The final publication is available at

http://dx.doi.org/10.1016/j.compedu.2012.12.015

The effects of the size and weight of a mobile device on an educational game

David Furió[1], Santiago González-Gancedo[1], M.-Carmen Juan[1,*], Ignacio Seguí[2], María Costa[2]

[1] 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, mcarmen@dsic.upv.es
[2] AIJU, Ibi, Alicante, Spain

Abstract

In this paper, we present an educational game for an iPhone and a Tablet PC. The main objective of the game was to reinforce children's knowledge about the water cycle. The game included different interaction forms like the touch screen and the accelerometer and combined AR mini-games with non-AR mini-games for better gameplay immersion. The main differences between the two devices were screen size and weight. A comparative study to check how these differences affect different aspects was carried out. Seventy-nine children from 8 to 10 years old participated in the study. From the results, we 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. There was only a statistically significant difference for the global score in which the iPhone was scored higher. We would like to highlight that the scores for the two devices and for all the questions were very high with means of over 4 (on a scale from 1 to 5). These positive results suggest that games of this kind could be appropriate educational games and that the mobile device used may not be a decisive factor.

1. Introduction

Teachers have always been seeking new ways to enhance students' understanding (Veenema & Gardner, 1996) and trying to find teaching methods using tools that can reach their students at multiple levels (Tan, Lewis, Avis & Withers, 2008). Since children may have a better learning experience when more senses are involved (Sandor & Klinker, 2005), play becomes an important activity to improve and develop children physically, mentally, emotionally, and socially. It is a natural way to learn because it uses all their senses to solve problems and to understand their environment (Rapeepisarn, Fung & Depickere, 2006). Play can help children experience a greater form of learning than when they are in classroom (Gee, 2003) and stimulates the children to understand new concepts that would otherwise be difficult to reproduce (Blecic, Cecchini, Rizzi & Tronfio, 2002; Squire, Barnett, Grant & Higginbotham, 2004). Play is an enjoyable activity that serves as a medium to develop capabilities, abilities, etc., through active participation (Blecic et al., 2002). Play is amusing and fun, and enjoyment is important when endeavoring to achieve learning goals, because what is enjoyably learned is less likely to be forgotten (Blecic et al., 2002). Thus, creative learning becomes a fun experience.

Educational games combine playing and learning though participative techniques that aim to develop children's knowledge and abilities. Educational games also help to motivate children to study. They can be used to reinforce and check knowledge acquired in the classroom. Psychologists and philosophers have studied the influence that playing games has on the learning process of children and they have concluded that entertainment was an important factor that helped to improve learning (Albert & Mori, 2001; Taran, 2005). Edutainment is a term that also merges education and entertainment that relies heavily on technology like video games. (Pan, 2006; Rapeepisarn et al., 2006). As some studies have demonstrated, digital educational games can hold significant benefits for children. They can improve knowledge and skills and also stimulate motivation and interest (Fisch, 2005; Rigas & Ayad, 2010; Shelton & Hedley, 2002). Games are linked to the four basic dimensions of children's development: psychomotor, intellectual, social, and affective-emotional

(Garaigordobil, 2005). Games can serve as tools for developing thinking skills. They can also work as learning sources. They can stimulate children's attention and memory and also support the development of language (Garaigordobil, 2005).

M-learning is a new educational model (Earnshaw, 2011). Even though there is no unified definition of m-learning and it depends on the researcher's point of view, m-learning is a term that usually refers to the use of mobile devices in education (Sharples, Corlett, &Westmancott, 2002). In other words, m-learning is learning services that can be used at any time and place through a simple mobile device (Brugnoli, Lorusso, Corsini, Bormida, Keefe, Ahonen, et al, 2004; Earnshaw, 2011). M-learning has some advantages over other learning methods. In m-learning, the devices used are small, portable and wireless. They make the educational process flexible and adaptable for students. They are usually cheaper than other devices like desktop computers and can be used anywhere at anytime (Georgiev, Georgieva, & Smrikarov, 2004; Jones & Jo, 2004; Earnshaw, 2011). Mobile technology is also becoming a potential platform for games (Facer, Joiner, Stanton, Reid, Hull & Kirk, 2004). Despite these advantages, m-learning also has some disadvantages (like the different screen sizes and the limited functionalities of the devices), which makes the design of applications more difficult than with other approaches. Furthermore, teacher may not be able to control students in the same way as they do in a classroom environment (Mahamad, Ibrahim & Taib, 2010). According to Earnshaw (2011), m-learning environments fall into two categories:

- Fieldwork: in this environment, the learner experiences physical world situations, using mobile device resources.
- Classroom: in this environment, virtual worlds are used to educate and engage learners via multimedia content.

Earnshaw (2011) suggested that a thorough m-learning system should have both of these environments and combine them. We also share this view. Several m-learning studies have been carried out with results that would indicate that the students who used this educational model gained more knowledge compared to those who used traditional methods (Thornton & Houser, 2005; Mcconatha, Praul, Chester, & Lynch, 2008). If all of these benefits are considered, mobile devices can become tools with a great potential for engaging students at multiple levels in their learning activities (Tan et al., 2008).

A commonly accepted definition of Augmented Reality (AR) defines it as a system that has three main features: 1) It combines physical and virtual objects; 2) It has real time interaction; and 3) It has 3-D registration (Azuma, 1997). In other words, AR allows the user to see the physical world with virtual objects superimposed upon the physical world, which supplements reality. In an ideal AR application, the physical and virtual objects could not be distinguished, seeming to coexist in the same space, thus achieving a total fusion of the two environments (Azuma, 1997). According to the virtuality continuum defined by Milgram and Kishino (1994) (Figure 1), Mixed Reality (MR) encompasses AR and Augmented Virtuality (AV). The virtuality continuum ranges from the completely physical to the completely virtual environment through AR and AV. Taking this continuum into account, MR can be used to refer any combination of physical and virtual elements. AR is appropriate where this combination is closer to the physical environment, that is, where there is a greater contribution from the physical environment. On the other hand, AV is appropriate where the combination is closer to the virtual environment, with a greater contribution from virtual elements.
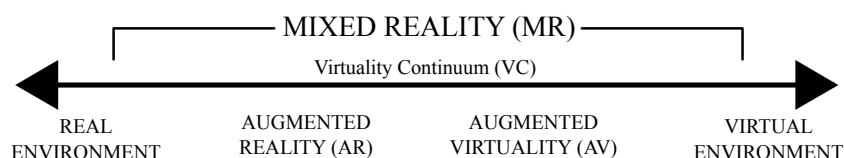

Figure 1. Representation of the virtuality continuum.

AR has matured to the point where it can be applied to a much wider range of application domains, and education is an area where this technology could be especially valuable (Billinghurst,

2002). AR is a technology that has the potential to engage and motivate learners to explore material from a variety of different perspectives (Kerawalla, Luckin, Seljeflot & Woolard, 2006). It has been demonstrated that AR systems usually impress those who have experienced it (Ardito, Costabile, Lanzilotti & Pederson, 2007; Billinghurst, Kato & Poupyrev, 2001) and that they are useful for teaching subject matters that students could not possibly experience first-hand in the physical world (Kerawalla et al., 2006). A number of experiments to gather empirical evidence as a basis for theoretical propositions and validation were conducted with results that showed that AR was a powerful and engaging visual and cognitive experience for users (Shelton & Hedley, 2004). Some advantages that AR can offer are the ability to transit smoothly between reality and virtuality and the use of a tangible interface metaphor that utilizes physical objects to manipulate virtual information (Billinghurst, 2002). Another advantage that AR offers is that participants can interact between real and virtual environments. That is something that cannot be done in virtual environments. As described in (Shelton & Hedley 2002), with AR, users do not have to use their imagination to envision what is happening: "With AR, there is no need to pretend an apple is the earth. There is the earth right there, positioned as an object before the users' eyes".

In this paper, we present a mobile AR game for learning about the water cycle, water composition, and water pollution. The game was developed for iPhone and Tablet PC devices. The game combines AR and non-AR mini-games. Using AR, the children explored a room looking for objects by focusing the device's camera on different markers. The non-AR mini-games did not use AR, but they used tactile or accelerometer capabilities. The non-AR mini-games usually consisted of children having to collect a certain number of objects to get to the next screen. We also present a study that compared which device was better suited for fun and learning and to determine if children acquired new knowledge after playing the game. Our game combined AR mini-games with non-AR mini-games for better immersion. The game also combined different forms of interaction, including the touch-screen and the accelerometer (Liarokapis & Newman, 2007).

The main objectives of this study were to determine if the game helped children in the learning process and to observe if one device had more influence than the other on the participants regarding the acquired knowledge, satisfaction, and interaction. The primary hypothesis was that there would be significant differences between playing in an iPhone or a Tablet PC device regarding acquired knowledge, satisfaction, and interaction, and that children would prefer the Tablet PC mainly due to its larger screen size. The secondary hypothesis was that there would be significant differences between the initial knowledge before playing the game and the knowledge acquired by playing the game using both devices.

The paper is organized as follows. Section 2 focuses on the state of the art in AR applications related to learning. Section 3 details the game design principles, describes the game, and briefly explains the software and hardware required to develop the application for the two devices. Section 4 presents the study. Section 5 presents the results, and Section 6 presents the discussion. Finally, in Section 7, a number of conclusions and areas for future research are identified.

2. State of the art
In the last few years, thanks to continuous improvements in technology, the production of mobile AR applications has increased. It has been demonstrated that mobile AR has multiple applications (Van Krevelen & Poelman, 2010) some of which are related to learning and edutainment. Several mobile AR and m-learning studies have been presented in this particular area, but not many have performed studies that analyzed their learning outcomes. There is also a lack of learning theories and research in developing educational games (Shaffer, Squire, Halverson & Gee, 2004).

2.1. M-learning and mobile AR learning studies focused on children
In 2003, Ketamo developed two slightly different geometry-learning games for a laptop PC and an adaptive geometry-learning game for a Compaq iPaq PDA (Ketamo, 2003a). Children were asked to find and mark the required polygon. A study with two stages was conducted. The first

stage included three groups of six-year-old children (two experimental groups of 21 and 20 participants, and a control group of 30 children). This phase compared one game with reflexive thinking, another game without reflexive thinking, and a traditional lesson. The second stage had only one experimental group made up of 17 participants, which was compared to the groups of the first stage. The results of the first stage showed that both games helped low-skilled students to reach the level of the average pupils. Ketamo concluded that this result could support the premise that computer-based teaching could help low-skilled students. The results of the second stage showed similar outcomes to those in the first stage, where low-skilled students benefitted more from the learning game. Specifically, the PDA version of the game achieved a double learning effect in the low-skilled group, while the average-skilled group did not have a significant improvement in the learning effect, similarly to the high-skilled group in the first stage. Despite the results, the second stage had a relatively small sample (almost half), which could lower the significance of the results. Moreover, no statistical measure was used to indicate the effect size. In 2010, Huang, Lin, and Chen developed a Mobile Plant Learning System (MPLS) for a PDA with a camera to facilitate botany learning among elementary school students (Huang, Lin & Cheng, 2010). A study was conducted with 32 participants with an average age of 11 years old. In this study, the children were organized into two groups in order to investigate the effectiveness of the new tool. One group used the MPLS and the other one used a guidebook. The study consisted in observing the characteristics of leaves and then attempting to find a plant. The results showed that more students in the MPLS group reported having positive perceptions about the outdoor learning activities. The analysis also revealed that there were significant differences in the students' attitudes in favor of the MPLS group. The authors concluded from the results that mobile technologies and an outdoor learning strategy are both useful tools in teaching children about plants. There were some problems observed in the trials; for example, the group who used PDAs experienced too much enthusiasm. This could lead to students not paying enough attention to the activities, thus giving more problems to teachers. This activity could have benefitted from a collaborative version in order to let the users interact with others, which could improve the learning activity (Johnson & Johnson, 1981; Jones, 2006). In 2011, Juan et al. (Juan, Furió, Alem, Ashworth & Cano, 2011) presented an AR mobile phone game for learning how to recycle. It was developed on a Nokia N95 8 GB. They compared the AR mobile game with a video game in a study with 38 children. The questionnaires used in the trials analyzed the participants' perceived knowledge of recycling, their behavior and attitudes towards recycling, and their perceived willingness to change the behavior. Other aspects investigated during the study were the ease of use, the engagement and fun, and the children's preference between the AR and non-AR games were. The results did not show significant statistical differences between the two games. Nevertheless, 69.4% of the participants preferred the AR game, which they perceived as being easy to use and more engaging and fun than the video-game. However, Juan et al. did not test whether or not the children learn with the game.

2.2. M-learning studies focused on high school students

There are also several other m-learning studies aimed at high school and university students to enhance learning. In 2009, Uzunboylu, Cavus, and Ercag presented a study to investigate the use of mobile phones to increase students' use of mobile technologies and to develop environmental awareness in students (Uzunboylu, Cavus & Ercag, 2009). Forty-one participants with an average age of 21.37 years old took part in the trials. The participants filled in a pretest and a posttest questionnaire. The results showed that the mobile devices engaged the participants and had a significant impact on students' attitudes toward maintaining environments. The study did not consider an experimental or control group to compare with the mobile group. In 2005, Schwabe and Göth (2005) presented a mobile game developed for a PDA to support the orientation days at a university. Twenty-two students from 19 to 25 years old participated in the study, where each participant was equipped with their own device. The evaluation showed that participants had fun with the game, but they could not demonstrate an enhancement of learning. However, this study did not use an experimental or control group to compare with the mobile game group. During the trials,

the children had to compete in two tasks at the same time. This could confuse the participants or could lead them to concentrate on only one of the tasks leaving the other task out.

2.3. Non-mobile AR learning studies focused on children

With regard to AR learning applications developed for PCs and laptop PCs, several works can be cited. In 2011, Connolly, Stansfield, and Hainey presented ARGuing, an Alternate Reality Game designed for a PC to increase the motivation of students in the learning of foreign languages (Connolly, Stansfield, & Hainey, 2011). Forty-five participants between 12 and 15 years old took part in the trials. Students played the game in the classroom or at home for 10 days. The study, which aimed at increasing the motivations of secondary school students in the learning of modern foreign languages, showed positive results regarding attitudes, motivation, and perceived learning with evidence suggesting that the system managed to deliver the motivational experience expected by the students. The participants complained about the amount of time involved in completing the tasks and the difficulty of some of the tasks. This might lead to a decrease of motivation. In 2007, Balog, Pribeanu, and Iordache presented the ARiSE project, which aimed at creating an AR technology in schools by adapting virtual showcases used in museums (Balog, Pribeanu & Iordache, 2007). Twenty students from 13 to 17 years old filled in a usability questionnaire. The study tested three types of interaction (pointing at a real object, selection of a virtual object, and selection of a menu item) in an AR platform formed by 4 independent modules organized around a table on which real objects were placed. While the results showed that the participants enjoyed interacting with the AR objects since the exercises were similar to computer games, the trial had a small population, which lowered the significance of the results. Also, the study did not compare the different interaction methods used in the trials. In 2008, Freitas and Campos presented *SMART*, which consisted of a TV-show style learning game. The system is composed of several racquets with 3D augmented reality markers, a web camera, a PC, and displays such as LCD or projectors (Freitas & Campos, 2008). Fifty-four students between 7 and 8 years old participated in the trials. The study compared a class using traditional methods and students who used the SMART system. The questionnaires focused on the knowledge questions and did not ask the participants about the usability or the engagement and fun, etc. of the system used. Analyses showed that SMART had better results than the traditional method in weak and average students, but the effect of the experimental system on good students was less noticeable than in the traditional class group. These results should be further examined, since Freitas and Campos did not examined whether or not there were statistical differences between the different groups and gender. In 2010, Hsiao presented a new approach to the implementation of AR in the educational environment by creating a Chemistry Augmented Reality Learning System (CARLS) (Hsiao, 2010). The participants consisted of 673 seventh-grade and eighth-grade students, aged between 13 and 14 years old. This system combined learning with three types of physical activity. In the study, the first three groups used the CARLS learning system, while a control group used a keyboard and a mouse to operate the computer. Each experimental group practiced a different type of AR physical activity. The study revealed that the students using all three types of physical activity together with CARLS had a significantly higher academic performance compared to the traditional ways. Despite these results, the author remarks that future researchers should emphasize more valuable characteristics of AR and that this study did not directly prove that any specific physical activity in CARLS improved any specific student's abilities. In 2006, Chen conducted a study to compare the use of AR and physical models in chemistry education (Chen, 2006). The application was developed for a laptop PC with a webcam. Four students participated in the trials. The study evaluated their perceptions regarding these two representations in learning about amino acids through interviews. From these interviews, it was inferred that students liked to manipulate AR by rotating the markers to see different orientations of the virtual objects. Their interactions with AR demonstrated that they tended to treat AR objects as real objects. However, due to the small sample, these findings should be corroborated with a larger sample. Apart from the interviews done, some pretests and posttests about the system developed might have helped the authors to obtain more feedback. In 2010, Chang, Lee, Wang, and Chen

presented RoboStage, a mixed-reality learning environment with robots to help students learn new words (Chang, Lee, Wang, & Chen, 2010). Thirty-six eighth-grade students participated in the study. Four groups were formed. Two of them completed the learning activities using an English textbook and the other two used RoboStage to complete them. The comparison between the two methods showed that RoboStage significantly improved the sense of authenticity of the task and also positively affected learning motivation and performance. The participants felt like they were putting language into real use when using virtual robots. Despite this, no significant differences in terms of learning new words between using the virtual and the mixed-reality environments were found.

2.4. Device or visualization comparison studies

A critical aspect in the evaluation of new devices is to determine which device is better suited for a specific task. In our study, we have checked learning outcomes and other aspects such as ease of use and satisfaction using the device to try to determine which device is better suited for learning. In a review of previous work, we have found studies that compare devices or different visualization methods to determine the most suitable device/visualization for facilitating the learning process (Ketamo, 2003a; Ketamo, 2003b; Klatzky, Wu, Shelton & Stetten, 2008; Schönborn, Bivall & Tibell, 2011; Hamidi, 2011; Fassbender, Richards, Bilgin, Thompson & Heiden, 2012). Following this trend, we compared two different devices. Ketamo (2003b) developed xTask, which is an adaptive learning environment whose purpose was to provide services for (web-based and mobile) teaching, studying, and learning processes. Ketamo compared the use of the xTask environment in PCs and mobile phones to see how it affected the quality and quantity of work. The results showed that mobile technologies might bring some added value to network-based learning, but they could not replace conventional computers. As described in section 2.1, Ketamo (2003a) also presented two geometry-learning games for a laptop PC and an adaptive geometry-learning game for a Compaq iPaq PDA. The results showed that the PDA version of the game achieved a learning effect in the low-skilled group that was two times better than the laptop PC. Klatzky et al. (2008) examined the impact of their augmented-reality visualization device (called "sonic flashlight" (SF)) on the learning of ultrasound-guided needle insertion, and compared it to conventional ultrasound (CUS). The results showed that the SF users learned more because the SF had higher accuracy and lower variability in aiming and endpoint placements than CUS. Hamidi et al. (2011) compared the training effects of interactive multimedia devices (interactive CDs) and non-interactive media ones (films) to see if learners learned faster when instructed through interactive multimedia in comparison with non-interactive media. They also studied whether learners learned with greater accuracy when instructed using interactive CDs rather than films. Their results showed that there was a significant difference between the two groups in learning speed (in favor of the interactive CD group) but there was not any significant differences between the two groups in learning accuracy and in the amount of material memorized. Fassbender et al. (2012) presented a study of a virtual history lesson with different background stimuli (music or no music) to determine the effect of music on how much players remembered from the presented information. Two different display systems were used: three monitors set side by side, and three semi-cylindrical big screens on which the images were projected. The children were split into four groups combining music or no music and the two types of visualization. The results of the experiment showed that the participants using the monitors remembered a significantly higher number of facts, particularly if no background music was played in the second half of the history lesson. However, the participants using the three big screens had a significantly higher recall of facts when they listened to music in the second half of the history lesson. Schönborn et al. (2011) explored students' interaction and learning with a haptic or no-haptic virtual model, representing the specific binding of two biomolecules. The results showed that the students in the haptic group obtained higher learning gains since they tended to engage fewer visual representational switches.

Based on the developments and studies mentioned in this section, we consider that the selection of a device depends on the learning tasks that the users are going to perform. In our study, we

developed a m-learning game that uses AR technology, and we compared two mobile devices that are capable of running applications with high graphic demand like games, but with different characteristics (screen size and weight) that could influence the motivation and, consequently, the learning outcomes and knowledge improvements.

3. Theoretical background

The game that we have developed is based on two learning theories of education in an attempt of providing a comprehensive learning experience to the children. This two learning theories are:

1) Gardner's theory of Multiple Intelligences (MI) (Gardner, 1983). The theory of multiple intelligences has become a catalyst and framework for many current educational strategies. According to Gardner, intelligence is not a unitary element, but it includes different and specific capacities. There are eight types of intelligence through which individuals approach problems and develop solutions. According to the MI theory, every person has at least eight forms of intelligence (linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, naturalistic, interpersonal, and intrapersonal), and technologies can help students to use several of them (Gardner, 2000). In Table 1, we detail the activities proposed in our game for seven of these forms of intelligences:

FORMS OF INTELLIGENCES	ACTIVITIES INCLUDED
Linguistic This is the competence to use words in an effective way, both oral and written expressions. It assumes having skill in the use of the syntax, phonetics, semantics and pragmatic uses of the language.	The children have to hear and, in some cases, read the directions that the guide character gives during the game, which allow the children to gradually understand the dynamics of the activity. It is a linguistic activity, like reading a book, listening to a story, ... The children have to use their language skills to get through the game.
Logical-mathematical This is to have skill in solving logical-mathematical problems.	Seven different problems arise that the children have to solve using logical thinking in order to progress in the game. One of the activities in the game helps the children to identify the chemical formula of water and asks them to take atoms in order (two hydrogen atoms and one oxygen atom) to form molecules of water.
Visual-spatial This includes the ability to perceive and represent the visual-spatial world accurately and to form and manipulate mental images.	We propose solving spatial problems through observation and perceptual stimulation of objects from different angles. To do this, four Augmented Reality games were used that allow playing with the physical space and virtual objects using 3D in concepts related to the water cycle. While the children play, they can visualize objects from different angles. Augmented Reality allows virtual objects in to be inserted in the real space so that children can play with the augmented space created.
Naturalistic This consists of understanding the natural world.	One of the learning objectives is to bring the water cycle game to children in the second cycle of primary education in a fun and attractive way. The learning contents are designed to help the children understand the water cycle.
Intrapersonal This develops individual and personal knowledge, identity construction and self-esteem.	The media and development of the game have been proposed for a single use in order to respect the times for personal learning, individual working knowledge construction and individual self-esteem.

FORMS OF INTELLIGENCES	ACTIVITIES INCLUDED
Musical This is composed of different skills: perception, performance and production. Music perception allows different meaning within a musical composition to be discerned.	In our game, music and sound effects have been introduced to allow the children to get into the gameplay and to reinforce feelings of accomplishment and self-esteem related to the upcoming game stages.
Bodily-kinesthetic This is linked to the ability to control our body.	The game is conceived as an exploration of space, where the children have to move in order to locate objects (augmented reality games).

Table 1. **Activities proposed in our game for seven forms of Gardner's theory of Multiple Intelligences.**

2) The second theory, Kolb's theory (1984) was also taken into consideration to design our game. Kolb formulated a model of experiential learning, which conceives learning as an ongoing process. Kolb outlines this process through a model called "the wheel of Kolb's learning". This learning model is developed in 4 stages, as shown in Figure 2.
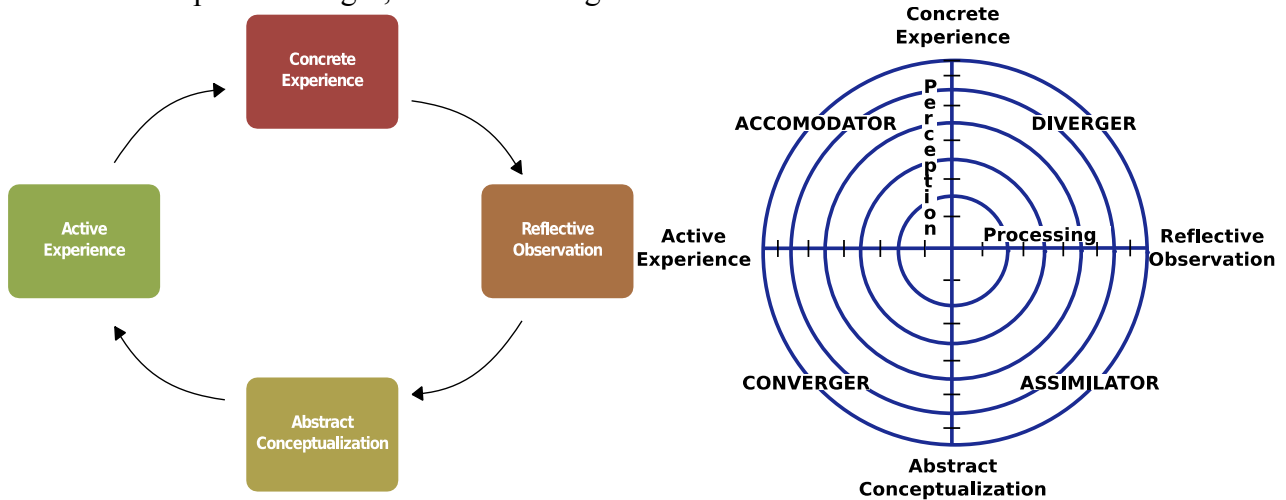

Figure 2. The wheel of Kolb's learning.

Our game was designed to initiate the learning wheel through a concrete experience. In the reflecting phase, the students think about what the have experienced (what have been seen, heard, or manipulated) and integrate this with their prior knowledge about the water cycle. These previous phases allow abstraction and generalization, adding meaning to the experience. The acquired concepts can be used actively in new situations (e.g. daily life).

4. Method
4.1. Design issues
    The subject of the game that was chosen was the water cycle. We selected this theme for multiple reasons. The first reason why we chose the water cycle was due to a former study. Professionals in education participated in that study to determine the subject preferences for educational computer games and their type for children (Furió et al., 2013). From the survey, we observed that "Nature" was one of the most preferred subjects. The second reason why we chose the water cycle theme was because this topic is covered in the Spanish law for primary education. The law establishes as an objective: "To know and value their natural, social, and cultural environment as well as the possibilities for caring for and preserving it". The law also establishes that the water cycle be included in the contents to be learned by children. It explicitly states that children have to know that water exists in different states and can be changed from one state to

another by means of an increase or decrease in temperature. It also states that children have to learn to respect, defend and improve the environment.

Several design characteristics were identified for virtual learning environments (Mueller & Strohmeier, 2010). Although they referred to virtual environments, we consider that they could also be applied to AR environments. Since "Interaction" was one of the highest rated characteristics, we took into account several design suggestions from different authors to improve the learner-system communication. Liarokapis and Newman, suggested to combine multiple interaction forms (Liarokapis & Newman, 2007). We used accelerometers, a touch screen, and tangible interaction. The accelerometers and the touch screen were used to move characters on the screen or to pick up objects during the game in order to complete certain tasks. There was tangible interaction because markers could be manipulated, for example, by rotating them or translating them with the hands. We also took into account studies that had provided other design principles and suggestions to enhance interaction (Koh, Duh & Gu, 2010). Sandor suggested mixing several input and output channels. Our game had camera, accelerometers, and a tactile screen as input channels, and videos, sounds, and graphics as output channels (Sandor & Klinker, 2005). In 2007, Henrysson and Billinghurst considered that the device camera could be used for a 6 Degrees of Freedom input by tracking real objects (Henrysson & Billinghurst, 2007). We used the iPhone and Tablet PC cameras for tracking markers, and children could select the virtual objects that appeared on them. We also tried to keep interaction techniques as user friendly as possible (Zhou, Duh & Billinghurst, 2008) since it is an important factor to take into account in order to provide an engaging gaming experience (Koh et al., 2010). We also tried to achieve a high degree of naturalness in interaction (Aliakseyeu, Subramanian, Martens & Rauterberg, 2002) by using two-handed interaction instead of one-handed interaction, since children use two hands when playing the mini-games. These mini-games had visual feedback: every action the children took was reflected on the screen. And we only used wireless mobile devices, so there were no other instruments that could cause annoyance to children.

Another important characteristic that obtained a high score in Mueller and Strohmeier's study was "Learning-process supportive" (Mueller & Strohmeier, 2010). In our game, the children are guided through a series of episodes that conform the water cycle in a structured order and the children are always aware of the episode that they are currently playing, and can understand the workflow of the whole experience. Education cannot be improved with only having technology (Fisch, 2005; Veenema & Gardner, 1996). Technologies that include a variety of media can help students to understand and learn concepts (Veenema & Gardner, 1996).

Another design aspect that we incorporated was the feedback from wrong answers. We did not reveal the correct answers after the children responded incorrectly (Fisch, 2005). Instead, the game told them the answer was not correct and the guide character reminded them of the task to perform. Other feedback that was added in order to help the children to complete the tasks was, for example, images of the objects they had to look for.

It is important to note that, in our game, the children used the camera in AR mini-games whose perspective was similar to a first-person perspective, making the children embedded agents, becoming part of the game (Dickey, 2005).

4.2. General description of the game

The game was aimed at reinforcing the learning of children about the subject of water. In this game, children could learn about the water cycle, water composition, and water pollution in a way similar to the way they studied it in school. They saw the major processes of the water cycle. They began viewing the evaporation phase, which was when the sun heated up water and turns it into vapor or steam. Then, this water vapor or steam lifted into the air. Next, they saw the condensation phase, which was produced when water vapor in the air got cold and changed back into liquid, forming clouds. The third water cycle phase they saw was precipitation, which occurred when the small drops that form the clouds got cold. The last water cycle phase was collection, which

occurred when water fell back to earth as precipitation. Apart from the water cycle, the children also learned what objects could pollute the water and what elements made up water.

Using AR, the children explored a room looking for the objects requested by the guide character, which in the game was a drop of water. To search for the objects, the children focused the device's camera on the different markers distributed around the room. Ten different AR markers were used and placed in the activity room, which was decorated with wall posters and images throughout the room for a more immersive experience, as showed in Figure 4. When the children selected an object, a message was displayed telling them whether the object picked up was the right one. Figures 5 and 6 show children playing the game. Non-AR mini-games (which in our application were virtual reality games) appeared in combination with some AR mini-games. The non-AR mini-games did not use AR, but they used tactile or accelerometer capabilities. The non-AR mini-games usually consisted of children having to collect a certain number of objects to get to the next screen. Between each mini-game (AR and non-AR), video and audio explanations were displayed, describing the rules and goals to complete the next mini-game. They also served as a reward for the children when they had completed a mini-game, showing what they had achieved. This way, we could link all the mini-games together in a continuous story thread.

There were seven mini-games in total, which are shown graphically in the flowchart in Figure 3. The game started with a video, which introduced the children to it. Next, they searched the room looking for the guide character, a drop of water. This activity served as a tutorial and was the first contact with our AR system. The children learned to focus and to select objects in the AR mini-games. After the guide character had been found, the first task required the children to form water drops from oxygen and hydrogen atoms. The atoms were represented as characters (different colored drops of water) with the letter 'H' or 'O' on their backs. The children first had to select two hydrogen atoms and then an oxygen atom. Then, a drop of water was formed. Once the children made three drops of water, they went to the first non-AR mini-game. Here, they had to collect twenty suns to evaporate the drops of water they had made before. This mini-game corresponds to the evaporation phase in the water cycle. The suns fell from the top of the screen and the children had to move a water drop that was at the bottom of the screen from left to right using the tactile screen capabilities of the devices. Next, in the second non-AR mini-game, the children had to place ten clouds over a mountain peak tilting the device to use the accelerometer capabilities. The third AR mini-game, which corresponded to the condensation phase of the water cycle, asked the children to find thermometers with low temperatures, which were blue in color. There were also thermometers with high temperatures, which were red in color. This way, they could cool the clouds and produce precipitation. Then, in the collection phase, the children had to collect the drops of water that were falling from the mountain peak (Figure 7). In order to complete the task, the children had to collect twenty drops of water by touching them. Finally, the water cycle was completed, but the guide character discovered a problem. A river was filled with objects that could be pollutants, so another request was made to the children. They needed to find and pick up the pollutant objects using the AR capabilities of the devices. All the mini-games contained a head-up display that showed the current status.

4.3. Hardware and Software

An iPhone 3GS with iOS version 4.3 and a HP Slate 500 Tablet PC with Windows 7 were used during the evaluation. Both devices had a built-in camera that allowed the real world to be recorded and displayed on the screen as well as an accelerometer. The dimensions of the iPhone were 11.55 x 6.21 x 1.23 cm. Its weight was 135 g. The dimension of the screen was 3.5". The dimensions of the Tablet PC were 23.40 x 15.00 x 1.47 cm. Its weight was 680 g. The dimension of the screen was 8.9".

To develop and run the iPhone game, we used the Xcode 4 IDE and the iPhone SDK 4.3. The software used to include AR capabilities on the iPhone was ARToolKitPlus 2.1.1. We used SIO2 1.4 version, a 2D/3D game engine with OpenGL-ES 1.1 capabilities designed for mobile devices (http://www.sio2interactive.com). The engine also provided the necessary functions, methods, and

scripts to import, load, and animate 3D scenes created in Blender, which was used in our game to create 3D objects.
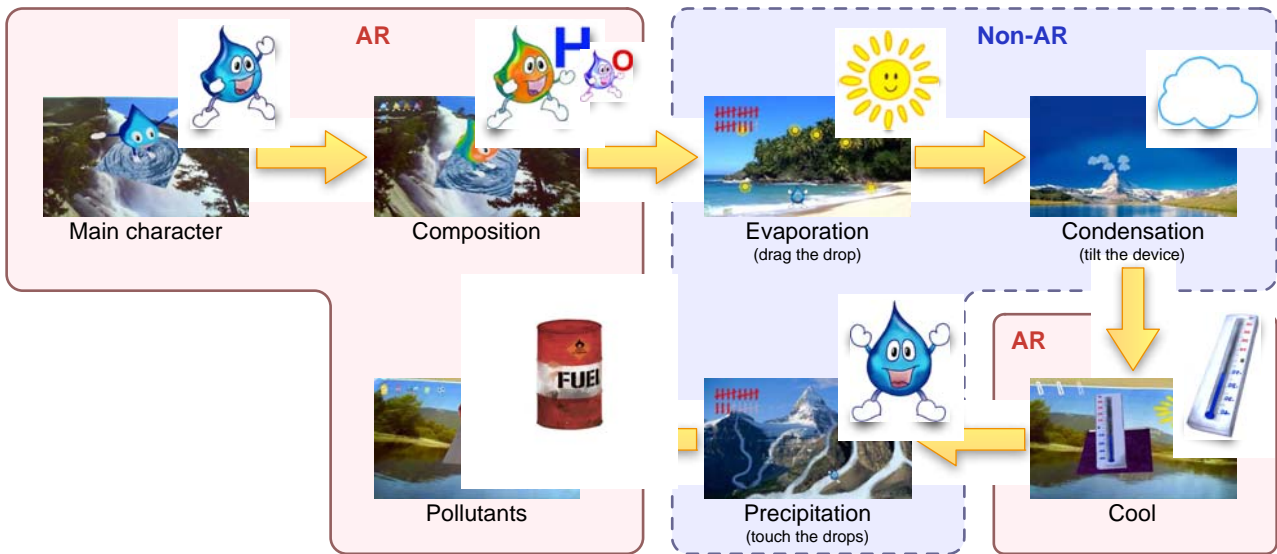

Figure 3. Flowchart of the AR and non-AR mini-games.

Figure 4. Activity room decoration.

Figure 5. Two children playing the game with the iPhone.

Figure 6. Child finding an oxygen drop.

Figure 7. Non-AR game with the Tablet PC.

To develop the game on the Tablet PC, we used the OpenSceneGraph (OSG) toolkit –version 2.9.5 with the FFmpeg plugin–. OpenSceneGraph let us import, animate, and render 3D objects with high performance in C++ language. This allowed us to create the 3D models in Blender and load them in our program. The FMOD sound library provided sound support. To include AR

capabilities, we used the OSG plugin osgART 2.0 RC 3, which used the ARToolKit library (Kato & Billinghurst, 1999) version 2.72.1. To protect both devices from shocks and falls, an external case was made (Figure 8). The dimension of the external case for the iPhone was 20.60 x 11.70 x 1.75 cm. Its weight was 127 g. The dimension of the external case for the Tablet PC was 32.80 x 22.00 x 2.10 cm. Its weight with the accelerometer included was 268 g. This external case means that the iPhone/Tablet PC are no longer common devices. They are modified devices, and hereafter they will be referred to as M_iPhone and M_Tablet PC. It also had a ribbon that children could put around their neck in order to provide more stability when holding the devices.

Figure 8. External case of the Tablet PC.

4.4. Participants

Seventy-nine children from 8 to 10 years old –with a mean age of 8.70±0.70– took part in the study: 42 boys (53%) and 37 girls (47%). These children were attending the summer school of the Technical University of Valencia (Escola d'Estiu).

4.5. Measurements

Three questionnaires were used for the validation. The first one was the pretest, and the other two were filled out after playing the first game and the second game. Table 5 in the appendix shows the relation of questions for each questionnaire. While most of the answers to the questions followed a Likert scale ranging from 1 to 5, others had categorical answers and, in some cases, hand-written answers.

The pretest (Q1) was composed of 6 questions designed to evaluate how much the children knew or remembered about water from school (composition (Questions #1, #2), cycle (Questions #3, #4, #6), and pollutants (Question #5)). An example of the questions was: *Do you remember what comprises water? a) Hydrogen and Oxygen; b) Potassium and sodium; c) The water is made up of nothing; d) I do not know / I do not remember.* These six questions about acquired knowledge were established by the pedagogues that participate in the project. These six questions were also ratified by a team of fifteen 3rd-grade teachers in order to adapt our assessment to the assessment that is carried out in the classroom on the subject. In order not to prolong the test phase the number of questions was limited to the minimum number of questions that the teachers agreed upon. The rest of the questions followed the same structure, with the exception of question 5, which was: *Check every object you think is a pollutant*. Then the participants had to check every object they believed to be a pollutant from a list of nine items. To make it friendlier to the children, an image accompanied every object. For the data analysis, this question was considered to be correct if all the objects were correctly checked. Apart from the knowledge questions, we also gathered gender, grade, and age from the participants.

The second questionnaire (Q2) was composed of 19 questions. The first 6 questions were the same ones the participants answered in Q1. This way we could observe if they had learned something about the water cycle they did not know before. The rest of the questions were about participant satisfaction and interaction with the game. One of the questions was: *Did you have fun playing the game? a) Very much; b) Quite a lot; c) Moderate; d) Not much; e) None*. Other questions were related to the AR part of the game, such as: *Did you like to see how objects appeared above the black squares (patterns)? a) Very much; b) Quite a lot; c) Somewhat; d) Not much; e) None*. In question 19, the participants had a chance to give a global score to the game.

The third questionnaire (Q3) was composed of 8 questions. Some of the questions were the same as in Q2, which allowed us to compare the two devices. In question 6, the children had to select which device they preferred to play the game: *What device did you like the best? a) iPhone; b) Tablet PC*. In question 7, the participants explained why they preferred one device over the other. In the last question, children described what they liked the most of the whole experience.

4.6. Procedure

The children who participated in this study were randomly assigned to one of two groups:

    A. The group that played the iPhone game first and then the Tablet PC game.

    B. The group that played the Tablet PC game first and then the iPhone game.

Both groups had a similar number of subjects: 41 children were assigned to group A, and 38 to group B. Figure 9 shows the procedure of both groups graphically. As can be observed, before playing any game, each child filled out the Q1 entry questionnaire and some instructions were given to the children about how to play the game. Then, the first group played the iPhone game. After completing the game, they answered the post-game questionnaire Q2. Then, these children played the Tablet PC game and filled out the ending questionnaire, Q3, when they had finished playing. The second group played the Tablet PC game first. After completing the game, they also filled out Q2 questionnaire. Then, these children played the iPhone game and also answered the Q3 questionnaire when they had finished playing. In our study, the content evaluation protocol was established in a way similar to the one applied to assess the contents in the classroom. In classrooms of the second cycle of the Spanish primary education, the usual established dynamic is to teach a subject and then evaluate the level of learning of the content.
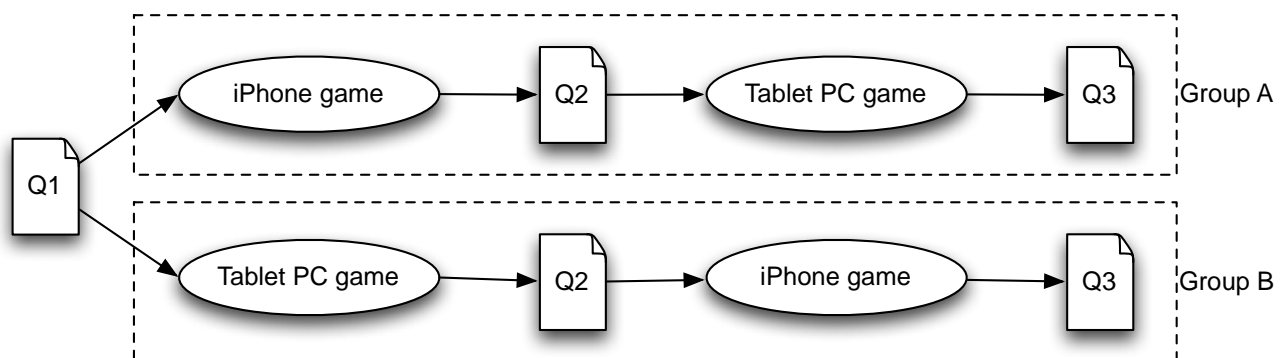

Figure 9. Study procedure. Two games and three questionnaires were used.

The evaluations took place in two mirrored rooms. Each room had two identical playing areas where two children (from each group) could play simultaneously and individually. The two playing areas had the same set of markers placed in the same positions. There was no interaction between the two children, and there was also a person with each child to guide them and to clarify the possible doubts during the whole activity. The questionnaires were filled out in the same room where the activities took place. To avoid the influences of answering the pretest on the results of the posttest, the children were encouraged to answer all questions without any pressure; if they did not know the answer it was considered to be completely normal, the children were not informed whether or not their answers on the pretest were correct or not. Thus, the children did not acquire any knowledge by answering the pretest; they only learned during the game.

5. Analysis

The statistical analysis that follows uses parametric tests, such as t-test and ANOVA, which are robust to violations of equality of variances if the ratio between the largest variance and the smallest variance is not larger than 4:1. Our analysis is within this margin with a ratio lower than 2:1. Moreover, the sample sizes have been balanced as described in Section 4.3, which also reinforces the validity of the tests. We also provide effect sizes (Cohen's d, generalized eta-squared, and Cramer's V) in order to have a measure that is independent from the sample sizes and to have a magnitude of the significant differences found, which is not measured by p-values.

The time that the children took to finish a game was an average of 14.62±2.03 minutes for the Tablet PC, and 13.3±1.96 minutes for the iPhone. Although the differences in the time of play were not perceived during the evaluations, we found that, surprisingly, there were significant differences between the two devices, with a medium effect size (t[124] = 3.67, p < 0.01, Cohen's d = 0.65). It is very reasonable that the bigger screens took more time, since the children had a bigger area to interact with and the movements had more amplitude, while with the smaller screen, the children had to focus their attention on a more limited area. Relating these times to the preferred device of the children (question #20 in the appendix), there were no significant differences in the time of those who preferred the M_Tablet PC and those who preferred the M_iPhone (t[54] = 1.18, p = 0.24, Cohen's d = 0.32).

The variable of knowledge was created to condense the six knowledge questions (Questions #1 to #6 in Table 5 in the appendix) by counting the number of correct answers. Figure 10 shows the box plot for the pretest and posttest scores indicating the high dominance of correct answers after playing the first game (posttest).
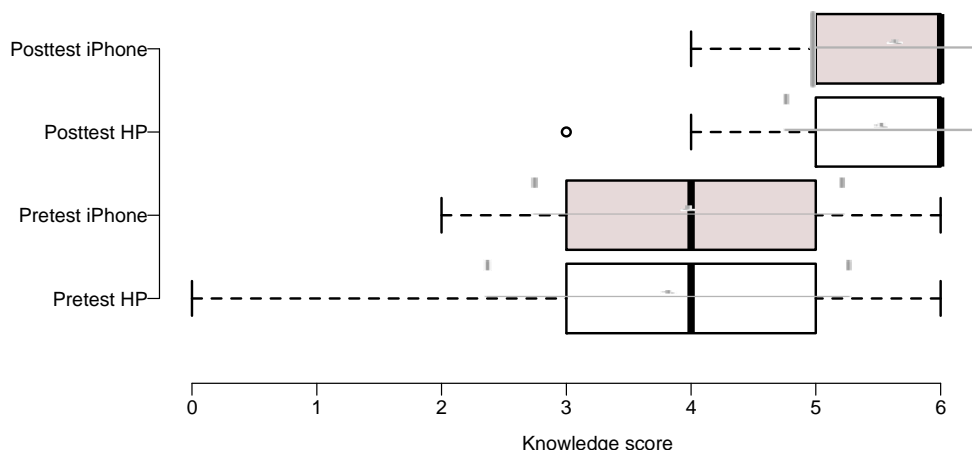

Figure 10. Box plots of the scores of the knowledge questions in the pretest (Q1) and the posttest (Q2) for the M_Tablet PC and the M_iPhone devices. The mean and standard deviation are represented in solid gray lines.

In order to have a complete understanding of the knowledge variable, we performed several statistical tests. To compare the effect of the game on this variable, a t-test was performed, which revealed that the scores in the posttest (mean 5.58±0.70) were significantly higher than in the pretest (mean 3.90±1.33) (t[78] = 11.21, p < 0.01, Cohen's d = 1.26). Another t-test was performed to see the differences of knowledge between the iPhone game (group A, mean 5.63±0.65) and the Tablet PC game (group B, mean 5.53±0.75). No significant differences were found between the two devices (t[77] = 0.67, p = 0.50, Cohen's d = 0.15).

We also performed a mixed design ANOVA test in order to take into consideration several factors simultaneously. The factors of grade and gender were between subjects because they were observed, but the factor of questionnaire (Q1-pretest or Q2-posttest) was within subjects because it contained repeated measures (first when children filled Q1 and second when the same children filled Q2). The factor of device used to play (M_iPhone, M_Tablet PC) was also taken into consideration as between subjects because children from the two groups were different. The effect

size used was the generalized Eta-squared (η_G^2) (Olejnik and Algina, 2003), which has been proven to be very suitable for mixed design analyses because it takes into account the repeated measures and the observed and manipulated factors (Bakeman, 2005). The result of the analysis in Table 2 showed that there were very significant differences in the grade and the questionnaire factors. The p-values in these cases were below 0.01, and the effect sizes revealed that the most influential factor was the questionnaire with an exceptionally large size, followed by the grade, which had a small-medium size. There was also some interaction between the grade and questionnaire factors, but it had a small effect size and we do not consider it relevant. It should be also noted that no significant differences were found between the devices.

Also, the perceived acquired knowledge by the children fitted the previous result. On a scale from 1 to 5, the score given to how much they thought they had learned (Question #12) was high after using the M_iPhone (mean 4.42±0.86) and after using the M_Tablet PC (mean 4.42±0.78), and there was no significant difference between the devices (t[76] = 0.02, p = 0.98, Cohen's $d < 0.01$).

Factor	d.f.	F	p		Effect size (η_G^2)
Grade	1	7.43	<0.01	***	0.06
Gender	1	0.63	0.43		<0.01
Device	1	0.70	0.41		<0.01
Questionnaire	1	124.68	<0.01	***	0.38
Grade:Questionnaire	1	5.15	0.03	**	0.03
Other interactions	1	<1.04	>0.31		<0.01

Table 2. Mixed design ANOVA for the knowledge scores. Significance: * p<0.1, ** p<0.05, *** p<0.01. N = 79.

Another way to study the knowledge variable is to use the average normal gain (Hake 1998). The average normal gain was created to evaluate the effectiveness of a course based on a pretest and a posttest. It allows measuring the degree of learning obtained through a particular mode of instruction (Mayo, 2007). It is defined as the ratio of the actual average gain to the maximum possible average gain (Savinainen & Scott, 2002):

$$\langle g \rangle = \frac{\langle Gain \rangle}{\langle Gain \rangle_{max}} = \frac{\% \langle posttest \rangle - \% \langle pretest \rangle}{100 - \% \langle pretest \rangle},$$

where $\% \langle pretest \rangle$ is the percentage of correct answers in the pretest and $\% \langle posttest \rangle$ is the percentage of correct answers in the posttest. The values of $\langle g \rangle$ can vary from 0 (null learning gain) to 1 (maximum learning gain), and Hake established three types of courses depending on this value: a) "High-g" courses ($\langle g \rangle \geq 0.7$), b) "Medium-g" courses ($0.3 \leq \langle g \rangle < 0.7$), and c) "Low-g" courses ($\langle g \rangle < 0.3$). In our case $\% \langle pretest \rangle = 63.6\%$ and $\% \langle posttest \rangle = 92.1\%$, and therefore the average normal gain in our game was $\langle g \rangle = 0.78$. This measure is an indicator of the high effectiveness our system had during the evaluation.

Apart from the knowledge variable, another variable of satisfaction was created to combine the answers of several questions (the mean of answers to Question #7, #8, #9, #16, #17, #18, and #19), giving us a measure of the degree of engagement and enjoyment with the game. The overall rating was very high (mean 4.77±0.23), indicating that the children were highly satisfied with the game. The variable was analyzed using a multifactorial ANOVA with the gender, grade and the device factors (Table 3). The high p-values and small effect sizes showed that there were no significant differences for any of the factors, including the device used. We could deduce that since the game was very appreciated and enjoyed very similarly by all the groups studied motivation to learn was increased by using the game.

Factor	d.f.	F	p	Effect size (η_G^2)

Grade	1	2.14	0.15	0.03
Gender	1	0.02	0.83	<0.01
Device	1	1.05	0.31	0.01
Grade:Gender	1	2.05	0.16	0.03
Other interactions	1	<0.09	>0.36	<0.01

Table 3. Multifactorial ANOVA for the variable of satisfaction. N = 79.

We also analyzed several questions (#7, #8, #16, #17, #18, #19) individually. We used t-tests in order to observe if there were statistical differences between the M_iPhone and the M_Tablet PC. As Table 4 shows, for the questions analyzed, the results of the two devices were similar with the exception of Question #19. The M_Tablet PC obtained a marginally higher result in engagement questions (Questions #7, #8, #9). In contrast, the M_iPhone obtained marginally higher results in ease-of-use questions (Questions #16, #17). The only statistical difference between the two devices was found in Question #19, where the children scored the M_iPhone game significantly higher than the M_Tablet PC game.

Questions	#7	#8	#16	#17	#18	#19
M_iPhone (mean±sd)	4.85±0.42	4.84±0.36	4.52±0.74	4.76±0.54	4.87±0.34	9.88±0.31
M_Tablet PC mean±sd)	4.89±0.31	4.87±0.34	4.29±0.77	4.61±0.59	4.87±0.34	9.62±0.75
t	0.49	0.34	1.38	1.18	0.03	2.04
p	0.63	0.73	0.17	0.24	0.97	0.04*
Cohen's d	0.11	0.08	0.31	0.27	0.01	0.46

Table 4. T-tests of several questions (#7, #8, #16, #17, #18, #19). The symbol * indicates significant differences; df = 77.

Another hint that determines if the game was satisfying for the children is their desire to continue playing the game (Question #13). A total of 97.37% of the children who played with the M_Tablet PC and 95.12% of the children who played with the M_iPhone would like to play the game again to learn about new subjects. Only one child did not want to play the game on the M_Tablet PC again, while only two children were undecided with regard to playing the game again on the M_iPhone.

Multiple t-tests were performed in order to study the AR experience that the children had while playing the game (Questions #9, #10, #14). As Table 5 shows, no statistically significant differences were found between the two devices for the first time the children played the game. Similar results were obtained for the second time they played the game (Table 6). This would indicate that the children who used the M_Tablet PC had a similar AR experience to the children who used the M_iPhone. We also performed ANOVA tests in order to see if other factors affected the AR questions. With regard to the engagement of AR in the game (Questions #9 and #10), the overall rate for the enjoyment of AR was very high (Question #9: 4.85±0.43; Question #10: 4.85±0.39). For Question #9, the ANOVA test showed that all the groups studied appreciated the fact that the objects appeared on the black squares in a similar manner in both devices. There were no significant differences for: grade ($F[1,71] = 0.22$, $p = 0.64$, $\eta^2_G < 0.01$), gender ($F[1,71] = 0.05$, $p = 0.82$, $\eta^2_G = 0.01$), and device ($F[1,71] = 0.76$, $p = 0.38$, $\eta^2_G < 0.01$). An ANOVA test of Question #10 showed similar results: grade ($F[1,71] = 0.09$, $p = 0.77$, $\eta^2_G < 0.01$), gender ($F[1,71] = 0.90$, $p = 0.35$, $\eta^2_G = 0.01$), and device ($F[1,71] = 0.58$, $p = 0.45$, $\eta^2_G < 0.01$). This result is very positive because it means that the use of AR can be spread over all of the factors studied without any restriction. For the AR ease of use (Question #14), the children found it easy to play AR games, and all the groups experienced AR gameplay similarly, since no significant differences were found in the grade, gender, group, the order in which the devices were used, or the device used. The results of the mixed design analysis are shown in Table 7, where the order and device factors were within subjects and the rest were between subjects. According to the p-values and the effect sizes, none of the factors had significant differences. Similarly to the previous result, this result is very

appreciated by the authors because it means that the introduction of AR was very positive for all the children and all the devices homogeneously.

	#9	#10	#14
M_iPhone (mean±sd)	4.80±0.46	4.88±0.33	4.39±0.77
M_Tablet PC (mean±sd)	4.89±0.39	4.82±0.46	4.03±0.88
t	0.94	0.70	1.95
p	0.35	0.49	0.05
Cohen's d	0.21	0.16	0.44

Table 5. T-tests of the AR questions in Q2 between the iPhone and Tablet PC devices. The symbol * indicates significant differences; df = 77.

	#9	#14
M_iPhone (mean±sd)	4.79±0.47	4.29±0.77
M_Tablet PC (mean±sd)	4.68±0.52	4.43±0.86
t	0.95	0.75
p	0.35	0.46
Cohen's d	0.21	0.17

Table 6. T-tests of the AR questions in Q3 between the iPhone and Tablet PC devices; df = 77.

Factor	d.f.	F	p	Effect size (η^2_G)
Grade	1	2.52	0.12	0.03
Gender	1	0.77	0.38	<0.01
Group	1	2.25	0.14	0.02
Order	1	2.24	0.14	<0.01
Device	1	1.22	0.27	<0.01
Interactions	1	<1.81	>0.15	<0.02

Table 7. Mixed design ANOVA for the ease of use of the AR system. N = 79.

A Chi-squared test for Question #20 revealed that the preference for the M_iPhone or the M_Tablet PC significantly differed between children who finished playing with one of the devices (X^2[1, N=79] = 12.08, p < 0.01, Cramer's V = 0.42). After analyzing the results, we could see that children tended to choose the device they had used the last time. This difference is shown graphically in Figure 11, where 71.43% of the children in group A chose the M_Tablet PC, and 70.27% of the children in group B chose the M_iPhone. Another test revealed that there were not significant differences in this tendency for the devices (X^2[1, N=79] = 0.05, p = 0.83, Cramer's V = 0.05). Therefore, we could conclude that there was a bias towards preferring the last device used, but the effect of that bias was equivalent for the M_Tablet PC and the M_iPhone. Thus, we could conclude that there was no significant difference in the preferences for the M_Tablet PC and the M_iPhone, and that the differences were due to the order of playing.

To interact with the device, the children thought that it was easier to use the accelerometer rather than to touch the tactile screen (Question #15), as Figure 12 shows. There were no significant differences in the proportions of the preferred interaction on the two devices (X^2[1, N=79] = 1.69, p = 0.19, Cramer's V = 0.12). However, we believe that the higher weight of the Tablet PC could have influenced this result decreasing the tactile rates because some children had some difficulties holding the device while touching the screen.

Question #11 evaluated the preferences of different types of technologies in the mini-games: AR, tactile screen, and accelerometers (Figure 13). There were no significant differences found in these three types of interaction when people used the M_Tablet PC and the M_iPhone (X^2[2, N=79] = 0.51, p = 0.77, Cramer's V = 0.08). In spite of the differences in the proportions, there were no significant differences in the proportions for the M_iPhone (X^2[2, N=79] = 3.59, p = 0.17) or the M_Tablet PC (X^2[2, N=79] = 2.21, p = 0.33).
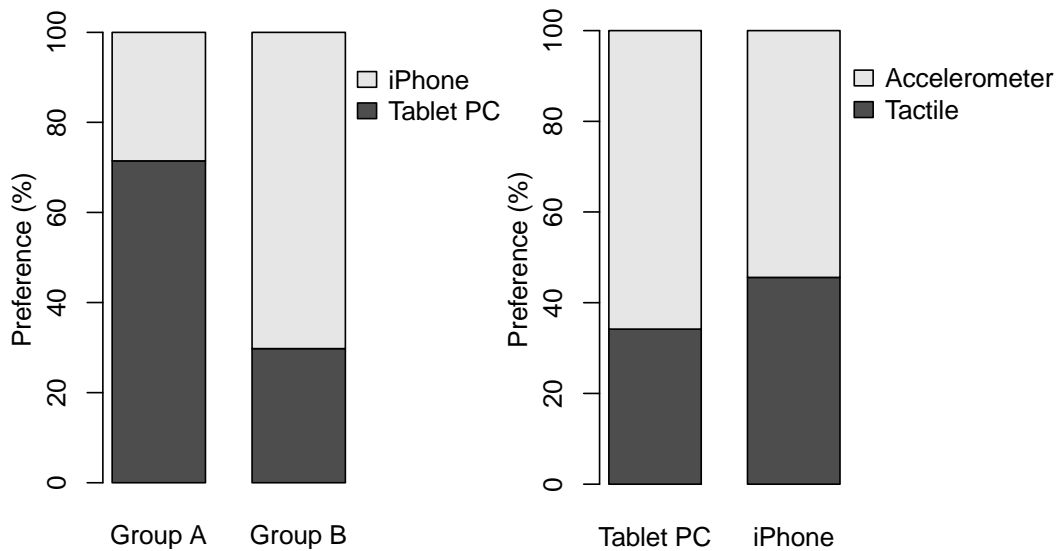
Figure 11. Preferences for the favourite device in groups A and B.

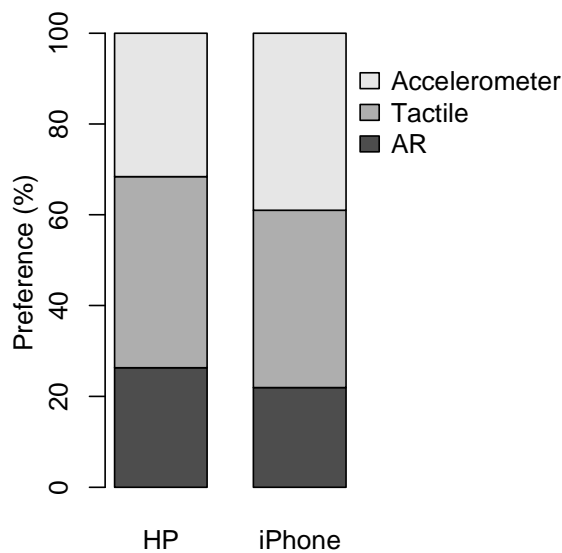Figure 12. Preferences for the easiest technology for each device.

Figure 13. Preferences for the interaction type for each device.

6. Discussion and implications

Educational games have been proven to be useful learning tools. Video games produce chemical changes in the brain that promote learning, and some studies have demonstrated that video games are better than a lecture (Mayo, 2007). Video games also have some advantages over other learning techniques or even traditional games. Video games have massive reach, and they appeal to children and adults and to men and women. They are effective learning paradigms. Data can be obtained from them in order to improve the games or collect information about the children. Although there are empirical studies that have already shown evidence of children's learning from educational computer games (Fisch, 2005), there are not many AR mobile games that have studies that corroborate it, and not all educational games are equally effective in transmitting knowledge. In order to maximize the effectiveness of the game, we have followed multiple design patterns from various authors. The theories by Kolb (Kolb, 1984) and Gardner (Gardner, 1983) were used for the learning design part. For the interaction and other game aspects, several suggestions and design principles from various authors were used. We have also taken into account the Spanish curricula in the design of our game.

6.1. Device comparison

We performed a study to find out which device was better suited as an educational game, the M_iPhone or the M_Tablet PC. We would like to point out that an iPhone and a Tablet PC were adapted adding an external case. These external cases protect the two devices from shocks and falls. However, at the same time, the addition of these external cases implies that the two devices are significantly modified. Therefore, our conclusions are related to these modified devices. From our point of view, these conclusions can be extrapolated to the original devices. However, a specific study on this topic should be carried out to corroborate this possible extrapolation.

We compared the devices regarding several aspect such as learning outcomes, satisfaction, engagement, interaction, and AR experience. No significant differences were found between the two devices, and they provided a similarly good experience. In this study, no evidence was found to say that the large screen of the M_Tablet PC (where children could see the virtual objects in greater detail) or the better handling of the M_iPhone were determinant factors. This could also be due to the fact that the two devices had a similar degree of naturalness (Aliakseyeu et al., 2002), with neither device being intrusive, making them both very easy to use. The study indicated that, despite the fact that the M_iPhone and the M_Tablet PC had different characteristics (screen size and weight), both devices were well suited for educational games.

As to what kind of interaction was considered to be easier, participants said that the accelerometer was easier to use than touching the tactile screen on both devices. Although there were no statistically significant differences between the devices, the preference for the accelerometer was higher with the M_Tablet PC. We believe that this was due to the fact that it was heavier than the M_iPhone and some children had some difficulties touching the screen with one hand while holding it with the other hand. Most of them rapidly found strategies to make the experience more comfortable by grabbing the M_Tablet PC with the whole arm, but some of them did not find it intuitive. This problem did not affect the M_iPhone, which was lighter than the M_Tablet PC.

There was not a clear preference of technology for the games that used the accelerometer, the tactile screen and AR; the children appreciated it similarly. As a consequence, we could conclude that there was a good balance in the use of AR games and non-AR games, which achieved an excellent level of satisfaction to the children when combined.

In the questionnaires, the participants showed a clear preference for the last device they had used. After analyzing this bias, we showed that it had a very similar magnitude in the case of the M_iPhone and the M_Tablet PC. Therefore, we concluded that by eliminating the effect of the order in which children played the game, they had no significant preference for either device.

The children spent more time using the M_Tablet PC than the M_iPhone due to the difference in the size of the area they had to focus on. However, although there are statistical significant differences in the time to play, we believe they are not meaningful, since it did not affect the device they preferred.

6.2. Use of the devices

With regard to the manipulation of the device, some 8-year-old children have a little trouble holding the Tablet PC since it was too big or too heavy for them. In some cases, they needed external help to point with the device in the AR mini-games. However, the vast majority of them managed to interact with the display properly within the first minute of use, finding a comfortable way to hold it (using the hands and the arms). At the other end of the age range, hardly any of the 10-year-old children had trouble using the device. Our recommendation is that applications with a Tablet PC should be carefully designed to have some pauses in order to allow the children to rest the device on a table or on their laps if they are sitting. While they are resting, children are shown a video explanation for the next game. In our game about the water cycle, there were non-AR games in between the AR games, such as the games of the sun, clouds and rain, where the children could lay the device on a table in order to avoid a prolonged use of AR games with the M_Tablet PC. This recommendation is still true for older ages, but the resting pauses could probably be shorter.

We tested the M_iPhone with both younger and older children, and we did not find any of these problems; this makes the M_iPhone more suitable for a wider range of ages. However, as children grow up, they usually become acquainted with more gaming devices and they expect more difficult challenges and more complex scenarios and advanced graphics. These requirements are more easily met on a M_Tablet PC, which has more capabilities.

With regard to the content that is the most suitable for AR educational games, we conducted a study to determine the subject preferences for an educational computer game for children ranging in age from 8 to 10 years old, but not limited to mobile devices (Furió et al., 2013). A total of 150 professionals participated in this study. The results indicated that 'nature and living organisms' and 'multiculturalism, solidarity and tolerance' were the two most preferred subjects. As mentioned in section 4.1., this was one of the reasons why we developed the game for learning the water cycle. This study could be used as a first recommendation for choosing the most suitable subject for an AR educational game. Apart from being useful to researchers, we consider that this study could also be valuable and useful for companies and training institutions.

In general, children playing AR games with handheld devices is metaphorically similar to children exploring the world and discovering new elements with a magnifying glass. In our game, the children discovered things in the world of nature. However, more studies should be carried out to help educators determine the best learning situations for using a mobile phone or a Tablet PC to improve student learning.

6.3. Acquired Knowledge

We performed a study to determine if children acquired new knowledge after playing the game with the devices. The children answered questions in pretest and posttest questionnaires. The results showed significant differences in both devices regarding the acquired knowledge, and we could state that playing with the game very positively affected the learning outcomes of the participants. These results revealed a high level of learning. This indicates that children remembered a lot of the knowledge transmitted in the game. For this reason, we can say that our game has been effective when it comes to transmitting knowledge in the short-term. From our point of view, having significant differences regarding the acquired knowledge is a very good result. This implies that the children had indeed learned new concepts after playing the game with the devices. This could imply that these kinds of games could be used as a way to reinforce the knowledge learned in school with the advantage that these games can be used in any place and at any time (Jones & Jo, 2004).

We also studied if there were differences in the acquired knowledge regarding the gender, grade and device used. No significant differences were found. This was an unexpected result because our first hypothesis was that children would prefer the Tablet PC mainly for its larger screen size. Although unexpected, it is an excellent result because it means that the game is well suited for these factors and the system can be used in more situations without many restrictions.

We also used the average normal gain (Hake, 1998) to have a measure of the degree of learning obtained. Our system achieved a "High-g" result, which indicates that interactive engaging systems like our game could be effective in enhancing concepts (Hake, 2002). Comparing our study to other studies like (Freitas & Campos, 2008), where they used SMART system (Section 2), we find that despite the fact that they obtained good results like 34%, 62% and 33% of average normal gain, our study obtained a higher $\langle g \rangle$ (78%). This reinforces the idea that our game has been proven to transmit the knowledge to the children successfully. We consider that the guidelines mentioned in Section 3.1 contributed to make the game more effective. Since in this work we have not validated this contribution, different principles should be used and data that is directly related to the design should be collected.

7. Conclusions

Following the design principles of various authors regarding learning theories and game design patterns, we have developed a game for iPhone and Tablet PC for learning about the water cycle. The game included multiple interaction forms (touch-screen interaction and accelerometer) and

combined AR mini-games with non-AR mini-games. We added an external case to the devices with the objective of facilitating manipulation.

We performed a study to check how the size and weight of the devices affect several aspects. Nowadays, children are used to playing with consoles and mobile devices (Beck & Wade, 2004, 2006; Prensky, 2001; Tapscott, 1998) such as their parents' mobile phones; in many cases the interaction of the devices was not new. While most of the children found it easy to manipulate the two devices, we found that, in order to improve the experience, different factors should be addressed, such as the weight or the touch screen sensitivity. Despite these small problems related to the ease of use, the results showed that both devices were equally suited for educational games. The different characteristics of the devices did not affect the variables analyzed. This is encouraging because it shows that children could adapt perfectly to different devices without major problems. In our opinion, the fact that there is no apparent difference between large and small size screens represents the most important finding of our study. We believe it is potentially important for educators to think about how best to provide the technology needed to support learning in schools.

The results showed that the game was really fun. The children scored the fun experienced with a mean of about 4.8 (on a scale from 1 to 5) for the two devices. For the AR experience, the analysis revealed that the children appreciated the AR experience scoring very high, means of over 4 (on a scale of 1 to 5) for all the related questions and for the two devices. For learning in the short-term, the results were very positive, showing that the children acquired new knowledge. With regard to the desire for learning new things at school using these devices, the means were very high with a mean of 4.87 for the two devices (on a scale from 1 to 5). This implies that the two devices would be well received for use in the class for learning purposes, and they could probably be helpful as a reinforcement in learning environments like classrooms. In our opinion, the designed game can be an important educational resource in the classroom because of its close relationship with the content of school curriculum and its highly motivational component as a tool to introduce or reinforce classroom content. Moreover, games of this kind can be used outside school because they only need a minimum setup with some markers printed on paper and the handheld device. The features of the game and its minimal requirements position it as a pervasive educational game with great potential. However, the use of this technology at school presents several drawbacks. First, there is a cost issue because each student has to use a device. Second, the markers should be distributed in the room. In current school rooms, it is not possible for 20 students to play the game at the same time in the same room.

With regard to future work, the game could be enhanced by adding other play modes such as multiplayer. With the multiplayer mode, we could make collaborative and competitive gameplay. Our game could also be improved graphically with the introduction of more advanced animations in the AR mini-games. From this experience and previous experiences with different types of devices, we firmly believe that the added external cases help children's manipulation. However, a formal study should be carried out to determine their benefits and drawbacks. Since the evaluation was made by filling out some questionnaires, making these more interactive by using the same devices in which they played the games would make the children more willing to fill them out. In this paper, we have compared two devices, but other comparisons are also possible; for example, using a control group in which the children learn the water cycle using traditional learning or carrying out practical experiments with heating and cooling water. Comparing the results of different comparative studies could bring interesting conclusions (e.g. the children perform/do not perform differently in two different comparatives). In another study that is in progress, we are currently comparing learning the water cycle using one of the devices and traditional learning in a classroom. Another possible future work is to consider long-term learning. To do this, two groups (with or without rehearsal) could be considered. With the work and ideas presented here, we hope to benefit the educational community that is technologically supported, especially with regards to m-learning, and new developments and validations.

References

Albert, D. and Mori, T. (2001). Contributions of cognitive psychology to the future of e-learning. Bulletin of the Graduate School of Education, Hiroshima University, Part I (Learning and Curriculum Development), 50 , 25-34.

Aliakseyeu, D., Subramanian, S., Martens, J.-B., Rauterberg, M. (2002). Interaction techniques for navigation through and manipulation of 2d and 3d data. In EGVE '02: Proceedings of the workshop on Virtual environments 2002, Eurographics Association, Aire-la- Ville, Switzerland, Switzerland, 179–188.

Ardito, C., Costabile, M.F., Lanzilotti, R., Pederson, T. (2007). Making dead history come alive through mobile game- play, Conference on Human Factors in Computing Systems 2007, 2249-2254.

Au, C., Entwistle, N. (1994). Memorisation with understanding in approaches to studying: cultural variant or response to assessment demands? European Association for Research on Learning and Instruction Conference.

Azuma, R. T. (1997). A Survey of Augmented Reality. Media, 6(4), 355-385.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3):379–384, 2005.

Balog, A., Pribeanu, C. & Iordache, D. (2007). Augmented Reality in Schools: Preliminary Evaluation Results from a Summer School. Proceedings of WASET International Conference on Technology and Education - ICTE 2007, 24, 114-117.

Beck, J. C., & Wade, M. (2004). Got game: How the gamer generation is reshaping business forever. Boston: Harvard Business School Press.

Beck, J. C., & Wade, M. (2006). The kids are alright: How the gamer generation is changing the workplace. Boston: Harvard Business School Press.

Billinghurst, M., Kato, H., Poupyrev, I. (2001). The MagicBook - Moving Seamlessly between Reality and Virtuality. Computer Graphics and Applications, 21(3), 2-4.

Billinghurst, M. (2002). Augmented reality in education. New Horizons for Learning, www.newhorizons.org/strategies/technology/billinghurst.htm

Blecic, I., Cecchini, A., Rizzi, P. Tronfio, G.A. (2002). Playing with automata. An innovative perspective for gaming simulation. Proc. 5th Int. Conf. on Cellular Automata for Research and Industry, 337-348.

Brugnoli, M., Lorusso, I., Corsini, F., Bormida G.D., Keefe, T., Ahonen, M. & Bo, G. (2005). MOBIlearn deliverable D12.1-Market study, in Brasher A. MacAndrew P. and Sharples M. (2005) Roadmap for further research on pedagogical issues," MOBIlearn/OU/WP4/4.3/1.0, 2004 [online] Available from:
http://www.mobilearn.org/download/results/public_deliverables/MOBIlearn_D4.3_Final.pdf
[Accessed: April 2012].

Chang, C-W., Lee, J-H., Wang, C-Y., and Chen, G-D. (2010). Improving the authentic learning experience by integrating robots into the mixed-reality environment. Computers & Education. 55 (4), 1572-1578.

Chen., Y.C. (2006). A study of comparing the use of augmented reality and physical models in chemistry education. In Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications (VRCIA '06). 369-372.

Connolly, T. M., Stansfield, M., Hainey, T. (2011). An alternate reality game for language learning: ARGuing for multilingual motivation. Computers & Education. 57 (1), 1389-1415.

Dickey, M.D. (2005). Engaging By Design: How Engagement Strategies in Popular Computer and Video Games Can Inform Instructional Design. Education Technology Research & Development, 53(2), 67–83.

Earnshaw, R. (2011). From E-Learning to M-Learning - the use of Mixed Reality Games as a new Educational Paradigm. International Journal of Interactive Mobile Technologies iJIM, 5(2), 17-25.

Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R. & Kirk, D. (2004). Savannah: mobile gaming and learning? Journal of Computer Assisted Learning, 20, 399–409.

Fassbender, E., Richards, D., Bilgin, A., Thompson, W. F., & Heiden, W. (2012). VirSchool: The effect of background music and immersive display systems on memory for facts learned in an educational virtual environment. Computers & Education, 58(1), 490-500.

Fisch, S.M. (2005). Making educational computer games "educational". Proc. Int. Conf. on Interaction Design and Children, ACM Press, 56-61.

Freitas, R., and Campos, P. (2008). SMART: a SysteM of Augmented Reality for Teaching 2nd grade students. In Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 2 (BCS-HCI '08), (2), 27-30.

Furió, D., González-Gancedo, S., Juan, M.C., Seguí, I., Costa, M. (2013). Evaluation of learning outcomes using an educational iPhone game vs. traditional game. Computers & Education, in press, http://dx.doi.org/10.1016/j.compedu.2012.12.001.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic.

Gardner, H. (2000). Can technology exploit our many ways of knowing?. Retrieved November 10, 2011., from http://www.howardgardner.com.

Garaigordobil, M. (2005). Importancia del juego infantil en el desarrollo humano, Aula de infantil, 25, 37-43.

Gee, J. P. (2003). What video games have to teach us about learning and literacy. Comput. Entertain. 1 (1), 20-20.

Georgiev, T., Georgieva, E., & Smrikarov, A. (2004). M-Learning - a New Stage of Е -Learning. (K. Boyanov, Et Al., Ed.)International Conference on Computer Systems and Technologies CompSysTech2004, 28, 1-5.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66(1), 64.

Hake, R. 2002. Lessons from the physics education reform effort. Conservation Ecology. 5(2), 28. [online] URL: http://www.consecol.org/vol5/iss2/art28/

Hamidi, F., Kharamideh, Z. M., & Ghorbandordinejad, F. (2011). Comparison of the training effects of interactive multimedia (CDs) and non-interactive media (films) on increasing learning speed, accuracy and memorization in biological science course. Procedia Computer Science, 3, 144-148.

Henrysson, A., Billinghurst, M. (2007). Using a mobile phone for 6 DOF mesh editing. The 7th ACM SIGCHI New Zealand Chapter's international Conference on Computer-Human interaction: Design Centered HCI (CHINZ '07), 254, 9-16.

Hsiao, K.F. (2010). Can we combine learning with augmented reality physical activity?. Journal of CyberTherapy and Rehabilitation, 3 (1), 51-62

Huang Y., Lin, Y., and Cheng, S. (2010). Effectiveness of a Mobile Plant Learning System in a science curriculum in Taiwanese elementary education. Computers & Education. 54 (1), 47-58.

Johnson, D., & Johnson, R. T. (1981). Effects of cooperative and individualistic learning experience on interethnic interaction. Journal of Educational Psychology, 73, 444–449.

Jones, V., & Jo, H. J. (2004). Ubiquitous learning environment: an adaptive teaching system using ubiquitous technology. In Proceedings of the 21st ASCILITE conference, Perth, Western Australia, 5–8 December, 2004, 468–474.

Jones, L. C. (2006). Effects of Collaboration and Multimedia Annotations on Vocabulary Learning and Listening Comprehension. CALICO Journal, 24(1), 33-58.

Juan, M.C., Furió, D., Alem, L., Ashworth, P., Cano, J. (2011). ARGreenet and BasicGreenet: Two mobile games for learning how to recycle, 19th International Conference on Computer Graphics, Visualization and Computer Vision'2011.

Kato, H., Billinghurst, M., Poupyrev, I., Tetsutani, T., Tachibana, K. (2001). Tangible Augmented Reality for Human Computer Interaction. The 17th Nicograph (in japanese), 39-44.

Kerawalla, L., Luckin, R., Seljeflot, S., Woolard, A. (2006). "Making it real": exploring the potential of augmented reality for teaching primary school science. Virtual Real. 10 (3), 163-174

Ketamo, H. (2003a). An Adaptive Geometry Game for Handheld Devices. Educational Technology & Society, 6(1), 83-95.

Ketamo, H. (2003b). xTask-an adaptable learning environment. Journal of Computer Assisted Learning, 19, 360–370.

Klatzky, R. L., Wu, B., Shelton, D., & Stetten, G. (2008). Effectiveness of augmented-reality visualization versus cognitive mediation for learning actions in near space. ACM Transactions on Applied Perception, 5(1), 1-23.

Koh, R.K.C., Duh, H.B.L., Gu, J. (2010). An integrated design flow in user interface and interaction for enhancing mobile AR gaming experiences, IEEE International Symposium On Mixed and Augmented Reality (ISMAR'10), 47 - 52.

Kolb, D. A. (1984). Experiential learning: Experience as a source of learning and development. Englewood Cliffs, NJ: Prentice-Hall.

Liarokapis, F., Newman, R.M. (2007). Design experiences of multimodal mixed reality interfaces. The 25th Annual ACM international Conference on Design of Communication (SIGDOC '07).34-41.

Mahamad, S., Ibrahim, M.N., Taib, S.M. (2010). M-Learning: A New Paradigm of Learning Mathematics in Malaysia. International journal of computer science & information Technology (IJCSIT), 2(4), 76-86.

Mayo, J. (2007). Games for science and engineering education. Commun. ACM 50 (7), 30-35.

Mcconatha, D., Praul, M., Chester, W., & Lynch, M. J. (2008). Mobile learning in higher education: an empirical assessment of a new educational tool. Educational Technology, 7(3), 15-21.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. IEICE Transactions on Information Systems, E77-D(12), 1-15.

Mueller, D., & Strohmeier, S. (2010). Design characteristics of virtual learning environments: an expert study. *International Journal of Training and Development*, 14(3):209–222.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. Psychological Methods, 8(4):434-447.

Pan, Z. (2006). Special issue on edutainment (E-learning and game). Computers & Graphics, 30, 1-2.

Prensky, M. (2001). Digital game-based learning. New York: McGraw-Hill.

Prensky, M. (2003). Digital game-based learning. Comput. Entertain 1 (1), 21-21.

Rapeepisarn, K., Wong, K.W., Fung, C.C., Depickere, A. (2006). Similarities and differences between "learn through play" and "edutainment". In Proceedings of the 3rd Australasian conference on Interactive entertainment (IE '06). Murdoch University, Murdoch University, Australia, Australia, 28-32.

Rigas, D., Ayad, K. (2010). Using edutainment in e-learning application: an empirical study. International Journal of Computers. 1 (4), 36-43.

Sandor, C., Klinker, G. (2005). A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality. Personal Ubiquitous Comput. 9 (3), 169-185.

Savinainen, A., Scott, P. (2002). The force concept inventory: A tool for monitoring student learning. Physics Education, 37(1), 45-52.

Schwabe, G., Göth, C. (2005). Mobile learning with a mobile game: design and motivational effects. Journal of Computer Assisted Learning, 3 (21), 204-216.

Shaffer, D.W., Squire, K. T., Halverson, R., & Gee, J. P. (2004). Video games and the future of learning. Phi Delta Kappan . http://www.academiccolab.org/resources/gappspaper1.pdf.

Sharples, M., Corlett, D., & Westmancott, O. (2002). The design and implementation of a mobile learning resource. Personal and Ubiquitous Computing, 6(3), 220–234

Shelton, B., & Hedley, N. (2002). Using augmented reality for teaching earth-sun relationships to undergraduate geography students. In The First IEEE International Augmented Reality Toolkit Workshop. Darmstadt Germany.

Schönborn, K. J., Bivall, P., & Tibell, L. A. E. (2011). Exploring relationships between students ' interaction and learning with a haptic virtual biomolecular model. Computers & Education, 57(3), 2095-2105.

Shelton, B. E., & Hedley, N. R. (2004). Exploring a cognitive basis for learning spatial relationships with augmented reality. Technology, Instruction, Cognition and Learning, 1(4), 323-357.

Squire, K., Barnett, M., Grant, J.M., Higginbotham, T. (2004). Electromagnetism supercharged!: Learning physics with digital simulation games. Proc. 6th Int. Conf. on Learning Sciences, ACM Press, 513-520.

Tan, K.T.W., Lewis, E.M., Avis, N.J., Withers, P.J. (2008). Using augmented reality to promote an understanding of materials science to school children. In ACM SIGGRAPH ASIA 2008 educators programme (SIGGRAPH Asia '08). ACM, New York, NY, USA, , Article 2 , 8 pages.

Tapscott, D. (1998). Growing up digital. The rise of the net generation. New York: McGraw-Hill.

Taran, C. (2005). Motivation Techniques in eLearning. Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on. 617-619.

Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. Journal of Computer Assisted Learning, 21(3), 217-228.

Uzunboylu, H., Cavus, N., and Ercag, E. (2009). Using mobile learning to increase environmental awareness. Comput. Educ. 52 (2), 381-389.

Van Krevelen, D.W.F., Poelman, R. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. The International Journal of Virtual Reality, 2 (9), 1-20.

Veenema, S., Gardner, H. (1996). Multimedia and Multiple Intelligences. The American Prospect, 7 (29).

Vitzthum. A. (2006). SSIML/AR: A Visual Language for the Abstract Specification of Augmented Reality User Interfaces, IEEE Symposium on 3D User Interfaces, 135- 142.

Zhou, F., Duh, H.B.L., Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR, 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'08), 193-202.

Appendix: Questions used in the evaluations

#	Q1	Q2	Q3	Question
1	x	x		Do you remember what comprises the water? a) Hydrogen and Oxygen b) Potassium and Sodium c) The water is made up of nothing d) I do not know / I do not remember
2	x	x		Of the components that you are going to read, which ones do not belong to water? a) Hydrogen b) Potassium c) Oxygen d) I do not know / I do not remember
3	x	x		Do you remember what helps water to evaporate? a) Cold b) Hot c) Movement d) I do not know / I do not remember
4	x	x		Once water drops are in the clouds, what do they need in order to go back down to the land or to the sea? a) Cold b) Hot c) Movement d) I do not know / I do not remember
5	x	x		Check every object you think is a pollutant. a) Oil can b) Tree c) Food can d) Fuel barrel e) Flower f) Wine bottle g) Water lily h) Deteriorated ball (An image accompanied each object)
6	x	x		Could you tell us how the water cycle is? a) First, water is liquid, then it evaporates, then it cools, and then it precipitates as rain or snow. b) First, it is in the clouds, then it cools and then it becomes liquid. c) First, it falls as snow, then it evaporates and then it goes to the sea. d) I do not know / I do not remember
7		x	x	Did you have fun?
8		x		Did you like that the main character guided you during the game?
9		x	x	Did you like to see how objects appeared on the black squares?
10		x		Did you like playing to the games that use what you found on the black squares?
11		x		What did you like the most? [AR/Tactile/Accelerometer]
12		x		Do you think you have learned new things?
13		x		Would you like to play again to learn about new subjects? [Yes/No/Maybe]

#	Q1	Q2	Q3	Question
14		x	x	Was it easy for you to find objects on the black squares?
15		x	x	What did you find the easiest? [Tactile/Accelerometer]
16		x		How easy did you find the game to play?
17		x		Did you understand the rules of the game?
18		x		Would you like to learn new things at school with this system?
19		x	x	Please, rate the game.
20			x	What did you like the most? [iPhone/Tablet PC]
21			x	Why? (referring to #20)

Table 5. Numbered questions and their appearance in each questionnaire. The first six questions have custom answers. The last two questions are hand-written. Answers in brackets are a summary of the possible choices (categorical data). The rest of answers follow a Likert scale.