driven by pre-defined goals leading to rewards or possible victory or loss conditions. Rules as the underlying concept define these elements. The communication between a game and its players is achieved through an interaction interface that enables the player to trigger the actions. In addition an audiovisual feedback is needed that informs the player about the state of the game. This is based on the game setting. The players' view of this setting can be from the first-person or third-person perspective. The configuration of these elements can change during the game, as new challenges might be introduced when entering a new level or a player might come across a puzzle which gets presented by changing to a different view in order to solve it. These varying sets of element are called modes and organized through a structure of the game. The structure can be supported by a story. A common way to structure a game is via level. Based on this model, we have decided upon the overall design of the game and its elements (Figure 3), which will be described in the following subsection.

2.3. Design based on Game Elements

For the pilot study the fundamental idea was to keep the design of the game as simple as possible. In further studies it is then planned to create different variations for the set of game elements and to compare their effects on the learning quality, which is referred to both players' user experience and learning effect. The decisions made for each element are described below (see also Figure 3).

The interaction was intended to be intuitive and based on familiar concepts. Since the game is supposed to

be played with a computer, a mouse is used for drag & drop activities and clicking on buttons to evoke the functionality assigned to them. The visual feedback was given through the computer screen and no audio was included due to the fact that the study was undertaken with multiple students simultaneously in the same room with no headsets available.

The gameplay was directly derived from the tasks students had to solve to practise the topic described earlier. They were supposed to become familiar with the use of lists by enhancing a given framework for processing strings character-wise and to develop methods that do simple checks on the string such as counting the number of letters it consists of. This was associated with the use of common programming concepts such as counter or flags. These tasks were directly used as challenges in the game. With the given framework, students were asked to count all or only defined letters in the string, check via flags if certain letters are included and to combine both concepts to identify the position of a specified letter in the string. The actions needed to solve these challenges were lines of code. To simplify the coding and avoid problems with the syntax, the required code lines were given to the player in the form of game bricks. A symbol on the brick indicated its functionality, for example "+1" stands for "counter = counter+1". The goal is to solve the task by adding the lines of code at the right position by placing the bricks. To enable the practice of using the bricks and to introduce new bricks, multiple levels with increasing difficulty were designed with each providing a new challenge. Solving a level is rewarded by unlocking the following one. Simple feedback on the correctness of the solution is given, similar to the testing of code in a programming environment. If the code works for the given string, it is accepted as correct even if it does not work in general. While we recognize that providing detailed feedback can be very useful for learning, it is a resource-demanding task to achieve at the current stage of the research project.

Since the intention was to focus on the programming concepts by dissociating them from coding, the visualisation is used to get this across. The basic framework of processing a string character-wise is illustrated through a pipeline. When pressing the provided "play" button an animation is started where one character after another of a presented string flies through. Prior to running the animation the game bricks resembling lines of code can be placed inside the pipeline and will then react to the characters flying past. An "if" brick redirects the character depending on the condition to an underlying field and then back to the pipeline. Required variables like counters or flags are presented below the pipeline with their current value.

Figure 3: Defined requirements from the design of the educational game based on the game elements model.

Since the player should be able to transfer the conceptual knowledge back to coding, a further game element is included which is not essential for the gameplay. This is an additional window showing the code behind the visualisation in the game. This code changes with the arrangement of the code bricks in the pipeline and during the animation the codeline processed at the time is highlighted through a frame.

3. METHOD

The study was conducted in the departmental computer labs. Each student was allocated to a computer workstation, where they could play the game individually. The test procedure was structured as six consecutive steps:

Step 1: Background Questionnaire: All participants were first asked how they assess their own programming skills, how often they play video games, what type of games they like to play, and any former experiences with educational games.

Step 2: Knowledge Test: Students were then given a short test to assess their pre-knowledge about the learning topic. In the first question students were presented with the code used to process a string character-wise and asked to explain it. The following three questions were related to the tasks in level one, two and six of the game. Students were expected to complete or produce Java code for their answers.

Step 3: How-To Video: A screen cast of the game was shown guiding through the completion of the first level as a short instruction on how to operate the game.

Step 4: Educational Game: The game was then started together with a screen cast recorder that captured the player's interaction with the game. While

playing, students did not receive any further help or instruction other than the one provided in the game.

Step 5: Knowledge Test: After playing the game, the students were asked to take the same knowledge test again that they had completed prior to playing the game (Step 2).

Step 6: Game Questionnaire: The student's perception of the game was gathered through a questionnaire asking for their level of agreement on the statements about the game on a five-point Likert scale. The 58 statements were developed by the authors and structured under topics like the use of game elements, the perception of entertainment and learning experiences. Space for comments on each topic was provided.

The study was undertaken with computer science students from two universities: one in England and the other in Germany; the questionnaires and game contents were translated for the latter. It was not intended to do a cross cultural study but to verify the generalisability of the results.

4. RESULTS & DISCUSSION

To evaluate the reliability of results for the target group "students with basic Java knowledge", two groups of students have been involved in the study. The findings are presented in subsection 4.1. Since the perceptions of both groups were very similar, especially with regard to those game-element-related statements, results from both groups were then considered together.

For most of the analysis, participants were divided into two groups with respect to some variables such as "university they study at", "how they rate

their own programming abilities" or "how often they play games". The answers of both groups were then compared to see if there were any significant differences. The non-parametric Wilcoxon rank sum test was used to evaluate the significance of the differences.

4.1. Homogeneity of Perception

In total the pilot study involved 50 participants, 31 from the UK and 19 from Germany, where an additional 6 from the UK were discarded because of their invalid responses. The two groups had a slightly different orientation: Computer Science in the UK and Digital Media in Germany; the average ages of the UK and German students were 20 and 23.4 years, respectively. They had a notable difference in the pre-knowledge tested (see subsection 4.3).

In comparing the responses from the students of both universities, only seven out of the 58 statements showed significant differences. The UK students expressed more interest in some of the suggested changes to the game, namely having time pressure to solve a level (median: UK=4, GER=2; $p<.01$), having a highscore (median: UK=4, GER=2; $p<.01$) and being in competition with other players (median: UK=4, GER=3; $p<.05$). They also agreed more that the game is enough to learn the topic (median: UK=3.5, GER=2; $p<.05$) as well as that the game aims to teach how to process a string character-wise (median: UK=4, GER=4; $p<.05$). This can be explained by the fact that the UK students had a higher pre-knowledge of the topic than the German students. Since the game focuses on teaching concepts like counter and flags by only introducing the framework of processing a string like a list as a given context, it is likely to be less noticed by students with lower pre-knowledge. Overall the German students claimed that the game was not enough to teach the topic. Having the knowledge of the topic could raise the interest for further challenges like time pressure or competition. But it might also result in a reduced need for game mechanics for learning, since the UK students showed more interest in having a pure simulation of how the code is processed instead of the game (median: UK=3, GER=2; $p<.01$). The German students disagreed that the graphical design was appealing (median: UK=4, GER=2; $p<.01$); it could be due to the fact that their study is more design oriented.

4.1.1. Pre-knowledge

Since the level of pre-knowledge is assumed to be the main reason for the differences detected between the two groups, this was chosen as a main differentiator. Participants were divided into two groups with one group consisting of 22 students who had perfect solutions in their first knowledge test (HPK=higher pre-knowledge) and the other group comprising 28,

who made at least some minor errors (LPK=less pre-knowledge).

Students with higher pre-knowledge found it easier to learn programming (median: HPK=4, LPK=3; $p<.01$) and had accurate self-perception by agreeing more strongly on the statements that they are good at programming (median: HPK=4, LPK=3; $p<.05$) and that they already knew everything the game aimed to teach them (median: HPK=5, LPK=3; $p<.01$). Students with less pre-knowledge agreed more strongly with the statements that the main goal of the game is to learn something (median: HPK=4, LPK=5; $p<.05$) and that they have learned something through the game (median: HPK=2.5, LPK=4; $p<.01$). They also agreed more strongly that it was easier to solve the questions in the knowledge test after playing the game (median: HPK=3, LPK=3; $p<.01$) and that the tasks in the game were clear (median: HPK=4, LPK=4.5; $p<.05$).

These findings show that overall students had a good understanding of their pre-knowledge and that having less pre-knowledge leads to a higher awareness of the educational value of the game.

4.2. Perception of the DEG

On average each participant spent 9:42 min on the game (range: 4:14 min to 34:06 min, SD=6:00). In general, most of the time was spent on the level six, where all introduced code bricks were combined and a new one was added as well. Therefore it was the most challenging level and might have benefited from an interim step. This could be the reason for 16% disagreement on whether it is of right level of difficulty; this was one of the weakest results about the approval of the design requirements being met. Only the requirement of providing enough feedback was worse with 20% of disagreement. Another element with slightly divergent perception was the Java code window. This is plausible as it is not necessary for the gameplay, but was only included for the learning purpose to demonstrate the link between the visualized programming in the game and the underlying Java code. All the other requirements had at most 10% disagreement on the statements claiming them to be met and more than 75% agreement except for the statement that unlocking new levels is a satisfying reward. Thus it can be claimed that overall the design requirements have been perceived as achieved by the participants. Furthermore, with 86% agreement and only 2% disagreement the majority of students indicated that they liked the game and about 94% agreed (43% strongly) that they thought that the game could improve the player's understanding of the topic that it aims to teach.

4.3. Learning Outcome

Since the learning topic of the DEG was taken from a course attended by the UK students, they were expected to be already familiar with it and to improve less than the German students. This could be verified by analysing the results of the first and second knowledge test to identify any improvement or deterioration for each student. For the evaluation of the results, it was focused on the overall correctness of the solution, by ignoring minor syntax errors.

It was found that 17 out of the 29 students from the UK (two missed to complete one of the tests) did have a perfect solution for the tasks in the first knowledge test, but only 5 out of the 19 students from Germany. The UK students showed a higher level of pre-knowledge than their German counterparts.

In the second (post gameplay) knowledge test three improvements as well as three deteriorations have been detected in the UK students. There was one clear improvement from weak or wrong answers to a nearly perfect solution and two from nearly perfect to perfect solutions. The deteriorations were from a perfect solution to forgetting one line of code and two students deteriorated on one question from an already wrong solution to even more errors. This might be due to less care was given when having to solve the same test again, and seemed not to be caused by the game. In contrast, eight German students improved and two deteriorated with both producing minor errors by changing to a slightly more game-like solution.

5. CONCLUSION & FUTURE WORK

The pilot study described in this paper exemplifies how the perception of a learner on a DEG can be assessed. A DEG was designed based on our model of game elements. Requirements were specified how these elements should be perceived by a player, for instance, challenges should have the right difficulty or new activities should be introduced with right frequencies. Examining the perception of two different groups of students, it was shown that the design requirements were widely met. The results between both groups mostly converged, strengthening their credibility.

It was also found that overall students did improve on the learning topic through playing the game. But students with higher pre-knowledge did show less improvement than students with lower pre-knowledge but the same number of deteriorations. This leads to the assumption that saturation was reached which might require a more extensive game than the simple one used for the study.

To gather understanding of how the learning quality is related to the design decisions made on basis of the game elements, multiple games will be implemented in next phases of our research project. Comparing these games might reveal which game elements should be used to integrate learning content into a DEG and how they can best support learning.

6. REFERENCES

- Amory et al. (1999) The use of computer games as an educational tool: identification of appropriate game types and game elements. *British Journal of Educational Technology*, 30, 311-321.
- Chong, Y.; Wong, M. and Thomson Fredrik, E. (2005) The Impact of Learning Styles on the Effectiveness of Digital Games in Education. In: *Proceedings of the Symposium on Information Technology in Education*, KDU College, Patailing Java, Malaysia.
- Dondlinger, M. J. (2007) Educational video game design: A review of the literature. *Journal of Applied Educational Technology*, 4(1), 21-31.
- Frazer, A.; Argles, D. and Wills, G. (2008) The Same, But Different: The Educational Affordances of Different Gaming Genres Advanced Learning Technologies. *IEEE International Conference on Advanced Learning Technologies*, Santander, Cantabria, Spain, July 1-5, 2008, 891-893. IEEE Computer Society Press.
- Fullerton, T. (2008) *Game design workshop: A play-centric approach to creating innovative games* (2nd edition). Morgan Kaufmann, Waltham, MA.
- Prensky, M. (2001) *Digital Game-Based Learning*. McGraw-Hill, New York.
- Prensky, M. (2005) Computer Games and Learning: Digital Game-Based Learning. In Raessens, J., Goldstein, J. (eds), *Handbook of Computer Game Studies*. The MIT Press, Cambridge.
- Rapeepisarn et al. (2008) The Relationship between Game Genres, Learning Techniques and Learning Styles in Educational Computer Games. *Proceedings of the 3rd international conference on Technologies for ELearning and Digital Entertainment*, Nanjing, China, June 25-27, 2008, 497-508. Springer-Verlag.
- Rollings, A. and Adams, E. (2003) *Andrew Rollings and Ernest Adams on Game Design*. New Riders Games.
- Svingby, G. and Nilsson, E. M. (2011). Research Review: Empirical Studies on Computer Game Play in Science Education. In P. Felicia (eds), *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*. IGI Global. Web.