The final publication is available at

http://dx.doi.org/10.1080/10494820.2015.1090455

Additional Information

Advanced Displays and Natural User Interfaces to support learning

Juan-Fernando Martín-SanJosé[a], M.-Carmen Juan[a*], Ramón Mollá[a], Roberto Vivó[a]

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

Advanced displays and Natural User Interfaces (NUI) are a very suitable combination for developing systems to provide an enhanced and richer user experience. This combination can be appropriate in several fields and has not been extensively exploited. One of the fields that this combination is especially suitable for is education. Nowadays, children are growing up playing with computer games, using mobile devices, and other technological devices. New learning methods that use these new technologies can help in the learning process. In this paper, two new methods that use advanced displays and NUI for learning about a period of history are presented. One of the methods is an autostereoscopic system that lets children see themselves as a background in the game and renders the elements in 3D without the need for special glasses; the second method is a frontal projection system that projects the image on a table in 2D and works similarly to a touch table. The Microsoft Kinect© is used in both systems for the interaction. A comparative study to check different aspects was carried out. A total of 128 children from 7 to 11 years old participated in the study. From the results, we observed that the different characteristics of the systems did not influence the children's acquired knowledge, engagement, or satisfaction. There were statistically significant differences for depth perception and presence in which the autostereoscopic system was scored higher. However, of the two systems, the children considered the frontal projection to be easier to use. We would like to highlight that the scores for the two systems and for all the questions were very high. These results suggest that games of this kind (advanced displays and NUI) could be appropriate educational games and that autostereoscopy is a technology to exploit in their development.

1. Introduction

The rapid development of technology has provided a lot of new and advanced systems that were unimaginable just a few years ago. Nowadays, the use of technological systems is common for daily tasks such as playing at home. The user increasingly expects to have an experience that is similar to the real world, which means having stereoscopic visualization and interacting naturally. For both of these, the user desires to wear as few devices and wires as possible. The user perceives the illusion of depth with stereoscopic visualization. To achieve stereoscopic visualization, three main technologies are used: passive, active, and autostereoscopic. The main difference between active/passive stereoscopy and autostereoscopy is that the autostereoscopic visualization generates the illusion of depth without the use of special glasses or other headgear. For natural user interaction, Microsoft Kinect© (Kinect) has been a revolutionary device. Kinect is widely used in video-games by connecting it to an Xbox console; nevertheless, it is also possible to develop Kinect programs for PCs. These possibilities have led the natural user interaction to be incorporated in a large number of different types of applications. However, advanced displays and Natural User Interfaces (NUI) have not been extensively exploited in learning environments. From our point of view, this technology is on the right track for being a good complement to the traditional educational approach.

In our systems, the Kinect device was used to recognize the user's gestures. The first system uses an autostereoscopic display as the visualization device, and it merges the

---

image from the real world captured by the camera with the virtual elements that are rendered in 3D. The second system consists of a projected surface that is used as an interactive table. Different technologies such as Augmented Reality (AR) have been used to develop educational systems (e.g. Furió et al., 2013). Taking into account Azuma's definition of AR, our autostereoscopic system cannot be considered an AR system; however, it shows the real world captured by the Kinect camera as the background and mixed virtual elements. In the two systems, the children use gestures for the interaction. The first difference between the two systems is that, in the autostereoscopic version, the children perform the gestures in the air, and in the frontal projection system, the gestures are performed over the table. The second difference is that in the autostereoscopic system, the visualization of the models is rendered in 3D, and the visualization of the projected system is not stereoscopic (2D). Using the combinations of autostereoscopic display+Kinect and projected surface+Kinect, we designed an educational game about historical ages. To our knowledge, this is the first time that these combinations have been used to develop a learning environment for children and have been compared. Our first idea was to compare the two systems without stereoscopy (vertical and horizontal projection/interaction), that is, only checking the effects of NUI (in the air vs. over a table). Our second idea was to add stereoscopy to the two systems. In that case, the two systems would have stereoscopy and NUI, and, therefore, only the effects of the difference in the interaction could be studied. However, we would like to study whether or not autostereoscopy provides positive effects on learning. For that reason, and with the aim of studying the potential of NUI as well, we decided to compare two configurations (frontal NUI+2D visualization vs. vertical NUI+3D visualization). Nevertheless, other possible comparisons could be considered. Several of them are proposed in the conclusion section as future work.

The first objective of this work was to develop two different systems that include advanced displays and NUI. The second objective was to carry out a study to find out which system was most appreciated and effective. The first of our three hypotheses is that the children will prefer the autostereoscopic system over the frontal projection system. The second of our hypotheses is that children will increase their knowledge about the subject treated in the game by using the two systems, and that the autostereoscopic system would lead to greater learning results. Our assumptions for formulating these hypotheses are the following:

- The autostereoscopic display provides an illusion of depth that could improve the immersion in the game. Pang et al. (2006) pointed out the positive effects of stereoscopic video for learning purposes. In their study, Arino et al. (2014) deduced that children perceive the illusion of depth in autostereoscopic displays.
- Since the autostereoscopic display is 46", the fact that the children can play video-games using such a big TV could make a deep impression on them and they might eager to start playing. While playing, the children can see themselves inside the game in the display, and this could give them a sensation of prominence that could encourage them and could influence their motivation and involvement in the game. Some studies have pointed out the positive relationship between satisfaction and learning outcomes (Lee et al., 2011; Shea et al., 2004).

The third hypothesis is that the frontal projection system will be easier to use. Our assumption for this hypothesis is that nowadays children are accustomed to using actual gadgets and peripherals that are controlled in the same way, which is much different than using a gesture-oriented autostereoscopic system. Most children of this generation

have grown up playing with electronic devices and computer games and have been surrounded by technology since they were born (Bekebrede et al., 2011).

2. Background

NUI are an important modality for human-computer interaction. According to Fishkin (2004), NUI facilitate the acceptance of an application by users. Meanwhile (Gope, 2011) stated that compared to many existing interfaces, hand gestures have the advantages of being easy to use, natural, and intuitive. Roman (2010) pointed out that "The mouse's days are numbered"; the current trend in new devices, games, and consoles is to get rid of all gamepads, joysticks, and other input methods. NUI allow users to be the controller themselves by detecting the position of the different parts of their body. In this paper, we present a comparison between two different NUI (touch over a surface vs. gestures in the air). The first part of this section focuses on previous works in which comparisons were performed where one of the systems used was similar to those presented in this paper.

First, several works have focused on touch table/screen devices for determining their usability and advantages over other devices. For example, Bhalla & Bhalla (2010) compared various touch screen technologies and concluded that touch screens have several advantages over other pointing devices, one of which is that they are easy to use mainly because they use direct manipulation. Buisine et al. (2007) studied the usability and usefulness of interactive touch table technologies vs. traditional paper-and-pencil to support group creativity in a mind-map application. Twenty-four users participated in their study. Their results showed no significant difference between the two methods regarding the ease of use. There was also no difference in idea production, but the touch table condition significantly improved both the subjective and the collaborative dimensions. Soro et al. (2011) compared the user's behavior in a task of pair programming, which was performed at a traditional desktop vs. a multi-touch table. Forty-four students between 20 and 35 years old participated in the study working in pairs. Their study showed that people performed significantly better at the multi-touch table than at the desktop, especially for exercises that involved cooperation, discussion, and exchange of communicational information. The conclusions of these works suggest that touch tables/screens are easy to use and promote collaboration.

Second, taking into account the features of Kinect as a device for NUI, several works involving adults have also studied the effectiveness and ease of Kinect with regard to other devices. For example, Juhnke (2013) compared the interaction in a medical imaging application using Kinect and using a mouse. Seventeen people between 22 and 45 years old participated in the study. The results showed that in both cases the participants were able to correctly identify the anatomy with an accuracy of 75%; however, those using the Kinect spent less time to complete the tasks. Libardi et al. (2014) also compared the Kinect with a mouse device for information visualization (e.g., cars). Their study involved twenty participants (teenagers and adults). From their results, they concluded that quantitatively (time, and number of operations) and qualitatively (physical effort and ease), the Kinect device was not as efficient as the mouse. However, the qualitative subjective measures pointed out a higher user satisfaction with respect to the convenience and the adequacy of Kinect. Moreover, the users declared a reasonably higher desire to substitute the mouse for the Kinect. They also identified several disadvantages such as that Kinect requires extra effort and attention from the users who are standing and have to make the right movement at the right time which is in contrast to the mouse interaction where the users are sitting and

only have to move one hand. Francese et al. (2012) compared a Wii Remote and the Kinect for 3D geographical mapping. Twenty-four people between 18 and 41 years old participated in the study. Hand gestures were used to navigate in a virtual environment by measuring yaw, pitch, and roll. In their study, the Kinect showed less variability in task performance and was less distracting. Tsai & Yen (2013) developed a cubic net assisted learning system for enhancing learners' spatial ability that used Kinect for the interaction. Ninety-eight students participated in the experiment. They were taking information technology and science courses at the university. The results showed that the usability of their system was suitable for learning and that it encourages the students to become active learners. Finally, the work that is most closely related to ours is the work by Tuveri et al.(2013), in which they compared the control of a planetarium using the Kinect and a multi-touch table. In that work thirteen users between 21 and 26 years old participated in the study. Even though their results did not reveal differences in the overall usability between the two versions, the means for the multi-touch table were slightly higher. However, the authors indicated that they found a difference in the perceived control of the application (which was higher in the multi-touch version) and in the perceived realism of the experience (which was higher in the Kinect version). With regard to interaction, our work is similar to Tuveri et al.'s work because both compare the interaction of a touch table with the Kinect. However, there are differences: their vertical system did not consider stereoscopy, their study did not consider learning outcomes, their sample size was smaller than ours and their participants were adults and ours were children.

Third, some studies involving children and systems that use Kinect can also be mentioned. A therapeutic modality for children with cerebral palsy using Kinect was presented by Luna-Oliva et al. (2013). In this study, a post-treatment and a follow-up assessment related to motor and process skills were performed. After 8 weeks of treatment, the results showed statistically significant differences between pre- and post-treatment. Hsu (2011) studied the potential of Kinect in education by carrying out a survey of Kinect tools related to interactivity, gestures, teaching, learning, and pedagogical background. These tools were oriented for children, like *Mikumikudance* or *Scratch*. Hsu stated that as a learning tool, Kinect has the affordances to create enjoyable, interesting interaction types to boost student motivation and to promote learning via its multimedia and multi-sensory capacity. Boutsika (2014) suggested using Kinect as a learning auxiliary tool for children with autism. Ten children with moderate autism participated in the study. The results showed that Kinect games enable children to work in teams, and this helps children to cooperate and gradually develop their oral expression. Wang & Cheok (2011) also presented a gaming platform with Kinect. Their game used mixed reality to support playful learning for children. The flexibility of this game stimulated imagination, enabled children to have more control, and allowed them to customize their gaming experience easily. De Greef et al. (2013) stated that one way to improve the engagement of physical therapy is to embed it into a game with the aid of Kinect. The aim of their work was to design Kinect-based games using full body interaction for children with mild motor disabilities.

Fourth, to our knowledge, very few autostereoscopic systems for learning purposes have been presented. One of these works is Arino et al.'s work (2014). Arino et al. carried out a study comparing Augmented Reality (AR) and Virtual Reality (VR) using an autostereoscopic display in which 39 children from 8 to 10 years old participated. In this study, no statistically significant differences were found between AR and VR. Nevertheless, the authors deduced that the children perceive the illusion of depth in autostereoscopic displays even though they only used autostereoscopy and did not

compare 3D vs. non-3D. AR has been used for developing educational applications; for example, the water cycle (Furió et al., 2013a), or multiculturalism, solidarity, and tolerance (Furió et al., 2013b). In the water cycle game (Furió et al., 2013a), the authors presented an educational game for an iPhone and a Tablet PC to reinforce children's knowledge about the water cycle. The effects of the size and weight of the mobile devices were compared. Seventy-nine children from 8 to 10 years old participated in the study. The authors observed that the different characteristics (screen size and weight) of the devices did not influence the children's acquired knowledge, engagement, satisfaction, ease of use, or AR experience. In Furió et al.'s (2013b) work, the authors compared an iPhone game with a traditional game. A total of 84 children ranging in age from 8 to 10 years old participated in their study. The authors did not find significant differences between the two groups for learning outcomes.

3. Systems Development

Custom software was required to develop the two systems. Custom hardware was also required for the frontal projection method. This section presents the design principles and a description of the game and briefly explains the software and hardware required to develop the game for the two systems.

3.1. Game Design

The subject chosen for the game was a historical timeline, with five historical ages (Prehistory, Ancient Times, the Middle Ages, the Early Modern Period and the Contemporary Period). The knowledge presented in the game is the same as what the children study at school. This knowledge was extracted from books used in the classroom. To design our game, several theories and guidelines were considered:

1) The experiential learning theory of Constructivism (Dewey, 1963).
2) The approach for the computer-supported group-based learning system proposed by Strijbos et al. (2004). Strijbos et al.'s 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). The following six steps are suggested for the design of a game:
   - Determine the learning objectives
   - Determine the expected (changes in) interaction
   - Select the task type
   - Determine whether pre-structuring is needed and how much
   - Determine group size
   - Determine how computer support can be applied (with, at, through)
3) The design guidelines for classroom collaborative games proposed by Villalta et al. (2011). Villalta et al.'s proposal considers the following features:
   - Interactivity and guidance
   - Mechanics linked to learning objectives
   - Clear narrative
   - Gradual increase in difficulty
   - Teacher mediation during the game
   - Organization of face-to-face interaction
   - Mechanics linked to collaboration
   - Adequate spatial distribution

4) The DPE framework (Winn, 2008) which considers the following layers:
- The learning layer
- The storytelling layer
- The gameplay layer
- The user experience layer

A more detailed explanation about the game, including theories, design guidelines, and how these theories and guidelines have affected the design of our game can be found in Martín-SanJosé et al. (2014a) and Martín-SanJosé et al. (2014b).

3.2. Description of the game

The aim of the game is to reinforce the learning of the concept of timeline, including its order, and the characteristics of each historical age. The game is divided into mini-games, several of which pertain to each time period on the timeline. There are also video and audio explanations at the beginning of the mini-games to introduce the historical ages and to give more detailed information.

In our study, the game had the same stages and order in both configurations. The children played the game from Prehistory to the present day. The children had to use their own hands to interact with the games, searching for shapes or pressing buttons by moving their hands to the active area. In the case of frontal projection, the buttons were

placed at the bottom of the screen (

Figure 2), half of the buttons appear on the left (for one child) and the other half of the buttons appear on the right (for the other child). In the autostereoscopic case, the buttons were placed on the sides of the display following the same idea; half of the buttons appear on the left side (for one child) and the other half of the buttons appear on the right side (for the other child) (Figure 3). For the visualization, the main difference between the two systems is that in the frontal projection, the visualization is in 2D (without illusion of depth), and in the autostereoscopic system, the visualization is in 3D. All the contents of the game are 3D models, except the videos. Even the buttons were modelled in 3D. Moreover, the 3D models have a spinning movement. In the case of the autostereoscopic system, this movement allows the 3D perception of the game elements to be increased. Figure 1 shows a 3D model of the medieval castle. The game consisted of seven mini-games that were distributed along the five historical ages mentioned above. The color code for the buttons used in the game was the following: yellow for unselected buttons, green for buttons selected correctly, and red for buttons selected incorrectly.

At the beginning, the children heard the voice of an avatar introducing them to the game. Once they were ready to start, they had to select the first historical age from the timeline, Prehistory, by pressing the correct buttons. After a video explanation of

Prehistory, they played two mini-games from this time period; the first consisted of finding some cave paintings. In this mini-game, an image of a cave is shown in which two paintings are hidden. The paintings can be identified because two faint silhouettes are shown. The children have to find the silhouette that is on their side by passing their hand over the painting area: in the autostereoscopic system, the children wave their hand over the painting area; in the frontal projection system, the children touch the painting area. This interaction is the same for the entire game. In the second mini-game, the children had to select a color that was used during the Prehistory period and leave an imprint of the shape of their hand in the cave. When all of this was done, they had to select the next historical age; this time, it was the turn for Ancient Times. In this mini-game, the children had to reconstruct a Roman city by placing different architectural elements such as the Roman circus in the city. A map without the architectural elements of an ancient city appears on the screen. The children have to reconstruct the city adding the architectural elements by selecting buttons that appear on their side. Afterwards, the game asked them some questions about the use of the buildings they had just used to construct the Roman city. In this mini-game, the buildings appear (rotating) in the center of the screen and the children have to choose the correct button from the buttons in their side. The next historical age was the Middle Ages. Here, the children had to build a medieval castle in the same way as the mini-game of the ancient times. After completing the Middle Ages, the children began the Early Modern Period, where they had to find three objects that Christopher Columbus used in his journeys to discover the American continent. . These objects are found when the children pass their hand over the object area. Once all of these objects were found, the children reached the final historical age (and last stage) of the game, the Contemporary Period. Finally, the children had to complete a puzzle that recreated the timeline. The timeline appears in the upper area of the game with holes where the historical time periods should be placed. Half of them appear on one side (for one child) and the other half appear on the other side (for the other child). The children have to take the historical time period and

drop it into the correct position.

Figure **4** shows how the buttons were located throughout the game; hand-shaped pointers for hand guidance are also shown. The avatar that guided the children during the whole game is represented by an alarm clock figure, which is shown in the upper-left corner. He guided the children telling them what they must do in each part of the mini-games.

3.3. Technical features

For the autostereoscopic system, the real world and users' gestures are captured by a Kinect device. The Kinect was placed in front of the 3D display, which was centered relative to the 3D display. Also, there were two numbered markers at a distance of about 2m. from the display to let the children know the surface area where they should stand. The autostereoscopic rendering was made possible by using an XYZ display. The specific model was XYZ3D8V46, which had a screen size of 46" and full HD resolution (1920×1080 pixels). The OpenSceneGraph toolkit 3.0.1 was used to render the 3D models and the virtual world. The autostereoscopic rendering was performed by using the Mirage SDK (www.mirage-tech.com). OpenNI and the Kinect drivers for Windows were used for registration and video capture.

For the frontal projection system works as a touch screen (

Figure 2). For the user interaction, a Kinect device and an InFocus IN1503 short throw projector were used for the projection. A table covered with a white cardboard was used for the projection area. A steel support was used to place the Kinect device and the projector vertically as shown in Figure 5. The table surface was used for capture and display at the same time. The programming language that was used to develop the game was C#. We also used the XNA Framework with the official Kinect drivers from Microsoft. Emgu.CV was used to manipulate complex graphics. GoblinXNA was used to display the 3D scene.

4. Description of the study

This section presents the characteristics of the children that played the game, the measurements that were used during the experiment, and the steps that were followed. As mentioned in the introduction section, the three hypotheses to corroborate in our study are the following:

1) The children will prefer the autostereoscopic system over the frontal projection system.
2) The children will increase their knowledge about the subject treated in the game by using the two systems, and that the autostereoscopic system would lead to greater learning results.
3) The frontal projection system will be easier to use.

4.1. Participants

A total of 128 children participated in our study. There were 67 boys (52.34%) and 61 girls (47.66%). They were between seven and eleven years old, and they had already finished their academic course between the second and fifth grades of primary school. The mean age was 8.96 ± 0.90 years old. All of the children belonged to the same

summer school and they all lived in a similar environment. The children had computers at home and they were used to playing computer and mobile games, mostly on weekends. Therefore, most of them had previous experience playing video games. Moreover, none of them had serious problems when using the two systems.

4.2. Measurements

To retrieve data for the analysis, three different questionnaires were used. There was a pre-test questionnaire with only thirteen questions of plain text related to knowledge. The knowledge questions were multiple‑choice with four, five, or six options. For the knowledge variable, these questions were counted as 0 for fail and 1 for success. This test was used to evaluate the children's knowledge before they started playing the games. There was a second post-test questionnaire. This questionnaire had the same thirteen questions from the pre-test, and thirteen additional questions related to different aspects including usability. By comparing the pre-test and this post-test, it was possible to determine if there had been an increase in knowledge. There was a last questionnaire that the children filled out once they had played with the two system configurations. This questionnaire was used to determine which of the two configurations they preferred. This questionnaire had nineteen questions; ten questions obtained information about the last configuration played and the last nine questions compared the two configurations.

4.3. Procedure

The participants were assigned to one of the following two groups:
- Group A: Participants that played with the autostereoscopic configuration first and afterwards played with the frontal projection configuration.
- Group B: Participants that played with the frontal projection configuration first and afterwards played with the autostereoscopic configuration.

The A and B groups were balanced by grouping the children into pairs (1 boy + 1 girl, 2 boys, 2 girls), with the same number of pairs for each combination. The participants filled out web-based questionnaires using a computer. The children did not complain about the number of questions or about having to use a computer to answer them. The following protocol was used:
1. A pair of children filled out the pre-test questionnaire (PreAuto for Group A and PreFrontal for Group B).
2. These children played one configuration (frontal projection or autostereoscopy). The children went through all historical ages with this configuration.
3. Then, they filled out the post-test questionnaire on-line (Pos1Auto for Group A, and Pos1Frontal for Group B).
4. Then, they played with the other configuration that they had not played with in step 2. For their second game, the children were asked which historical time period was their favorite. That favorite historical age was the only mini-game that they played with the second configuration. This implies that the time spent with the second configuration was shorter.
5. Finally, they filled out the final questionnaire (Pos2Auto for Group A, and Pos2Frontal for Group B).

5. Results

The data from the study were analyzed using the statistical open source toolkit *R*.

5.1. Learning outcomes

Several *t*-tests were performed to determine if there were significant differences in the knowledge acquired. In these tests, the knowledge variable was analyzed, which took into account all of the knowledge questions and represents the number of correct answers. The knowledge variable was compared in the Pre and Pos1 questionnaires.
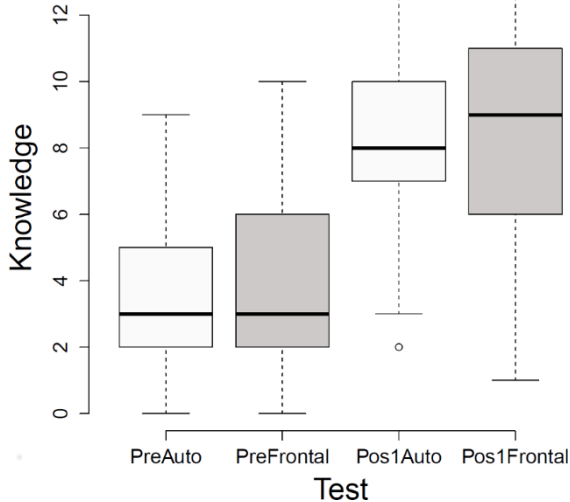

Figure 6 shows the box plot for the scores before and after playing with the first game (group A − autostereoscopic configuration, and group B − frontal projection). A high dominance of correct questions after playing the first game over the pre-test can be observed. All *t*-tests are shown in the format: *(statistic[degrees of freedom], p-value, Cohen's d)*; and ** indicates the statistical significance at level α=0.05. A paired *t*-test between PreAuto (mean 3.54±1.93) and Pos1Auto (mean 8.18±2.53) showed that there was a statistically significant difference (*t*[66]=-18.02, *p*<0.001**, Cohen's *d*=2.20). Another paired *t*-test revealed that there was a statistically significant difference between the ratings of the knowledge variable in PreFrontal (mean 3.79±2.62) and Pos1Frontal (mean 8.62±3.35) (*t*[60]=-14.85, *p*<0.001**, Cohen's *d*=1.90). To determine whether or not there was difference between the initial knowledge of the two groups, an unpaired *t*-test was performed between the knowledge variable in PreAuto (mean 3.54±1.93) and the knowledge variable in PreFrontal (mean 3.79±2.62) (*t*[126]=-0.61, *p*=0.542, Cohen's *d*=0.11). These results revealed that there was no statistically significant difference between the knowledge in the two pre-tests. To determine whether or not there was difference between the acquired knowledge in the two groups, an unpaired *t*-test was performed between the knowledge variable in Pos1Auto (mean 8.18±2.53) and the knowledge variable in Pos1Frontal (8.62±3.35) (*t*[160]=-0.84, *p*=0.401, Cohen's *d*=0.15), which also revealed that there was no statistically significant difference between the acquired knowledge using the two systems.

A multifactorial ANOVA test was also performed to take into consideration several factors simultaneously (age, game and, gender). The results showed that there were statistically significant differences for only the Age factor (*F*[4,110]=14.92, *p*<0.001**, η^2 =0.351) and not for the Game factor (*F*[1,110]=1.15, *p*=0.797, η^2=0.014) or the Gender factor (*F*[1,110]=3.58, *p*=0.061, η^2=0.031), or for the interactions among the factors. The generalized eta-squared effect sizes revealed that Age was the most influential factor. A Tukey post-hoc test showed that the acquired knowledge was significantly different between children of ages 7 and 8, 7 and 9, 7 and 10, 7 and 11, 8

and 10, 8 and 11, and 9 and 10. For the *knowledge* variable, interaction plots show the knowledge the children had after playing the first game, between gender and game
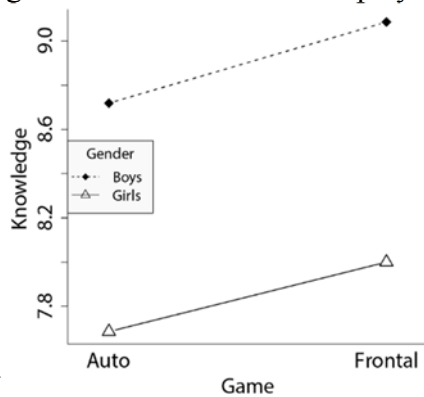
factors (

Figure 7) and gender and age factors (Figure 8). From these figures, it can be observed that, on average, boys had more knowledge than girls after playing the first game; and the score means at older ages were higher than at younger ages with significant differences among the age groups.

To complete this analysis, the dichotomous Rasch model was used (Rasch, 1960). This model measures a person's latent trait level from a probabilistic perspective. The probability of a user answering a question correctly relies on the user's underlying ability and the difficulty of question. A graphical model check of this analysis was performed, where the questions were grouped by raw scores and the ones which are higher than the mean are separated from the ones which are lower. The red lines represent the confidence bands. The results of the questions for both groups are shown in
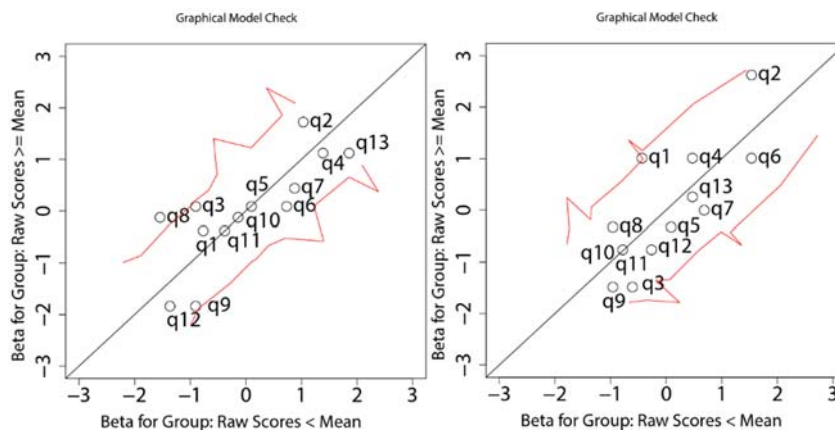

Figure 9. Every question was inside the confidence bands, except Q8 for the autostereoscopic group. This indicates that Q8 is an easy question. Based on these results, it can be concluded that the questions are appropriate for the assessment of the acquired knowledge for both configurations. In order to check the goodness of fit of the Rasch model, the test proposed by Andersen (1973) was used. In our study, this test offered the values, LR-value=20.231, df=12, p=0.063, which fit the Chi-squared distribution. Therefore, in our study the Rasch model is true.

5.2. System comparison outcomes

Several non-parametric tests were performed for our Likert questions (the Mann-Whitney U test for unpaired questions and the Wilcoxon Signed-rank sum test for paired questions) to determine if there were statistically significant differences in the opinions of the children depending on which game configuration was played first. First,

the data of the children that played the autostereoscopic system first versus the children that played the frontal projection system first were analyzed (Table 1). Then the scores of each child playing with one system first and later with the other (Pos1Auto versus Pos2Frontal (Table 2), and Pos1Frontal versus Pos2Auto (Table 3) were also compared. These tables only show the questions where statistically significant differences were found. From the analysis of Q14 (*How much fun did you have? [1-5]*), no statistically significant differences were found. Nevertheless, when the same child played first with one of the two games, he/she scored the first time statistically significantly higher than the second time. The analysis of Q16 (*How difficult was the game? [1-5]*) showed that the children that played with the autostereoscopic system gave a statistically significant higher score to the ease of use than the children who played with frontal projection. However, when playing the second time, the results showed that there was a statistically significant difference in favor of the system played last. Our explanation for these results is that the second time they played, they found the game easier because they had already played before and they already knew what they had to do in the game even though the interaction was not exactly the same. Something similar happened with Q18 (*Selecting the elements/options of the game was: [1-5]*). The first time the children played, no statistically significant differences were found; however, the second time they played, they gave a statistically significant higher score to the second system used. Statistically significant differences were found in the autostereoscopic vision-oriented questions Q23 (*Evaluate the sensation of viewing the castle. Did it look like it were coming out of the screen? [1-7]*) and Q24 (*Did you think you were able to touch the castle? [1-7]*). The analysis of these questions revealed that when playing with the frontal projection system first and with the autostereoscopic system second, the significance was relevant; the significance was not as high as when playing with autostereoscopic first and with the frontal projection second. This means that the children noticed a great change in the visualization system when changing from non-3D to 3D, and they didn't notice this when changing from 3D to non-3D. We can confirm that with autostereoscopy the children had the feeling of being able to touch the 3D elements like the medieval castle. Finally, another test was made for Q25 (*Score the game from 1 to 10*). The results of this question showed that there were only statistically significant differences in favor of the autostereoscopic system when it was played first.

5.3. Satisfaction outcomes

In order to measure the satisfaction that the children had while playing the game, Fisher's exact test was performed for each satisfaction question. These satisfaction questions were answered after playing the second time. The tests revealed that there was only a statistically significant difference for Q27 ($p=0.035**$), where children who played the autostereoscopic game first chose "*autostereoscopic*" and children who played the frontal projection game first chose "both". After analyzing the results, we could see that the children tended to choose the system they had used the first time. Our explanation for this result is that the second time they play, they play a short version of the game, and that the first time they play, everything is new; this impresses them to a greater extent. Table 4 shows the percentages of children's preferences for different questions.

In order to determine which of all the mini-games was liked the most, Q26 was asked after playing with the game for the first time. In that question, the children could select the mini-games they preferred, and they could select more than one option. The mini-game with the highest score was Prehistory (find cave painting and place an imprint of the shape of your hand in the cave) with 66.39% of the votes. The second highest was

Ancient Times with 55.73%. The next preferred mini-game was the Middle Ages (build a medieval castle) with 53.27%. Following, 47.13% of children selected the Contemporary Period (solve the timeline puzzle) as one of their favorites. Finally, the mini-game with the least votes was the Early Modern Period (find objects used by Christopher Columbus) with 39.75% of votes. A correlation analysis was performed to determine whether or not there was dependency between mini-game preferences and age or gender; the results showed that there was no dependency between those factors.

5.4. Subjective considerations of the person in charge of the study

The following conclusions were deduced from the observation sheets filled out by the person in charge of the study. In general, the children used both systems easily and the difference in the time they spent learning how to interact with the systems was not significant. Once they had learned how to play with the first system, the second system was practically already learned and they already knew what they had to do. However, the children could easily activate buttons with the frontal projection system. In contrast, the children spent more time activating a button in the autostereoscopic system, and they sometimes made some weird movements in order to activate them (e.g., moving their hand back towards their body instead of moving forwards). Our argument is that the interaction in the frontal projection system is similar to pressing a button on a mobile device and children are used to using this type of interaction.

In the autostereoscopic system, the children are placed in front of the display and they see themselves inside. This is similar to being in front of a mirror with special features. For children, being together in pairs inside the game is a fun experience. Moreover, the children could move the elements that appeared in front of them with their hands; before placing the elements in the right place, they could play with them, move them through the space, put them above their heads or over the heads of their teammate.

With regard to the 3D perception in the autostereoscopic system, the 3D was so perceptible that most of the children tried to touch the elements in the air because they thought that those elements were outside of the display.

6. Conclusions

In this paper, advanced displays and Natural User Interfaces were used to develop two learning environments for children (NUI+3D visualization vs. NUI+2D visualization). The two different configurations were developed with the background of an educational game based on historical ages. We compared the two configurations. The autostereoscopic configuration allows the users to have a complete experience (3D visualization and natural interaction) without having to carry devices or wires on their bodies. In the autostereoscopic configuration, the children could see themselves in the autostereoscopic display, and the game was controlled by gestures. In contrast, the frontal projection configuration simulated a touch table (2D) in which the children did not carry devices or wires on their bodies. To our knowledge, this is the first time these system combinations have been compared, especially for education. With regard to our first hypothesis (the children will prefer the autostereoscopic system over the frontal projection system), the children liked both configurations equaly (45%), followed by the autostereoscopic system (40%), and then the Frontal Projection (14%). From the percentages, we can affirm that this hypothesis has been corroborated (both + autostereoscopy > both + frontal projection).

The second of our hypotheses was that children would increase their knowledge about the subject of the game by using the two systems, and that the autostereoscopic system would lead to better learning results. Comparing their initial knowledge and their knowledge after playing, statistically significant differences were obtained, which corroborates the first part of the second hypothesis. Differences in age, gender, and which system was played first were also considered. These results indicate that systems of this type can facilitate learning outcomes to a greater extent, especially for older children (in our case 11-year-olds). However, there was no statistically significant difference between the acquired knowledge using the two systems. Therefore, the second part of our second hypothesis (the autostereoscopic system will obtain better learning results) was not corroborated. Although unexpected, it is an excellent result because it means that the game is well suited for learning outcomes and that the two systems can be used for this purpose. For depth perception, the results showed that the illusion of depth (Q23) was mainly perceived and appreciated, being more evident when the children played with the autostereoscopic system after playing with the frontal projection system. The results revealed that autostereoscopy gave the children the feeling of being able to touch the 3D elements (Q24). From our point of view, these results are important and can be exploited for the development of educational games. For ease of use, in the two related questions (Q16 & Q18), the medians were equal or more than 4 on a scale from 1 to 5. This indicates that the two systems are easy to use. These results are in line with previous works (Buisine et al. (2007), Bhalla & Bhalla (2010), Tsai & Yen (2013)). A statistically significant difference was found in which the autostereoscopic system was scored higher (Q27). However, when the children were asked explicitly about the easiest system to use (Q28), they preferred the frontal projection system (41%), followed by both systems (31%), and the autostereoscopic system (27%). Moreover, the person in charge of the activity perceived that the children could more easily activate buttons with the frontal projection system than with the autostereoscopic system. Therefore, from the results and our observations during the activity, we consider that the frontal projection system is easier to use, which corroborates our third hypothesis (the frontal projection system will be easier to use). Our conclusion is in line with the work of Tuveri et al. (2013) in which they also perceived a difference in favour of the touch table. Our opinion is that, in the frontal projection system, the children interacted easily and fast simply by placing their hands over the buttons. However, the autostereoscopic system requires more effort and attention. Users are standing and they have to perform the right movement in the air. This observation is in line with one of the disadvantages pointed out by Libardi (2014). Nevertheless, more studies should be carried out to assure that the frontal projection system is the easiest to use. With regard to the topic of the game, some of the knowledge questions revealed that data like dates or the names of historic events are the most difficult for children to remember.

Based on our study, we believe that using natural gesture interaction and having stereoscopic vision without wearing devices or wires provides an enhanced and richer user experience that is metaphorically similar to the real-world experience. In this situation, the selection of elements is done by using your hands and interacting by yourself.

To date, we have compared two systems, but for future work other comparisons are also possible; for example, comparing the two systems with or without stereoscopy, comparing the PC version vs. tablet/smartphone versions, or using a control group in which the children learn about the same period of history using traditional learning. In the autostereoscopic system, the children can see themselves inside the game in the

display. We believe this has contributed to having a richer experience. However, a formal study should be carried out to corroborate this hypothesis. Other future work could include involving children in the design of the game. In this study, the children were not involved in the design. However, it would have been positive to involve the children in the design phase allowing them to contribute to this process. Two possibilities for their involvement are the following: informant design (Scaife & Rogers, 1999), in which children contribute to the design, but are not considered as design partners; or participatory design and cooperative enquiry (Guha et al., 2005), in which children have a more relevant role. For the evaluation, we mainly used questionnaires, but other evaluation methods could also be used, such as drawing intervention (Mazzone et al., 2007), in which the children have to draw anything related to the task accomplished. Advanced displays and NUI in the field of education are in their earliest stages, but they could be a very great addition to the learning process for different topics and for different communities. We also believe that the educational field can be improved with the use of stereoscopy.

References

Andersen, E. B. (1973). A goodness of fit test for the rasch model. *Psychometrika*, *38*(1), 123–140. doi:10.1007/BF02291180

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.

Bhalla, M. R., & Bhalla, A. V. (2010). Comparative Study of Various Touchscreen Technologies. *International Journal of Computer Applications IJCA*, *6*(8), 12–18. doi:10.5120/1097-1433

Boutsika, E. (2014). Kinect in Education: A Proposal for Children with Autism. *Procedia Computer Science*, *27*, 123–129. doi:10.1016/j.procs.2014.02.015

Buisine, S., Besacier, G., & Najm, M. (2007). Computer-supported creativity: Evaluation of a tabletop mind-map application. *Lecture Notes in Computer Science*, *4562*, 22–31. doi:10.1007/978-3-540-73331-7_3

De Greef, K., van der Spek, E. D., & Bekker, T. (2013). Designing Kinect games to train motor skills for mixed ability players. *Games for Health*, 197–205. doi:10.1007/978-3-658-02897-8_15

Dewey, J. (1963). *Experience and Education*. New York: Collier.

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, *8*(5), 347–358. doi:10.1007/s00779-004-0297-4

Francese, R., Passero, I., & Tortora, G. (2012). Wiimote and Kinect: gestural user interfaces add a natural third dimension to HCI. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 116–123).

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. doi:10.1016/j.compedu.2012.12.001

Gope, D. C. (2011). Hand Gesture Interaction with Human-Compute. *Global Journal of Computer Science and Technology*, *11*(23), 3–12.

Guha, M. L., Druin, A., Chipman, G., Fails, J. A., Simms, S., & Farber, A. (2005). Working with young children as technology design partners. *Communications of the ACM*, *48*(1), 39–42. doi:10.1145/1039539.1039567

Hsu, H.-M. J. (2011). The Potential of Kinect in Education. *International Journal of Information and Education Technology*, *1*(5), 365–370. doi:10.7763/IJIET.2011.V1.59

Juhnke, B. J. (2013). *Evaluating the Microsoft Kinect compared to the mouse as an effective interaction device for medical imaging manipulations*. Master Thesis. Iowa State University.

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. *Internet and Higher Education*, *14*, 158–163. doi:10.1016/j.iheduc.2011.04.001

Libardi, R. M. O., Rodrigues, J. F., & Traina, A. J. M. (2014). Design and evaluation case study: Evaluating the Kinect device in the task of natural interaction in a visualization system. *International Journal of Human Computer Interaction (IJHCI)*, *5*(1), 1–20. Retrieved from http://www.cscjournals.org/library/manuscriptinfo.php?mc=IJHCI-89

Luna-Oliva, L., Ortiz-Gutiérrez, R. M., Cuerda, R. C. la, Piédrola, R. M., Alguacil-Diego, I. M., Sánchez-Camarero, C., & Martínez Culebras, M. del C. (2013). Kinect Xbox 360 as a therapeutic modality for children with cerebral palsy in a

school environment: A preliminary study. *NeuroRehabilitation*, *33*(4), 513–521. doi:10.3233/NRE-131001

Martín-SanJosé, J.-F., Juan, M.-C., Gil-Gómez, J.-A., & Rando, N. (2014a). 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. (2014b). Playful interaction for learning collaboratively and individually. *Journal of Ambient Intelligence and Smart Environments*, *6*(3), 295–311. doi:10.3233/AIS-140257

Mazzone, E., Xu, D., & Read, J. C. (2007). Design in Evaluation: Reflections on designing for children's technology. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers, Volume 2* (pp. 153–156).

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Roman, D. (2010). Interact naturally. *Communications of the ACM*, *53*(6), 12. doi:10.1145/1743546.1743552

Scaife, M., & Rogers, Y. (1999). Kids as informants: Telling us what we didn't know or confirming what we knew already. In A. Druin (Ed.), *The design of children's technology* (pp. 27–50). San Francisco, CA: Morgan Kaufmann.

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 Elrbaum Associates.

Soro, A., Iacolina, S. A., Scateni, R., & Uras, S. (2011). Evaluation of User Gestures in Multi-touch Interaction: a Case Study in Pair-programming. In *Proceedings of the 13th international conference on multimodal interfaces* (pp. 161–168).

Strijbos, J. W., Martens, R. L., & Jochems, W. M. G. (2004, May). Designing for interaction: Six steps to designing computer-supported group-based learning. *Computers & Education*. doi:10.1016/j.compedu.2003.10.004

Tsai, C.-H., & Yen, J.-C. (2013). The Development and Evaluation of a Kinect Sensor Assisted Learning System on the Spatial Visualization Skills. *Procedia - Social and Behavioral Sciences*, *103*, 991–998. doi:10.1016/j.sbspro.2013.10.423

Tuveri, E., Iacolina, S. A., Sorrentino, F., Spano, L. D., & Riccardo Scateni, R. (2013). Controlling a planetarium software with a Kinect or in a multi-touch table: a comparison. In *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI (CHItaly'13)* (p. Article N. 6).

Villalta, M., Gajardo, I., Nussbaum, M., Andreu, J. J., Echeverría, A., & Plass, J. L. (2011). Design guidelines for Classroom Multiplayer Presential Games (CMPG). *Computers and Education*, *57*(3), 2039–2053. doi:10.1016/j.compedu.2011.05.003

Wang, X., & Cheok, A. D. (2011). ClayStation: a mixed reality gaming platform supporting playful learning for children. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology - ACE '11* (p. 69). Lisbon, Portugal: ACM Press. doi:10.1145/2071423.2071509

Winn, B. (2008). The design, play, and experience framework. In R. E. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (pp. 1010–1024). IGI Global.

Table 1: Medians for questions of the Pos1Auto and Pos1Frontal questionnaires, Mann-Whitney *U* test analysis, and *r* effect size

#	Pos1Auto	Pos1Frontal	*U*	*Z*	*p*	*r*	Range
Q16	4	4	2347.5	2.045	0.041**	0.182	[1-5]
Q23	6	5	2495.5	2.602	0.009**	0.232	[1-7]

Table 2: Medians for questions of the Pos1Auto and Pos2Frontal questionnaires, Wilcoxon Signed-rank sum test analysis, and *r* effect size

#	Pos1Auto	Pos2Frontal	*W*	*Z*	*p*	*r*	Range
Q14	5	5	160	2.255	0.034**	0.199	[1-5]
Q16	4	5	160	-3.704	<0.001**	0.320	[1-5]
Q18	4	5	91	-2.863	0.004**	0.249	[1-5]
Q25	10	10	211	2.325	0.018**	0.201	[1-10]

Table 3: Medians for questions of the Pos1Frontal and Pos2Auto questionnaires, Wilcoxon Signed-rank sum test analysis, and *r* effect size

#	Pos1Frontal	Pos2Auto	*W*	*Z*	*p*	*r*	Range
Q14	5	5	19.5	-2.165	0.040**	0.196	[1-5]
Q16	4	5	453.5	4.698	<0.001**	<0.001**	[1-5]
Q17	5	5	141	2.424	0.015**	0.015**	[1-5]
Q18	4	4	231	2.830	0.004**	0.004**	[1-5]
Q23	5	6	503.5	3.264	<0.001**	<0.001**	[1-7]
Q24	4	6	646	3.292	<0.001**	<0.001**	[1-7]

Table 4: Children's preferences in percentages. The highest score for each question is highlighted in bold type

#	Question text	Autost.	Frontal P.	Both	None
Q27	Which system did you like the most?	40	14	**45**	1
Q28	Which system was the easiest to use?	27	**41**	31	1
Q29	Which system was the most comfortable?	28	35	**36**	1
Q30	Which did you control better?	28	**44**	26	2
Q31	In which system were the images viewed better?	**42**	26	31	1
Q32	Would you recommend any of these systems to friends?	21	11	**67**	1
Q33	Which system was the most fun?	31	15	**54**	0
Q34	Would you like to use any of these systems at school?	41	13	**45**	1

Figure 1: 3D model of the medieval castle

Figure 2: Pressing a button in frontal projection

Figure 3: Children playing with the autostereoscopic system

Figure 4: Button disposition in the autostereoscopic configuration for the Ancient Times period

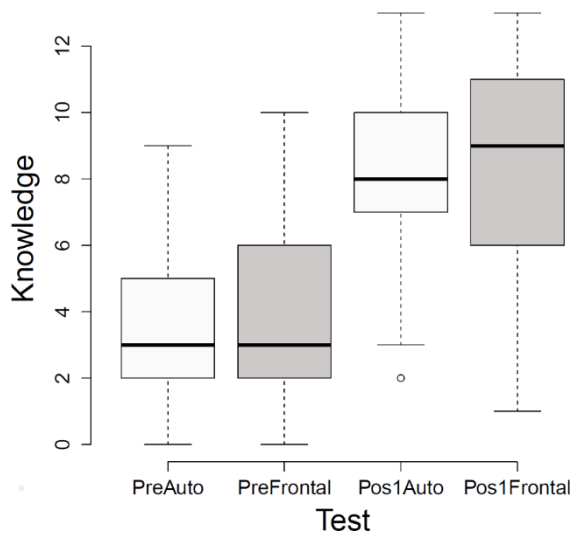Figure 5: Two children playing with the frontal projection configuration

Figure 6: Scores of the knowledge variable in the Pre and Pos1questionnaires for the Autostereoscopic system and for Frontal Projection. This box plot shows how the children's correct answers are grouped into quartiles, and the median is indicated with a thick solid line
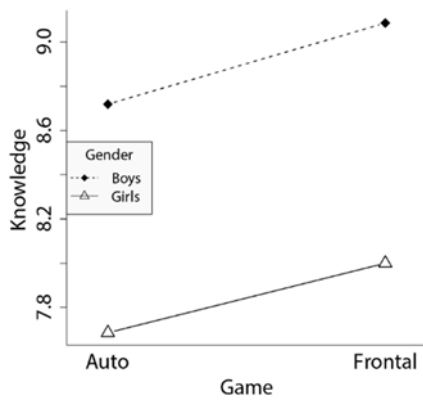

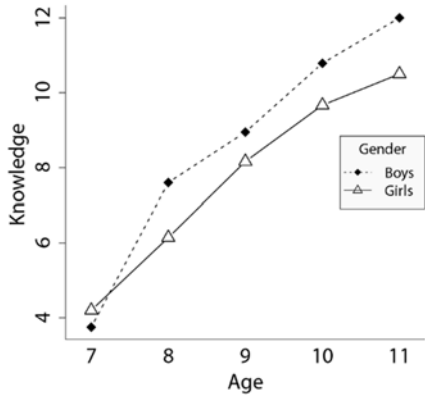Figure 7: The mean knowledge after playing the first game for gender and game factors

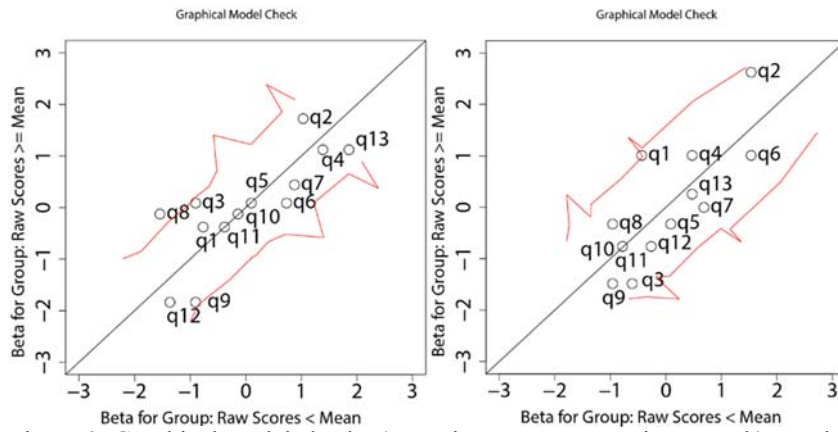Figure 8: The mean knowledge after playing the first game for gender and age factors

Figure 9: Graphical model check. a) For the autostereoscopic group. b) For the frontal projection group