Handheld Augmented Reality in education

Universidad Politécnica de Valencia
DSIC

Master thesis
Inteligencia Artificial, Reconocimiento de Formas
e Imagen Digital

Santiago González Gancedo

Supervised by:
M. Carmen Juan Lizandra

July 2012

This thesis has been partially supported by the Spanish APRENDRA project
(TIN2009-14319-C02-01), and by the Ministry of Science and
Innovation (MICINN) of Spain (grant BES-2010-035118).

To Adriana

Contents

1 Introduction **3**
1.1 Motivation . . . 3
1.2 Scientific goals and research hypotheses . . . 5
1.3 Thesis structure . . . 6

2 State of the art **7**
2.1 Introduction to Augmented Reality . . . 7
2.2 AR in education . . . 8
2.3 MRLE . . . 10
2.4 Tactile and tangible interfaces in AR . . . 11

3 *Study 1:* Mixed Reality Learning Environments **13**
3.1 Introduction . . . 13
3.2 The game . . . 14
3.2.1 Game design . . . 14
3.2.2 Description of the game . . . 17
3.3 Materials and methods . . . 19
3.3.1 Hardware . . . 19
3.3.2 Software . . . 22
3.4 Design of the evaluations . . . 22
3.4.1 Participants . . . 22
3.4.2 Measurements . . . 22
3.4.3 Procedure . . . 24
3.5 Results . . . 24
3.5.1 Learning Outcomes . . . 24
3.5.2 Satisfaction Outcomes . . . 26
3.5.3 Interaction . . . 28
3.6 Conclusions . . . 29

4 *Study 2:* Impact of handheld devices in AR for children **31**
4.1 Introduction . . . 31
4.2 Design of the evaluations . . . 32
4.2.1 Participants . . . 32
4.2.2 Measurements . . . 32
4.2.3 Procedure . . . 32
4.3 Results . . . 34

4.4 Conclusions 36

5 *Study 3:* Tactile and tangible interfaces **39**
5.1 Introduction 39
5.2 Experiment 40
5.2.1 Description of the experiment 40
5.2.2 Apparatus 41
5.3 Design of the evaluations 42
5.3.1 Participants 42
5.3.2 Measurements 42
5.3.3 Procedure 43
5.4 Results 44
5.4.1 Time analysis 44
5.4.2 Satisfaction 46
5.4.3 Interaction 47
5.5 Conclusions 48

6 Concluding remarks **51**
6.1 Conclusions 51
6.2 Scientific contributions 52
6.3 Future work 52

Bibliography **54**

Definitions and abbreviations

Throughout this thesis we use a series of terms that have a specific meaning. Next, there is the list of definitions and abbreviations ordered alphabetically.

Augmented Reality (AR): It is a specific type of MR where most of the information is real and virtual objects are coherently located onto the real scene. See Section 2.1.

Edutainment: Refers to the combination of education and entertainment. It aims to engage students in the learning process throughout playful activities.

Handheld AR: Refers to the use of handheld devices to interact with AR applications.

Handheld device: It is typically an electronic device that is portable and is intended to be held only with one or two hands. It can be used for m-learning, entertainment, video games, serious games and many other areas. Two of the most common handheld devices are mobile phones and Tablet PCs, that nowadays usually have tactile screens.

Head-Mounted Display (HMD): Device that allows the rendering of computer generated imagery in a display close to the eyes. It can allow the visualization of AR from the user's perspective.

Human-Computer Interaction (HCI): It is a very active research area where the interaction with computers is studied, and usually involves highly multidisciplinary studies. It also refers to any kind of possible interaction and communication between a machine and a person.

M-learning: Mobile learning. It involves a particular educational model where information is ubiquitous and uses powerful handheld devices to provide rich multimedia experiences.

Mixed Reality (MR): It refers to the synthesis of virtual and real imagery that creates a combined scene of virtual and real information in any kind of proportion. See Section 2.1.

Mixed Reality Learning Environment (MRLE): Refers to that use MR or AR under an educational framework.

Tablet PC: Handheld device that has a relatively large screen around 10" and nowadays typically has a great computational power and the screen is tactile.

Tactile User Interface (TacUI): It is an interface where the interaction is performed using a tactile surface, which usually is also a screen so that the user can push directly on the object being displayed. This term was defined by the authors of this thesis and is not usually found in the literature.

Tangible User Interface (TUI): This interface involves an interaction that is performed through the manipulation of physical objects.

Virtual Reality Learning Environment (VRLE): Refers to that use a fully VR system under an educational framework.

Chapter 1

Introduction

1.1 Motivation

Teachers are continuously trying to understand better the process of education to enhance students' comprehension, and there is a very active body of research in finding better teaching methods using tools that can reach their students at multiple levels (Veenema and Gardner, 1996; Tan et al., 2008). The learning experience is a key point in education, and it has been shown that it can become more meaningful when more senses are involved (Sandor and Klinker, 2005). From the children's perspective, a very exciting way of using their senses is through play, which is an important activity to improve and develop children physically, mentally, socially, and emotionally. Play can be used effectively to make use of all of children's senses to solve problems and to understand their environment in a natural way (Rapeepisarn et al., 2006). Play can help children experience a more meaningful form of learning than other traditional approaches (Gee, 2003) and stimulates them to understand new concepts that would otherwise could be found too difficult to grasp (Blecic et al., 2002; Squire et al., 2004). Playful activities serve as a medium to develop capabilities and abilities through active involvement in an amusing way, and enjoyment is important when endeavoring to achieve learning goals, since enjoyability takes an important role in creating mindful learning (Blecic et al., 2002).

Educational games integrate learning in a playful environment through participative techniques that help children to develop their knowledge and abilities. They also increase motivation towards study and can also be used to reinforce knowledge acquired in the classroom. Psychologists and philosophers have studied the influence of playing games on the learning process concluding that entertainment is an important factor that helps improving learning (Albert and Mori, 2001). *Edutainment* is a term where education and entertainment converge, and it relies heavily on technology like video games (Pan, 2006; Rapeepisarn et al., 2006). As some studies have shown children can benefit significantly from digital educational games, since they can improve their knowledge and skills, and are also stimulated motivated (Fisch, 2005; Rigas and Ayad, 2010; Shelton and Hedley, 2002).

M-learning is a research area that involves a new educational model (Fotouhi-Ghazvini et al., 2011). Even though there is no a consensus on the definition of m-learning, it is a term that usually refers to the use of mobile devices in an educational context (Sharples et al., 2002). The new educational model of 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, and they are usually cheaper than other devices like desktop computers

(Georgiev et al., 2004; Jones and Jo, 2004). Thus, it incorporates services that can be used at any time and place thanks to the great portability of the devices (Fotouhi-Ghazvini et al., 2011). Moreover, mobile technology is also starting to have a widespread use as a platform for games (Facer et al., 2004). Despite these advantages, m-learning also has some downsides such as the different screen sizes and the limited functionalities of the devices, which makes the design of applications more difficult than with other approaches in education and entertainment. Furthermore, teachers may not be able to control students in the same way as they do in a classroom environment (Mahamad et al., 2010). Fotouhi-Ghazvini et al. (2011) classify m-learning systems under two different environments:

- **Fieldwork:** the learner experiences real world situations with the help of mobile device resources to explore it, where contents are designed to be simple and effective.
- **Classroom:** virtual worlds are used to educate and engage learners with multimedia content displayed on mobile devices.

They argued that a m-learning system benefit from the combination of both environments. Several m-learning studies have reported success in some areas using this educational model with which students gained more knowledge with this system than with other traditional models (Thornton and Houser, 2005; Mcconatha et al., 2008). Therefore, taking in consideration all of the the advantages described, mobile devices have a big potential for engaging students at multiple levels when they are in an educational context (Tan et al., 2008).

Augmented Reality (AR) is a technology that makes possible to interact with the world in a very unique way that was never before possible (Billinghurst, 2002). AR combines the real world with synthetic information overlaying multimedia onto it. Unlike *Virtual Reality* (VR), where the user is completely immersed in a different world, AR supplements reality in such a way that virtual imagery is attached to real world objects. A more formal definition of AR can be found in Section 2.1. AR has matured to the point where it can be applied to a wide range of application domains, and education is an area that could specially benefit from this technology (Kaufmann and Schmalstieg, 2003). AR has the potential to highly engage and motivate students to explore the world from a variety of different perspectives (Kerawalla et al., 2006). It has been demonstrated that AR systems usually makes a big impact to those who have experienced it (Ardito et al., 2007; Billinghurst et al., 2001) and that they are useful for teaching subjects that students could not easily experience first-hand in the real world otherwise since the virtual content can be built about a very wide variety of topics (Kerawalla et al., 2006). There exists a large body of research that gather empirical evidence as a basis for theoretical propositions and validation where the results showed that AR was a powerful and engaging visual and cognitive experience for students (Shelton and Hedley, 2004). AR has the interesting ability of offering a smooth transit between reality and virtuality in a seamlessly way, and it also uses a tangible interface metaphor where the user manipulates physical objects that are related to virtual information (Billinghurst, 2002). Furthermore, AR offers mixed interactions between real and virtual environments, which is something that can not be done in other traditional virtual environments. Some works have stressed the visual power of AR that helps students 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" (Shelton and Hedley, 2002).

Handheld devices are becoming tools of widespread use for m-learning since they have recently adopted technologies such as a variety of sensors, communication technologies and powerful processors that allow rich multimedia applications (Economides and Nikolaou, 2008). Recent handheld devices such as smartphones and Tablet PCs are capable of rendering advanced real-time graphics and capture video with an incorporated webcam, allowing AR applications, what is frequently known as *handheld AR*. The study of handheld AR from a *Human-Computer Interaction* (HCI) perspective is very interesting. It is known that AR supports the use of *Tangible User Interfaces* (TUI) (Ullmer and Ishii, 2000), and handheld AR also allows other forms of interaction thanks to the tactile capabilities of the devices, which has not been thoroughly studied yet.

All the mentioned literature has motivated our study, which focuses in pushing forward the HCI aspects of AR applied to education. We make an emphasis in understanding the particularities of working with children and how they deal with this kind of technology that presumably can improve the way they interact with the world.

1.2 Scientific goals and research hypotheses

Following what we have glimpsed in the motivations, the most general objective of the research presented in this document is to **explore the possibilities of AR in a learning environment and the implications in interaction**. To achieve this long-term objective, we have established a series of goals for the realization of this thesis:

- Develop a playful activity that engages and motivates children.
- Design the activity with an educational background.
- Explore the use handheld AR interaction with the children.
- Study the consequences of using different handheld devices in children.
- Study the introduction of AR and VR in learning environments.
- Study the use of tactile and tangible user interfaces with children.
- Use a multimodal interface providing several ways of interaction.
- Test the systems with a minimum number of children.
- Measure the educational outcomes, satisfaction and interaction issues in our experiments.
- Provide a thorough statistical analysis of the results.

In order to accomplish these goals we conducted three studies, each of whom had a particular objective and hypothesis that we want to test.

Explore the advantages of AR and VR systems in education (*Study 1*). The first objective that we considered was to question the need of AR in this context. The advantages of AR and VR have been explored before, but both systems have not been compared in an educational perspective, what would fall under the classification of *Mixed Reality Learning Environment* (MRLE) rather than the traditional *Virtual Reality Learning Environment* (VRLE). From the conclusions of the literature, we designed an experiment to study the introduction of AR in a VRLE. Our initial hypothesis was that a **MRLE would not improve the educational outcomes, but it would increase engagement and motivation.** Since the content of the game we developed was similar in MRLE and VRLE, we did not expect them to learn more, but to increase their satisfaction.

Compare different handheld devices (*Study 2*). There is a wide variety of handheld devices used in m-learning, but those with AR capabilities can be divided in smartphones and Tablet PCs. Providing guidelines and prior experiences for future researches that studied the advantages and disadvantages of each device was one of our objectives. In this thesis we put emphasis in the satisfaction and interaction aspects, and our initial hypothesis was that **the Tablet PC would be more engaging than the smartphone, and the interaction would be easier**. We based this hypothesis in that the bigger screen of the Tablet PC allows the information to be clearer rendered and the movements of the user's hands do not need to be so precise than with the smartphone.

Compare tactile and tangible user interactions (*Study 3*). AR has been extensively studied in relation to TUIs. However, handheld AR brings the possibility of incorporating a TacUI that can be also used as another input channel in educational environments, games and many other areas. We carried out an experiment from an HCI perspective that compared how children manage the TacUI of the device and the TUI we built. Our initial hypothesis was that **the TacUI would be found more easy and satisfactory to use than the TUI**, since we believe *a priori* that TacUIs feel very natural to use in handheld devices.

1.3 Thesis structure

The rest of this thesis is structured as follows.

Chapter 2 shows the state of the art, reviewing the most relevant literature relative to this work in Augmented Reality, education, learning environments and interfaces.

Chapter 3 describes the first study, where a system was built and evaluated with children under different circumstances of learning environments (MRLE and VRLE).

Chapter 4 describes the second study, where the same system was evaluated using two different handheld devices: a smartphone and a Tablet PC.

Chapter 5 describes the third study in which a tactile interface is compared with a tangible interface and the implications are analyzed.

Chapter 6 finalizes the work with the conclusions and future work, and shows the publications derived from this thesis.

Chapter 2

State of the art

2.1 Introduction to Augmented Reality

Augmented Reality (AR) is the combination of real and virtual imagery. Milgram and Kishino (1994) defined the *reality-virtuality continuum* (Figure 2.1), which establishes a gradation between *Virtual Reality* (VR), where the user is completely virtually immersed, and the real environment, where there is no virtual component. *Mixed Reality* (MR) is the most global concept since it includes the full range of this continuum between VR and the real environment. AR systems lay inside MR where most of the information is real and there are some virtual elements included that supplements it. Azuma (1997) defined three characteristics of AR systems:

- Combines real and virtual
- Are interactive in real time
- Are registered in 3-D

These characteristics have the great advantage of defining AR independently from the hardware used to provide it. Thus, we can find a variety of visualization systems that determine different metaphors of *Human-Computer Interaction* (HCI) where users can interact in an AR environment, such as *Head-Mounted Displays* (HMD), projection displays and handheld devices. Although in recent years there have been advances in all the mentioned visualization systems, handheld devices (specifically smartphones and more recently Tablet PCs) are probably what have popularized most the use of AR taking it to the everyday life, since they allow a *video see-through* metaphor without the need of wearing HMD (Azuma et al., 2001), which can be annoying and is usually not recommended for prolonged use.

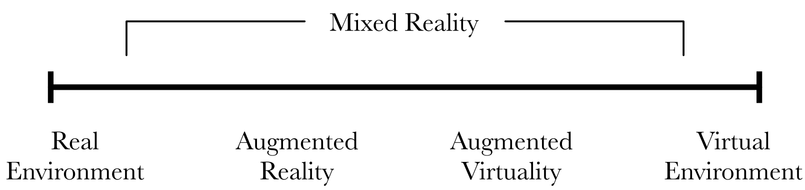

Figure 2.1: Milgram's Reality-Virtuality continuum.

There are plenty of applications for AR, such as geographical systems, archeology, entertainment, medicine, assembly assessment, psychology, education, etc. In the next sections we will focus in the field education, and we will also see some aspects of HCI in AR through handheld devices, tactile interfaces and tangible interfaces.

2.2 AR in education

The research in handheld AR applications has increased highly in the last years thanks to the continuous improvements in the hardware that includes plenty of sensors and the recognition toolkits that have been optimized for the computational power of mobile devices. It has been shown that mobile AR has a wide variety of applications (Van Krevelen and Poelman, 2010) some of which are related to learning and edutainment. Several AR and m-learning studies have been presented in this particular area, but not many studies have analyzed the learning outcomes. Some authors have even signaled a lack of learning theories and research in many of the educational games found in the literature (Shaffer et al., 2004).

In 2003, Ketamo developed two slightly different games for learning geometry comparing the use of a laptop PC and a Compaq iPaq PDA (Ketamo, 2003). Children were asked to find and mark the required polygon. A total of 88 students participated in the study, divided in two stages with three groups of children in the first one who tested the PC application and one group in the second stage where children tested the PDA application. The results from the two stages showed that both games helped specially low-skilled students to reach the level of the average pupils. Ketamo concluded that this results of the experience supported the premise that computer-based teaching could help low-skilled students. More concretely, 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, similarly to the high-skilled group in the first stage. However, 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 help botany learning among elementary school students (Huang et al., 2010). A study was conducted with 32 participants with an average age of 11 years old. They were divided into two groups in order to investigate the effectiveness of the system compared to the traditional guidebook. 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 participants also noted some difficulties with user interface. They had to spend too much time keying and using the UI in order to record their learning experiences. User interfaces are important when designing the application since a badly designed UI could make the application too tedious to play. 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. It could lead to distractions for students and some difficulties for the teachers.

In 2011, Juan et al. (Juan et al., 2011b) 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 knowledge that the children perceived, their behavior and their attitudes towards recycling, as well as 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 significant 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 the children improved their knowledge in the subject thanks to the game.

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 et al., 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 presented a mobile game developed for a PDA to support the orientation days at a university (Schwabe and Göth, 2005). 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. The authors also noted that the interface was not user-friendly, adding unnecessary difficulty to the gameplay.

With regard to AR learning applications developed for PCs and laptop PCs, several works can be cited. 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 et al., 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 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 et al., 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.

2.3 MRLE

Education is a field of research that can benefit extraordinarily from technology. The VRLE is well established as an educational tool (Lee et al., 2010), while AR is not as spread yet. AR could be a strong complement to this traditional approach to education, and learners could enrich their experience with a MRLE (Pan et al., 2006) in the classroom.

Many VRLE applications have been developed. In 2010, Yang, Chen and Jeng developed a video-capture VR system in a classroom environment for English learning (Yang et al., 2010). The system was tested on 60 students, divided in an experimental group and a control group that experienced traditional english learning. The experimental group perceived visual and auditory feedback, and they could interact physically with the virtual environment instead of the traditional avatar. The authors measured the differences in learning and found that both systems achieved a similar level of learning in the immediate posttest, but interestingly another delayed test showed a significantly higher english knowledge in the experimental group. Moreover, the motivation was also measured, and the results showed that the children rated the game higher in almost all items. The authors concluded that the system effectively enhanced student's learning motivation and assisted in English long-term learning.

In 2009, Lee, Wong and Fung developed a VRLE application to study the effectiveness on students, emphasizing the positive academic effects (Lee et al., 2009). The subject to be studied was frog dissection, and they used the VR program *V-Frog*. A total of 431 students between 15 ant 17 years old participated in the study who were divided in the experimental group and the traditional classes group using Power Point slides with the biology teacher. The authors found a significant difference in the academic performance, perceived learning and satisfaction. Students showed a better attitude towards using the VR application. The authors advise that VRLE should not be the panacea, but they affirm that VRLE should become part of the everyday life.

We can compare the previous VRLEs examples chosen previously to the next AR systems used in classes. 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 et al., 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 2008, Freitas and Campos presented SMART, which consisted of a TV-show style learning game that was composed of several racquets with 3D augmented reality markers, a web camera, a PC, and displays such as LCD or projectors (Freitas and 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.

While some researches in the field of education state that VR is "extremely close to reality" (Inoue, 2007, p. 1), such is not still the case of AR. We attribute this to the fact that it is very modern in comparison to VR, and to the few easy to use, low cost, updated AR tools for educators that can be found. Handheld devices offer excellent capabilities to be used for education, and have a great perspective of future (Billinghurst and Henrysson, 2006). In our opinion, Tablet PCs can be an excellent tool that helps AR to have good acceptance as a complement to VRLE. MRLEs should be a natural step towards computer-aided education in the class. Tablet PCs are usually equipped with several sensors and mechanisms for rich Human-Computer Interface. They generally include camera, tactile screen and inertial measure units such as accelerometer and gyroscope. Thus, they can be used in a wide range of education areas, reducing the cost of custom hardware.

2.4 Tactile and tangible interfaces in AR

There exists a large body of research TUIs in AR applied to education education. In 2011, Sayed et al. presented ARSC, a low cost visualization tool of 3D objects that could be used online and offline for different subjects reducing the learning time (Sayed et al., 2011). The authors used desktop computers with web cameras to visualize objects on markers that students could move with their hands as a TUI. They tested the system with students between 10 and 17 years old, and 89% of them were satisfied with it. The ARSC set decreased the expenses and increased the visualization ability of the students.

In 2002, Sharlin et al. presented cognitive cubes, a TUI system that allowed the manipulation of cubes for the assessment of building 3D figures (Sharlin et al., 2002). This system did not work with AR. The users had to physically connect the cubes to form a shape that was sent through a wire to the computer. The system was tested on adults and compared to paper-and-pencil 3D spatial assessment, and it was concluded that for the cubes improved flexibility, reliability, sensitivity of cognitive ability. In what could be considered as an evolution of the cognitive cubes metaphor, Juan et al. (2010) combined AR with tangible cubes, in which each side there was a marker and no physical connections were needed (Juan et al., 2010). The users wore a Head-Mounted Display (HMD) to visualize the AR scenes, freely manipulating the cubes with both hands. The evaluations compared the AR that displayed videos to traditional cubes with images on the sides, on the subject of endangered animals. The authors tested the system with children from 7 to 12 years old, and they concluded that despite the HMD being uncomfortable, children enjoyed more the AR system. This thesis brings interesting conclusions that could be applied to the continuation of the research on cognitive cubes and other similar metaphors that use tangible markers with video see-through visualization.

In 2008, Kim and Maher studied the use of TUIs on spatial cognition, comparing it to a traditional graphical user interface with keyboard and mouse (Kim and Maher, 2008). The experimental study was based on the organization of an interior design office and tested with architecture students. The tangible setup consisted of markers placed on a table representing parts of the office and a vertical monitor with an AR system, while the traditional system was a CAD program. The authors found that TUIs changed users' spatial cognition and affected the design process.

In our research, we have paid special attention to previous work in TacUIs and handheld devices. In 2006, Wagner, Schmalstieg, and Billinghurst presented the collaborative handheld AR game Virtuoso for learning history of art (Wagner et al., 2006). In the game, players had to sort a collection of artworks by date of creation in three different ways: using a paper, a PC and a PDA. The authors did not find significative differences in educational outcomes, but users preferred paper and PDA over the PC to have more working space to collaborate, and they also preferred the AR PDA game over the paper game to have different points of view .

In 2007, Schmalstieg and Wagner (Schmalstieg and Wagner, 2007) presented *Studierstube* as a complete handheld AR framework with a case study where students from 12 to 15 years old used handheld devices to explore historical artifacts in the environment of a museum. The students had to select the items on their screens once they had found them, and some pieces of information and multimedia was displayed. The results were very satisfying, and they were very motivated and wanted to extend the game to other exhibitions.

Chapter 3

Study 1: Mixed Reality Learning Environments

In this chapter, we present a novel study that emphasizes the use of AR as a natural complement for the VRLE model, towards a general acceptance of MRLE in the classroom. Handheld devices help this scheme serving as general purpose computers available for use by other applications. AR has not been explored deeply enough to have full acceptance of use in the classroom. We present an application in which a Tablet PC was used to evaluate our game, working with multimodal interaction provided by a tactile screen and an accelerometer. It can be played in two modes: combining AR and non-AR (NAR), and full NAR. Seventy three children of primary school tested the system. For the learning outcomes, there were no statistically significant differences between both modes, but the AR mode enhanced highly user satisfaction and engagement. This confirms our hypothesis that AR can be an excellent complement to VRLE for the use in the classroom.

3.1 Introduction

As far as we are concerned, there is still too much separation between the fields of AR research and education inside the classroom. We believe that the latter could benefit from the great contributions of handheld AR applications. This study has the objective of highlighting this gap and serving to future researchers as a reference to deepen in this direction.

Our initial hypothesis is that AR can be a good complement to VRLE for educational purposes to improve learning. With this study, we want to prove or reject this hypothesis.

In this research, we propose the case of use of a game. Many computer games have been developed for learning purposes, but very few perform a deep analysis, as several researchers have highlighted (Connolly et al., 2011; Freitas and Campos, 2008; O'Neil et al., 2005). Some researchers have also pointed out the lack a coherent theory of learning and underlying body of research in the development of educational applications (Shaffer et al., 2004).

In this study, we present a handheld game that not only uses AR, but also combines it with non-AR (NAR) parts –including video games– as a case of VRLE. Video games is a subject widely studied previously. For desktop computers, different subjects can be learnt such as volcanoes (Woods et al., 2004), dinosaurs (Bimber et al., 2001), the relation between the earth and the sun (Shelton and Hedley, 2002), mathematics and geometry (Kaufmann, 2004), how to play billiards (Larsen

Figure 3.1: Child playing one of the AR games, catching the main character water drop.

et al., 2005), organic chemistry (Fjeld et al., 2007), or endangered animals (Juan et al., 2011a). For handheld devices, several educational AR applications have also been presented. For example, for learning heritage temples (Wang et al., 2009), math and literacy skills (O'Shea et al., 2009), or how to recycle (Juan et al., 2011b).

The innovate aspect of this study relies on the research of a scenario with practical usability in the classroom. We emphasize the convergence between AR and VRMLE to form a MRLE, as a very suitable tool for educators that can improve the outcomes of VRLE and support meaningful learning. We also believe that Tablet PCs are very appropriate to put MRLE into practice, as they are affordable, allow AR applications and provide multimodal interaction.

3.2 The game

To study MRLEs we decided to design and build a game that incorporates AR and VR in a single game. In this section we explain the design principles and educational background that we incorporated to the game and a description of its functionalities and phases.

3.2.1 Game design

The game that we have developed is themed on the water cycle. This subject was chosen given a former study in which professionals in education were consulted to determine the subject preferences for educational computer games and their type for children. From the survey, we found that "Nature" was one of the most preferred subjects. Furthermore, we chose the water cycle theme because this topic is covered in the author's country primary education law. Some works have pointed out the importance of considering national curricula to develop educational computer games (De Freitas and Oliver, 2006; Lai-Chong Law et al., 2008). We have taken into account the national curricula for our game, as stated by the national primary education law in the Royal Decree 2211/2007, on July the

12th. In the first section *Contents* of second cycle of primary education, it establishes: "The water cycle. Explanation of the water existence in different states and how it can change from one state to another through heating or cooling" (BOE, 2007, p. 31501). And in the 6th section, it establishes: "Pollutants production, contamination and its environmental impact" (BOE, 2007, p. 31502).

We followed some of the design characteristics that were identified for virtual learning environments by Mueller and Strohmeier (2010), which referred to virtual environments, but can also be applied to AR systems. One of the highest rated characteristics was "Interaction", so 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 and Newman, 2007). Therefore we used accelerometers, a tactile screen, and tangible interaction. The accelerometers and the tactile screen were used to move characters in the screen or to pick objects during the game in order to complete some tasks. There was tangible interaction because markers could be rotated and translated with the hands.

Other design principles and suggestions were also followed in order to enhance interaction (Koh et al., 2010), and it has also been recommended mixing several input and output channels (Sandor and Klinker, 2005). In our game had camera, accelerometers, and a tactile screen as input channels, and videos, sounds, and graphics as output channels. Henrysson and Billinghurst considered tracking real objects using a camera to have a 6 Degrees of Freedom input (Henrysson and Billinghurst, 2007). We used the camera of the device 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 et al., 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 et al., 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 since every action of the children had a reaction in the game. We only used wireless mobile devices, so children were not annoyed by any other instrument.

The characteristic of "Learning-process supportive" was also highly rated in Mueller and Strohmeier's study (Mueller and Strohmeier, 2010). As some studies have pointed out, education cannot be improved with only having technology (Fisch, 2005; Veenema and Gardner, 1996), but technology can include a variety of media that helps understanding and learning concepts (Veenema and Gardner, 1996). In our game, for example, 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).

The educational context in which the design of our game is supported consists of two learning theories, which are described next.

Gardner's Multiple Intelligences

Gardner's theory of Multiple Intelligences (Gardner, 1983) has become a catalyst and a framework for many current educational strategies. According to Gardner, intelligence is not a unitary element, but it includes different and specific ways of learning and processing information. Kolb described eight types of intelligence through which individuals approach problems and develop solutions. According to the Multiple Intelligences theory, a person has at least eight forms of intelligence (linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, naturalistic, interpersonal, and intrapersonal), and a correct use of technology can help their development (Gardner, 2000). Next we detail the activities proposed in our game for seven of these forms of intelligences, describing the theoretical meaning of the intelligence and how it has been incorporated to our game.

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.
In our game, children have to hear and, in some cases, read the directions that the guide character (a drop of water) 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.
In our game, seven different problems arise that the children have to solve using logical thinking in order to progress in the game. The proposal to form water molecules in their scientific formulation stands out from the rest: H2O. The children have to identify the formula for water and take the atoms in order (two hydrogen atoms and one oxygen atom).

Visual-spatial. It includes the ability to perceive and represent the visual-spatial world accurately and to form and manipulate mental images.
We propose solving spatial problems in our game 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. It consists of understanding the natural world.
One of the learning objectives in our game 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 form of intelligence develops individual and personal knowledge, identity construction and self-esteem.
In our game, 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.

Musical. It 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. It 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).

Kolb's Experiential Learning Theory

For the design of the game, the experiential learning theory of Kolb was used, which states that "learning is the process whereby knowledge is created through the transformation of experience" (Kolb, 1984, p. 38). As shown in Figure 3.2, the experiential learning consists of four phases: a

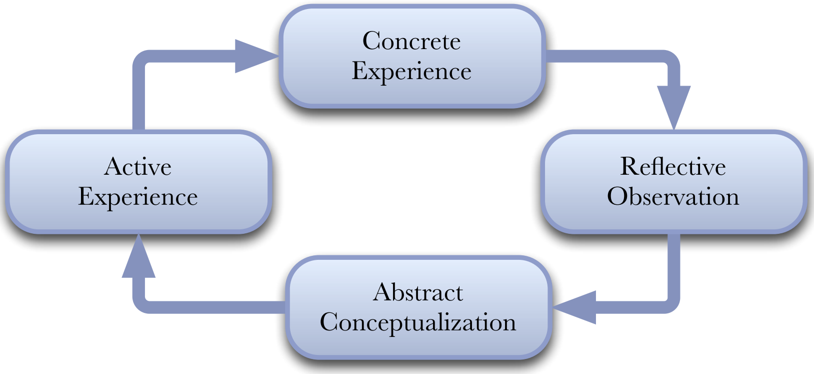

Figure 3.2: The wheel of Kolb's learning.

concrete experience (do), a reflective observation (observe), an abstract conceptualization (think), and an active experience (plan or testing in new situations).

In Kolb's theory, the wheel that forms the four phases can be started from any point, which can lead to different learning processes. 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, and the children can use what they learned in their daily life.

As an example of the application of this theory to our game, the player assumes the mission of completing the cycle of water and she has to perform different activities. The player has the concrete experience of collecting suns to raise the temperature, with a reflecting observation with the feedback of the game, and creating an abstract conceptualization as a result of receiving all the information, in this case, evaporation. And what is very important, the student has an active experience using the game.

3.2.2 Description of the game

The aim of the game was to reinforce the learning of children in the subject of water, including its composition, the water cycle and water pollution. These topics were shown in the same way they had been studied at school. They saw the major processes of the water cycle. They started with the evaporation phase, when the sun heated up water and turned 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 formed 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 the water composition and about water pollutants.

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

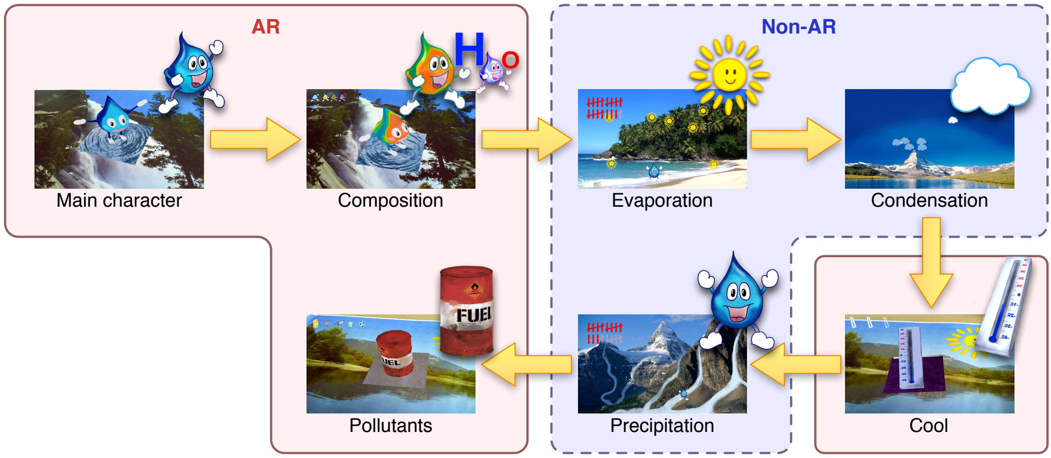

Figure 3.3: Flowchart of the AR and non-AR mini-games.

the room for a more immersive experience. When the children selected an object, a message was displayed telling them whether the object picked up was the right one. Figure 3.5 shows a child 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 3.3. 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 3.4). 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

Figure 3.4: Child playing the mini-game of the rain.

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.

3.3 Materials and methods

For the evaluation of our study, we needed to build both custom hardware and software to provide a multimodal interface. This section explains the hardware built and the software developed.

3.3.1 Hardware

For the evaluations we used the handheld device *HP Slate 500 Tablet PC*. The dimensions of the device were 23x15x1.5 cm, and it had a weight of 0.68 kg. It used an Intel Atom Processor Z540 working at 1.86 GHz, and a 2 GB DDR2 SDRAM memory cell. For the graphics, it operated with a built-in GPU Intel Graphics Media Accelerator 500. Among its interfaces we could find a 8.9" capacitive touch-screen and an outward facing 3 Megapixel camera capturing at 30 fps.

The HP Slate 500 had a pre-installed 32 bit version of Windows 7 Professional. There were severe performance issues with the graphic controllers used for the integrated GPU. In our early tests running the basic OpenSceneGraph application *osgviewer* with a model consisting in about 8000 polygons, the frame rate was about 5 *frames per second* (fps). After some web research, we found that some low power devices using Windows suffer from this problem. It is due to the inadequate adaptation of the default drivers to the hardware, which in our case was the combination of Itel GMA

Figure 3.5: Child starting the mini-game of the evaporation.

500 graphics card and the Z5xx processor with chipset US15W. We found reports of solving these problems replacing the original drivers by a more specific version from Intel Embedded Graphics Drivers: IEGD 10.4[1]. These drivers were developed for Windows XP (SP3), but worked correctly on Windows 7 with some minor malfunctions. This upgrade made our graphic tests on the device work properly, which afterwards run about 50 fps.

To provide a multimodal interaction we decided to use accelerometers to be able to provide an interaction based on tilting the device. It is well known that the accelerometers are devices that produce a measure that is proportional to the tilt when they are static, correlated to the cosine of the angle they form with the vector of the gravity of the Earth. It is not possible, therefore, to correctly measure precisely the tilt of the device with accelerometers when the device is itself moving and accelerating, since the gravitational component would not be possible to isolate from the final component measured. A solution would be to use gyroscopes and fuse it to the accelerometer signal. However, four our purposes using only the accelerometer provides a sufficient approximation, since the interactions we have designed for the game do not require strong movements of the device.

This Tablet PC was one of the very first modern and powerful devices of its generation. It was advertised that it incorporated accelerometers, but it was only available for HP built-in applications, and the architecture was completely closed since no information or drivers were provided. HP did not even use the standard way of reporting sensors, using Windows Sensors and Services, so there was not any way to use the accelerometers from our code. Posterior devices typically provide adequate mechanisms for developers to use any kind of sensors. Nowadays there are several alternatives that are known not to suffer from this problem, such as the iPad 2 –that now incorporates a camera– or the Samsung Galaxy Tab 10.1, which runs Android OS. Therefore it was necessary to install external hardware.

[1] http://edc.intel.com/Software/Downloads/IEGD/

Figure 3.6: Internal view of the case built, with the Tablet PC, the USB accelerometer and the neck ribbon.

The accelerometer board *1056 - PhidgetSpatial 3/3/3* from *Phidgets* [2] was used to measure the acceleration accurately in two axes. It was connected via USB to the Tablet PC. The dimensions were 36x31x6 mm, which made it to be easily enclosed with the Tablet PC. The need for this device was due to technical problems with the built-in accelerometer, for which HP did not provide any usable driver. At the time we bought this device, it was one of the very few available in the market. Nowadays there are several alternatives that are known not to suffer from this problem, such as the iPad 2 –that now incorporates a camera– or the Samsung Galaxy Tab 10.1.

Both the Tablet PC and the accelerometer board were protected by a case built *ad-hoc*, shown in Figure 3.6. It was designed in 3D CAD and created using Selective Laser Sintering. Several primer layers were applied before the final painting to obtain the prototype. The case protects all the devices and also places the accelerometer in a fixed position with the axes aligned to the Tablet PC orientation. The case was adapted to hold it with both hands in landscape orientation, and a neck ribbon was fastened at the bottom to have a third pivot point apart from the hands. The ribbon is attached to the case at two points to increase stability when holding the device, which was specially important for children. We had two identical devices available to use: one blue and other white.

[2] http://www.phidgets.com/

3.3.2 Software

To develop the system we decided to use OpenSceneGraph (OSG) toolkit[3] version 2.9.5 to use its high capabilities to import, animate and render 3D objects with high performance in C++ language. This allowed us to build the 3D models in Autodesk 3ds Max and Blender, export them to *3ds* format, and load them in our program. Sound support was provided by the FMOD sound library, and the videos were rendered using with the FFmpeg plugin.

The registration was made using the OSG plugin osgART 2.0 RC 3, which used the ARToolKit library (Kato and Billinghurst, 1999) version 2.72.1. This plugin provided simple access to the camera, and to certain OSG nodes which applied the corresponding transformation matrices associated to markers when they were recognized.

Phidgets provided a free multiplatform library with an excellent API to access data from multiple programming languages, including C/C++, which we used. Accelerometer measures could be easily taken in real time.

3.4 Design of the evaluations

The developed game described in the previous section was extensively played by a group of children who tested all the possibilities offered. This section explains in detail the participants, the measurements and the procedure carried out during the evaluations.

3.4.1 Participants

A total of 73 children from 8 to 10 years old –with a mean age of 9.07 ± 0.65– took part in the study: 37 boys (50.68%), and 36 girls (49.31%). They were attending the Escola d'Estiu (Summer School) at the Polytechnical University of Valencia.

3.4.2 Measurements

Five questionnaires were used for the validation. The first one was the pretest, and the other four were the combinations of playing AR or NAR, for the first or the second time. Each child had to fill in only three of them, as shown in Figure 3.7. Table 3.1 shows the questionnaires and the questions in them.

The pretest (QPre) was composed of six questions designed to evaluate the level of children's knowledge and remembrance about water (composition, cycle and pollutants) from school lessons. These six questions about acquired knowledge were agreed upon by a team of 15 teachers from 3rd grade 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. One example of question was: *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 don't know / I don't remember*. Other questions followed the same structure of multiple-choice answer with the exception of question 5, which was: *Mark all the objects you think are pollutants*. Participants had to check every object they believed to be pollutant from a list of nine objects. To make it more friendly to the children, an image was presented with every object. For the data analysis, this question was considered correct if all the objects were correctly checked or unchecked. Besides, the questionnaire collected gender, age and the course they had finished.

[3] http://www.openscenegraph.org/

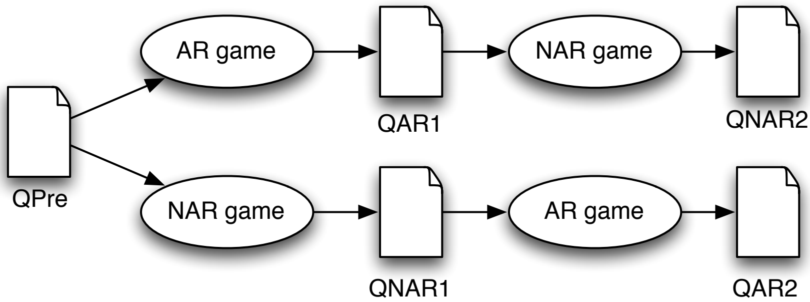

Figure 3.7: Scheme of the games and questionnaires where each children followed one path.

#	QPre	QAR1	QNAR1	QAR2	QNAR2	Question
1	•	•	•			Do you remember what comprises water?
2	•	•	•			Of the components that you are going to read which ones do not belong to water?
3	•	•	•			Do you remember what helps water to evaporate?
4	•	•	•			Once water drops are in the clouds, what do they need to go down to the land or to the sea?
5	•	•	•			Check every object you think is pollutant.
6	•	•	•			Could you tell us how the water cycle is?
7		•	•	•	•	Did you have fun?
8		•	•			Did you like that the guide character guided you during the game?
9		•		•		Did you like to see how objects appeared on the black squares?
10		•		•		Did you like playing to games that use what you found on the black squares?
11		•		•		What did you like the most? [AR/Tactile/Accelerometer]
12		•	•			Do you think you have learned new things?
13		•	•			Would you like to play again to learn about new subjects? [Yes/No/Maybe]
14		•		•		Did you find it easy to find objects on the black squares?
15		•	•			What did you found the easiest? [Tactile/Accelerometer]
16		•	•	•	•	How easy to play did you find the game?
17		•	•			Did you understand the rules of the game?
18		•	•			Would you like to learn new things at school with this system?
19		•	•	•	•	Please, rate the game.
20				•	•	What did you like the most? [AR/NAR]
21				•	•	Why? (referring to #20)

Table 3.1: 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.

The questionnaires QAR1 and QNAR1 were similar. Both of them contained the knowledge questions from QPre to be able to compare learning before and after the AR and NAR games. The rest of the questions were about participants satisfaction with the game. Some of them followed a Likert scale, presented as: *a) Very much; b) Quite a lot; c) Somewhat; d) Few; e) Nothing*, with a numerical equivalency linearly ranged from 5-a) to 1-e). Other questions did not have a possible numerical representation, as they had to choose among technologies (e.g. *What did you like the most? a) Searching objects with the camera; b) Games that used the tactile screen; c) Games that used the accelerometer*).

The questionnaires QAR2 and QNAR2 were designed to compare AR and NAR games preferences. QAR2 shared the AR specific questions from QAR1. The last question was subjective and children could express their feelings about the game.

3.4.3 Procedure

The children who participated in the experience were randomly assigned to one of two situations:

Group A) Children who played the AR game first and then the NAR game.

Group B) Children who played the NAR game first and then the AR game.

Both groups were counterbalanced: 38 children were assigned to group *A*, and 35 to group *B*.

Before playing any game, every children filled in the QPre questionnaire. Next, the first group of children played the AR game. After completing the game, they filled in the QAR1 questionnaire. Then, these children played the NAR game and filled in QNAR2 when finished playing. The second group, instead, played the NAR game first. After completing the game, they filled in QNAR1 questionnaire. Then, these children played the AR game and filled in QAR2 when they had finished. Two children, one of each group, could play simultaneously, since we had two equal devices available.

The questionnaires were filled in in the same room where the activities took place. Two people were in the activities room to clarify doubts from the participants while they were playing the game.

In order to make the experience more immersive, the room was decorated with wall posters and images according to the subject. For the AR game, the play area had water related prints where the markers were placed. These prints were carefully chosen to avoid possible false positives from the AR marker recognizer.

3.5 Results

After the evaluation with the children, all data was transcribed to electronic format and analyzed with the statistical open source toolkit *R*.

3.5.1 Learning Outcomes

Several t-tests were performed to check if there were significant differences in the degree of acquired knowledge during the game. The test was performed over the *knowledge* variable, which condenses all knowledge questions (number 1 to 6 of QPre in Table 3.1) counting the number of correct answers. This variable was compared among the questionnaires QPre, QAR1 and QNAR1. All t-tests in the text are showed in the format: *(statistic[degrees of freedom], p-value)*, and ** indicates statistical significance at level $\alpha = 0.05$.

#	Pretest	AR 1st	t	p
1	0.49 ± 0.50	0.97 ± 0.17	5.67	<0.001**
2	0.74 ± 0.44	0.97 ± 0.17	3.17	0.003**
3	0.94 ± 0.23	0.94 ± 0.23	0	1
4	0.71 ± 0.45	0.97 ± 0.17	3.43	0.002**
5	0.40 ± 0.49	1.00 ± 0.00	7.14	<0.001**
6	0.89 ± 0.32	0.97 ± 0.17	1.36	0.18

Table 3.2: Means and standard deviations of questions of the Pretest and AR game played first, and t-test analysis. d.f.=34.

#	Pretest	NAR 1st	t	p
1	0.34 ± 0.47	1.00 ± 0.00	8.44	<0.001**
2	0.50 ± 0.50	0.95 ± 0.22	5.47	<0.001**
3	0.92 ± 0.27	0.97 ± 0.16	1.00	0.32
4	0.61 ± 0.49	0.95 ± 0.22	3.95	<0.001**
5	0.50 ± 0.50	0.97 ± 0.16	5.77	<0.001**
6	0.84 ± 0.36	0.89 ± 0.31	0.81	0.42

Table 3.3: Means and standard deviations of questions of the Pretest and NAR game played first, and t-test analysis. d.f.=37.

From a paired t-test we obtained that the ratings of the knowledge in QPre (mean 4.17 ± 1.21) were significantly different from QAR1 (mean 5.83 ± 0.45) ($t[34] = -8.09, p < 0.001^{**}$). Another paired t-test revealed that the ratings of the knowledge in QPre (mean 3.71 ± 1.25) were also significantly different from QNAR1 (mean 5.74 ± 1.25) ($t[37] = -9.28, p < 0.001^{**}$). Moreover, a third unpaired t-test shows that there were no differences between knowledge in QAR1 and QAR2 ($t[71] = -0.82, p = 0.42$). These results can be interpreted as that there was a significant amount of knowledge acquired in both modes, and both achieved a similar amount. Besides, the subjective thoughts of children about what they had learnt followed a similar structure. After playing AR and NAR for the first time, the scores of the perceived acquired knowledge were high, and both groups did not have statistically diferences, as revealed from comparing question 12 in Table 3.5.

We can see the detailed paired t-test for each question between QPre and QAR1 (Table 3.2), between QPre and QNAR1 (Table 3.3), and between QAR1 and QNAR1 (Table 3.4). In both cases, questions 3 and 6 were the only ones that did not have statistical differences. They were probably found easier than the rest by the children.

The factor of school year has also been studied. There were no evidences of significant differences on acquired knowledge between 3rd course students (mean 5.87 ± 0.50) and 4th course students (mean 5.80 ± 0.40) ($t[33] = 0.43, p = 0.67$). In addition, there were no findings of statistical differences in the factor of gender. Males (mean 5.78 ± 0.53) and females (mean 5.88 ± 0.32) had very similar scores ($t[33] = -0.68, p = 0.50$).

The motivation towards learning using the device seemed to be very high, as reveals the analysis of question 13. After playing the AR game first, there was a 91.4% of the children who would like to play again with a different subject, a 8.6% who had doubts, and no one who would not. These statistics remain similar after playing the NAR game first, where the 92.1% of them would like to

#	AR 1^{st}	NAR 1^{st}	t	p
1	0.97 ± 0.17	1.00 ± 0.00	1.04	0.30
2	0.97 ± 0.17	0.95 ± 0.22	0.51	0.61
3	0.94 ± 0.23	0.97 ± 0.16	0.66	0.51
4	0.97 ± 0.17	0.95 ± 0.22	0.51	0.61
5	1.00 ± 0.00	0.97 ± 0.16	0.96	0.34
6	0.97 ± 0.17	0.89 ± 0.31	1.29	0.20

Table 3.4: Means and standard deviations of questions of the AR and NAR games played in the first place, and t-test analysis. d.f.=71.

#	AR 1^{st}	NAR 1^{st}	t	p
7	4.86 ± 0.35	4.84 ± 0.36	0.18	0.86
8	4.66 ± 0.47	4.66 ± 0.53	0.01	0.99
12	4.09 ± 1.00	4.34 ± 0.98	1.09	0.28
16	4.43 ± 0.65	4.39 ± 0.81	0.19	0.85
17	4.77 ± 0.42	4.68 ± 0.52	0.78	0.44
18	4.89 ± 0.40	4.92 ± 0.27	0.44	0.66
19	4.76 ± 0.60	4.75 ± 0.39	0.17	0.86

Table 3.5: Means and standard deviations of questions of amusement and satisfaction of the AR and NAR games played in the first place, and t-test analysis. d.f.=71.

play again, and the 7.9% were not sure.

3.5.2 Satisfaction Outcomes

In order to determine if the experiment influenced participants in regard to the level of amusement experienced and satisfaction between the two modes of game, the variable *satisfaction* was created. It condensed all information in the concerning questions, giving each one of them the same weight. An unpaired t-test showed that the satisfaction ratings in QAR1 (mean 4.73 ± 0.47) were not statistically different from QNAR1 (mean 4.66 ± 0.63) ($t[71] = 0.28, p = 0.78$). The detailed results of each question are shown in Table 3.5, where several paired t-tests were performed. There were no significant differences in any isolated question; the scores were very high for both games, from what we can say they enjoyed the experience in both cases and were engaged to learn in a similar manner.

In order to determine the effect of the order in which both game modes was played, several tests were conducted. Table 3.6 shows the unpaired t-tests for AR games played the first and the second time. These shared questions measure the general satisfaction with the game –questions 7, 16 and 19–, and also the specific thoughts about AR –questions 9, 10 and 14–.

AR had excellent acceptance, and children enjoyed it very much. We can see the high values of question 9, asking if they liked this technology, and question 10, asking if they liked the games that used AR. From the notes of our staff that was with the children during the evaluations, we obtained very good impressions, specially in the AR part, which was more encouraging to be a dynamic activity.

Table 3.7 shows the unpaired t-tests for NAR games played the first and the second time with

#	AR 1^{st}	AR 2^{nd}	t	p
7	4.86 ± 0.35	4.82 ± 0.39	0.47	0.64
9	4.86 ± 0.35	4.82 ± 0.39	0.47	0.64
10	4.77 ± 0.48	4.89 ± 0.31	1.29	0.20
14	3.86 ± 0.72	4.00 ± 1.00	0.68	0.50
16	4.43 ± 0.65	4.50 ± 0.94	0.37	0.71
19	4.76 ± 0.30	4.78 ± 0.32	0.25	0.80

Table 3.6: Means and standard deviations of questions of the AR game played first and NAR in second place, and t-test analysis. d.f.=71.

#	NAR 1^{st}	NAR 2^{nd}	t	p
7	4.84 ± 0.36	4.86 ± 0.35	0.18	0.86
16	4.39 ± 0.81	4.51 ± 0.55	0.72	0.47
19	9.50 ± 0.78	9.41 ± 0.83	0.45	0.66

Table 3.7: Means and standard deviations of questions of the NAR game played first and NAR in second place, and t-test analysis. d.f.=71.

questions about satisfaction. The results showed no significant differences between both game modes. From this data, we can deduce that the order of playing did not significantly affect the scores of the games.

The preferences for AR or NAR were measured in question 20 (Figure 3.8). After playing in AR mode first and then NAR, 65.7% of the children chose AR over NAR, while 94.7% of them chose the same when playing AR in second place. The augment in the proportions was statistically significant ($\chi^2[1,73] = 9.90, p = 0.002$), but the difference of proportions between AR and NAR in the worst case (AR first) was also significant ($\chi^2[1,73] = 6.91, p = 0.008$), which means that AR was preferred even if we do not take into account the effects of the order of the games. Relating this to results from question 11, we can see that a trend in which AR was preferred over other alternatives, and this is emphasized when it is the last technology to be used.

Among the reasons given in question 21 to prefer AR over NAR, we could find:

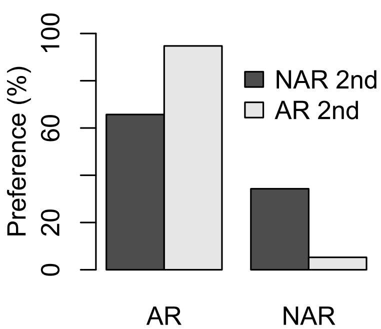

Figure 3.8: Preferences after the last game played (question 20).

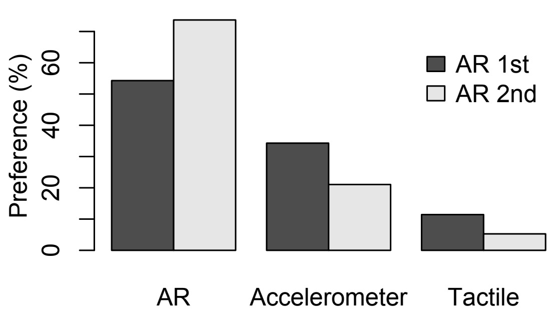

Figure 3.9: Preferences after the AR game played the first and the second time (question 11).

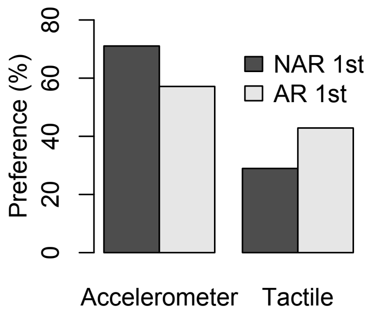

Figure 3.10: Preferences after the first game played (question 15).

- "I can see things of the game in the real world."
- "Because you learn a lot."
- "Because it was more entertaining and fun."
- "You had to move and you were not seated all the time."
- "Because it was easier."

The reasons given in question 21 to prefer NAR can be summarized in:

- "Because I like to play seated."
- "Because it was more comfortable."

3.5.3 Interaction

The preferences of the different types of interaction were measured in question 11, considering the accelerometer games, tactile games and AR games using the device in a video see-through configuration. The question was asked after playing games in AR mode (results in Figure 3.9). When AR was played in first place, the percentages of preferred technology was: AR (54.3%), accelerometer (34.3%) and tactile (11.4%), but when AR was played the last, the order was: AR (73.7%), accelerometer (21.1%) and tactile (5.3%). AR was always the favorite, and it raised when it was played in last place. The proportions of AR in both cases were significantly different at level $\alpha = 0.1$ ($\chi^2[1,73] = 2.99, p = 0.08$). We believe this difference is due to the big impact AR caused to the children. As a consequence, accelerometer and tactile proportions dropped similarly without significant differences.

Question 15 asked about how easy it was to use the tactile screen and the accelerometer after each game the first time they were played (Figure 3.10). On one hand, the accelerometer was the most appreciated, and was higher NAR (71.1%) than in AR (57.1%). On the other hand, the tactile screen raised from NAR (28.9%) to AR (42.8%). In the NAR mode the proportions between both interaction methods were statistically significant at level $\alpha = 0.01$ ($\chi^2[1,73] = 13.47, p < 0.001$), but in AR mode, no statistical differences were found ($\chi^2[1,73] = 1.43, p = 0.23$) and the differences may be due to chance. Therefore we can deduce that the use of AR stimulated the use of tactile interaction to the point that it was not significantly different from the accelerometer interaction.

3.6 Conclusions

In this work we have presented a study that emphasizes the use of AR as a natural complement for VRLE, which is MRLE. Educational AR researches typically focuses in Human-Computer Interfaces and usability, and the applications are frequently intended for museums. To our knowledge, AR has not yet been seriously studied as a complement to VRLE to be used in the classroom for a long-term use. We believe that MRLEs are very suitable for the classroom and we encourage future educational researches to take it into consideration.

Handheld devices, and Tablet PCs in particular are an excellent tool for MRLEs. They commonly provide all the sensors needed to build interactive applications and multimodal interfaces, such as camera, tactile screen and inertial measure units. This makes them to be an exceptional tool that can be used for a wide spectrum of interactive educational applications with more possibilities and versatility than standard desktop applications. Moreover, Tablet PCs are usually comparable in prize to desktop computers, both being low cost solutions.

We developed a MRLE application for primary school children, which consisted of a game about water that included multiple interaction forms (tactile screen and accelerometer). It could be played in a combined mode with AR and NAR mini-games, or in full NAR mode.

After playing to the game, children's knowledge was statistically higher than in the pretest, but no significant differences were found between AR and NAR modes. However, the AR mode enhanced user satisfaction and engagement highly. This confirms our initial hypothesis that AR is an excellent complement for VRLE and that it can improve some of its outcomes. In our game, this combination of AR and NAR games throughout the story thread and the links between them like the main character was very appreciated.

Playing in AR mode caused very good impression to the children, who improved motivation and were encouraged to be more dynamic during the activity, not being so perceivable in the NAR mode. We could see the big impact that AR caused, as the underlying interaction –AR in a handheld device as a video see-through– was significantly more preferred than accelerometer and tactile screen, specially when it was played in second place.

There was a significant preference for the accelerometer interaction over the tactile interaction in NAR mode. However, the introduction of AR encouraged tactile interaction, and the difference was no longer significant.

During the evaluations we observed that a percentage of the children tended to touch the capacitive screen with their fingernails, causing an apparent malfunction that would not have happened with a resistive screen. Although capacitive screens have better acceptance in the general public than resistive screens, the latter should not be trivially discarded when developing applications for education purposes with primary school children.

With regard to future work, we believe that more research in this direction is needed. MRLE in the field of education is still in an early stage. It could be used for many subjects, including Natural Science, Mathematics, History, Technology and outdoor activities. Furthermore, more engaging games and serious applications that use different input channels (AR marker tracking, tactile screen, accelerometer, etc.) could be developed with current handheld devices.

Chapter 4

Study 2: Impact of handheld devices in AR for children

In this chapter, we present an AR game for a smartphone (iPhone) and a Tablet PC, designed to reinforce children's knowledge about the water cycle. The game included different interaction forms like a tactile screen and accelerometers, and during the gameplay AR mini-games were combined with non-AR mini-games for better immersion. We present a study from a HCI perspective to determine the differences in satisfaction and interaction caused by the use of the two handheld devices. Seventy-nine children from 8 to 10 years old participated in the study. The two devices were enjoyed and the children were very motivated. However, despite that our initial hypothesis was that the Tablet PC would be more appreciated, from the results we observed that the different characteristics (screen size and weight) of the devices did not influence significantly children's engagement and satisfaction. We analyze the reasons of these results and the conclusions from this experiment.

4.1 Introduction

In this study, we present a mobile AR game for learning about the water cycle, water composition, and water pollution. The content of the game is the same as in the first study, described in Section 3.2, but contrary to that study where we compared two versions of the game, we are now only interested in using the original version that contains a mixture of AR and VR mini-games. In this case, the game was compared using two handheld devices with different characteristics: a smartphone (iPhone) and the Tablet PC used in the first study, described in Section 3.3.1. 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 usually consisted of children having to collect a certain number of objects using the sensors in the device to get to the next screen. Our game combined AR mini-games with non-AR mini-games for better immersion. Furthermore, the game combined different forms of interaction, including the tactile screen and the accelerometer in the multimodal interface, as it has been suggested by some authors (Liarokapis and Newman, 2007).

The main objective of this study is 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

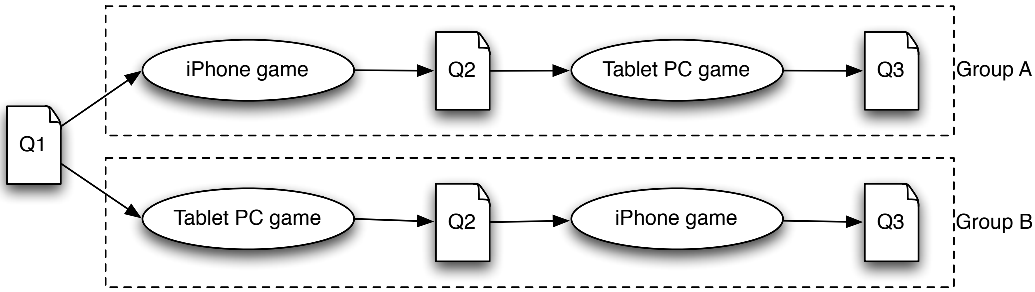

Figure 4.1: Scheme of the games and questionnaires where each children followed one path.

PC device regarding acquired knowledge, satisfaction, and interaction, and that children would prefer the Tablet PC mainly due to its larger screen size. Based on the developments and studies mentioned in the state of the art, in Section 2.2, to our knowledge our work is the first one that compares two different mobile devices with different characteristics (screen size and weight) to see its impact on children's knowledge reinforcement, satisfaction and interaction.

4.2 Design of the evaluations

4.2.1 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 a summer school They were attending the Escola d'Estiu (Summer School) at the Polytechnical University of Valencia..

4.2.2 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 as shown in Figure 4.1. Table 4.1 shows the relation of questions for each questionnaire. The questions are very similar than the used in the first study, and we refer the reader to Section 3.4.2, for a detailed explanation. The differences in this study are the distribution of questionnaires.

The pretest (Q1) was composed of six 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)). The second questionnaire (Q2) was composed of 19 questions. The first six questions were the same ones the participants answered in Q1 to be able to compare them. The rest of the questions were about participant satisfaction and interaction with the game. The third questionnaire (Q3) was composed of eight 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.

4.2.3 Procedure

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

#	Q1	Q2	Q3	Question
1	•	•		Do you remember what comprises water?
2	•	•		Of the components that you are going to read which ones do not belong to water?
3	•	•		Do you remember what helps water to evaporate?
4	•	•		Once water drops are in the clouds, what do they need to go down to the land or to the sea?
5	•	•		Check every object you think is pollutant.
6	•	•		Could you tell us how the water cycle is?
7		•	•	Did you have fun?
8		•		Did you like that the guide character guided you during the game?
9		•	•	Did you like to see how objects appeared on the black squares?
10		•		Did you like playing to games that use what you found on the black squares?
11		•		What did you like the most? [AR/Tactile/Accelerometer]
12		•		Do you think you have learned new things?
13		•		Would you like to play again to learn about new subjects? [Yes/No/Maybe]
14		•	•	Did you find it easy to find objects on the black squares?
15		•	•	What did you found the easiest? [Tactile/Accelerometer]
16		•		How easy to play did you find the game?
17		•		Did you understand the rules of the game?
18		•		Would you like to learn new things at school with this system?
19		•	•	Please, rate the game.
20			•	What did you like the most? [AR/NAR]
21			•	Why? (referring to #20)

Table 4.1: 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.

Group A) Children who played the iPhone game first and then the Tablet PC game.

Group B) Children who 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 4.1 shows the procedure of both groups graphically. As can be observed, before playing any game, each child filled out the pretest Q1 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 posttest 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, and 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 the classroom of the second cycle of the primary education in the country of the authors of this work, the usual established dynamic is to teach a subject and then evaluate the level of learning of the content.

Two mirrored rooms were used for the evaluations. Each room had two identical playing areas where two children (one of each group) could play simultaneously but individually. There was no interaction between them, and there was also a person with each child to guide them and to

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 4.2: Multifactorial ANOVA for the variable of satisfaction. N = 79.

clarify the possible doubts during the whole activity. The questionnaires were filled out in the same room where the children played. The children were encouraged to answer all questions without any pressure to avoid the influences of answering the pretest on the results of the posttest. If they did not know the answer it was considered to be completely normal, the children were not informed whether their answers on the pretest were correct. Thus, the children did not acquire any knowledge by answering the pretest; they only learned during the game.

4.3 Results

The variable *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 of this variable 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 4.2). As we can see, there are p-values and small effect sizes, which means 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.

With regard to the introduction of AR in the game (question 10), the overall rate for the enjoyment of AR games was very high (mean 4.85 ± 0.39), and an ANOVA test showed that all the groups studied appreciated it in a similar manner without significant differences: grade ($F[1,71] = 0.09, p = 0.77, \eta_G^2 < 0.01$), gender ($F[1,71] = 0.90, p = 0.35, \eta_G^2 = 0.01$), and device ($F[1,71] = 0.58, p = 0.45, \eta_G^2 < 0.01$). 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.

We also measured how easy it was for the children to use the AR system (question 14). The global rating was high (mean 4.29 ± 0.83), which means that children found it easy to play AR games, and all the groups experienced it 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 4.3, 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.

A Chi-squared test for question 20 revealed that the preference for the iPhone or the Tablet PC significantly differed between children who finished playing with one of the devices ($\chi^2[1,79] = 12.08, p < 0.01,$ Cramer's $V = 0.42$). After analyzing the results, we could see that children tended

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 4.3: Mixed design ANOVA for the ease of use of the AR system. N = 79.

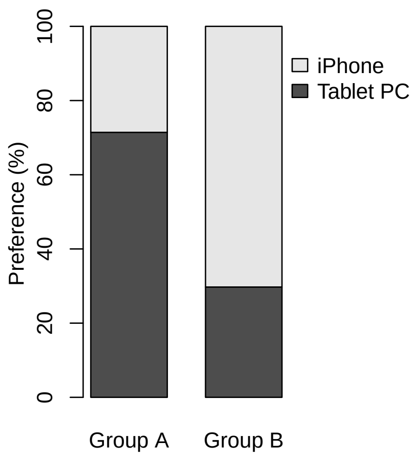

Figure 4.2: Preferences for the favourite device in groups A and B.

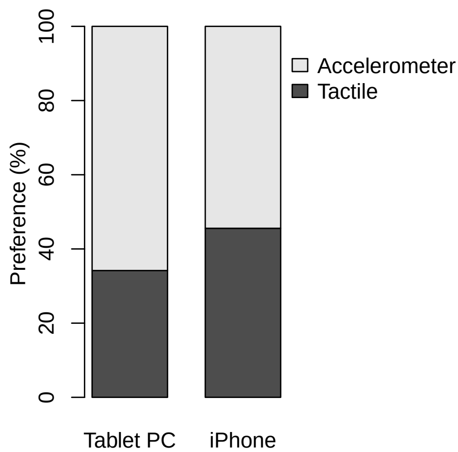

Figure 4.3: Preferences for the easiest technology after playing each device.

to choose the device they had used the last time. This difference is shown graphically in Figure 4.2, where 71.43% of the children in group A chose the Tablet PC, and 70.27% of the children in group B chose the iPhone. Another test revealed that there were not significant differences in this tendency for the devices ($\chi^2[1,79] = 0.05, p = 0.83,$ 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 Tablet PC and the iPhone. Thus, we could conclude that there was no significant difference in the preferences for the Tablet PC and the 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 we can see in Figure 4.3. There were significant differences in the proportions when children played the Tablet PC ($\chi^2[1,79] = 14.6, p < 0.001, h = 0.64$), but not after playing the iPhone ($\chi^2[1,79] = 0.911, p = 0.34, h = 0.176$). Comparing both devices, we found no significant differences in the proportions of the preferred interaction on the two devices ($\chi^2[1,79] = 1.69, p = 0.19,$ 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 4.4). There were no significant differences found

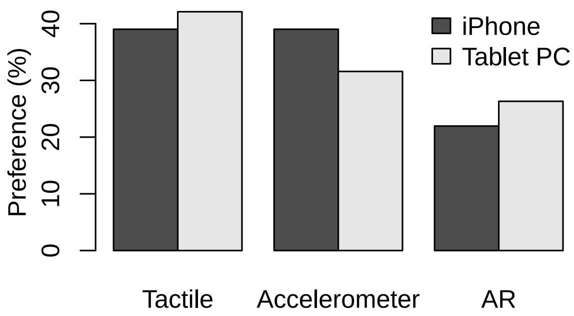

Figure 4.4: Percentages of preference of the favorite interaction for the two devices used.

in these three types of interaction when people used the Tablet PC and the iPhone ($\chi^2[2,79] = 0.51, p = 0.77$, Cramer's $V = 0.08$). In spite of the differences in the proportions, there were no significant differences in the proportions for the iPhone ($\chi^2[2,79] = 3.59, p = 0.17$) or the Tablet PC ($\chi^2[2,79] = 2.21, p = 0.33$).

4.4 Conclusions

Nowadays, children are used to playing with consoles and mobile devices (Beck and Wade, 2006; Prensky, 2003) such as their parents' mobile phones; in many cases the interaction of the devices was not new. While most of the children thought that it was 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 transmitting knowledge to the children. The study indicated that, despite the fact that the iPhone and the Tablet PC had different characteristics (screen size and weight), both devices were well suited for educational games. This is encouraging because it shows that children could adapt well 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.

When asked what 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 Tablet PC. A possible explanation is the fact that the Tablet PC was heavier than the iPhone and some children experienced some difficulties touching the screen with one hand while holding it with the other hand at the same time, but it depended mainly on the constitution of the child. Most of them rapidly found strategies to make the experience more comfortable by grabbing the Tablet PC with the whole arm, but some of them did not find it intuitive. This problem did not affect the iPhone, which was lighter than the 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 iPhone and the 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.

There are several paths to follow for future work. In this study we have compared two devices, but other comparisons are also possible; for example, using a control group in which the children learn the subject by 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). Also, it seems very likely that the testing with older children would decrease the noise in data due to the weight of the Tablet PC and the results would be more radically different, obtaining more significant differences in the analyses.

Chapter 5

Study 3: Tactile and tangible interfaces

This chapter presents a comparative study between tangible user interfaces (TUIs) and tactile user interfaces (TacUIs) in handheld AR, with a contribution to the state of the art in HCI oriented to education. While TUIs work with the manipulation of physical objects, TacUIs work with virtual representations of them. In our evaluations to compare these two interactions with primary school children, we found that the TacUI was the fastest for completing the task. The TacUI was found easier to use by the children, although the TUI was found more solid and less slippery. Our conclusions should be of interest not only to educational researchers, but also to the general HCI community working on tangible and tactile interfaces. Furthermore, we provide a statistical analysis using the effect size generalized eta-squared η^2_G, which has a great interest as it allows the comparison among many different researches. We encourage future researches to include it as well.

5.1 Introduction

Augmented Reality is a field with a great potential in education, providing a very stimulating environment that not only helps visualizing 3D objects, but also enhances highly the motivation and enjoyability of the students. AR has evolved fast in the recent years, specially concerning to the area of Human-Computer Interaction (HCI), and designing user interfaces that help the students visualize and explore AR environments.

Tangible User Interfaces (TUIs) deal with the manipulation physical objects. The hardware used to work with TUIs and AR in education is very frequently based on markers printed on books, such as the well-known MagicBook (Billinghurst et al., 2001), or on other objects, such as cards (Juan et al., 2011a). These systems have the great advantage of working with a video see-through metaphor that provides a high sense of presence and immersion with a direct view from the user perspective. However, they frequently use Head Mounted Displays (HMD) or devices built ad-hoc, what is an inconvenience for their use in the classroom, where only non-specialized equipment should be used (Inoue, 2007).

Tactile User Interfaces (TacUIs), in opposition to TUIs, do not deal with physical objects, but with metaphors or pictures of them displayed on a sensible surface, and therefore the differences in the user interaction and perception may be drastically different. TacUIs can be divided in tactile screens (capacitive or resistive) and projection based (Jones et al., 2010). As we will shortly describe, we will use handheld devices in this research. Consequently, we contextualize this work under tactile screens.

The introduction of powerful handheld devices such as mobile phones with graphical capabilities and PDAs has allowed a new exciting research in education and edutainment with AR (Billinghurst and Henrysson, 2006). In the recent years, smartphones and Tablet PCs with plenty of sensors that allow multimodal interfaces have popularized AR allowing video see-through interaction without the annoyances of a HMD. In this context, TacUIs play an important role in handheld interaction, as they contribute to AR providing a new interaction channel. These devices, specially Tablet PCs, are excellent candidates to replace desktop PCs in schools thanks to their popularity and for being affordable, multipurpose and widespread.

The relation between TUIs and TacUIs in handheld AR has not been examined previously to our knowledge. In this work we present a novel study that compares these two different forms of user interaction. Based on related work studied, we developed an experiment based on one of the most common uses of AR in the schools consisting of visualizing 3D objects from different points of view. We measured and analyzed the differences from an objective perspective (time spent by the user to interact with the system) and from the user's subjective perspective (satisfaction with the system and the interaction). We tested our experiment with 51 children from 8 to 10 years old. The conclusions of this work should be of interest not only to educational researchers trying to search better forms of interaction for their students, but also to the general HCI community that work in tangible or tactile interfaces.

To the best of our knowledge, this is the first HCI work that presents the effect size η^2_G (Olejnik and Algina, 2003) in the statistical analysis, a measure to enhance ANOVA analyses contributing in a better understanding of data. In addition, this measure is very appropriate to compare a wide variety of studies and statistical designs, including repeated measures. The ability of comparison is very limited in other effect size measures. We believe that the calculation of effect sizes should be a common practice in HCI, and we encourage researchers to use η^2_G in their future analysis, what could result in higher uniformity to compare studies.

5.2 Experiment

This section explains the experiment created to compare TUIs and TacUIs with handheld devices and the apparatus used and created for this purpose.

5.2.1 Description of the experiment

The experiment we implemented was useful to compare a TacUI and a TUI under very similar conditions. Since it was aimed to children, we developed a game that made necessary to use the two interactions to achieve the same objective. We designed the game paying a special consideration the related educational work so that the interaction metaphors were similar to current researches and the conclusions of this work may have a direct impact on future improvements.

Therefore, in our game the handheld device is used in a video see-through configuration to visualize a marker with an augmented object on it. In order to perform a fair comparative we decided to modify the AR registration system, as we can see in Figure 5.1. In terms of geometric operations, the translation of the object is determined by the AR system in the traditional way, but the rotation on the Z axis is determined dynamically by the TUI or the TacUI.

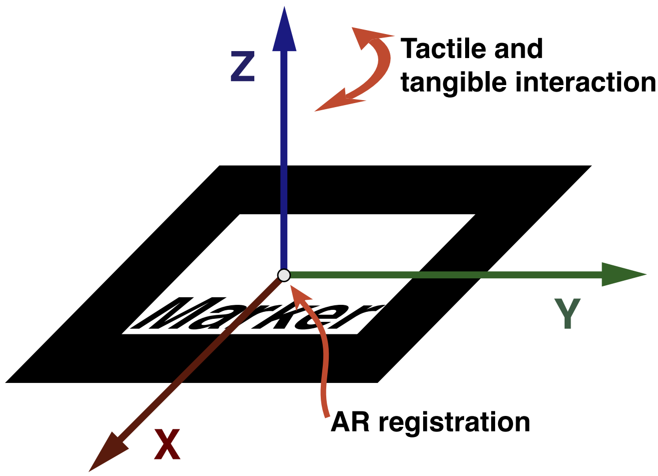

Figure 5.1: Modifications of the AR system to work with TUI and TacUI.

The game can be played in two modes. In the tangible mode the child only uses the handheld device merely as a visualization system, and she can control a TUI device to rotate the marker. We built this interface that allowed easily the rotation of the marker but not the translation. This avoids the possible distraction of holding the marker with the hands, and the user can focus on the task she has to do. In the tactile mode, the marker was static, and the interaction was only through the tactile screen. The user can touch and move the finger horizontally in the screen, but vertical displacements do not have any effect. The handheld device was in a fixed position using a stand to have a better comparison between the TUI and the TacUI, and the manipulation of the device did not interfere on the time spent using the interface.

The AR system placed a model of the Taj Mahal with several objects in different parts of the virtual scene located in such a way that it was necessary to rotate the object to see the objects hidden in the back. The objective of the game was to count the number of objects in the scene. Since children were studying nature and water in other activities, the objects were drops of water represented as characters. We also added paper with grass printed to make the experience more appealing to the children when they saw the video see-through. The setup of the game can be seen in Figure 5.2.

5.2.2 Apparatus

We used a commercial handheld device for the TacUI to use multipurpose, non-specialized hardware easily available by schools. The hardware for the TUI was aimed to compare the two interactions in very similar terms.

Hardware for the TacUI. We used the handheld device *HP Slate 500 Tablet PC*. It has a dimensions of 23x15x1.5 cm, and a weight of 0.68 kg. The processor works at 1.86 GHz, and it has a memory of 2 GB. The tactile screen is capacitive and has a diagonal of 8.9". For the AR we used the 3 Megapixel camera. We installed a case to protect the device during the evaluations, but it is completely optional and had no effect on the study, since the device was mounted on a stand that fixed its position and it was not necessary to hold it.

Figure 5.2: Set used for the evaluation consisting on a TUI and a TacUI in the handheld device.

Hardware for the TUI. We built a home-made plastic rotatory base resembling to a gearwheel with four teeth that could be manipulated to rotate it comfortably. the marker was placed in the centre.

Software. We decided to use OpenSceneGraph (OSG) toolkit 2.9.5 to develop the system to take advantage of its high capabilities to import, animate and render 3D objects with high performance in C++ language. The AR registration was achieved with the OSG plugin osgART 2.0 RC 3, that used the ARToolKit library (Kato and Billinghurst, 1999), version 2.72.1. osgART provided simple access to the camera, and to certain OSG nodes that applied the corresponding transformation matrices associated to markers when they were recognized.

5.3 Design of the evaluations

The developed game described in Section 5.2 was tested by a group of children. This section details the participants, the measurements and the procedure designed for the evaluations.

5.3.1 Participants

A total number of 51 children from 8 to 10 years old –with a mean age of 8.65 ± 0.74– took part in the study. The gender distribution was: 29 boys (57%), and 22 girls (43%). (More information about the participants will be provided after the anonymous review).

5.3.2 Measurements

Each child tested the two types of interaction (TacUI and TUI) so that they could compare them at the end. This leads to paired samples and repeated measures in the statistical analysis. In each game the time to achieve the task was measured, and the child had to fill in a questionnaire afterwards. Consequently, we used four different questionnaires (QTc1, QTc2, QTn1, and QTn2; *Tc* stands

#	TUI		TacUI		Question
	QTn1	QTn2	QTc1	QTc2	
1	•	•	•	•	Did you enjoy playing the game?
2	•		•		Did you like how the Taj Mahal appeared on the black square?
3	•		•		Would you like to play again to search other objects? [Y/N/M]
4	•	•	•	•	How easy did you find the game to play?
5	•	•			Would you like to use the wheel control in other games?
6	•	•	•	•	Please, rate the game.
7		•		•	What did you find the easiest? [Tactile/Tangible]
8		•		•	What did you like the most? [Tactile/Tangible]
9		•		•	Why? (referring to #7 and #8)

Table 5.1: Numbered questions and their appearance in each questionnaire. The last question is hand written. Answers in brackets are a summary of the possible choices (categorical data). The rest of the answers follow a Likert scale.

for tactile and *Tn* for tangible) during the evaluations, which contained the proper questions to be asked after using each type of interaction in first or second place. However, each child had to fill in only two questionnaires, one for each interaction. We can see the list of all the questions and the organization of the questionnaires in Table 5.1.

The questionnaires for the first game, QTc1 and QTn1, are similar. Both have satisfaction measures, and ask about AR and about the interaction method. The questionnaires of the games played in second place, QTc2 and QTn2, include questions to compare the interaction methods. Some of the questions follow a Likert scale presented as: *a) Very much; b) Quite a lot; c) Somewhat; d) Few; e) Nothing*, with a numerical equivalency linearly ranged from 5-a) to 1-e). Other questions have categorical answers, as they ask to choose among several options (#3, #7 and #8). The last question is hand written and gives the children an opportunity to express their opinion about the game openly.

5.3.3 Procedure

The children were divided in two groups depending on the order in which they would play with both interfaces:

Group A) Children who used the TacUI first and then the TUI.

Group B) Children who used the TUI first and then the TacUI.

The children were randomly assigned to the groups and counterbalanced: 26 children were assigned to group A and 25 to group B. During the evaluations, only one child was playing at a time. She played two times to the same game of counting objects, changing the interaction form. The time to complete the task was recorded for posterior analysis. After playing each game, the child answered a questionnaire. We can see a graphic that summarizes the process in Figure 5.3. A member of our staff guided the child during the process and explained her the tasks to do. After the measures, all data was analyzed sing the open source statistical toolkit R[1].

[1] http://www.r-project.org/

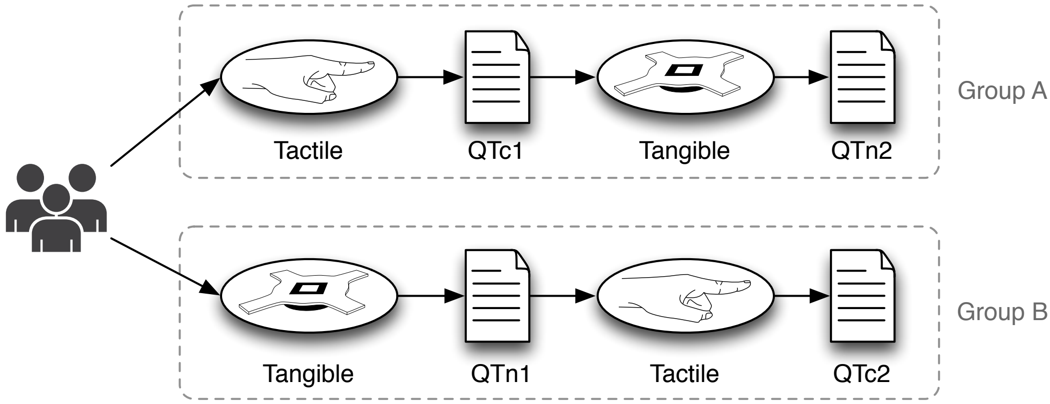

Figure 5.3: Questionnaires and games that children played in groups A and B.

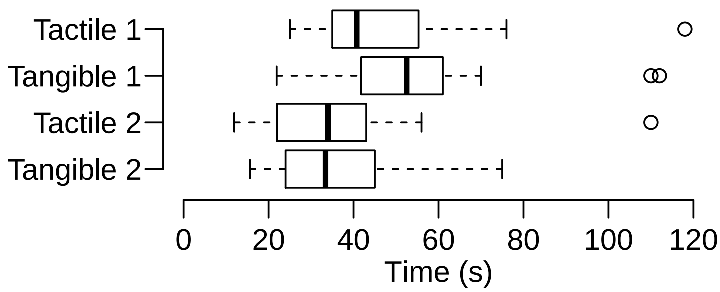

Figure 5.4: Box plot for the time to complete the task in the four combinations of tactile and tangible interfaces played the first and the second time.

5.4 Results

The first step to analyze the results was the identification the possible outliers. From a visual inspection at the box plot of the time to complete the task, shown in Figure 5.4, we detected four outlier points. These outliers correspond to three different participants, one of whom scored a very high value with the two interfaces he played. These participants were two boys and one girl, and all of them were 8 years old. To avoid possible misleading conclusions, their times and questionnaires answered were completely excluded from the analysis.

5.4.1 Time analysis

The time to complete the task of counting the objects in the scene was measured and studied as an independent variable. Comparing this variable in the first interaction that the two groups used, we found that the time spent when using the TacUI (mean 43.21 ± 13.12) was statistically significantly lower than when using the TUI (mean 50.30 ± 13.16), as an unpaired t-test revealed ($t[46] = 1.83, p = 0.04,$ Cohen's $d = 0.53$). In this analysis we follow the guidelines of Cohen to interpret the effect size, which state the approximate meaning of d: 0.2 small effect size, 0.5 medium effect size, and 0.8 large effect size (Cohen, 1988, pp. 25–26). Therefore, the interaction factor had

Factor	d.f.	F	p	Effect size (η^2_G)
Interaction	1	4.94	0.03**	0.04
Order	1	33.13	<0.01***	0.20
Age	2	0.23	0.79	<0.01
Gender	1	0.49	0.49	<0.01
Group	1	0.09	0.76	<0.01

Table 5.2: Summary of the mixed design ANOVA analysis for the time to finish the task. *$p < 0.1$; **$p < 0.5$; ***$p < 0.01$. N = 48.

a medium effect on the time the children spent to complete the task when they played to it for the first time. A medium effect size was interpreted by Cohen as a magnitude perceivable by the "naked eye". Consequently, we can conclude that the TacUI was significantly faster to use than the TUI, not having in consideration any other factor.

We also performed a mixed-design ANOVA analysis to take into account the two interactions the same child used, and to consider several factors simultaneously. In a mixed-design ANOVA, some factors are between-subjects (i.e. all of the measures come from different subjects or it is an observed factor like the age or gender) and other factors are within-subjects (i.e. some measures are repeated from the same subject). In our analysis we took into account the factors of age, gender, group (A or B), interaction (TUI or TacUI) and order (1st game or 2nd game). Therefore, the factors of age, gender and group were between-subjects, and the factors of interaction and order were within-subjects.

To perform the analysis we did two different mixed-design ANOVAs, one with the interaction factor as within-subject and the age, gender and group factors as between-subjects, and another one with the order factor as within-subject and the age, gender and group factors as between-subjects. This separation in two different tests is due to the strong relation between the interaction factor and the order factor. These two factors contain the same subjects but classify them according to different criteria, and since the mixed-design ANOVA needs that each subject has a measure of all combinations of the levels of the within factors, it is not possible to use the two factors simultaneously, but they can be used independently.

Nonetheless, thanks to the design of our variables, this separation has no effects in our analysis because the two within-subjects factors classify the subjects independently from the between-subjects factors. Thus, the two mixed-design ANOVAs obtain the same results for the factors of age, gender and group, and this is the reason why we present the results of the two analysis in the same table that summarizes the two statistical tests (Table 5.2). In this table we can see that there are significant differences in the order and interaction factors. As both factors have two levels only, a simple post-hoc analysis comparing the mean of each level revealed that the users took less time to complete the task using the TacUI than using the TUI with a significant difference. Similarly, we concluded that the users took more time to complete the task the first time they played than the second time. No significant differences were found for the age, gender and group factors.

To enhance our analysis with a measure independent from the sample size, we also calculated the effect size used using the generalized eta squared, η^2_G (Olejnik and Algina, 2003). This effect size measure has been proven to be more suitable for mixed-design analyses than other traditional alternatives such as the eta squared, η^2, or the partial eta squared, η^2_P, because it can consider the repeated measures of some subjects along the within-subjects factors. Olejnik and Algina stated that, unlike η^2 and η^2_P, η^2_G is a measure consistent across a wide variety of different statistical designs.

#	Factor	d.f.	F	p	Effect size (η_G^2)
1	Interaction	1	0.19	0.67	<0.01
4	Interaction	1	0.10	0.75	<0.01
6	Interaction	1	2.08	0.16	<0.01
1	Order	1	2.96	0.09*	0.02
4	Order	1	0.00	1.00	<0.01
6	Order	1	1.41	0.24	<0.01

Table 5.3: Summary of the mixed design ANOVA analysis for the common questions. *$p < 0.1$; **$p < 0.5$; ***$p < 0.01$. N = 48.

Bakeman also stated that it seemed appropriate to use Cohen's guidelines for η^2 for multi-factor analysis: 0.02 small effect size, 0.13 medium effect size and 0.26 large effect size (Bakeman, 2005). These guidelines for multi-factor analysis (Cohen, 1988, pp. 413–414) should not be confused with the guidelines for one-way analysis (Cohen, 1988, pp. 285–288), which have lower values.

Unfortunately, it is very infrequent to find effect sizes reported by researchers in the field of HCI. To our knowledge this is the first study in the field to report an analysis using η_G^2, and it was not possible for us to compare our study with previous results. Our results in Table 5.2 show a small to medium effect size for the interaction factor, which is consistent with the previous result that analyzed the same factor only for the first game played, revealing that the second game introduced noise in the analysis. Nevertheless, a small effect size "in practice represents the true order of magnitude of the effect being tested" (Cohen, 1988, pp. 413–414). There is also a medium to large effect size for the order factor, revealing an important learning from the first task to the second task, which was an expected effect in this experiment.

5.4.2 Satisfaction

The answers from the questionnaires were also studied. Table 5.3 shows the results for the interaction and order factors (other factors were also studied but are not shown since no relevant results were found) for questions 1 (enjoyability), 4 (ease of play) and 6 (global game rate). All of these questions have very high scores: #1 4.60 ± 0.64, #4 4.58 ± 0.66 and #6 4.64 ± 0.49. Moreover, there are no significant differences in questions 4 and 6 for the studied factors. The only arguable point is in question 1, where there is a relatively high p-value and a small effect size. A post-hoc analysis reveals that the possible difference would mean that the second game was more enjoyed than the first game. We believe that the difference is meaningful and not only caused partially by chance, but also by the order factor, as we can see from the answers to the question 3. In this question, 93% of the children said they would like to play again after the first game, and we could see during the evaluations that they were very excited to play the second game, what in our opinion has affected the results in question 1.

The analysis of the interaction factor in question 3 shows that all of the 25 children who played the TacUI first would like to play again, while 20 children would also like to play again after the TUI, and 3 were undecided. However, here are no evidences of significant differences over the interaction factor ($p = 0.10$, Fisher's exact test), which means that both interactions engaged the children highly.

The children appreciated the AR very much, as the high scores to the question 2 reveal for

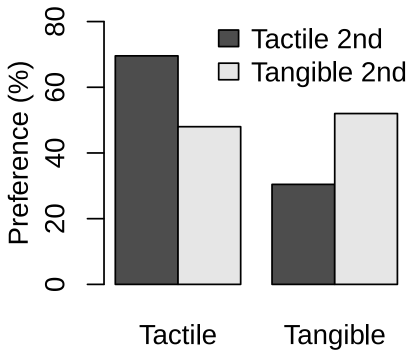

Figure 5.5: Percentages of selection for the easiest interaction (question 7).

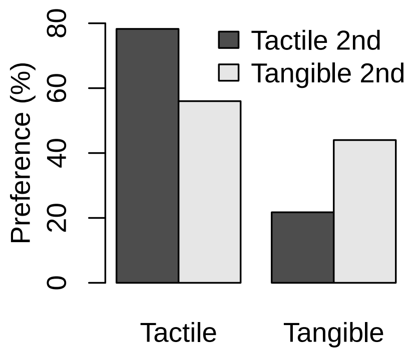

Figure 5.6: Percentages of selection for the preferred interaction (question 8).

the tactile game (mean 4.80 ± 0.40) and for the tangible game (mean 4.91 ± 0.28). It is very interesting to note that no significant differences were found between both games ($t[46] = 1.10, p = 0.28$, Cohen's $d = 0.32$), which is an excellent result because it means that the use of AR was very similarly enjoyed with both interactions.

Since the TUI was built ad-hoc, we asked the children if they would like to use the tangible controller in other games (question 5) as a measure of their attraction and evaluation. The scores to this question were very high when the TUI was the first interaction (mean 4.65 ± 0.70) or the second one (mean 4.84 ± 0.37). No significant differences were found ($t[46] = 1.15, p = 0.25$, Cohen's $d = 0.33$), from what we can deduce that the TacUI did not influence the appreciation of the tangible game.

5.4.3 Interaction

We also compared the two interactions studied in this work from the children's point of view after having tested both of them. The ease of use and preferable form of interaction is analyzed next.

Question 7 asked the children to choose the type of interaction they thought it was easier to use. As we can see in Figure 5.5, there is the effect that children chose more the last game they used, which is a common tendency that we have experienced in previous studies with children. When the TacUI was the last one to be used there was a 70% of children who preferred it to the TUI, but when the TacUI was the last one to be played, the TUI preferences dropped to 48%. When the children used the TacUI in second place, there were significant differences in the proportions of TacUI and TUI ($\chi^2[1,48] = 5.57, p = 0.02, h = 0.80$) with a large effect size (Cohen, 1988, pp. 184–185). This result evinces the preference of the TacUI. The difference is, however, no longer significant when the TUI was used last ($\chi^2[1,48] < 0.01, p = 1.00, h = 0.08$). From the difference in the significances we can conclude that the tactile screen caused a more powerful impact on the children.

The preference for the two types of interaction was measured in question 8. Figure 5.6 shows graphically the results in percentages. It is apparent that the situation from question 7 is present in this case as well. The most extreme proportions are found in the tactile game played the second time, where the tactile game was chosen by the 78% of the children, and the tangible game by the 22%. In such case there are significant differences and a very large effect size ($\chi^2[1,48] = 12.52, p < 0.01, h = 1.20$). When the tangible game was played last, the differences are not significant, but there is a medium effect size, revealing more apparent differences than in question 7 ($\chi^2[1,48] = 1.76, p = 0.18, h = 0.51$).

The reasons given by the children in question 9 on why to prefer the TacUI can be catalogued in the next:

- “Because it is easier”
- “Because it is faster to touch the screen”
- “Because it is funnier”
- “Because it is more comfortable”

The reasons given by the children that preferred TUI can be catalogued in the next:

- “Because it is slower to turn the wheel”
- “Because my fingers slipped in the screen”
- “Because it was the first time I used something like that”

Given the answers, it seems that children of these ages are not very used to tactile screens, and some of them found it a little difficult to control their actions in the screen. In addition, we saw that some of them tended to touch the screen with the fingernails, which was unfruitful since the screen was capacitive. The TUI offered them a more solid interface that they could grab. Despite this, most of the children preferred the TacUI.

5.5 Conclusions

We have presented a study that compares TacUIs and TUIs in handheld AR. Our study was centered in the field of education, but many conclusions can be extrapolated to other areas of HCI. To our knowledge, this is the first study that analyses these two interactions and compares them in this way.

We designed an system to compare TacUIs and TUIs under very similar circumstances. The handheld device was used in a video see-through mode to visualize the augmented world, taking some of the advantages of HMDs, but resulting in a much more comfortable way to the children. This metaphor was chosen for being one of the most used in education. Our proposal is very well suited for schools, since handheld devices are affordable and multipurpose, so they can replace desktop PCs and enhance their value contributing to create richer multimedia experiences for the students.

In the experiment we evaluated, the TUI consisted on a rotatory base on top of which the AR marker was placed. The children could rotate the marker to see the 3D objects from different perspectives. This interaction was compared to the TacUI, where the rotation was calculated from the horizontal movements of the finger on screen. We firstly measured the differences between the two interactions from an objective perspective, analyzing the differences on the time the users spent to achieve the objective of counting the hidden objects on the scene. The tests indicate that it took significantly less time to the children to use the TacUI than the TUI.

As for the children’s perspective, we could see a very high level of satisfaction with the game using the two interfaces, and no significative differences were found. The children were very engaged with the game. All of them wanted to play again after using the TacUI, and most of them after using the TUI. However, there was no evidence that the interface used influenced this engagement. The AR system appealed the children very much without being affected by the interaction method.

The children seemed to prefer the TacUI mainly for being easier and faster to use, and some of them also thought it was more comfortable. On the contrary, the TUI was preferred in some cases, surprisingly for very similar reasons. These children found the screen too fast to use and too slippery, and they preferred the more stable interface that the TUI was. In general, more children preferred the TacUI, although this difference was not always statistically sustainable. Despite preferring the TacUI, the appreciation of the TUI was very high, so we can discard that the preference of TacUI is due only to disliking the alternative. Finally, it is very possible that in older ages we would find a high increase on preference of the TacUI, as the children's physical skills improve and do not have troubles with the screen.

In our statistical analysis we have used the effect size η^2_G, which provides a better understanding of the meaning of the significative differences that may be obtained analyzing the p-values in ANOVA analyses. We encourage researchers to use this measure, that could bring more uniformity to the studies, making them more easily comparable.

With regard to future work, we think we have now the basis to open the study to more sophisticated interactions. We find suggestive to investigate the possibilities in terms of interaction of using handheld devices more dynamically in activities that require more movement and holding the devices with the hands.

Chapter 6

Concluding remarks

6.1 Conclusions

In this thesis we have carried out a research about the implications of handheld devices and AR in an educational context and the consequences in HCI. Through three studies we have fulfilled the principal and the secondary objectives.

A playful activity was developed, themed in natural sciences, concretely in the water cycle, composition and pollution. The design was carried out following the theories of Gardner and Kolb as the educational background with great success. The motivation of children was very high, and they learned much while playing. We programmed the game and evaluated it in handheld devices, making use of accelerometers, the tactile screen, and the camera, thus providing a multimodal interface that appealed to the children. The game had AR mini-games and non-AR mini-games as specified, and the consequences of using a MRLE or a VRLE were studied concluding that the introduction of AR was an important step for current VRLEs used in the classroom. Also, two different handheld devices were tested and compared when children played the game with them, and we saw that both of them were similarly effective. Furthermore, a study that compared tactile and tangible interaction was presented, providing data to support that it could be preferable to use tactile interfaces in some situations. All the studies presented a decent number of subjects and several measures were taken, such as learning outcomes, satisfaction and interaction. Finally, we presented a statistical analysis of the data captured and provided a critical discussion.

In the first study we compared a MR version of the game and a VR version, in a learning environment. The children who tested the two systems learned much, but in a similar amount. This evinces that technology itself is not enough to improve education. We saw our hypothesis accomplished when we saw the great results in motivation. The main conclusion we can get is that AR can improve VRLE, since it enhances the learning experience, but it does not imply a radical change in education.

We also tested the game in the second study comparing two handheld devices with different sizes and weights. Despite our initial hypothesis, the Tablet PC would be more engaging than the smartphone, and the interaction would be easier, we found that in our experience, both devices were similarly appreciated. We perceived in children an extraordinary ability to adapt to any kind of environment.

In the third study, we compared Tactile User Interfaces (TacUIs) with a Tablet PC and Tangible User Interfaces (TUIs) with an interface built *ad hoc*. The interface represented a very common setup used in AR, which is to show 3D objects so that the student can explore it to his will. This traditional setup is based on TUIs, and we showed that TacUIs can be a better alternative to consider. However, the technical details of the device in relation to the age of the students should be of further research, since we found some children with problems using the TacUI.

From a global perspective, we can see that handheld devices have a great potential in educational environments. Given our experience in the three studies, an interesting conclusion that can be reached is that providing a multimodal interface to the children seems to be favorable for engagement.

6.2 Scientific contributions

There is a number of publications that have been produced as a result of the work for this thesis, which are shown next:

1. **González-Gancedo, S.**, Juan, M.-C., Seguí, I., Rando, N., and Cano, J. (2012). Towards a mixed reality learning environment in the classroom. In *International Conference on Computer Graphics Theory and Applications*, GRAPP '12, pages 434–439, Rome, Italy. SciTePress.

    - Conference ranking: CORE/ERA A

2. **González-Gancedo, S.**, Juan, M.-C., Seguí, I., Abad, F. (2012). Tactile and tangible interfaces in handheld AR for children. In *Proceedings of the 26th BCS Conference on Human-Computer Interaction*, BCS-HCI '12, Birmingham, United Kingdom. British Computer Society.

    - Conference ranking: CORE/ERA A
    - This paper was sent, and is still under review.

3. Furió, D., **González-Gancedo, S.**, Juan, M.-C., Seguí, I., Costa, M. and Cano, J. Learning the water cycle using Augmented Reality and mobile devices. Evaluation of learning outcomes using an iPhone and a Tablet PC. *Computers & Education*.

    - Journal ranking: JCR-SCI, Impact Factor: 2.617. (first quartile under COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS)
    - This paper is already under the second review.

6.3 Future work

The first step that we plan to do is to test the game in a real school and compare the learning outcomes with the traditional way of teaching it. This could lead to interesting conclusions mainly in the educational aspect of the system rather than in HCI, and could also help us to establish design guidelines for the content.

Since one of the many objectives of education is to remember things, it should be proved that a system like ours is effective in the long term. We have studied the short term memory, since the pretest and the posttest were close in time, and testing the knowledge several months later would be a more powerful indicative of success. However this would be a challenge because the environment would not be so controlled, and probably the sample size should be larger. In this line our results are encouraging because children were highly motivated, and there are some researches that showed that motivation can stimulate memorization not in the short term, but in the long term (Yang et al., 2010).

Another very interesting research following Gardner's theory could be using the multimodal interface of handheld devices to provide a personalized learning experience that allowed children learn a subject using the form of intelligence they preferred. We have seen that many children liked using the accelerometer, where they had to move themselves, but it is still necessary a deep study to see the plausibility of this research.

Bibliography

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., and Rauterberg, M. (2002). Interaction techniques for navigation through and manipulation of 2d and 3d data. In *Proceedings of the workshop on Virtual environments 2002*, pages 179–188. Eurographics Association.

Ardito, C., Costabile, M. F., Lanzilotti, R., and Pederson, T. (2007). Making dead history come alive through mobile game-play. In *CHI '07 extended abstracts on Human factors in computing systems*, CHI EA '07, pages 2249–2254, New York, NY, USA. ACM.

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6):34–47.

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

Balog, A., Pribeanu, C., and Iordache, D. (2007). Augmented reality in schools: Preliminary evaluation results from a summer school. In *Proceedings of the World Academy of Science, Engineering and Technology*, ICTE 2007, pages 114–117, Nice, France.

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

Billinghurst, M. (2002). Augmented reality in education. *New Horizons for Learning*, 12.

Billinghurst, M., Kato, H., and Poupyrev, I. (2001). The magicbook-moving seamlessly between reality and virtuality. *Computer Graphics and Applications, IEEE*, 21(3):6–8.

Billinghurst, M. N. and Henrysson, A. (2006). Research Directions in Handheld AR. *Int Journal of Virtual Reality*, 5(2):251–58.

Bimber, O., Fröhlich, B., Schmalstieg, D., and Encarnação, L. M. (2001). The virtual showcase. *IEEE Computer Graphics & Applications*, 21(6):48–55.

Blecic, I., Cecchini, A., Rizzi, P., and Trunfio, G. (2002). Playing with automata. an innovative perspective for gaming simulation. In Bandini, S., Chopard, B., and Tomassini, M., editors, *Cellular Automata*, volume 2493 of *Lecture Notes in Computer Science*, pages 337–348. Springer.

BOE (2007). ORDER ECI/2211/2007, which establishes the curricula and regulates the Primary School order. *BOE*, 173:31487–31566.

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

Chen, Y. (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*, pages 369–372. ACM.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum, 2nd edition.

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

De Freitas, S. and Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education*, 46(3):249–264.

Dickey, M. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development*, 53(2):67–83.

Economides, A. and Nikolaou, N. (2008). Evaluation of handheld devices for mobile learning. *International Journal of Engineering Education*, 24(1):3.

Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R., and Kirk, D. (2004). Savannah: mobile gaming and learning? *Journal of Computer Assisted Learning*, 20(6):399–409.

Fisch, S. M. (2005). Making educational computer games “educational”. In *Proceedings of the 2005 conference on Interaction design and children*, IDC ’05, pages 56–61, New York, NY, USA. ACM.

Fjeld, M., Fredriksson, J., Ejdestig, M., Duca, F., Bötschi, K., Voegtli, B., and Juchli, P. (2007). Tangible user interface for chemistry education: comparative evaluation and re-design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI ’07, pages 805–808, New York, NY, USA. ACM.

Fotouhi-Ghazvini, F., Earnshaw, R., Moeini, A., Robison, D., and Excell, P. (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.

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, pages 27–30, Swinton, UK, UK. British Computer Society.

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

Gardner, H. (2000). Can technology exploit our many ways of knowing? *The digital classroom: How technology is changing the way we teach and learn*, pages 32–35.

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., and Smrikarov, A. (2004). M-learning – a new stage of e-learning. In *International Conference on Computer Systems and Technologies – CompSysTech'2004*, volume 28, pages 1–5.

Henrysson, A. and Billinghurst, M. (2007). Using a mobile phone for 6 dof mesh editing. In *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*, volume 254 of *CHINZ '07*, pages 9–16. ACM.

Hsiao, K. (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.

Inoue, Y. (2007). Concepts, applications, and research of virtual reality learning environments. *International Journal of Human and Social Sciences*, 2(1):1–7.

Jones, B., Sodhi, R., Campbell, R., Garnett, G., and Bailey, B. (2010). Build your world and play in it: Interacting with surface particles on complex objects. In *9th IEEE International Symposium on Mixed and Augmented Reality*, ISMAR '10, pages 165–174. IEEE.

Jones, V. and Jo, J. (2004). Ubiquitous learning environment: An adaptive teaching system using ubiquitous technology. In *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference*, pages 468–474.

Juan, M. C., Carrizo, M., Giménez, M., and Abad, F. (2011a). Using an augmented reality game to find matching pairs. In *Proceedings of the 19th International Conference on Computer Graphics, Visualization and Computer Vision*, pages 59–66.

Juan, M. C., Furió, D., Alem, L., Ashworth, P., and Cano, J. (2011b). ARGreenet and BasicGreenet: Two mobile games for learning how to recycle. In *Proceedings of the 19th International Conference on Computer Graphics, Visualization and Computer Vision'2011*, pages 25–32.

Juan, M. C., Toffetti, G., Abad, F., and Cano, J. (2010). Tangible cubes used as the user interface in an augmented reality game for edutainment. In *Proceedings of the 2010 10th IEEE International Conference on Advanced Learning Technologies*, ICALT '10, pages 599–603, Washington, DC, USA. IEEE Computer Society.

Kato, H. and Billinghurst, M. (1999). Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality '99*, pages 85–94. IEEE.

Kaufmann, H. and Schmalstieg, D. (2003). Mathematics and geometry education with collaborative augmented reality. *Computers & Graphics*, 27(3):339–345.

Kaufmann, M. (2004). Geometry education with augmented reality. *PhD dissertation thesis*.

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

Ketamo, H. (2003). An adaptive geometry game for handheld devices. *Educational technology & society*, 6(1):83–95.

Kim, M. J. and Maher, M. L. (2008). The impact of tangible user interfaces on spatial cognition during collaborative design. *Design Studies*, 29(3):222–253.

Koh, R., Duh, H., and Gu, J. (2010). An integrated design flow in user interface and interaction for enhancing mobile ar gaming experiences. In *IEEE International Symposium On Mixed and Augmented Reality-Arts, Media, and Humanities (ISMAR-AMH)*, pages 47–52. IEEE.

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

Lai-Chong Law, E., Kickmeier, D., and Holzinger, A. (2008). Challenges in the development and evaluation of immersive digital educational games. *Lecture Notes in Computer Science*, 5298:19–30.

Larsen, L., Jensen, R., Jensen, K., and Larsen, S. (2005). Development of an automatic pool trainer. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ACE'05, pages 83–87. ACM.

Lee, E. A.-L., Wong, K., and Fung, C. (2009). Learning effectiveness in a desktop virtual reality-based learning environment. In *Proceedings of the 17th International Conference on Computers in Education, Asia-Pacific Society for Computers in Education, Hong Kong*.

Lee, E. A.-L., Wong, K., and Fung, C. (2010). How does desktop virtual reality enhance learning outcomes? a structural equation modeling approach. *Computers & Education*, 55(4):1424–1442.

Liarokapis, F. and Newman, R. (2007). Design experiences of multimodal mixed reality interfaces. In *Proceedings of the 25th annual ACM international conference on Design of communication*, pages 34–41. ACM.

Mahamad, S., Ibrahim, M. N., and 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.

Mcconatha, D., Praul, M., Chester, W., and 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. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12):1321–1329.

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

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

O'Neil, H., Wainess, R., and Baker, E. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal*, 16(4):455.

O'Shea, P., Mitchell, R., Johnston, C., and Dede, C. (2009). Lessons learned about designing augmented realities. *International Journal of Gaming and Computer-Mediated Simulations*, 1(1):1–15.

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

Pan, Z., Cheok, A., Yang, H., Zhu, J., and Shi, J. (2006). Virtual reality and mixed reality for virtual learning environments. *Computers & Graphics*, 30(1):20–28.

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

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

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

Sandor, C. and Klinker, G. (2005). A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality. *Personal and Ubiquitous Computing*, 9(3):169–185.

Sayed, N. A. E., Zayed, H. H., and Sharawy, M. I. (2011). ARSC: Augmented reality student card. *Computers & Education*, 56(4):1045 – 1061.

Schmalstieg, D. and Wagner, D. (2007). Experiences with handheld augmented reality. In *6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '07, pages 3–18.

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

Shaffer, D. W., Squire, K. T., Halverson, R., Gee, J. P., et al. (2004). Video games and the future of learning. *Phi delta kappan*, 87(2):104.

Sharlin, E., Itoh, Y., Watson, B., Kitamura, Y., Sutphen, S., and Liu, L. (2002). Cognitive cubes: a tangible user interface for cognitive assessment. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, CHI '02, pages 347–354, New York, NY, USA. ACM.

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

Shelton, B. and Hedley, N. (2002). Using augmented reality for teaching earth-sun relationships to undergraduate geography students. In *1st IEEE International Augmented Reality Toolkit Workshop*, page 8. IEEE.

Shelton, B. and Hedley, N. (2004). Exploring a cognitive basis for learning spatial relationships with augmented reality. *Technology, Instruction, Cognition and Learning*, 1(4):323–375.

Squire, K., Barnett, M., Grant, J. M., and Higginbotham, T. (2004). Electromagnetism supercharged!: learning physics with digital simulation games. In *Proceedings of the 6th international conference on Learning sciences*, ICLS '04, pages 513–520. International Society of the Learning Sciences.

Tan, K. T. W., Lewis, E. M., Avis, N. J., and 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, pages 1–8, New York, NY, USA. ACM.

Thornton, P. and Houser, C. (2005). Using mobile phones in english education in japan. *Journal of Computer Assisted Learning*, 21(3):217–228.

Ullmer, B. and Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Systems Journal*, 39(3.4):915 –931.

Uzunboylu, H., Cavus, N., and Ercag, E. (2009). Using mobile learning to increase environmental awareness. *Computers & Education*, 52(2):381–389.

Van Krevelen, D. and Poelman, R. (2010). A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, 9(2):1–20.

Veenema, S. and Gardner, H. (1996). Multimedia and multiple intelligences. *American Prospect*, 7(29):69–76.

Wagner, D., Schmalstieg, D., and Billinghurst, M. (2006). Handheld AR for Collaborative Edutainment. *Advances in Artificial Reality and TeleExistence*, pages 85–96.

Wang, K.-H., Chen, L.-C., Chu, P.-Y., and Cheng, Y.-M. (2009). A study on the design of augmented reality user interfaces for mobile learning systems in heritage temples. In *Proceedings of the 3rd International Conference on Virtual and Mixed Reality: Held as Part of HCI International 2009*, VMR '09, pages 282–290, Berlin, Heidelberg. Springer-Verlag.

Woods, E., Billinghurst, M., Looser, J., Aldridge, G., Brown, D., Garrie, B., and Nelles, C. (2004). Augmenting the science centre and museum experience. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, GRAPHITE '04, pages 230–236, New York, NY, USA. ACM.

Yang, J. C., Chen, C. H., and Jeng, M. C. (2010). Integrating video-capture virtual reality technology into a physically interactive learning environment for english learning. *Computers & Education*, 55(3):1346–1356.

Zhou, F., Duh, H., and Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, pages 193–202, Washington, DC, USA. IEEE Computer Society.