

Document downloaded from:

<http://hdl.handle.net/10251/65383>

This paper must be cited as:

Martín San José, JF.; Juan Lizandra, MC.; Seguí, I.; García García, I. (2015). The effects of computer-based games and collaboration in large groups vs. collaboration in pairs or traditional methods. *Computers and Education*. 87:42-54.
doi:10.1016/j.compedu.2015.03.018.



The final publication is available at

<http://dx.doi.org/10.1016/j.compedu.2015.03.018>

Copyright Elsevier

Additional Information

The effects of computer-based games and collaboration in large groups vs. collaboration in pairs or traditional methods

Juan-Fernando Martín-SanJosé¹, M.-Carmen Juan^{1*}, Ignacio Seguí², Inmaculada García-García³

¹ Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera, s/n. 46022 Valencia, Spain

² AIJU, Ibi, Alicante, Spain

³ Departamento de Sistemas Informáticos y Computación, Universitat Politècnica de València, Camino de Vera, s/n. 46022 Valencia, Spain

Abstract

As new technologies have emerged in the last few years, the learning process has been changing. New and powerful e-learning systems are being developed and new teaching methods can be used in classrooms. In this paper, we present a computer-based game with an educational background that is played on a large-size tabletop display. The game can be used as reinforcement for educational content related to historical ages. The game uses natural interaction. A study to compare the traditional learning method with a collaborative learning method using the game was carried out. A group of up to 12 children could learn together using the game. The experience of children in large groups was also compared with the experience of children playing in pairs. One hundred children between 8 and 11 years old participated in the study; they were divided into three groups (LGroup, Pairs, TClass). When the pre-test and the post-test results were compared, it was shown that the children learned the contents in all three groups. The results also showed that there were statistically significant differences between the traditional method and the game played in a large group in favour of children who played the game in the large group. The knowledge acquired was independent from gender and age. There were no statistically significant differences between learning in large groups or learning in pairs. In both cases, the children expressed their satisfaction for the game and found it easy to use. Therefore, playing games of this type collaboratively in large groups or in pairs can be a valuable learning method that can be combined with traditional methods.

Keywords: computer-based games, collaborative learning, large groups, interactive learning environments, media in education

1. Introduction

Most children of this generation have grown up playing with electronic devices, computer games, and using advanced technologies having been surrounded by technology since they were born (Bekebrede et al., 2011; Tapscott, 1998). Technology manufacturers nowadays consider the children to be a significant market target and they are focusing on them as such (Read &

*Corresponding autor: M.-Carmen Juan

Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera, s/n. 46022 Valencia, Spain, e-mail: mcarmen@ai2.upv.es, Phone: +34 963879720, Fax: + 34 963877359

Markopoulos, 2013). This might create expectations about the learning environment that can be found in the classroom (Oblinger, 2004); however, the education of this generation is still using old methods (Beck & Wade, 2006; Prensky, 2001). New technologies can be recognized to be learning tools that help young children learn in relevant ways (Couse & Chen, 2010; Gimbert & Cristol, 2004). Today's children handle digital information, communicate with others via mobile technologies, and play more games than previous generations (Beck & Wade, 2006). Game-based learning might be a more appropriate approach for teaching and engaging the children in a more successful way than traditional learning methods (Prensky, 2001). This may also affect motivation, which is related to the children's willingness to participate in tasks and activities. According to Malone and Lepper (Malone & Lepper, 1987), there are seven factors that promote motivation: challenge, curiosity, control, fantasy, competition, cooperation, and recognition. Many of these are present in games (Prensky, 2001). Games of this kind also provide an argument for learning in a collaborative mode, which offers benefits over individual learning (Johnson & Stanne, 1986; Slavin, 1988). Hogle (Hogle, 1996) suggested that there were some advantages to learning with video games such as trigger instinct, proving high-level thinking, practice and feedback, and memory reserving.

Natural User Interfaces (NUI) have not been widely used in learning environments like classrooms and schools. NUI could be a good complement to the traditional approach to education. Currently, there is a trend to eliminate all gamepads, joysticks, and other input methods and to use detection of the position of different parts of the user's body to control the device. According to Fishkin et al. (Fishkin, 2004), NUI facilitate the acceptance of an application by users. Interactive tables are also enjoyable and engaging to use (Rick et al., 2009) and promote equity of participation (Harris et al., 2009) and encourage learning (Falcão & Price, 2009; Jamil et al., 2011).

Previous works have compared collaborative and individual learning, such as Sung & Hwang's work (Sung & Hwang, 2013). Three different groups participated in Sung & Hwang's study: experimental group that learned using a collaborative educational computer game with a repertory grid approach (three or four students made up a team); a control group (Group A) that learned using conventional collaborative game-based learning (three or four students made up a team); and a control group (Group B) that learned using an educational computer game individually. From their results, it was found that the learning achievements of the experimental group were significantly better than those of the students in control groups A and B. The collaborative educational game with a repertory grid approach also benefited the students by enhancing their learning motivation.

Although previous works (e.g., (Sung & Hwang, 2013)) have compared collaborative and individual learning and have also studied the positive influence of new technologies, to our knowledge, this study is the first work in which a collaborative computer-based game played on a large-size tabletop display and involving up to 12 students at the same time has been compared with traditional learning. This work also compares the game played in large groups with playing in pairs.

Our first hypothesis is that children who learn by playing a computer-based game collaboratively on a large-size tabletop display and involving up to 12 students will obtain significantly higher learning outcomes than those who learn by attending a traditional class.

Our second hypothesis is that children who learn by playing collaboratively a computer-based game collaboratively on a large-size tabletop display and involving up to 12 students will obtain significantly higher learning outcomes than those who learn by playing with the same game, but playing in pairs.

Our third hypothesis is that children who learn by playing a computer-based game collaboratively played on a large-size tabletop display and involving up to 12 students will show significantly higher satisfaction than those who learn by attending a traditional class.

Our fourth hypothesis is that children who learn by playing a computer-based game collaboratively played on a large-size tabletop display and involving up to 12 students table will show significantly higher satisfaction and usability than those who learn by playing with the same game, but playing in pairs.

The paper is organized as follows. Section 2 focuses on work in game applications related to learning in school environments. Section 3 describes the game. Section 4 presents the description of the study. Section 5 presents the results, and section 6 presents the discussion. Section 7 presents the conclusions and future research.

2. Related work

Recent studies have taken NUI interaction into account. One device that facilitates this interaction is Microsoft Kinect© (Kinect). For example, it has been used for physical rehabilitation (Chang et al., 2011; Lange et al., 2011), for navigating with Google Earth (Kamel Boulos et al., 2011), or for videoconferences in which depth perception was added to the participants (DeVincenzi et al., 2011). Other studies have used Kinect for learning purposes using NUI to compare two different types of learning itineraries: the linear learning itinerary (LLI) and the flexible learning itinerary (FLI) (Martín-SanJosé et al., 2014). The comparisons showed there were no statistically significant differences between the two learning itineraries. Many teachers have found that the systems that use interactive whiteboards are a great motivating teaching tool (Rudd, 2007; Warwick & Kershner, 2008). Lien et al. (Lien et al., 2012) developed an L-shape platform where the students could learn by moving their limbs in an easy way.

With regard to studies in which traditional learning methods are compared with learning using new technologies, several works can be cited. Girard, Ecalle & Magnan (Girard et al., 2013) analysed game-based learning tools versus traditional approaches. Their conclusion was that the users of game-based tools had the same learning improvement as with traditional approaches. In 2013, Al-Qahtani & Higgins investigated the effects of e-learning, blended learning, and classroom learning (Al-Qahtani & Higgins, 2013). Three groups were involved: the first group was taught by e-learning, the second group by blended learning (which combines e-learning and traditional teaching), and the third group was taught using the traditional teaching method. The sample consisted of 148 students: 43 students in the first group, 55 students in the second group, and 50 students in the third group. The Ethics unit from an Islamic Culture course was selected for the study and adapted for the online course (the e-learning method). The results showed that there was a statistically significant difference between the blended learning method and the other two methods. However, no significant difference was found between the e-learning group and the traditional learning group. Furió et al. (Furió et al., 2014) compared the learning effectiveness and satisfaction of children using an iPhone game vs. the traditional classroom lesson for learning the water cycle. Thirty-eight children from 8 to 10 years old participated in the study. The children made significant learning gains about the water cycle regardless of the method used. Even though the results showed that the iPhone method achieved higher knowledge results than the traditional classroom lesson, no statistically significant differences were found between the iPhone and the classroom lesson. Chuang et al. (Chuang et al., 2014) studied how happiness can improve learning performance by using interactive systems. The proposed system was a somatosensory system, which is a system that is made up of a number of different receptors like thermoreceptors, photoreceptors, mechanoreceptors, or chemoreceptors like the human body. With this system, users receive abundant feedback in the activity since the responding is direct. The main goal of the users was to solve a puzzle in two modes: scoring mode and timing mode. From the results, Chuang et al. determined that the learners maintained a positive attitude when using their approach for learning purposes. Of all the factors they studied, enjoyment was the most significant one.

With regard to studies in which collaborative vs. individual learning is compared, several works can also be cited. For a single individual, it is hard to achieve complex tasks without the support of other people. Therefore, working together is nowadays highly valued in the workplace (Barron, 2000) and links group members together (Johnson et al., 1998). When compared with the individual learning working collaboratively in groups has been demonstrated to improve students' critical thinking skills, social skills, self-esteem, and problem-solving skills (Gokhale, 1995; Li, 2002) and even improve learning outcomes (Neo, 2003). This collaborative learning combined with the use of new technologies can benefit both students and teachers,

when these technologies are used as communication, repository, or documentation tools (Kaptelinin, 1999).

With regard to the choice of the group size, according to Rau & Heyl (Rau & Heyl, 1990), smaller groups (i.e., three) contain less diversity and may lack divergent thinking styles and varied expertise that help to animate collective decision making. Conversely, in larger groups it is difficult to ensure that all members participate. A review of previous studies shows that the groups are normally made up of between 2 and 6 students. Elices *et al.* (Elices et al., 2002) presented 5 studies of children between 10 and 13 years old. The children had to carry out a task, either individually or in pairs. In the work of Kirschner et al. (Kirschner et al., 2009), the participants in the collaborative group worked in groups of three. The participants were high-school students and the learning problems were related to the field of biology. Gokhale (Gokhale, 1995; Li, 2002) used a group of four. Twenty-four students participated in the collaborative learning group at the college level, and the subject content was series and parallel circuits. A traditional methodology that was comprised of lectures and worksheets was used. Neo (Neo, 2003) used groups of 4-6 people. The students were undergraduates in their first year at the university. These students had to create a collaborative learning environment for a design project. The class was an 8-week course and the students were given lectures, tutorials, and labs.



Figure 1. Room and material used for the study



Figure 2: Children playing the game

3. The game

Our game combines NUI and Frontal Projection. It simulates a tactile screen. The required elements are a sort throw projector and a white table. The screen area was a normal table, which was covered with a piece of cardboard. To achieve the user interaction, the Kinect device was used. It was located next to the projector. The projector and the Kinect device were located above the table. A steel support was built to hold the two devices (Figure 1).

The game was designed based on the experiential learning theory of Constructivism (Dewey, 1963; von Glasersfeld, 1981). As a computer-supported group-based learning system, we also took into account the approach proposed by Strijbos et al. (2004). This approach consists of five elements: three elements are shown as dimensions (learning objectives, task type, and level of pre-structuring); two elements are shown in terms of discrete categories (group size and computer support). We followed the six steps suggested (Strijbos et al., 2004) for the design. The design guidelines for classroom collaborative games proposed by Villalta et al. (2011) were also taken into account. A more detailed explanation about the game design can be found in Martín-SanJosé, Juan, Gil-Gómez et al., (2014) and Martín-SanJosé, Juan, Torres, and Vicent (2014).

The game focuses on learning about historical ages: Prehistory, Ancient Times, the Middle Ages, the Early Modern Period, and the Contemporary Period. In this game, the children play the role of a time traveller, who must make a journey from the past to the present while learning about the most important events that have happened in history. The main goal of the game was to reinforce the learning process of the children in the subject of history.

Each historical age includes several mini-games where the children are given some audio and video explanations telling how the life was during that historical age, thereby introducing them to the historical ages. The audio also explains to the children what they must do in the related mini-games. In the upper left corner there is an avatar character with the shape of an alarm clock that describes the different historical ages to the children and tells them how to play each mini-game.

Since the game was developed using NUI, in each mini-game, the children had to use their own body as the controller, activating buttons by moving their hands around the active area that was projected on a white table as a screen. The game can be played by one or two players at a time. When there is only one player, the player is placed in front of the table in the middle and uses both hands to select the buttons on the right/left sides. When there are two players, one player is placed in front of the table on the left side and the other player is placed on the right

side. The child on the left side must use his/her left hand, and the child on the right side must use his/her right one. This placement helps the children to interact with the buttons close to them. In an attempt to balance the participation of the two children in the pair mode, half of the interactions with the game and half of the answers to the questions should be performed by each child in order to reach the next level. If one of the children is much more active than the other, the most active child could have more interactions than the other child. To limit this situation to the maximum, the buttons for each child are on his/her side and one child would have to invade the physical space of the other child in order to perform that selection. When there are more than two players, they must play collaboratively in pairs by turn, the children have to be placed around the table where the game is played (Figure 1). Since there are 6 different activities possible (5 historical ages and a final puzzle), one pair of children interacts physically with the game while the rest of the other children are around the table. The children that are interacting are responsible for that historical age, but they have to speak with all of their classmates in order to complete the activity. The children collaborate to answer correctly in order to pass to the next historical age. Since the game is collaborative and not competitive, the children do not compete to be the first to answer. If the group is composed of ten children, one of the pairs plays the contemporary period game and also completes the puzzle. If the number of children is not even, then, one of the children plays twice. Figure 2 shows a group of children playing the mini-game related to Ancient Times. The detailed play-through of the game is as follows:

- Prehistory: in this historical age, the children were introduced to the lifestyle of the cavemen and the colors that they used for creating the cave paintings. In the game, the children had to find some cave paintings in a rock-styled wall and use the colors that the cavemen used to color the paintings. After that, the children had to place one of their own hands on the wall using the color they had selected.
- Ancient Times: in this part of the game, the children had to locate some of the most important buildings or constructions of a Roman city (aqueduct, Roman road, amphitheatre, and Roman circus), after that, the children were asked some questions about the use of each construction.
- The Middle Ages: since this is the historical age of the castles, here the main goal of the children was to build a medieval castle from scratch. The game gave the children a description of some of the parts of a castle and when they gave the correct answer, the parts were placed in the correct position and the whole castle is built. These parts were the moat, the barbican, the defensive towers, and the keep.
- The Early Modern Period: when the children reached this stage, they entered into the sailors' historical age and they were asked about the instruments that Christopher Columbus used to perform the discovery of the American continent. These include maps, the astrolabe and the compass.
- The Contemporary Period: at the end of the game, the children arrived to nowadays. A video was shown to the children explaining this historical age.
- After visiting these historical ages, the children had to complete a puzzle by ordering all the historical ages they had visited on a timeline.

4. Description of the study

This section presents the characteristics of the children that played the game, the measurements used during the experiment, and the steps followed.

4.1 Participants

A total of 100 children participated in the study. Two different schools were involved. At the first school, a total of 82 children participated in the study. There were 43 boys (52.44%) and 39 girls (47.56%). They were between eight and eleven years old. The mean age was 8.83 ± 0.72 years old. The children were in 3rd grade (40) and 4th grade (42). There were four classes. Two for 3rd grade and two for 4th grade. These children were attending the Engeba school of Valencia (Spain). At this first school, the children participated by playing the game in large groups and by attending to a traditional class.

At the second school, a total of 18 children participated. There were 13 boys (72.22%) and 5 girls (27.78%). They were between eight and nine years old, and they had already finished the third grade of primary education. The mean age was 8.72 ± 0.46 years old. These children were attending the summer school of the Technical University of Valencia (Escola d'Estiu). At this second school, the children participated by playing the game played in pairs.

4.2 Measurements

To retrieve the data for the analysis, 5 different questionnaires were used. There was a pre-test (A) with only thirteen knowledge questions. This pre-test was designed to measure the knowledge the children had before playing the game or receiving the traditional learning method in the classroom. Afterwards, the knowledge questions were also asked in the B and C questionnaires. Satisfaction and usability questions were also included in these questionnaires. Finally, the children filled out two more different post-tests (D, E) once they received the two kinds of learning. These questionnaires include satisfaction and usability questions and also questions to determine which of the two learning methods they preferred. The questions asked in the questionnaires are detailed in Table 3.

4.3 Procedure

For the comparison between learning in large groups and learning in a traditional class, there were 82 participants divided into four classes. These four classes belonged to two groups for 3rd grade (40 children from 3A and 3B) and two groups for 4th grade (42 children from 4A and 4B). The children in 3A and 4B were assigned to Group α . The children in 3B and 4A were assigned to Group β . The students that had played with the game were also divided into two groups with a maximum of 12 students. The school does not have a policy of dividing the children into different classes based on their academic achievement. Therefore, a priori, we consider the two classes for 3rd grade and the two classes for 4th grade (i.e., third grade 3A and 3B, and fourth grade 4A and 4B), to be homogeneous in age. To corroborate this homogeneity, the previous knowledge of the two classes in each grade was analyzed in Section 5.1.

The procedure to compare between learning in large groups and learning in a traditional class can be summarized as follows:

- Group α : Participants that played the learning game in large groups first and afterwards received traditional learning in the classroom.
- Group β : Participants that received traditional learning in the classroom and afterwards played the learning game in large groups.

Figure 3 shows graphically the procedure for both groups. The following protocol was used:

1. The children filled out the pre-test questionnaire (A).
2. The children from Group α played the learning game in groups collaboratively and children from Group β received traditional learning in the classroom.
3. Then, all of the children filled out the first post-test questionnaire (B for Group α and C for Group β).
4. Afterwards, they received the other learning method: the children from Group α received traditional learning in the classroom and children from Group β played the learning game in groups collaboratively.
5. Finally, all of the children filled out the final questionnaire (D for Group α and E for Group β).

The traditional classroom lesson was given in a traditional classroom. The traditional classroom lesson had the same learning content as the game. The children were not informed about whether or not their answers on the pre-test were correct. Thus, the children did not acquire any knowledge by answering the pre-test; they only learned during the game.

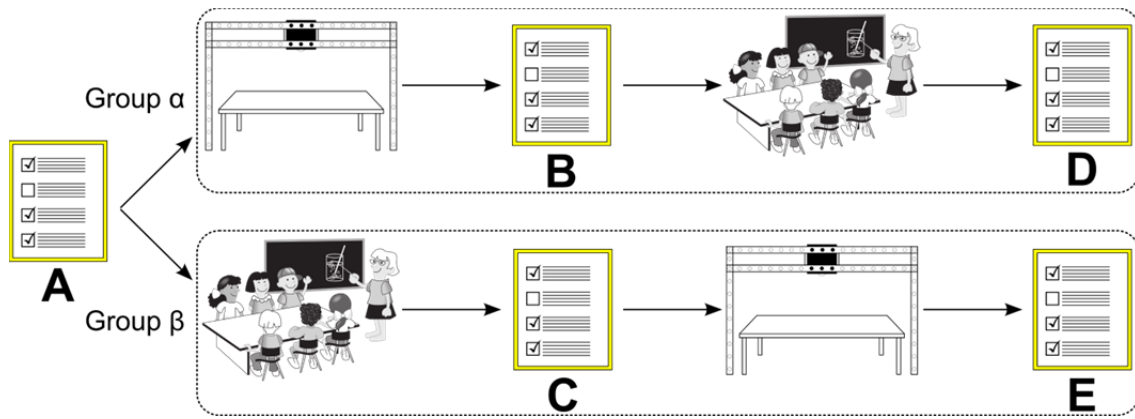


Figure 3: Study procedure

For the pair mode (18 participants), the protocol was the following:

1. The children filled out the pre-test questionnaire (A).
2. The children played the learning game in pairs.
3. The children filled out the post-test questionnaire (B).
4. Afterwards, the children filled out the final questionnaire (D, but only the questions not included in B).

The homogeneity of the pair mode and the 3A group (the large group used for the comparison) was also analyzed and is included in Section 5.1.

5. Results

This section presents the analysis of the data collected from our study to corroborate our four hypotheses. In order to explore means, standard deviations, and other measurements, an initial descriptive analysis was carried out. Then, data normality was checked and, the pertinent statistical tests were carried out based on those results. Before using inferential tests, Shapiro-Wilk and Anderson-Darling tests were performed to check data normality. Both tests reported that our data did not fit the normal distribution. For this reason, the tests used were non-parametric (the Mann-Whitney U test for unpaired data, the Wilcoxon Signed-rank sum test for paired data, and the Kruskal-Wallis test instead of ANOVA). The data from the study were analyzed using the statistical open source toolkit R.

5.1 Learning outcomes

In order to measure how much the children had learned, the knowledge variable was analyzed before playing (Pre-test) and after playing (Post-test). The knowledge variable was created to condense the thirteen knowledge questions by counting the number of correct answers. They were multiple-choice questions with 4 to 6 possible answers that were scored as success or fail. As an example, one of these questions was: *Ancient Times started with the: a) The invention of the wheel; b) The invention of writing; c) The discovery of America; d) The fall of the Roman Empire; e) The invention of the compass.*

Large Group vs. Traditional Class

Before using inferential tests, Shapiro-Wilk and Anderson-Darling tests were performed to check data normality. Both tests reported that our data did not fit the normal distribution. For this reason, the tests used were non-parametric (the Mann-Whitney U test for unpaired data, the Wilcoxon Signed-rank sum test for paired data, and the Kruskal-Wallis test instead of ANOVA). The descriptor of each group is presented in the format (*median; interquartile range*), and all tests are presented in the format (*statistic U/W, normal approximation Z, p-value, r effect size*); ** indicates the statistical significance at level $\alpha = 0.05$.

To determine whether or not there were differences between the initial knowledge of those two groups, an unpaired test was performed between the knowledge variable in PreLGroup (3; 2) and the knowledge variable in PreTClass (2; 1) ($U = 1031, Z = 1.833, p = 0.067, r = 0.202$).

These results revealed that there were no statistically significant differences between the knowledge of the two groups in the pre-test. We also checked if there were differences between the initial knowledge of the children from each class (3A, 3B, 4A, and 4B). The results of these tests were that there were no statistically significant differences between the classes 3A (3; 1.25) and 3B (2; 2) ($U = 276.5$, $Z = 1.422$, $p = 0.159$, $r = 0.219$), or between the classes 4A (3; 2) and 4B (3; 1) ($U = 224.5$, $Z = 0.105$, $p = 0.927$, $r = 0.016$).

To determine whether or not there were differences between the initial knowledge and the knowledge after playing the game, a paired test was performed between PreLGroup (3; 2) and PostLGroup (8; 5) that showed that there were statistically significant differences between the scores that the children obtained before and after playing the game in large groups ($W = 2.5$, $Z = -5.481$, $p < 0.001^{**}$, $r = 0.605$). Another paired test revealed that there were also statistically significant differences between the scores of PreTClass (2; 1) and PostTClass (4; 4) ($W = 41.5$, $Z = -4.639$, $p < 0.001^{**}$, $r = 0.512$), which refers to before and after receiving the traditional learning lesson in the classroom. Finally, to determine whether or not there were differences between the two learning methods, an unpaired test was performed between PostLGroup (8; 5) and PostTClass (4; 4) ($U = 1226.5$, $Z = 3.603$, $p < 0.001^{**}$, $r = 0.398$). These results showed that there were statistically significant differences in favour of the game learning method playing in large groups. Figure 4 shows the box plot for the scores before and after playing for the group of children that played the game in a large group and the children that received the traditional class. Therefore, our first hypothesis has been corroborated.

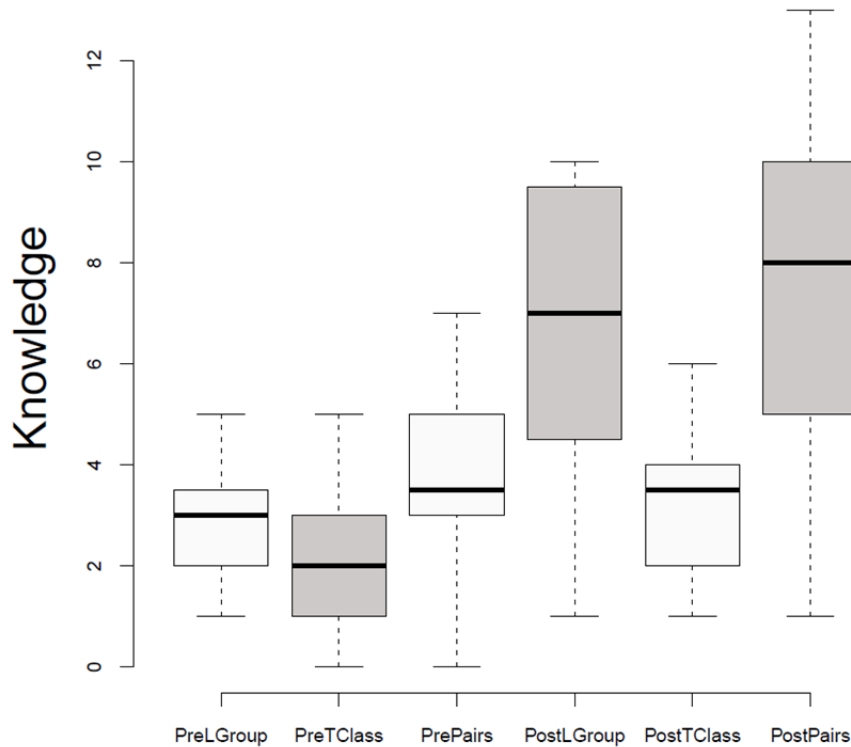


Figure 4: Scores of the knowledge variable before and after receiving either the game method (played in a large group or in pairs) or the traditional method.

The Kruskal Wallis test was also performed to take into consideration several factors in the study: Gender, Age and Learning method. The results, which shown in Table 1, revealed that the knowledge acquired was independent from the Gender and the Age factors, and a significant effect of the Learning Method on knowledge ($\chi^2[1] = 12.978$, $p < 0.001^{**}$, $r = 0.398$) was found. The Age factor obtained a very small p -value, although we cannot assume statistically significant differences at level $\alpha = 0.05$.

Table 1

The Kruskal-Wallis tests for the knowledge variable

Factors	Kruskal-Wallis χ^2	d.f.	p -value
---------	-------------------------	------	------------

Gender	0.403	1	0.525
Age	5.938	2	0.051
Learning Method	12.978	1	< 0.001**

For the knowledge variable, Figure 5 shows the interaction plot between the Gender and Learning Method factors. It can be seen graphically that the acquired knowledge was very similar for boys and girls regardless of the learning method they used. Figure 6 shows the interaction plot between the Gender and Age factors, where children from the 9-year-old and 10-year-old groups scored better than the children from the 8-year-old group, but not significantly better.

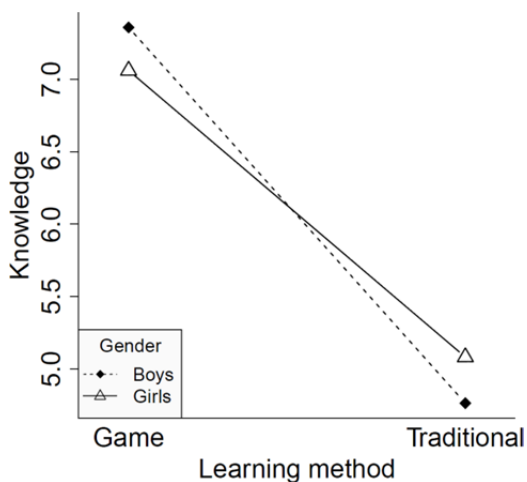


Figure 5: Interaction by gender for each questionnaire

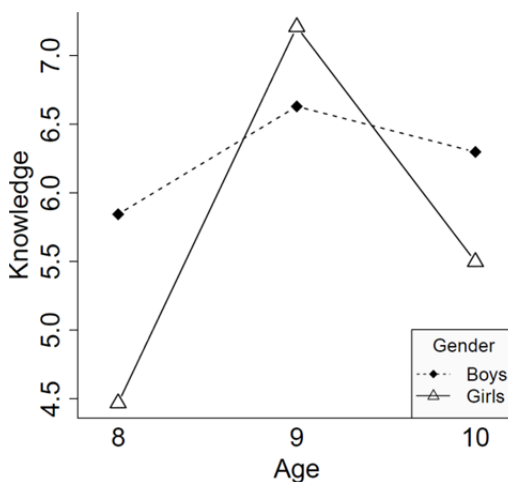


Figure 6: Interaction by gender for each age group questionnaire

Large Group vs. Pairs

To check the homogeneity between the children that played the game in a large group and the children that played the game in pairs, an unpaired Mann-Whitney U test was performed for their initial knowledge. No statistically significant difference was found between PrePairs (3.5; 1.75) and Pre3A (3; 2) ($U = 285.5$, $Z = -1.419$, $p = 0.159$, $r = 0.185$). After playing the game, the results of the comparisons between PostPairs (8; 4.75) and Post3A (8; 5) indicated no statistically significant differences ($U = 372$, $Z = -0.098$, $p = 0.928$, $r = 0.013$). Figure 4 shows the box plot for the scores before and after playing for the children that played the game in a large group and the children that played in pairs. Therefore, our second hypothesis has not been corroborated.

5.2 Rasch model analysis

In order to perform a qualitative analysis, the dichotomous Rasch model proposed by Georg Rasch was used. This model measures a person's latent trait level from a probabilistic perspective (Rasch, 1960). The probability of a user answering a question correctly relies on the user's underlying ability and the difficulty of question (Fischer, 2006).

Figure 7 shows the Item Characteristic Curve (ICC) for each question. The latent dimension shows the ability of the children measured in the interval $[-4, 4]$, with 0 being a child with medium ability. The curve indicates the probability that a child of each ability has to correctly answer a question. The dotted lines represent the medium values for each axis (0 for ability and 0.5 for probability). Figure 7.a shows the ICC for the group of children who played the game first. It can be observed that, the hardest question for this group was Q13, where it was necessary for a child to have an ability value of 2 in order to have a probability of 0.5 to answer the question correctly. The easiest question was Q3, where a child with an ability value of -1 was enough to have a probability of 0.5 to answer correctly. The most balanced question in this

group was Q1, which needed an ability value of 0 (the medium value) to have a probability of 0.5.

On the other hand, Figure 7.b shows the group of children who received the traditional class first. It can be observed that the order of the questions had changed with respect to the other group. In this case, the hardest question was Q3 (which was the easiest question in the other group) and the easiest question was Q12. The most balanced question was again Q1. From these figures, it can be observed that the way the questions are distributed is not balanced in the traditional class group, which means that the contents were not assimilated as well as in the group that played with the game.

A graphical model check was also performed. The questions were grouped by raw scores and the ones that were higher than the mean were separated from the ones that were lower. The red lines represent the confidence bands. The results of the questions for the group of children who played the game first are shown in the graph in Figure 8.a. For this group, every question was inside the confidence bands. In the case of the group of children who received the traditional class first, (shown in Figure 8.b), the questions were more scattered; there are even some questions outside of the confidence bands (Q8 and Q10). This means that since the questions were the same for both groups, the group that played the game first was better prepared to answer them than the traditional class group.

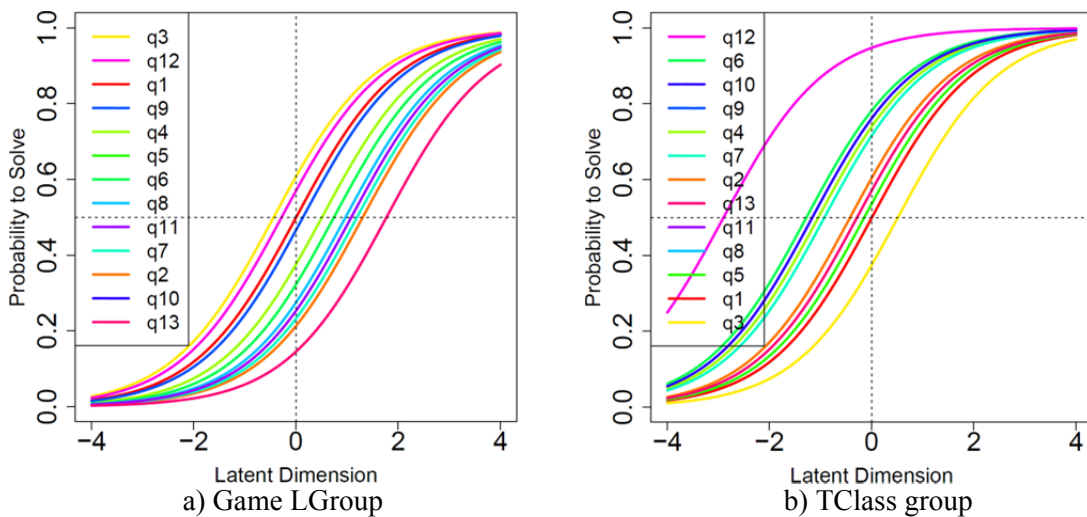


Figure 7: Item Characteristic Curve (ICC) for all questions

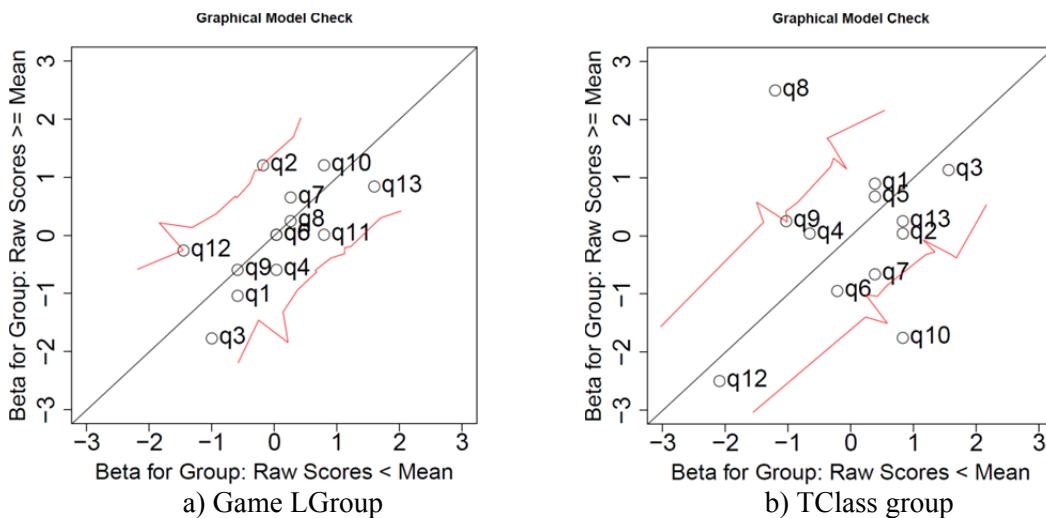


Figure 8: Graphical model check for both groups

In order to visually check the children and the questions, a Person-Item Map was plotted, where the estimated ability of the child and the question difficulty measures are placed side by side in one vertical dimension. The questions appear in order of difficulty. The Person-Parameter Distribution, which is at the top of the graph, is a distribution of the children's abilities. For the game group, Figure 9.a shows the distribution of the children with respect to the questions, and the most of the children are located to the right of the most difficult questions. This means that these children were well prepared to answer those questions. In contrast, in the traditional class group, (shown in Figure 9.b), most of the children are distributed between questions Q12 (the easiest) and Q6 (the second easiest). This means that most of the children were not as well prepared to answer the questions as the children from the other group.

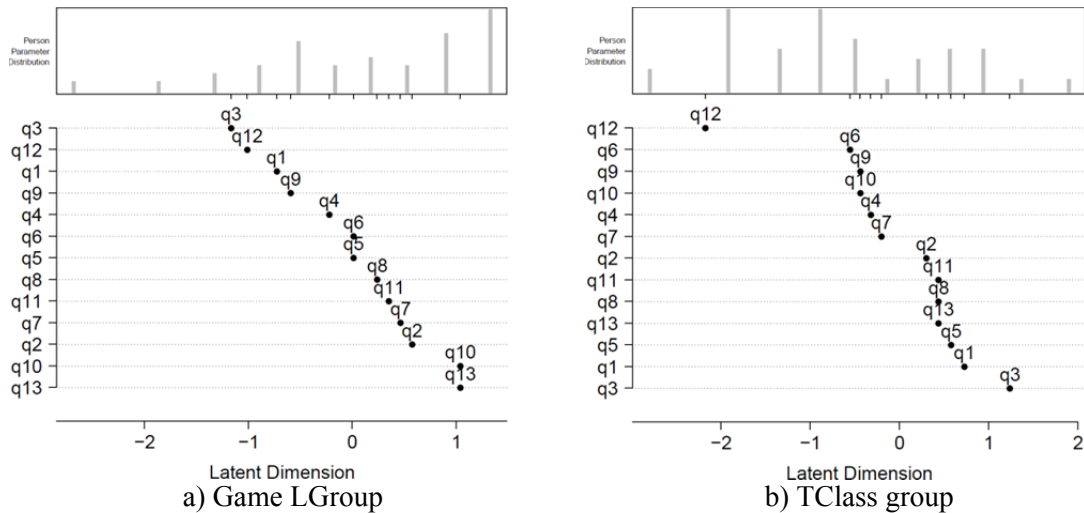


Figure 9: Person-Item Map for both groups

Finally, the test proposed by Andersen (Andersen, 1973) was used in order to check the goodness of fit of the Rasch model. This test is based on a comparison between the difficulties estimated from different score groups and over-all estimates, resulting in a conditional likelihood ratio. Andersen stated that 2 times the logarithm of this ratio is χ^2 -distributed when the Rasch model is true. In our study, this test offered the values, LRvalue = 35.274, $df = 12$, $p = 0$, which fit the Chi-squared distribution. Therefore, the Rasch model is true in our study.

5.3 Satisfaction and usability outcomes

Large Group vs. Traditional Class

Several non-parametric tests (the Mann-Whitney U test and the Wilcoxon Signed-rank sum test) were performed for the Likert questions to determine whether or not there were statistically significant differences for the satisfaction and usability questions (Table 3), depending on which learning method they received first (Q14-Q16, Q18).

In the first place, we define a satisfaction variable that combines the answers of questions related to satisfaction (Table 3). There were no statistically significant differences between the perceived satisfaction playing the game in large groups or the traditional class, ($U = 206$, $Z = 0.202$, $p = 0.853$, $r = 0.032$).

Second, we consider the individual questions and the unpaired data that show the differences between the different groups of children (i.e., the children who received the traditional class first compared with the children who played the game in a large group first). For all the combinations (LGroup first vs. TClass first; LGroup first vs. LGroup second; TClass first vs. TClass second; TClass second vs. LGroup second), there was only a statistically significant difference when the children who played the game in a large group first are compared with those who played the game in a large group second (LGroup first vs. LGroup second), and for the fun perceived in favour of the children who played the game after receiving the traditional classroom lesson, who had a lot more fun ($U = 619$, $Z = -3.078$, $p = 0.003^{**}$, $r = 0.342$).

Third, we consider the individual questions and the paired data that show the differences between the two learning methods (Large Groups vs. Traditional Class). This means that the answers from the children who played the game in a large group first were compared with their answers after receiving the traditional class and vice versa. From these results, no statistically significant differences were found for Q14-Q16, Q18.

Fourth, we consider the individual questions and the questions that are directly related to the game (Q17, Q19-Q24). This comparison includes the answers given by the children who played the game first and the children who played the game second. Statistically significant differences were found only when asking about being able to touch the castle (Q17, $U = 483$, $Z = -3.282$, $p < 0.001$, $r = 0.367$) and when asking about recommending the game to friends (Q19, $U = 676$, $Z = -2.011$, $p = 0.045$, $r = 0.223$). In both cases, these differences were in favour of the children who had played the game second. This indicates that the children were more impressed by the game after receiving the same content from a traditional class.

Although few statistical significant differences were found for the satisfaction questions between the Large Group and the Traditional Class, if the means are considered, in almost all of the questions the means were higher in favor of the Large Group.

The children were also asked which learning method they liked the most (Q26), why (Q27), and what they liked the most about the experience (Q28). The resulting percentages for Q26 are presented in Figure 10. Separated by three different groups (group, grade and gender), it can be observed that in all cases, a high percentage of the children preferred the game over the traditional class. Some of the answers they gave when they were asked why were the following: “because there were funny things”, “because I learned a lot”, “because it has been more fun than in the classroom”, “because we had to use electronic devices”, “because it was entertaining”, or “because with this game you can learn fast”.

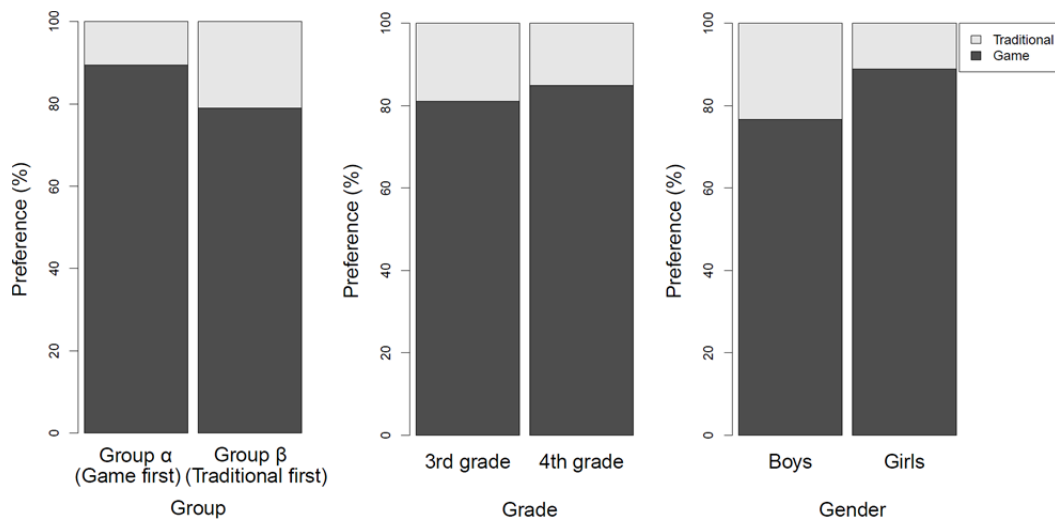


Figure 10: The children’s vote for the preferred learning method from the different groups, grades, and ages

When they were asked about what they liked the most about the experience (Q28) some of the answers were the following: “to touch the table”, “the games”, “to learn the colors the cavemen used”, “the Roman city”, “the cave paintings”, “the projection”, or “Prehistory”. In the last question (Q29), the children were asked if they would like to use this game at school to learn other things. For the group of children who played the game first, all of the children from 3rd grade voted “yes” (100%), while 81.81% of the children from 4th grade voted “yes” and 18.18% voted “no”. For the group of children who played the game second, all of the children (3rd grade and 4th grade) voted “yes” (100%).

Since there were no statistically significant differences between the Large Group and the Traditional Class, these results indicate that our third hypothesis has not been corroborated, even though there is considerable evidence of the satisfaction of the children with the game.

Large Group vs. Pairs

As in the previous analysis, the data of the children that played the game in a large group were compared with those who played the game in pairs. The results for the individual questions are shown in Table 2. These results showed the following: the children who played in a large group found it easier to learn and scored the game higher; they would recommend the game to their friends; and they also appreciated the appearance of the game to a great extent than the children who played in pairs. Even though no statistically significant differences were found between the Pairs and the LGroup for the satisfaction and usability questions, if the means are considered, they were higher in favour of the LGroup in all of the questions. Moreover, if the satisfaction variable is considered, there was a statistically significant difference in favour of the LGroup ($U = 355.5$, $Z = 5.197$, $p < 0.001^{**}$, $r = 0.843$). A similar trend in favor of the LGroup is observed when the usability variable is considered, ($U = 272$, $Z = 2.751$, $p = 0.006^{**}$, $r = 0.446$). These results show that our fourth hypothesis has been corroborated.

Table 2

Medians and Interquartile range of the PostPair and PostLGroup B questionnaires, the Mann-Whitney U test analysis, and r effect size

#	PostPair	$\mu_{PP} \pm \mu_{PP}$	PostLGroup B	$\mu_B \pm \mu_B$	U	Z	p	r
Q14	5;0	4.78±0.42	5;0	4.90±0.44	151	-1.445	0.156	0.234
Q15	5;1	4.44±0.68	5;0.25	4.75±0.43	140	-1.406	0.165	0.228
Q16	4;1	3.89±0.74	5;1	4.60±0.58	89	-2.878	0.004**	0.467
Q18	10;1	9.28±1.04	10;0	10.0±0.0	104.5	-2.959	0.003**	0.486
Q19	4;1.75	3.89±1.41	5;0	4.90±0.44	90.5	-3.275	0.001**	0.531
Q20	4.5;1	4.17±1.21	5;0.25	4.70±0.56	131.5	-1.660	0.100	0.269
Q21	4.5;1	4.11±1.24	5;1	4.55±0.74	144	-1.203	0.236	0.195
Q22	5;1	4.50±0.83	5;0	4.85±0.36	144	-1.420	0.162	0.230
Q23	4;2.5	3.50±1.34	5;0	4.95±0.22	55.5	-4.227	<0.001**	0.686
Q24	4;1.75	4.00±1.00	5;0	4.80±0.51	94	-2.927	0.004**	0.475

6. Discussion

A collaborative game-based learning system that is played on a large-size tabletop display was developed. Our system allows interaction on a table and, as previous works have shown, interaction using interactive tables can promote playfulness (Mansor et al., 2009; Marco et al., 2009), are enjoyable to use (Do-Lenh et al., 2009), and can facilitate the retention of the knowledge acquired (Albert & Mori, 2001; Taran, 2005).

The idea of the tabletop display of our system is similar to those works that have used Microsoft Surface (e.g. (Cao et al., 2010)) or other configurations (e.g. (Hwang et al., 2013)) for collaborative purposes. The main advantage of using material like ours is that the display area is not so limited as in Microsoft Surface. We used a sort throw projector (In Focus IN1503) that can produce a brightness of 3000 ANSI lumens using a resolution of 1280×800 pixels and can generate an image of 177×111 cm. at a throw distance of 140 cm. This projection area can also be increased by using more than one projector. Therefore, our system can be adapted to the requirements of different learning situations (e.g. varying numbers of students or classroom size).

Thanks to our system, several children can play with the game at the same time, thereby saving resources and time and encouraging the children to act collaboratively, which links group members together (Johnson et al., 1998). Moreover, the group size could be quite large (up to twelve). This group size is considerably larger than the group sizes used in previous studies (Sung & Hwang (2013), Hwang et al. (2013), Cao et al. (2010), and Jamil et al. (2011)). In Sung & Hwang's work, three or four students work as a team. In Hwang et al.'s work, three students participated in each group. In other works (Cao et al. (2010) and Jamil et al. (2011)),

the students played in groups of 4. In contrast, Rau & Heyl (Rau & Heyl, 1990) stated that it is difficult to ensure that all members in large groups participate. However, our study demonstrates that it is possible for all of the members in a large group to participate in the learning activity and improve learning outcomes.

From the results, the children learned by playing our game collaboratively in large groups or in pairs. The statistically significant differences between the children's knowledge before and after playing the game proved that games of this kind are suitable for transmitting knowledge. These results are in line with other studies that have demonstrated the following: the use of new technologies for learning purposes can improve the learning process (Furió et al., 2013a, 2013b); games have good learning effects (Egenfeldt-Nielsen, 2007); and playing collaboratively improves learning outcomes (Neo, 2003; Sung & Hwang, 2013).

When the results (after receiving the learning methods) for playing in large groups were compared with attending the traditional class, statistically significant differences were found in favour of the game. If the Person Parameter Distribution from the Rasch model is taken into account, it can be observed that the game provides more ability to correctly answer the learning questions. This ability was higher in the children who played the game. Therefore, it can be concluded that our game facilitates the learning process when played in large groups, which corroborates our first hypothesis: the children who learn by playing a computer-based game collaboratively played on a large-size tabletop display and involving up to 12 students will obtain significantly higher learning outcomes than those who learn by attending a traditional class. However, no statistically significant differences were found between playing in large groups and playing in pairs, which does not corroborate our second hypothesis. Although unexpected, this is a good result because it implies that children can learn by playing in large groups or in pairs without having statistically significant differences in the acquired knowledge. This result offers the possibility of using the game for groups of different sizes without leading statistically significant differences in the acquired knowledge.

The children expressed their interest in incorporating systems of this type for learning different types of content. For the satisfaction between the Large Group and Traditional Class, if the means are considered, the means were higher in favour of the Group for almost all of the questions. However, there were no statistically significant differences. These results do not corroborate our third hypothesis. However, if the Large Group vs. the Pairs is considered, there were statistically significant differences for the satisfaction in favour of the Large Group, corroborating our fourth hypothesis for satisfaction. Moreover, the means for the satisfaction variable were high in all of the cases. Therefore, the children were satisfied with our game. Other authors have studied the positive relationship between satisfaction and learning outcomes (e.g. (Lee et al., 2011; Shea et al., 2004)). Since the children were satisfied with our game, this satisfaction could positively influence their learning outcomes.

Three of the questions (Table 3) were related to usability and the means were very high (greater than 4.5 on a scale from 1 to 5 in the children that played in large groups) indicating that the children considered the game to be very easy to use. The children in the Pairs group scored lower, existing statistically significant differences in favor of the Large Group, corroborating our fourth hypothesis for the usability. However, the means for the Pairs group were also quite high (almost 4 on a scale from 1 to 5). In addition, the people observing the participants during the game stated that a great majority of users did not have any problems interacting with the system. Several authors have considered usability or perceived ease of use as an important technical factor that affects educational effectiveness (Jones et al., 1999; Mayes & Fowler, 1999). Sun et al. (2008) pointed out that learning systems that are easy to use help students to focus their attention on the learning content, and they are more motivated to learn. All of these arguments suggest that our game does help students focus their attention on the learning content, especially when they are playing in large groups.

7. Conclusions

This work presents a collaborative learning game that is played on a large-size tabletop display. The game uses Natural User Interfaces as the interaction method, which allows the children to

control the game with their own hands without using any other external device. From the results, we can affirm that the children acquired new knowledge in the short-term by playing our game collaboratively in large groups and that this increase in knowledge was statistically significantly higher than a traditional method. However, the acquired knowledge was not statistically significant different from playing in pairs.

The cost of building the structure is minimal. The Kinect is not an expensive device; its cost is similar to a basic web camera. The other devices (the projector and the computer) are materials that are commonly found in school. Therefore, cost is not a problem for the incorporation of the system in schools. From our point of view, using systems of this kind combined with traditional learning could be a valuable way to reinforce concepts while students have fun.

In this paper, we have compared two learning methods (playing a game collaboratively on a large-size tabletop display and attending a traditional classroom lesson); however, for future work other comparisons are also possible (e.g., comparing the game played on a large-size tabletop display and playing the game in a PC with typical interaction (mouse) and visualization (monitor)). The game itself could be used for academic assessment, and it would provide immediate feedback to both students and teachers. We have only studied the increase in knowledge in the short-term, but possible future work could study long-term learning. Making the game more customizable could improve the experience; for example, adapting the game difficulty to students in different academic grades and also allowing teachers to be more involved in the game.

Acknowledgements

- This work was funded by the Spanish Ministry of Economy and Competitiveness (MINECO) through the APRENDRA project (TIN2009-14319-C02-01).
- We would like to thank the following for their contributions:
 - Noemí Rando, Encarna Torres, Sonia, Juan Martínez, José Antonio Gil, Lucian Alexandrescu, and M. José Vicent for their help.
 - The Engeba school of Valencia (Spain).
 - The teachers who collaborated in the study.
 - The children who participated in the study.

References

- Albert, D., & Mori, T. (2001). Contributions of Cognitive Psychology to the Future of e-Learning. *Bulletin of the Graduate School of Education*, 50, 25–34.
- Al-Qahtani, A. A. Y., & Higgins, S. E. (2013). Effects of traditional, blended and e-learning on students' achievement in higher education. *Journal of Computer Assisted Learning*, 29(3), 220–234. doi:10.1111/j.1365-2729.2012.00490.x
- Andersen, E. B. (1973). A goodness of fit test for the rasch model. *Psychometrika*, 38(1), 123–140.
- Barron, B. (2000). Achieving Coordination in Collaborative Problem-Solving Groups. *Journal of the Learning Sciences*, 9(4), 403–436. doi:10.1207/S15327809JLS0904_2
- Beck, J., & Wade, M. (2006). *The Kids are Alright: How the Gamer Generation is Changing the Workplace*. Boston: Harvard Business School Press.
- Bekebrede, G., Warmelink, H. J. G., & Mayer, I. S. (2011). Reviewing the need for gaming in education to accommodate the net generation. *Computers & Education*, 57(2), 1521–1529.

- Cao, X., Lindley, S. E., Helmes, J., & Sellen, A. (2010). Telling the whole story: anticipation, inspiration and reputation in a field deployment of TellTable. In *Proceedings of the 2010 ACM conference on computer supported cooperative work* (pp. 251–260). Savannah, Georgia, USA.
- Chang, Y.-J., Chen, S.-F., & Huang, J.-D. (2011). A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in Developmental Disabilities, 32*(6), 2566–2570.
- Chuang, C.-H., Chen, Y.-N., Tsai, L.-W., Lee, C.-C., & Tsai, H.-C. (2014). Improving Learning Performance with Happiness by Interactive Scenarios. *The Scientific World Journal, 2014*(Article ID 807347), 12 pages. doi:10.1155/2014/807347
- Couse, L. J., & Chen, D. W. (2010). A Tablet Computer for Young Children? Exploring Its Viability for Early Childhood Education. *Journal of Research on Technology in Education, 43*(1), 75–98.
- DeVincenzi, A., Yao, L., Ishii, H., & Raskar, R. (2011). Kinected conference: augmenting video imaging with calibrated depth and audio. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 621–624).
- Dewey, J. (1963). *Experience and Education*. New York: Collier.
- Do-Lenh, S., Kaplan, F., & Dillenbourg, P. (2009). Paper-based Concept Map: the Effects of Tabletop on an Expressive Collaborative Learning Task. In *The 23rd BCS conference on Human Computer Interaction* (pp. 149–158). doi:10.1145/1671011.1671028
- Egenfeldt-Nielsen, S. (2007). *Beyond Edutainment: Exploring the Educational Potential of Computer Games*. Continuum Press.
- Elices, J. A., Caño, M. Del, & Verdugo, M. A. (2002). Interacción entre iguales y aprendizaje. Una perspectiva de investigación. *Revista de Psicología General y Aplicaciones, 55*(3), 421–438.
- Falcão, T. P., & Price, S. (2009). What have you done! The role of “interference” in tangible environments for supporting collaborative learning. In *Proceedings of Computer Supported Collaborative Learning 2009* (pp. 325–334).
- Fischer, G. H. (2006). Rasch Models. In C. R. Rao & S. Sinharay (Eds.), *Psychometrics* (Vol. 26, pp. 515–585). Handbook of Statistics. Elsevier. doi:10.1016/S0169-7161(06)26016-4
- Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal Ubiquitous Computing, 8*(5), 347–358.
- Furió, D., González-Gancedo, S., Juan, M. C., Seguí, I., & Costa, M. (2013a). The effects of the size and weight of a mobile device on an educational game. *Computers & Education, 64*, 24–41.
- Furió, D., González-Gancedo, S., Juan, M. C., Seguí, I., & Rando, N. (2013b). Evaluation of learning outcomes using an educational iPhone game vs. traditional game. *Computers & Education, 64*, 1–23.

- Furió, D., Juan, M.-C., Seguí, I., & Vivó, R. (2014). Mobile learning vs. traditional classroom lessons: a comparative study. *Journal of Computer Assisted Learning*, In press.
- Gimbert, B., & Cristol, D. (2004). Teaching curriculum with technology: Enhancing children's technological competence during early childhood. *Early Childhood Education*, 31(3), 207–216.
- Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, 29(3), 207–219. doi:10.1111/j.1365-2729.2012.00489.x
- Gokhale, A. A. (1995). *Collaborative Learning Enhances Critical Thinking*. Digital Library and Archives of the Virginia Tech University Libraries. Retrieved from <http://scholar.lib.vt.edu/ejournals/JTE/jte-v7n1/gokhale.jte-v7n1.html>
- Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., & Rogers, Y. (2009). Around the table: are multiple-touch surfaces better than single-touch for children's collaborative interactions? In *Proceedings of Computer Supported Collaborative Learning 2009* (pp. 335–344).
- Hogle, J. G. (1996). *Considering Games as Cognitive Tools: In Search of Effective "Edutainment."*. University of Georgia Department of Instructional Technology. Retrieved from <http://eric.ed.gov/?id=ED425737>
- Hwang, W. Y., Shadiev, R., Huang, Y. M., Cai, Y. T., Yang, Y. S., & Su, J. H. (2013). Effects of drag-and-response interaction mechanism of multi-touch operated tabletop technology on users' awareness and collaborative performance. *Computers and Education*, 67, 130–141. doi:10.1016/j.compedu.2013.03.004
- Jamil, I., O'Hara, K., Perry, M., Karnik, A., & Subramanian, S. (2011). The Effects of Interaction Techniques on Talk Patterns in Collaborative Peer Learning around Interactive Tables. In *Proceedings of Conference on Human Factors in Computing Systems* (pp. 3043–3052). ACM Press. Retrieved from <http://www.cs.bris.ac.uk/Publications/Papers/2001336.pdf>
- Johnson, D., Johnson, R., & Holubec, E. (1998). *Cooperation in the classroom*. Boston: Allyn & Bacon.
- Johnson, D. W., & Stanne, M. (1986). A comparison of computer-assisted cooperative, competitive, and individualistic learning. *American Educational Research Journal*, 23(3), 382–392.
- Jones, A., Scanlon, E., Tosunoglu, C., Morris, E., Ross, S., & Butcher, P. (1999). Contexts for evaluating educational software. *Interacting with Computers*, 11(5), 499–516.
- Kamel Boulos, M. N., Blanchard, B. J., Walker, C., Montero, J., Tripathy, A., & Gutierrez-Osuna, R. (2011). Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation. *International Journal of Health Geographics*, 10(45), 1–14.
- Kaptelinin, V. (1999). Learning Together: Educational Benefits and Prospects for Computer Support. *Journal of the Learning Sciences*, 8(3-4), 499–508. doi:10.1080/10508406.1999.9672078

- Kirschner, F., Paas, F., & Kirschner, P. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, 25(2), 306–314.
- Lange, B., Chang, C. Y., Suma, E., Newman, B., Rizzo, A. S., & Bolas, M. (2011). Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 2011, pp. 1831–1834).
- Lee, S. J., Srinivasan, S., Trail, T., Lewis, D., & Lopez, S. (2011). Examining the relationship among student perception of support, course satisfaction, and learning outcomes in online learning. *The Internet and Higher Education*, 14(3), 158–163.
doi:10.1016/j.iheduc.2011.04.001
- Li, Q. (2002). Exploration of collaborative learning and communication in an educational environment using computer-mediated communication. *Journal of Research on Technology in Education*, 34(4), 503–516.
- Lien, C.-L., Huang, C.-Y., Wang, C.-Y., & Chen, G.-D. (2012). Using Kinect to track learning behavior of students in the classroom as video Portfolio to enhance reflection learning. In *Proceedings of the 20th International Conference on Computers in Education (ICCE '12)*.
- Malone, T., & Lepper, M. (1987). Making Learning Fun: A Taxonomy of Intrinsic Motivations for Learning. In R. Snow & M. Farr (Eds.), *Aptitude, Learning, and Instruction Volume 3: Conative and Affective Process Analyses* (pp. 223–253). New Jersey: Hillsdale.
- Mansor, E. I., Mansor, E. I., De Angeli, A., De Angeli, A., de Bruijn, O., & de Bruijn, O. (2009). The fantasy table. *Proceedings of the 8th International Conference on Interaction Design and Children - IDC '09*, 70–79. doi:10.1145/1551788.1551801
- Marco, J., Cerezo, E., Mazzone, E., & Read, J. C. (2009). Bringing Tabletop Technologies to Kindergarten Children. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (pp. 103–111).
- Martín-SanJosé, J.-F., Juan, M.-C., Gil-Gómez, J.-A., & Rando, N. (2014). Flexible learning itinerary vs. linear learning itinerary. *Science of Computer Programming*, 88, 3–21.
doi:10.1016/j.scico.2013.12.009
- Martín-SanJosé, J. F., Juan, M. C., Torres, E., & Vicent, M. J. (2014). Playful interaction for learning collaboratively and individually. *Journal of Ambient Intelligence and SmartEnvironments*, 6(3), 295e311.
- Mayes, J. T., & Fowler, C. J. (1999). Learning technology and usability: a framework for understanding courseware. *Interacting with Computers*, 11(5), 485–497.
doi:10.1016/S0953-5438(98)00065-4
- Neo, M. (2003). Developing a collaborative learning environment using a web-based design. *Journal of Computer Assisted Learning*, 19(4), 462–473. doi:10.1046/j.0266-4909.2003.00050.x
- Oblinger, D. G. (2004). The Next Generation of Educational Engagement. *Journal of Interactive Media in Education*, 8, 1–18.

- Prensky, M. (2001). *Digital Game-Based Learning*. California: McGraw-Hill.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.
- Rau, W., & Heyl, B. (1990). Humanizing the college classroom: Collaborative learning and social organization among students. *Teaching Sociology*, 18, 141–155.
- Read, J. C., & Markopoulos, P. (2013). Child-computer interaction. *International Journal of Child-Computer Interaction*, 1(1), 2–6.
- Rick, J., Harris, A., Marshall, P., Fleck, R., Yuill, N., & Rogers, Y. (2009). Children designing together on a multi-touch tabletop: an analysis of spatial orientation and user interactions. In *Proceedings of the 8th International Conference on Interaction Design and Children* (pp. 106–114).
- Rudd, T. (2007). *Interactive whiteboards in the classroom*. Bristol: Futurelab Report: IWBs 2007.
- Shea, P., Fredericksen, E., Pickett, A., & Pelz, W. (2004). Faculty development, student satisfaction, and reported learning in the SUNY learning network. In T. Duffy & J. Kirkley (Eds.), *Learner-centered theory and practice in distance education* (pp. 343–377). Mahway, NJ: Lawrence Erlbaum Associates.
- Slavin, R. (1988). Cooperative learning and student achievement. In R. Slavin (Ed.), *School and classroom organization*. Hillsdale, New Jersey: Erlbaum.
- Strijbos, J. W., Martens, R. L., & Jochems, W. M. G. (2004). Designing for interaction: Six steps to designing computer-supported group-based learning. *Computers & Education*, 42(4), 403–424.
- Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education*, 50(4), 1183–1202.
- Sung, H.-Y., & Hwang, G.-J. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, 63, 43–51. doi:10.1016/j.compedu.2012.11.019
- Tapscott, D. (1998). *Growing up digital. The rise of the net generation*. New York: McGraw-Hill.
- Taran, C. (2005). Motivation Techniques in eLearning. In *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies* (pp. 617–619). IEEE Computer Society.
- Villalta, M., Gajardo, I., Nussbaum, M., Andreu, J. J., Echeverría, A., & Plass, J. L. (2011). Design guidelines for Classroom Multiplayer Presential Games (CMPG). *Computers & Education*, 57(3), 2039–2053.
- Von Glasersfeld, E. (1981). Einführung in den Radikalen Konstruktivismus. In P. Watzlawick (Ed.), *Die Erfundene Wirklichkeit* (pp. 16–38). Munich: Piper.

Warwick, P., & Kershner, R. (2008). Primary teachers' understanding of the interactive whiteboard as a tool for children's collaborative learning and knowledge building. *Learning, Media and Technology*, 33(4), 269–287. doi:10.1080/17439880802496935

Appendix

Satisfaction and usability questions

To determine their satisfaction and the perceived usability, the children were asked the following questions. The TQ column shows the type of questions: SA indicates Satisfaction, US indicates usability, and PRE indicates preference

Table 3: Satisfaction questions and the questionnaires they appeared in.

#	TQ	B	C	D	E	Question
Q14	SA	•	•	•	•	How much fun did you have? [1-5]
Q15	SA	•	•	•	•	How much did you learn? [1-5]
Q16	US	•	•	•	•	How difficult was it to learn “The timeline”? [1.Very difficult / 2.Difficult / 3.Normal / 4.Easy / 5.Very easy]
Q17	SA	•			•	Did you think you were able to touch the castle or the bridge? [1 - 7]
Q18	SA	•	•	•	•	Score the game/activity from 1 to 10 [1-10]
Q19	SA	•			•	Would you recommend this game to friends? [1-5]
Q20	US	•			•	Did you understand the rules of the game? [1-5]
Q21	US	•			•	Selecting the answers was: [1.Very difficult / 2.Difficult / 3.Regular / 4.Easy / 5.Very easy]
Q22	SA	•			•	How much did you like the images in the game? [1-5]
Q23	SA	•			•	How much did you like the Clock Avatar (Mr. Tic-Tac)? [1-5]
Q24	SA	•			•	How much did Mr. Tic-Tac help you during the game? [1-5]
Q25	PRE	•			•	Which of all the mini-games did you like the most? [Prehistory / Ancient Times / the Middle Ages / the Early Modern Period / the Contemporary Period]
Q26	PRE			•	•	Which of the following learning methods did you like the most? a) Frontal projection (the game) b) Traditional class
Q27	PRE			•	•	Why? (Referring to Q26)
Q28	PRE			•	•	What did you like most about the experience?
Q29	SA	•			•	Would you like to use this game at school to learn other things?